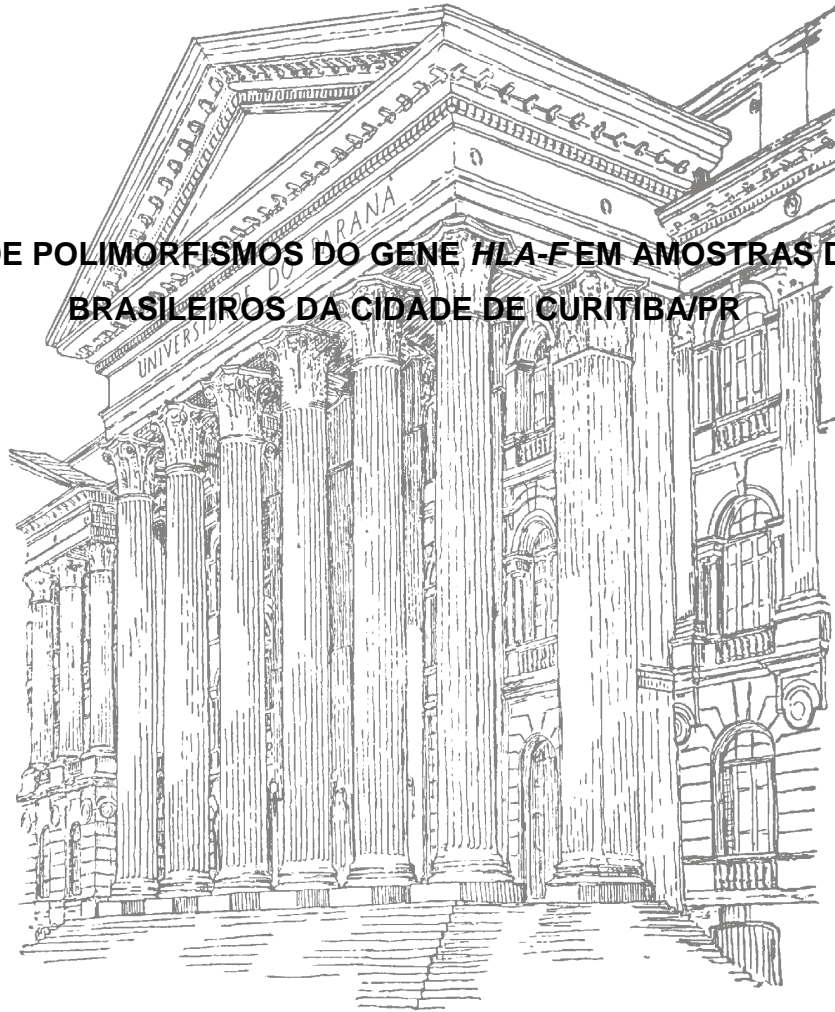


UNIVERSIDADE FEDERAL DO PARANÁ

LUIS FELIPE SANTOS MANVAILER

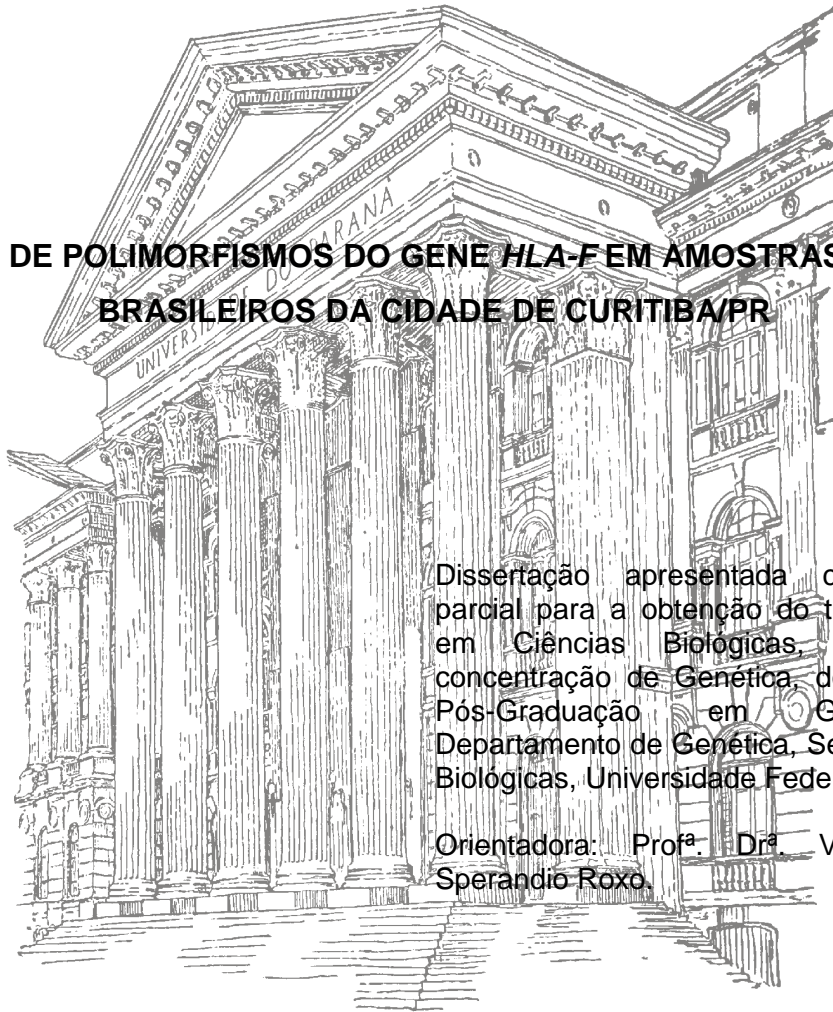
**ANÁLISE DE POLIMORFISMOS DO GENE *HLA-FEM* AMOSTRAS DE EURO-
BRASILEIROS DA CIDADE DE CURITIBA/PR**



CURITIBA
2012

LUIS FELIPE SANTOS MANVAILER

**ANÁLISE DE POLIMORFISMOS DO GENE *HLA-F* EM AMOSTRAS DE EURO-
BRASILEIROS DA CIDADE DE CURITIBA/PR**



Dissertação apresentada como requisito parcial para a obtenção do título de Mestre em Ciências Biológicas, na área de concentração de Genética, do Programa de Pós-Graduação em Genética, do Departamento de Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná.

Orientadora: Profa. Dra. Valéria M. M. Sperandio Roxo.

CURITIBA
2012



Ministério da Educação
UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA



PARECER

Os abaixo-assinados, membros da Banca Examinadora da defesa de dissertação de Mestrado, a qual se submeteu **LUIS FELIPE SANTOS MANVAILER**, para obtenção do título de Mestre em Genética pela Universidade Federal do Paraná, no Programa de Pós-Graduação em Genética, são de parecer que se confira ao candidato o conceito "A".

Secretaria da Coordenação do Programa de Pós-Graduação em Genética do Setor de Ciências Biológicas da Universidade Federal do Paraná.

Curitiba, 27 de fevereiro de 2012

Professora Doutora Maria Luiza Petzl-Erler
Dep. Genética/UFPR – Membro Titular

Professora Doutora Karin Braun-Prado
Dep. Patologia Básica/UFPR – Membro Titular

Professora Doutora Valéria Maria Munhoz Sperandio Roxo
Dep. Genética/UFPR – Orientadora e Presidente da Banca

Visto

Professor Doutor Ricardo Lehtonen Rodrigues de Souza
Coordenador do Programa de Pós-Graduação em Genética



AGRADECIMENTOS

Ao longo do desenvolvimento deste trabalho muitas pessoas estiveram envolvidas direta ou indiretamente e, indubitavelmente, sem o apoio destas a realização deste estudo passaria por caminhos mais tortuosos. No entanto a participação de algumas pessoas merecem um agradecimento especial por fornecerem subsídios decisivos para a realização desta dissertação, desde recursos materiais, financeiros, estendendo-se aos conhecimentos que foram passados adiante e ao apoio emocional. Primeiramente em especial à prof^a. Dr^a. Maria da Graça Bicalho por ter me aceitado como mestrando pela equipe de pesquisa LIGH. Não poderia deixar de agradecer à prof^a. Dr^a. Valéria Sperandio Roxo por me orientar, auxiliar e apoiar neste trabalho que realizamos em conjunto de forma muito harmoniosa. À prof^a. Dr^a. Maria Luiza Petzl-Erler e à prof^a. Dr^a. Karin Braun Prado por terem aceitado fazer parte da banca examinadora. Também agradeço aos colegas de laboratório que fazem parte da equipe de pesquisa e rotina, em especial à Dr^a. Priscilla Fanini Wowk e à Sibelle Botogosque Mattar que muito me auxiliaram desde o início até a conclusão da dissertação, sendo pessoas-chave para a logística e execução da prática laboratorial. Sem dúvida minha família exerceu papel fundamental ao longo desta caminhada, fornecendo crucial apoio nas minhas decisões e planos. É claro que não poderia deixar de agradecer a todos os momentos de felicidade passados juntos aos meus amigos, pessoas imprescindíveis em minha vida. E finalmente, gostaria de agradecer ao indispensável apoio financeiro da CAPES e da aliança LIGH-FUNPAR.

RESUMO

HLA-F é um gene pertencente ao complexo principal de histocompatibilidade (MHC) não clássico. Este gene codifica moléculas MHC de classe Ib com distribuição restrita e menos variações nucleotídicas que genes MHC de classe Ia. Dos 22 alelos registrados no banco de dados IMGT apenas quatro codificam proteínas que diferem em suas estruturas primárias. A fim de estimar o genótipo e as frequências alélicas, este estudo teve como foco as regiões previamente descritas do gene *HLA-F* que são codificadoras de proteínas. A genotipagem foi feita através de sequenciamento (SBT). A amostra foi composta por 199 doadores de medula óssea sem relações de parentesco entre si e que fazem parte do Registro Nacional de Doadores de Medula Óssea (REDOME), euro-brasileiros, provenientes da região Sul do Brasil. Cerca de 1673 pares de base foram analisados. O alelo mais frequente foi o *HLA-F*01:01* (87.19%), seguido por *HLA-F*01:03* (12.31%), *HLA-F*01:02* (0.25%) e *HLA-F*01:04* (0.25%). Significantes desequilíbrios de ligação foram verificados entre alelos do gene *HLA-F* e alelos de genes HLA de classe I e II. Este é o primeiro estudo sobre polimorfismos do gene *HLA-F* em uma amostra da população euro-brasileira contribuindo para a caracterização genética da região Sul do Brasil.

Palavras-chave: Euro-brasileiros, *HLA-F*, desequilíbrio de ligação, polimorfismos, população.

ABSTRACT

HLA-F is a non-classical major histocompatibility complex (MHC) gene. It codes class Ib MHC molecules with restricted distribution and less nucleotide variations than MHC class Ia genes. Of the 22 alleles registered on the IMGT database only four alleles encode for proteins that differ in their primary structure. To estimate genotype and allele frequencies, this study targeted on known protein coding regions of the *HLA-F* gene. Genotyping was performed by Sequence-Based Typing (SBT). The sample was composed by 199-unrelated bone marrow donors from the Brazilian Bone Marrow Donor Registry (REDOME), Euro-Brazilians, from Southern Brazil. About 1673 bp were analyzed. The most frequent allele was *HLA-F*01:01* (87.19%), followed by *HLA-F*01:03* (12.31%), *HLA-F*01:02* (0.25%) and *HLA-F*01:04* (0.25%). Significant linkage disequilibrium (LD) was verified between *HLA-F* and HLA classes I and II alleles. This is the first study regarding *HLA-F* polymorphisms in a Euro-Brazilian population contributing to the Southern Brazilian genetic characterization.

Keywords: Euro-Brazilians, *HLA-F*, linkage disequilibrium, polymorphisms, population.

SUMÁRIO

1. Introdução	1
1.1. A importância dos estudos populacionais	1
1.2. História, função e organização genômica do MHC	2
1.3. Estrutura do gene <i>HLA-F</i>	7
1.4. Estrutura proteica HLA-F	8
1.5. Região promotora, regulação e expressão do <i>HLA-F</i>	9
1.6. Polimorfismos	15
1.7. O <i>HLA-F</i> e a evolução do MHC	20
2. Objetivos	22
2.1. Objetivo geral	22
2.2. Objetivos específicos	22
3. Justificativa	23
4. Artigo	24
5. Discussão	40
6. Referências Bibliográficas	44

LISTA DE ILUSTRAÇÕES

Figura 1. Representação esquemática do cromossomo 6 humano, com as regiões do complexo de Histocompatibilidade (MHC) e o sistema gênico HLA	3
Figura 2. Representação do reconhecimento do complexo HLA-peptídeo pela célula T	4
Figura 3. Estrutura de uma molécula HLA de classe I	5
Figura 4. Estrutura de uma molécula HLA de classe II	6
Figura 5. Estrutura da região promotora dos genes <i>HLA-E</i> , <i>-F</i> e <i>-G</i> e as sequências de nucleotídeos dos elementos reguladores conservados dos módulos de indução e SXY	9
Figura 6. Comparação entre a região promotora do <i>HLA-F</i> com regiões homólogas dos genes <i>HLA-A2</i> , <i>HLA-G</i> e <i>HLA-E</i>	12
Figura 7. Representação esquemática dos elementos reguladores e proteínas ligantes dos sítios promotores de <i>HLA-E</i> , <i>HLA-F</i> e <i>HLA-G</i>	13
Figura 8. Sequência nucleotídica do éxon 3 do gene <i>HLA-F</i>	17
Figura 9. Representação do gene <i>HLA-F</i>	18
Figure 1. Geographic location of Curitiba, the capital city of Paraná state. The maps are not drawn to scale. The geographical coordinates of Curitiba are 25°25'42" S, 49°16'23" W	35
Figure 2. Graphical overview of LD between <i>HLA-F</i> alleles and <i>HLA-A</i> , <i>-B</i> , <i>-DRB1</i> alleles. The comparisons with significant LD ($p < 0.01$) are discriminated by superscripted indexes: (1) indicates $D' = 1$, (2) indicates $D' = -1$, (3) D' between 0 and 1, (4) indicates D' between 0 and -1. The numbers given in the cells are the corresponding r^2 estimates	37

LISTA DE TABELAS

Tabela 1. Variação na sequência promotora em genes HLA não clássicos de classe I	16
Tabela 2. Alelos codificantes de proteínas do gene <i>HLA-F</i>	19
Table 1. <i>HLA-F</i> allele and genotype frequencies in the Euro-Brazilian population	38
Table 2. Comparison among Euro-Brazilian and Chinese <i>HLA-F</i> allele frequencies	38
Table 3. Haplotype frequencies of HLA-F-A, -B and -DRB1 in Euro-Brazilians	39
Tabela 3. Comparação das frequências alélicas entre amostras de diferentes populações	40

LISTA DE SIGLAS E ABREVIATURAS

- APC – Células apresentadoras de antígeno
- dNTPs – Desoxinucleotídeos trifosfatos
- EVT – Citotrofoblasto extraviloso
- HLA – Antígenos Leucocitários Humanos
- IMGT – *The international ImMunoGenetics information system*
- ILT – Transcritos semelhantes a imunoglobulina
- kDa – Quilodalton
- Kb – Quilobase
- µL – Microlitro
- MHC – Complexo Principal de Histocompatibilidade
- mM – Milimolar
- RT-PCR – Reação em cadeia da polimerase com transcrição reversa
- PCR – Reação em cadeia da polimerase
- SNP – Polimorfismo de nucleotídeo único
- T CD4⁺ – Linfócito T que expressa a molécula de superfície CD4
- T CD8⁺ – Linfócito T que expressa a molécula de superfície CD8
- TCR – Receptor de linfócitos T
- TAP – Transportador associado ao processamento de antígenos

1. Introdução

1.1. A importância dos estudos populacionais

Com o progressivo avanço das tecnologias aliadas à genética molecular, nas últimas décadas foram acumulados dados de diversas formas de variação no DNA, sobretudo, em populações humanas. Com esses dados foi possível, entre outras aplicações, melhor compreensão da evolução humana. Ademais, outros aspectos importantes são elucidados através de estudos populacionais, tais como, a origem e história demográfica das populações, a ação de fatores evolutivos em sequências genômicas e até mesmo estendendo-se à base genética de patologias (CAVALLI-SFORZA, 2003).

A caracterização da diversidade alélica e haplotípica nos principais genes do sistema HLA constitui um poderoso instrumento em estudos populacionais (ARNAIZ-VILLENA et al., 2001). Por fazer parte de um sistema altamente polimórfico e ter como característica forte desequilíbrio de ligação, conseqüentemente baixa taxa de recombinação, os diferentes locos do sistema HLA tem sido amplamente utilizados como marcadores genéticos em estudos populacionais. Diferentes populações apresentam padrões distintos de frequências alélicas e haplotípicas, e isto tem sido utilizado em diversos estudos para elucidar as origens das populações e suas relações evolutivas (PIMTANOTHAI et al., 2001; LUO et al., 2002; PIANCATELLI et al., 2004; AYED et al., 2004). Populações com origens em comum apresentam mais similaridade entre suas frequências alélicas acerca dos genes que fazem parte do sistema HLA, sendo determinados alelos característicos de certas populações de acordo com suas respectivas origens geográficas (MIDDLETON, MANAVALAN e FERNANDEZ-VINA, 2004). Além disso, inúmeros haplótipos do sistema HLA tem sido relacionados com a distribuição geográfica coincidente com a existência de relações passadas entre as populações (ARNAIZ-VILLENA et al., 2001; Sanchez-Velasco et al., 2003).

1.2. História, função e organização genômica do MHC

Os vertebrados desenvolveram ao longo de sua história evolutiva um refinado sistema imunológico. Os principais genes que fazem parte deste sistema estão localizados na região denominada Complexo Principal de Histocompatibilidade (CPH), mais conhecida pela sigla MHC, proveniente do inglês *Major Histocompatibility Complex*. Este complexo de histocompatibilidade foi primeiramente descrito em meados dos anos 40 através da observação da rejeição ou aceitação de transplantes de tecidos entre camundongos alogênicos ou singênicos (SNELL, 1948). Os homólogos murinos do Antígeno Leucocitário Humano (HLA – *Human Leukocyte Antigen*) de classe I são codificados pelos genes presentes na região denominada Histocompatibilidade-2 (H-2) localizada no cromossomo 17 destes organismos (KLEIN, 1975). Em contrapartida, nos humanos esta região está localizada no braço curto do cromossomo 6, precisamente na região 6p21.31, sendo denominada de Antígeno Leucocitário Humano. Esta denominação surgiu devido aos anticorpos, originalmente usados para identificar as moléculas humanas de MHC, reagirem com os glóbulos brancos do sangue, os leucócitos, mas não com as hemácias, que não possuem moléculas MHC (BREUNING et al., 1977). É estimado que o sexto cromossomo humano representa cerca de 6% do genoma e que o mesmo é constituído por 166.880.988 pares de base (MUNGALL et al., 2003). Nos humanos o Complexo Principal de Histocompatibilidade abrange cerca de 3,6 milhões de bases (3,6 Mb) de extensão e são conhecidos neste sistema aproximadamente 240 locos (SHIINA, INOKO e KULSKI, 2004), dos quais aproximadamente 128 destes foram relacionados a funções diversas no sistema imune de maneira funcional e 96 são pseudogenes (KLEIN e SATO, 2000; The MHC Sequencing Consortium, 1999). Estudos de densidade gênica estimam que há uma distância média de 16 mil bases (16 Kb) entre os diversos genes expressos e os pseudogenes localizados em sua extensão (The MHC Sequencing Consortium, 1999). A transmissão de alelos de cada gene *HLA* é realizada em blocos denominados haplótipos. Cada indivíduo herda um haplótipo paterno e outro materno, ou seja, um bloco de alelos *HLA* de cada progenitor. Isto deve-se a outra característica intrínseca do MHC, a baixa taxa de recombinação. Estima-se uma frequência de recombinação de 4% entre os pontos

extremos do MHC (GERAGHTY et al., 1992). O complexo MHC abrange vários genes *HLA* e outros genes não pertencentes ao sistema HLA (figura 1). O sistema gênico *HLA* é conhecido como o mais polimórfico do genoma humano (MUNGALL et al., 2003).

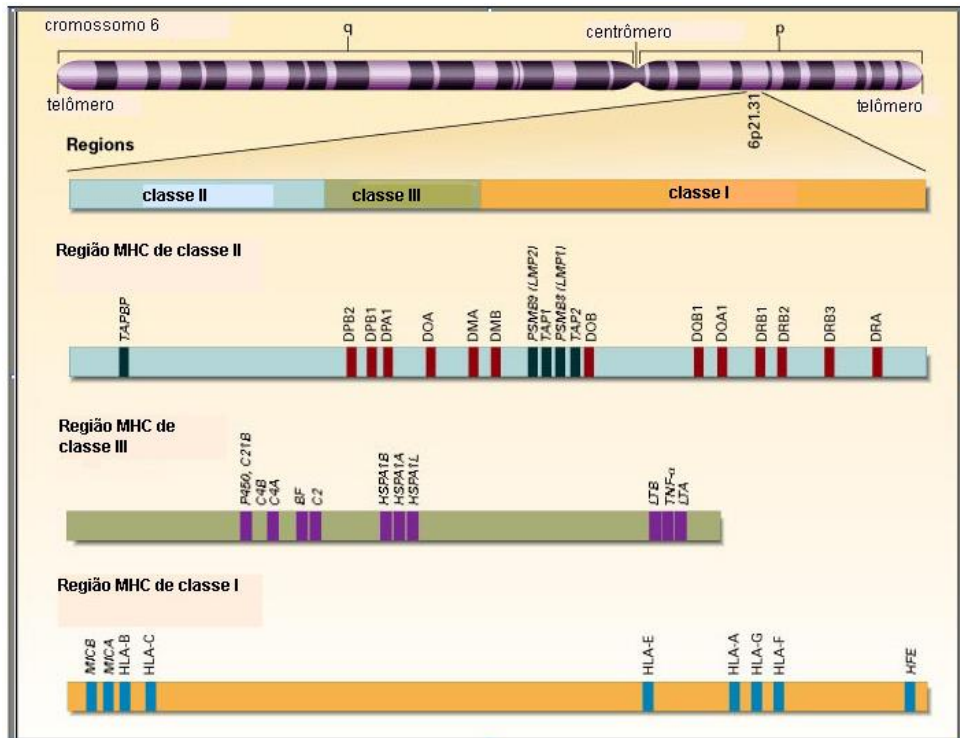


Figura 1. Representação esquemática do cromossomo 6 humano, com as regiões do complexo de Histocompatibilidade (MHC) e o sistema gênico HLA. Em laranja, está representada a região de classe I, em azul a região de classe II e em verde a região de classe III. Fonte: KLEIN e SATO, 2000.

Cada molécula MHC consiste em uma fenda extracelular que tem a capacidade de ligar-se aos peptídeos antigênicos, seguida de domínios semelhantes a imunoglobulinas (Igs), e permanece ancorada na membrana celular por meio de domínios transmembrânicos e citoplasmáticos (McDEVITT, 2000). Estas estruturas apresentam antígenos aos linfócitos T, desencadeando uma série de sinalizações que resultam em uma resposta imune. Os linfócitos T possuem em sua superfície os receptores de células T, também denominados TCR, do inglês *T Cell Receptor*. Estas estruturas são responsáveis pelo reconhecimento dos antígenos apresentados pelas células apresentadoras de antígenos, APCs – *Antigen Presenting Cells* (figura 2).

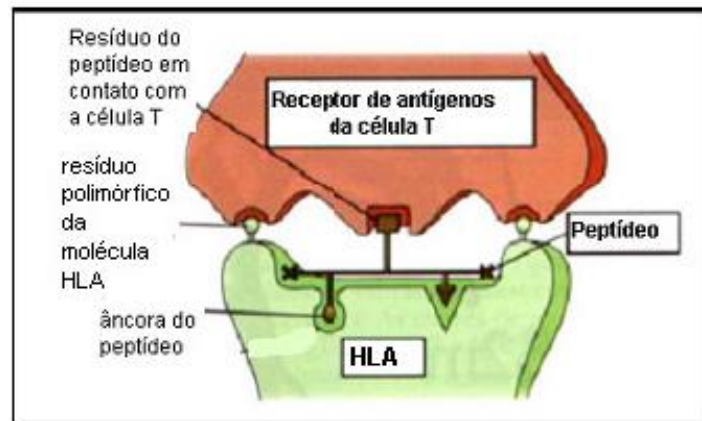


Figura 2. Representação do reconhecimento do complexo HLA-peptídeo pela célula T. O esquema demonstra uma molécula HLA (em verde) ligando e exibindo um peptídeo (em marrom) a um receptor de células T (em vermelho), que reconhece dois resíduos polimórficos da molécula HLA e um resíduo peptídico. Fonte: ABBAS e LICHTMAN, 2005.

O MHC humano é dividido em três classes gênicas de acordo com sua função, estrutura e distribuição nas células: Classe I, Classe II e Classe III.

A região de classe I é a mais telomérica em relação às outras regiões MHC e é composta por mais de 122 locos e 41 deles são genes que se expressam e transcrevem algum produto funcional (SHIINA, INOKO e KULSKI, 2004). Fazem parte desta região os genes *HLA-A*, *-B*, *-C*, *-E*, *-F*, *-G*, *-H*, *-J*, *-K*, *-L*, *-N*, *-P*, *-S*, *-T*, *-U*, *-V*, e *-W* (LAWLOR et al., 1990; GERAGHTY et al., 1990; GERAGHTY, 1993; CAMPBELL e TROWSDALE, 1993; BODMER et al., 1995; HORTON et al., 2004). Esta região ainda pode ser dividida em moléculas de HLA I clássicas (HLA-Ia) e não clássicas (HLA-Ib). Os genes *HLA-A*, *HLA-B* e *HLA-C* representam as moléculas HLA I clássicas e são os mais conhecidos. As respectivas proteínas são expressas na superfície de quase todas as células nucleadas (ZAVAZAVA, 1998; SULLIVAN et al., 2006). Por outro lado, os genes *HLA-E*, *HLA-F* e *HLA-G* fazem parte do grupo HLA-Ib e codificam moléculas encontradas em tecidos fetais e em tecidos mais restritos em indivíduos na fase adulta (GERAGHTY, 1993; PYO et al., 2006). Além das diferenças funcionais em relação aos genes clássicos deste grupo, *HLA-E*, *HLA-F* e *HLA-G* apresentam relativamente poucas variações nucleotídicas em suas sequências codificantes, distribuição muito restrita e baixos níveis de expressão (Le BOUTEILLER, 1997). Os locos considerados pseudogenes são *HLA-H*, *HLA-J*, *HLA-K*, *HLA-L*, *HLA-N*, *HLA-P*, *HLA-S*, *HLA-T*, *HLA-U*, *HLA-V*, *HLA-W* (CAMPBELL e TROWSDALE, 1993; HORTON et al., 2004). Os genes *HLA* de classe I codificam para a formação da cadeia polipeptídica alfa de receptores glicoproteicos de membrana, constituídos por uma região extracelular de ligação a antígenos, uma

região transmembrana e um segmento citoplasmático (The MHC Sequencing Consortium, 1999). As proteínas HLA de classe I (figura 3) contêm uma porção extracelular constituída por heterodímeros de uma cadeia pesada α de 44 kDa ($\alpha 1$, $\alpha 2$ e $\alpha 3$), ligada não-covalentemente à cadeia leve $\beta 2$ -microglobulina de 12 kDa sendo a última molécula codificada por um gene de mesmo nome localizado no cromossomo 15 (ZAVAZAVA, 1998; ADAMASHVILI, et al., 2005).

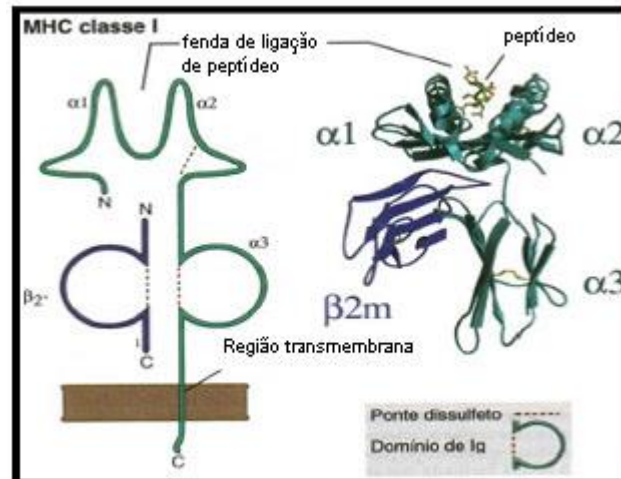


Figura 3. Estrutura de uma molécula HLA de classe I. Fonte: ABBAS e LITCHMAN, 2005.

As variações nucleotídicas nos genes de classe I localizam-se essencialmente nos éxons 2 e 3, os quais codificam para as estruturas moleculares $\alpha 1$ e $\alpha 2$ do receptor de membrana onde se localiza o sítio de ligação aos antígenos (MALISSEN, MALISSEN e JORDAN, 1982). A expressão de moléculas HLA na superfície celular enfatiza seu importante papel na apresentação do peptídeo antigénico para células T e início da resposta imune mediada por linfócitos T. As moléculas HLA de classe I normalmente apresentam peptídeos derivados de proteínas citosólicas para linfócitos T $CD8^+$. Através desta interação é possível detectar alterações de expressão celular as quais podem ocorrer devido a infecções virais ou desenvolvimento de células tumorais. Como resultado, os linfócitos T $CD8^+$ são capazes de reconhecer e atacar as células sinalizadas, combatendo o avanço da infecção ou tumor (DOHERTY e ZINKERNAGEL, 1975; TABAYOYONG e ZAVAZAVA, 2007).

A capacidade das APCs de apresentarem antígenos às células T é devida à presença de proteínas especializadas que são codificadas por genes presentes no MHC de classe II (ALAM et al., 1996). A região de classe II está localizada mais

centromericamente quando comparada com as regiões de classe I e III. Já foram descritos aproximadamente 34 locos na região de classe II (KLEIN e SATO, 2000). Os genes mais estudados contidos nesta região são: *HLA-DP*, *HLA-DQ* e *HLA-DR* codificando proteínas de classe II de mesmo nome. Estas proteínas (figura 4) são formadas também por heterodímeros, como as proteínas de classe I, todavia estes heterodímeros são constituídos por duas cadeias α de 32 kDa ($\alpha 1$ e $\alpha 2$) e duas cadeias β de 29 kDa ($\beta 1$ e $\beta 2$) (ZAVAZAVA, 1998; ADAMASHVILI et al., 2005). A expressão de HLA-II é restrita às células apresentadoras de antígenos (APCs), às células endoteliais e linfócitos T ativados (ZAVAZAVA, 1998; SULLIVAN et al., 2006).

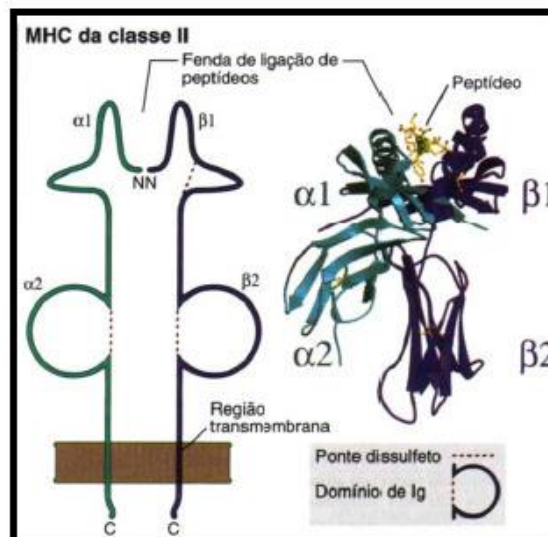


Figura 4. Estrutura de uma molécula HLA de classe II. Fonte: ABBAS e LITCHMAN, 2005.

No éxon 2 destes genes concentra-se a maior parte das variações nucleotídicas. Isto pode ser explicado pelo fato deste éxon codificar para a estrutura extracelular do receptor de membrana, responsável pela ligação aos antígenos. Estes receptores de membrana são constituídos por uma porção extracelular, uma região de ancoragem transmembrana e um segmento citoplásmático. Eles são responsáveis por reconhecer a presença de elementos de origem exógena e estimular os linfócitos T CD4⁺, desencadeando uma resposta imunitária contra agentes infecciosos (DOHERTY e ZINKERNAGEL, 1975; GATTI e PIERRE, 2003; TABAYOYONG e ZAVAZAVA, 2007).

A região de classe III localiza-se entre as regiões de classe I e II e estende-se por aproximadamente 900 mil bases (900 Kb), codificando mais de 70 polipeptídeos diferentes (HAUPTMANN e BAHRAM, 2004). Ao contrário das regiões de classe I e classe II onde existem dezenas de pseudogenes, a região de classe III possui apenas dois (The MHC Sequencing Consortium, 1999). De acordo com SHIINA, INOKO e KULSKI (2004) esta é uma região de alta densidade gênica, apresentando cerca de um gene a cada 14.500 bases. Nesta região estão presentes tanto genes que atuam no processo de ativação do sistema complemento (*BF*, *C2*, *C4A* e *C4B*) quanto genes que codificam citocinas como *TNFA* e *LTA* (MATSUZAKA et al., 2001).

1.3. Estrutura do gene *HLA-F*

Em 1990 GERAGHTY et al. caracterizaram por clonagem uma sequência genômica pertencente ao grupo de HLA de classe Ib denominada de *HLA-5.4*. Esta denominação primária foi designada devido a este gene estar localizado em um fragmento de restrição de 5.4 Kb (5.400 pares de base) gerado pela enzima de restrição *Hind III*.

Este gene é constituído por aproximadamente 3.550 pares de base, sendo este número variável entre os 22 alelos por conter deleções em alguns deles. O *HLA-F* consiste em oito éxons e sete íntrons, porém, a região referente ao sétimo éxon pode ser excisado do transcrito maduro. Isto deve-se à uma mudança de apenas um nucleotídeo na região 3' do sexto íntron alterando o sítio de *splicing* *GT/AG* altamente conservado em mamíferos. Os primeiros seis éxons desta sequência têm uma organização similar à organização do *HLA-A2*, e os domínios extracelular e transmembrana das proteínas codificadas, pelo então denominado *HLA-5.4*, são similares àqueles dos antígenos de HLA clássicos. O segmento citoplasmático desta sequência aparenta ser consideravelmente mais curto que das outras moléculas de HLA classe I clássicas.

A região 3'UTR do *HLA-F* é a mais distinta quando comparada com as regiões correspondentes de todos os outros genes de classe I. Os primeiros 32 pares de base desta região são homólogos aos outros genes de classe I, todavia, o restante desta sequência diverge completamente. Esta sequência única contém dois sinais adicionais de polinucleotídeos "A" localizados a 90 bases a jusante do códon

de parada. Ainda é sugerido que esta região seja homóloga a uma segunda família multigênica distinta (GERAGHTY et al., 1990).

1.4. Estrutura proteica HLA-F

A proteína HLA-F é altamente conservada no domínio $\alpha 3$, assim como as demais proteínas de classe I, presumivelmente devido à importância desta região na interação com a $\beta 2$ -microglobulina. Os domínios $\alpha 1$ e $\alpha 2$ apresentam 52-76% de homologia com os domínios correspondentes de outras moléculas de HLA classe I. Similarmente a comparações entre proteínas *HLA-A*, *-B* e *-C*, os domínios $\alpha 1$ e $\alpha 2$ da proteína HLA-F divergem igualmente das proteínas *HLA-A2*, *HLA-B58*, *HLA-C3* e *HLA-G*. Isto contrasta com *HLA-E*, que apresenta significativamente maior divergência em $\alpha 2$ que em $\alpha 1$ (GERAGHTY et al., 1990).

Em 1985 LOPEZ DE CASTRO et al. identificaram regiões hipervariáveis dentro dos domínios $\alpha 1$ e $\alpha 2$ em proteínas HLA de classe Ia. A maioria dos resíduos variáveis estão agrupados nas regiões de posições 57 a 83 em $\alpha 1$ e 143 a 171 em $\alpha 2$. Do mesmo modo, as substituições encontradas na proteína HLA-F estão agrupadas dentro destas regiões. Ambos os pares de cisteína encontrados em $\alpha 2$ e $\alpha 3$, conservados entre todos antígenos de classe I, estão presentes em posições idênticas na proteína HLA-F. Apenas um sítio de glicosilação aminoterminal, Asn-86, é encontrado na região $\alpha 1$. Se o éxon sete for excisado, o segmento citoplasmático apresentará uma cadeia de dezesseis aminoácidos a menos quando comparado à mesma região nas demais moléculas de HLA. Em 1990, GERAGHTY et al. demonstraram que a proteína HLA-F apresenta massa molecular de 41.000 daltons (41 kDa), cerca de 2.000 daltons (2 kDa) a menos que HLA-A, HLA-B e HLA-C.

Nas proteínas HLA de classe I são encontrados, em geral, dez resíduos completamente conservados, destes, cinco são tirosinas, evidenciando o sítio de reconhecimento do antígeno. Cinco destes dez resíduos são diferentes na proteína HLA-F. As mudanças significativas incluem a substituição de Tyr por Phe na posição 22, Glu por Gly na posição 26, e uma Arg por Tyr na posição 84. Apenas quatro dos resíduos conservados são tirosinas em HLA-F, diferente das outras moléculas HLA em que são encontradas cinco tirosinas. Em contraste, dos dez resíduos conservados nas moléculas HLA de classe Ia, apenas oito são conservados em

HLA-E, e nove são conservados em *HLA-G*, em ambos os casos incluindo todas as cinco tirosinas (TRANSY et al., 1987).

Os resíduos conservados na hélice longa do domínio $\alpha 1$ são também conservados em *HLA-F*, exceto uma substituição na posição 61 de uma Asp por Glu. A hélice longa do domínio $\alpha 2$ apresenta maior variabilidade, incluindo uma substituição de Lys por Phe na posição 146 e uma Ala por Glu na posição 150. Quando analisamos esta mesma região na proteína *HLA-E*, esta apresenta aproximadamente a mesma extensão de divergência, com substituição de uma Gly por Asp na posição 162, além de uma substituição de Arg por His na posição 169 (GERAGHTY et al., 1990).

1.5. Região promotora, regulação e expressão do *HLA-F*

Ao longo da região 5'UTR dos genes de classe I, clássicos e não clássicos, há pelo menos cinco sequências promotoras que são importantes na regulação da transcrição destes genes (figura 5). Quando comparadas as regiões 5' dos genes *HLA-F*, *HLA-A2* e *HLA-G*, são encontrados aproximadamente 850 pares de base de extensão em homologia. Contudo, o gene *HLA-E* compartilha homologia de apenas 190 pares de base a partir do códon de início, ATG, ao sítio de homologia A, a partir do qual esta sequência diverge completamente das outras. Já foram identificadas as sequências CAAT e TCTAA na região promotora do *HLA-F*, homologia encontrada na maioria dos outros genes HLA de classe I, com exceção do *HLA-E*, que apresenta a sequência mais típica TATAA (GERAGHTY et al., 1990).

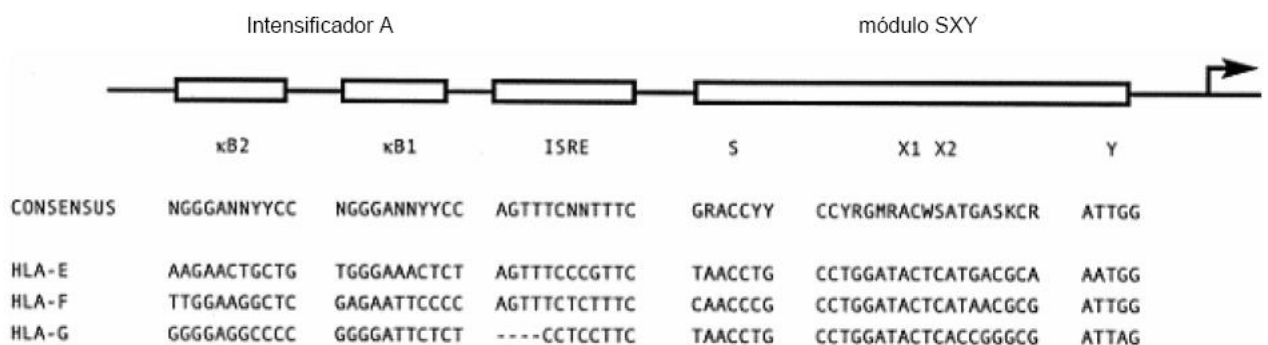


Figura 5. Estrutura da região promotora dos genes *HLA-E*, *-F* e *-G* e as sequências de nucleotídeos dos elementos reguladores conservados dos módulos de indução e SXY. Fonte: GOBIN e VAN DEN ELSEN, 2000.

Duas sequências reguladoras foram identificadas em genes de classe I em murinos (TRANSY et al., 1987) sendo também encontradas em genes *HLA* de classe I, em humanos. O sítio de homologia A é uma sequência palindrômica que estende-se por 13 pares de base e serve como um sítio de ligação para um fator proteico que interage com o promotor *H-2K* (BALDWIN e SHARP, 1987; ISRAEL et al., 1987), sendo a variante desta homologia também encontrada em região promotora do gene *β 2-microglobulina* (ISRAEL et al., 1987). Ambas as sequências dos sítios promotores dos genes *H-2K* e *β 2-microglobulina* apresentam acentuador, ou *enhancer*, e atividade de ligação proteica. Por outro lado, a sequência do promotor *Q10*, que possui duas substituições de bases correspondentes a esta região homóloga com *H-2K*, não dispõe de quaisquer das atividades de acentuador ou ligação proteica. No gene *HLA-F*, em relação aos murinos, a região de homologia A é alterada em duas posições, em uma delas ocorre uma substituição de G por A na posição três que é idêntica ao encontrado na sequência *Q10* (KIMURA et al., 1986). O gene *HLA-G* apresenta uma deleção de parte desta sequência ao passo que no gene *HLA-E* não é encontrada sequência semelhante em sua região 5'. O sítio de homologia B é constituído por uma sequência de 11 pares de base, sendo esta sequência encontrada em *HLA-A*, *HLA-B* e *HLA-C*. Em um fragmento contendo uma sequência idêntica nos promotores murinos *H2-K*, *H2-D* e *H2-L* foi evidenciada atividade acentuadora (KIMURA et al., 1986). O promotor do gene *HLA-F* apresenta uma variação de base nesta sequência. Esta substituição na região de homologia B é distinta das outras duas substituições encontradas no *HLA-G* assim como no *HLA-E*.

A montante da região promotora do gene *HLA-F* há uma sequência que consiste em 25 purinas com três repetições em tandem de GGGGGA. Quando comparadas as mesmas regiões em *HLA-A2* e *HLA-G*, estas também são ricas em purina ao contrário do que ocorre em *HLA-E*. As sequências de *HLA-A2* e *HLA-G* são quase idênticas umas as outras, porém, são mais curtas, diferem na composição de bases e não possuem a estrutura repetitiva detectada no gene *HLA-F*. A sequência mais marcante na região 5' do *HLA-F* é uma extensão contínua de 184 resíduos de purina estendendo-se da posição 110 à posição 293. Esta região contém 47 ocorrências de trinucleotídeos GAA e 15 de GAG, incluindo GAA repetindo-se 21 vezes em tandem (GERAGHTY et al., 1990).

O gene *HLA-G* não contém um trecho similar de sequência. Apesar da homologia com *HLA-F* em ambos os lados 5' e 3' da sequência 1 (figura 6), a sequência correspondente em *HLA-G* é bastante reduzida, 20 pares de bases, e enquanto contém 15 purinas consecutivas, não há repetições similares de trinucleotídeos. Ambos os alelos *HLA-C*01* e *HLA-C*02* apresentam uma extensão de 10 purinas, também sem repetições de trinucleotídeos, em posição análoga (GUSSOW et al., 1987).

Há evidências que apoiam o compartilhamento de mecanismos regulatórios na expressão dos genes *HLA-E* e *HLA-F*. Parte destes atuam via módulo SXY (figura 7), pois, as expressões do gene são favoráveis para *HLA-E* e *HLA-F* em células que expressam moléculas de HLA de classe II (ROUSSEAU et al., 2004).

Em contraste com o *HLA-E*, a expressão dos genes *HLA-F* e *HLA-G* é muito restrita em condições não patológicas. A expressão do gene *HLA-F* é predominantemente intracelular, e confinado às células B e tecidos tais como tonsilas em adultos, timo e fígado fetal (HOULIHAN et al., 1992; WAINWRIGHT, BIRO e HOLMES, 2000). Em 1992, HOULIHAN et al. descreveram a expressão dos genes *HLA-E*, *HLA-F* e *HLA-G* em células de fígado fetal, ressaltando a expressão do *HLA-F* no primeiro e segundo trimestre, distintamente de *HLA-E* que só foi constatada a expressão no segundo trimestre, e *HLA-G* que foi constatada a expressão apenas no primeiro trimestre. Ainda neste trabalho foi sugerido que os genes *HLA* de classe Ib parecem exibir padrões distintos de expressão ao longo do desenvolvimento do fígado humano, sugerindo uma possível correlação entre a existência de linfócitos T γ/δ no fígado humano (APARICIO et al., 1989), com a expressão de moléculas *HLA* de classe Ib neste tipo celular e a possível apresentação de antígeno destas moléculas não clássicas aos linfócitos T γ/δ (STROYNOWSKI, 1990). Posteriormente foi constatado que os tetrâmeros de HLA-F ligam-se a receptores inibidores ILT2 e ILT4, sugerindo que a proteína codificada pelo gene *HLA-F* pode ligar-se a peptídeos e atingir a superfície celular. Por consequência, *HLA-F* poderia também modular a função imunológica de células efectoras (LEPIN et al., 2000). De fato em 2003 ISHITANI et al. descreveram a primeira expressão do antígeno HLA-F na superfície celular, além de registrarem o primeiro tipo celular a expressar simultaneamente todos os antígenos não clássicos de classe I, as células trofoblásticas. Ainda em 2003 ISHITANI et al. afirmaram que

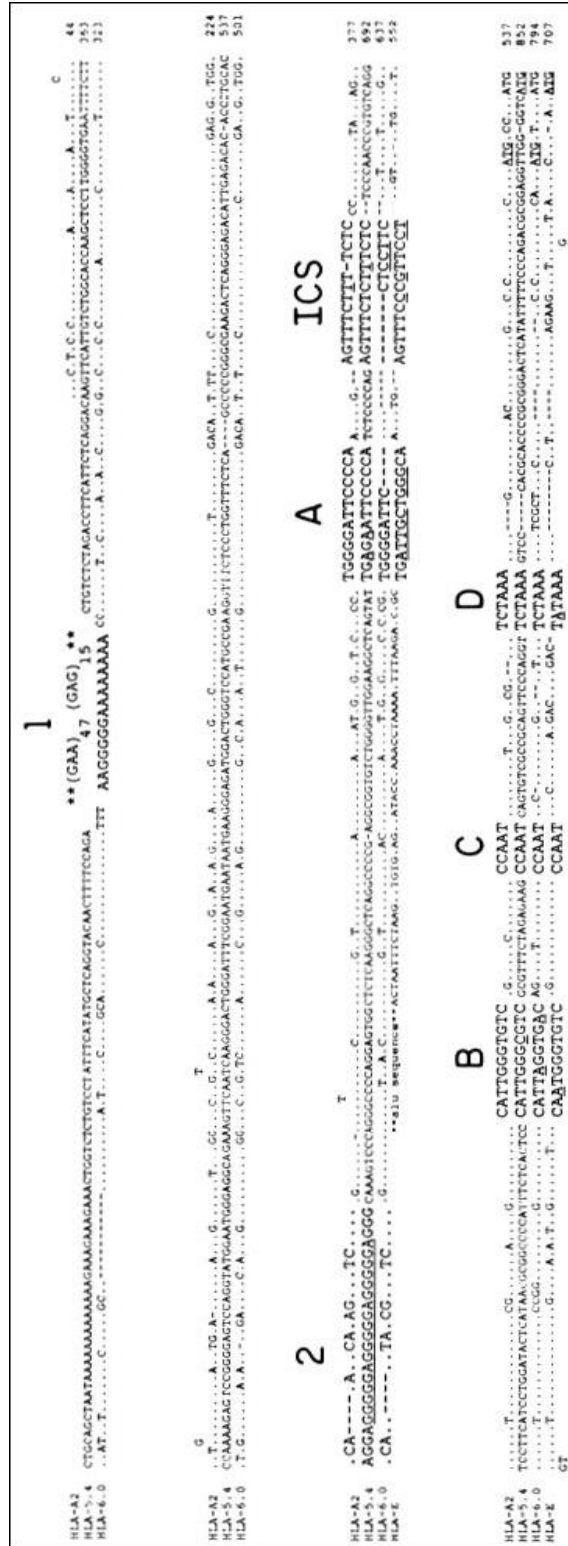


Figura 6. Comparação entre a região promotora do *HLA-F* com regiões homólogas dos genes *HLA-A2*, *HLA-G* e *HLA-E*. Fonte: GERAGHTY et al., 1990.

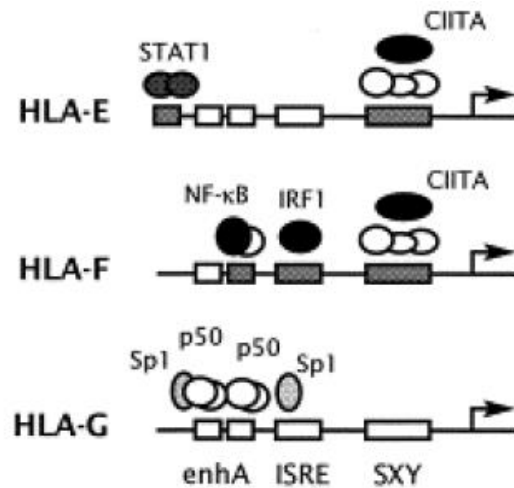


Figura 7. Representação esquemática dos elementos reguladores e proteínas ligantes dos sítios promotores de *HLA-E*, *HLA-F* e *HLA-G*. Os módulos conservados são os módulos de indução contendo sítios ligantes para NF- κ B, Sp1 e IRF1. O módulo SXY consiste dos blocos S, X1, X2, e Y que formam a região de ligação para o complexo de multiproteínas consistindo de RFX, CREB/ATF e NFY os quais mediam a transativação induzida por CIITA. Fonte: GOBIN e VAN DEN ELSEN, 2000.

os padrões obtidos de expressão na superfície celular fornecem a evidência de que o *HLA-F* exerce um papel fundamental na resposta imune materno-fetal. Realmente, em 2006, SHOBU et al. verificaram que há um aumento nos níveis da expressão do *HLA-F* durante o segundo trimestre de gestação, indicando que este gene tem um papel no desenvolvimento fetal. SHOBU et al. em 2006 ainda concluíram que o aumento das concentrações das proteínas HLA-E e HLA-F estão correlacionados com o aumento do crescimento fetal, sugerindo que estas proteínas tenham um importante papel na estabilização de um ambiente ótimo que permita o crescimento adequado do feto. Esta estabilização seria alcançada com um mecanismo de tolerância gerado pela expressão dos três genes de classe Ib, *HLA-E*, *HLA-F* e *HLA-G*, mesmo que um desses genes em questão não seja expresso, todavia, ainda há a manutenção da tolerância materno-fetal. Ainda há a hipótese da necessidade do *HLA-F* ser expresso na superfície celular junto com *HLA-E* e *HLA-G* para que ocorra a devida comunicação celular, fundamental na imunologia da gestação (LEE, ISHITANI e GERAGHTY, 2010). Em contrapartida, NAGAMATSU et al. (2006) encontraram resultados divergentes. Foi detectada, intracelularmente, a presença de moléculas HLA-F em células trofoblásticas no primeiro trimestre e em EVT's no segundo e terceiro trimestre. Os mesmos resultados foram obtidos quando utilizada a linhagem celular JEG3. Apesar de confirmar a presença de HLA-F em células do trofoblasto extraviloso fetal, não foi constatada a expressão destas moléculas na superfície celular. De acordo com os autores, esta divergência pode ser atribuída à diferença de estratégia experimental, a qual a citometria de fluxo tende ser mais

rigorosa, em questão de distribuição celular, que análise imunohistoquímica. Ainda foi verificado em cultura de EVTs que a expressão intracelular de HLA-F foi progressivamente aumentada, concomitantemente com a expressão de HLA-G na superfície celular. Embora este estudo ter sugerido, pela falta de expressão na superfície celular, que HLA-F não contribua diretamente na reação imune entre trofoblastos e células do sistema imune materno, a molécula HLA-F tem um papel fundamental na interface materno-fetal, podendo alterar a função do HLA-G (NAGAMATSU et al., 2006). APPS et al., 2008b também não verificaram a expressão de HLA-F na superfície de leucócitos decíduais, tampouco na superfície de células trofoblásticas fetais do primeiro trimestre de gestação, afirmando que, apesar de outros estudos apresentarem resultados opostos, só há evidências convincentes de que apenas três moléculas HLA de classe I são exibidas pelas células trofoblásticas extravilosas fetais: HLA-C, HLA-E e HLA-G (KING et al., 2000; APPS et al., 2007, 2008a).

Outra questão de extrema relevância é a maneira pelo qual o gene *HLA-F* é exportado para a membrana citoplasmática. Primeiramente, em 2000, WAINWRIGHT, BIRO e HOLMES, detectaram evidências de uma associação física entre a proteína HLA-F e TAP. Três anos após estas evidências, em 2003 LEE e GERAGHTY afirmaram que a expressão na superfície celular da proteína HLA-F é parcialmente independente de tapasina e completamente independente de TAP em linfócitos B e em linhagens de monócitos. Esta afirmação foi resultante dos experimentos que detectaram que a expressão de HLA-F na superfície celular não era reduzida em uma linhagem celular mutante que era TAP negativa. Porém, foi observado que apesar de não haver diferença nos níveis de expressão, ficou evidente que a expressão do HLA-F na superfície celular e subsequente infrarregulação ocorrem mais cedo que em células que não apresentam a mutação TAP⁻ (LEE, ISHITANI e GERAGHTY, 2010). Em 2006, BOYLE et al. identificaram a necessidade da modificação na cauda citoplasmática da molécula proteica HLA-F para sua expressão na superfície celular. É essencial que o único aminoácido codificado no éxon 8, uma valina, esteja na porção C-terminal para a exportação do retículo endoplasmático para a superfície celular. Outras moléculas do sistema imune como CD8 α , pro-TGF α , MTI-MPP, entre outras, também requerem uma valina na posição C-terminal a fim de serem capturadas pelas vesículas COPII para o transporte retículo endoplasmático-complexo de Golgi (BRILEY et al., 1997; URENA

et al., 1999; IODICE, SARNATARO e BONATTI, 2001). Apesar de todas as moléculas de classe I trafegarem pelo complexo de Golgi antes de alcançarem a superfície celular, a concentração de HLA-F nesta organela é única quando comparada com outras moléculas de classe I. Isto pode sugerir uma função alternativa, tal como monitorar infecções no aparelho de Golgi (BOYLE et al., 2006). BOYLE et al. (2006) ainda ressaltam que o padrão divergente de tráfego à superfície celular sugere que a função da molécula HLA-F pode ser independente da apresentação de peptídeo aos linfócitos T. Há evidências que sugerem que a proteína HLA-F é expressa independentemente da presença de peptídeo e que uma interação física específica às cadeias pesadas do MHC-I (MHC-I HC – *Major Histocompatibility Complex Class I Heavy Chains*) tem um papel fundamental na expressão dessas cadeias em linfócitos ativados. GOODRIDGE et al. (2010) sugerem que a molécula HLA-F possa servir como chaperona para estabilizar a expressão das cadeias pesadas MHC-I na ausência de peptídeo. Esta função de escolta poderia operar em ambos os sentidos, trazendo as cadeias pesadas MHC-I para a superfície e as internalizando depois de os sinais apropriados terem sido devidamente enviados.

1.6. Polimorfismos

Alguns dos SNPs identificados nas sequências *HLA-F* foram encontrados nas sequências promotoras. Um SNP foi identificado no sítio kB2 do acentuador A (GOBIN et al., 1998), enquanto um outro SNP, relativamente frequente, ocupou uma posição central no módulo SXY (GOBIN e VAN DEN ELSEN, 2000; ROUSSEAU et al., 2004). No último caso, todas as amostras que apresentavam como polimorfismo uma timina na posição -144 tinham 12 pb deletados no íntron 3, na posição 1.193, e todas estas amostras eram do alelo *HLA-F*01:01:02*. Essas variações na sequência promotora estão mais detalhadas na tabela 1.

Através da técnica denominada de Reação em Cadeia da Polimerase - Polimorfismo de Conformação de Fita Simples ou PCR-SSCP, do inglês *Polymerase Chain Reaction - Single-Strand Conformational Polymorphism*, UCHIGIRI et al. em 1997 descreveram um alelo nulo de *HLA-F* (figura 8). A mutação sem sentido foi identificada no éxon 3, uma troca de C por A no terceiro nucleotídeo do codon 28 deste éxon. Esta mutação foi confirmada com o auxílio da técnica de PCR-RFLP, do

Tabela 1. Variação na sequência promotora em genes HLA não clássicos de classe I. Fonte: PYO et al., 2006. Adaptado.

	Posições dos sítios regulatórios	SNPs	Referências
HLA-F			
Intensificador A	κB2	-226 to -216	-222 A to G (Gobin et al. 1998; Gobin e van den Elsen 2000)
	κB1	-209 to -199	-
Módulo SKY	S	-171 to -165	- (Gobin e van den Elsen 2000; Rousseau et al. 2004)
	X1X2	-149 to -131	-144 A to T ^b
	Y	-112 to -108	-
HLA-G			
Elementos de resposta trofoblasto-específicos	TSRE	-1201 to -1100	-1179 G to A (Ober et al. 2003; Solier et al. 2001)
			-1155 G to A
			-1140 T to A
			-1138 A to G
Elementos de resposta estimulados por interferon / Região flanqueadora de sítio ativado por interferon gama	ISRE/GAS	-754 to -735	-762 T to C (Ober et al. 2003; Solier et al. 2001)
			-725 C to G
			-716 G to T
			-
Elemento de proteína de choque térmico	HSE	-487 to -475	-477 G to C (Ober et al. 2003; Solier et al. 2001)
	HSE		-486 C to A

inglês *Restriction Fragment Length Polymorphism*, usando a enzima de restrição *RsaI*. Do total de 177 indivíduos apenas 2 apresentaram, em heterozigose, o alelo nulo resultando em uma frequência genotípica de 0,56%. UCHIGIRI et al., ainda em 1997, estimaram que o alelo nulo em homozigose estaria presente em 1 indivíduo a cada 30.000 na população japonesa. Ainda foi especulado que a proteína truncada resultante da mutação sem sentido, quando em heterozigose, não traria maiores complicações biológicas, pois, o outro alelo resultaria em uma proteína normal e compensaria a função da outra molécula truncada.

Posteriormente em uma amostra de 50 indivíduos, também da população japonesa, foi revelada uma transversão no nucleotídeo 1243 do gene *HLA-F*. A mutação identificada foi uma substituição sinônima de C por G no éxon 2, portanto, não modificando o aminoácido alanina presente no resíduo 67 do domínio α1 da proteína HLA-F madura (KUNISHIMA et al., 1999). Esta substituição foi inicialmente identificada pela técnica de PCR-SSCP e confirmada por PCR-RFLP e então sequenciamento. Dos 50 indivíduos analisados, 10 eram heterozigotos para a mutação 1243C/G, sendo os 40 indivíduos restantes homozigotos 1243C. A frequência alélica calculada foi de 0,9 para 1243C e 0,1 para 1243G. Outro dado relevante neste estudo foi a identificação do desequilíbrio de ligação entre os genes *HLA-A31* e *HLA-F*. A distribuição de *HLA-F* 1243G esteve presente em 77,8% dos indivíduos positivos para o alelo *HLA-A31*. KUNISHIMA et al. (1999) ainda

analisaram a presença do alelo *HLA-F* 1243G em outros 20 indivíduos homocigotos para o gene *HLA-A31*. A mutação em homocigose foi identificada em 11 indivíduos, por outro lado, apenas 3 indivíduos foram identificados como heterocigotos 1243G/C. A frequência alélica em indivíduos *HLA-A31* positivos foi de 0,375 para *HLA-F* 1243C e 0,625 para *HLA-F* 1243G.

										10									20	
HLA-F exon3	Gly	Ser	His	Thr	Leu	Gln	Gly	Met	Asn	Gly	Cys	Asp	Met	Gly	Pro	Asp	Gly	Arg	Leu	Leu
D78595	GGG	TCT	CAC	ACC	CTC	CAG	GGA	ATG	AAT	GGC	TGC	GAC	ATG	GGG	CCC	GAC	GGA	CGC	CTC	CTC
										28										
HLA-F exon3	Arg	Gly	Tyr	His	Gln	His	Ala	Tyr												
D78595	CGC	GGG	TAT	CAC	CAG	CAC	GCG	TAC	GAC	GGC	AAG	GAT	TAC	ATC	TCC	CTG	AAC	GAG	GAC	CTG
HLA-F exon3	CGC	TCC	TGG	ACC	GCG	GCG	GAC	ACC	GTG	GCT	CAG	ATC	ACC	CAG	CGC	TTC	TAT	GAG	GCA	GAG
D78595	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
HLA-F exon3	GAA	TAT	GCA	GAG	GAG	TTC	AGG	ACC	TAC	CTG	GAG	GGC	GAG	TGC	CTG	GAG	TTG	CTC	CGC	AGA
D78595	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
HLA-F exon3	TAC	TTG	GAG	AAT	GGG	AAG	GAG	ACG	CTA	CAG	CGC	GCA								
D78595	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Figura 8. Sequência nucleotídica do éxon 3 do gene *HLA-F*. Em destaque a transversão C/A que gera uma proteína truncada. Os traços indicam semelhança com o nucleotídeo indicado acima. Fonte: UCHIGIRI et al., 1997.

A mais marcante característica da proteína codificada pelo alelo *HLA-F*01:01:02:01* é o tamanho reduzido da cauda citoplasmática. A porção reduzida em 48 pares de base deve-se à excisão do éxon 7 do transcrito maduro em detrimento de um polimorfismo de apenas um nucleotídeo. Este polimorfismo de único nucleotídeo (SNP - *Single Nucleotide Polymorphism*) localiza-se na região terminal do íntron 6 substituindo um G por A. Em mamíferos o sítio de excisão conservado é determinado pelos pares de nucleotídeos GT-AG, portanto, a substituição do segundo G que determina o sítio de excisão, resultando em um par GT-AA, tem como consequência a excisão de toda a região delimitada até o próximo sítio de *splicing*, que estará localizado na porção terminal do íntron seguinte. Em 1990 GERAGHTY et al. já haviam descrito esta variação, entretanto, este alelo foi caracterizado “oficialmente” em 2004 por HE et al. em uma amostra da população chinesa, de etnia denominada Han, através da técnica de PCR por transcrição

reversa, ou RT-PCR (*Reverse Transcription-PCR*). Estes resultados foram consistentes com um estudo prévio no qual havia sido detectada a redução no tamanho da cauda citoplasmática da proteína HLA-F (LURY, EPSTEIN e HOLMES, 1990). Em 2006 PYO et al. também identificaram esta variação do gene *HLA-F* por sequenciamento, todavia, em uma amostra de origem afro-americana.

Outros polimorfismos, incluindo inserções e deleções, foram identificados neste gene, porém, as modificações que serão mais enfatizadas serão as não sinônimas por alterarem a estrutura proteica e potencialmente a função da molécula.

Três novas variantes codificadoras foram identificadas por PYO et al. em 2006, das quais todas continham substituições não sinônimas (figura 9). A variante *HLA-F*01:02* apresentou uma substituição de C (Ala) por T (Val) na sequência sinal situada no éxon 1, posição 38, código 13, enquanto o alelo *HLA-F*01:03* e *HLA-F*01:04* apresentaram as substituições de T (Ser) na posição 814 (código 272) foi trocada por uma C (Pro) no éxon 4, e uma substituição de C (Pro) para A (Gln) no éxon 2, posição 212, código 71, respectivamente. A tabela 2 reúne os polimorfismos em éxons descritos até então para o gene *HLA-F*.

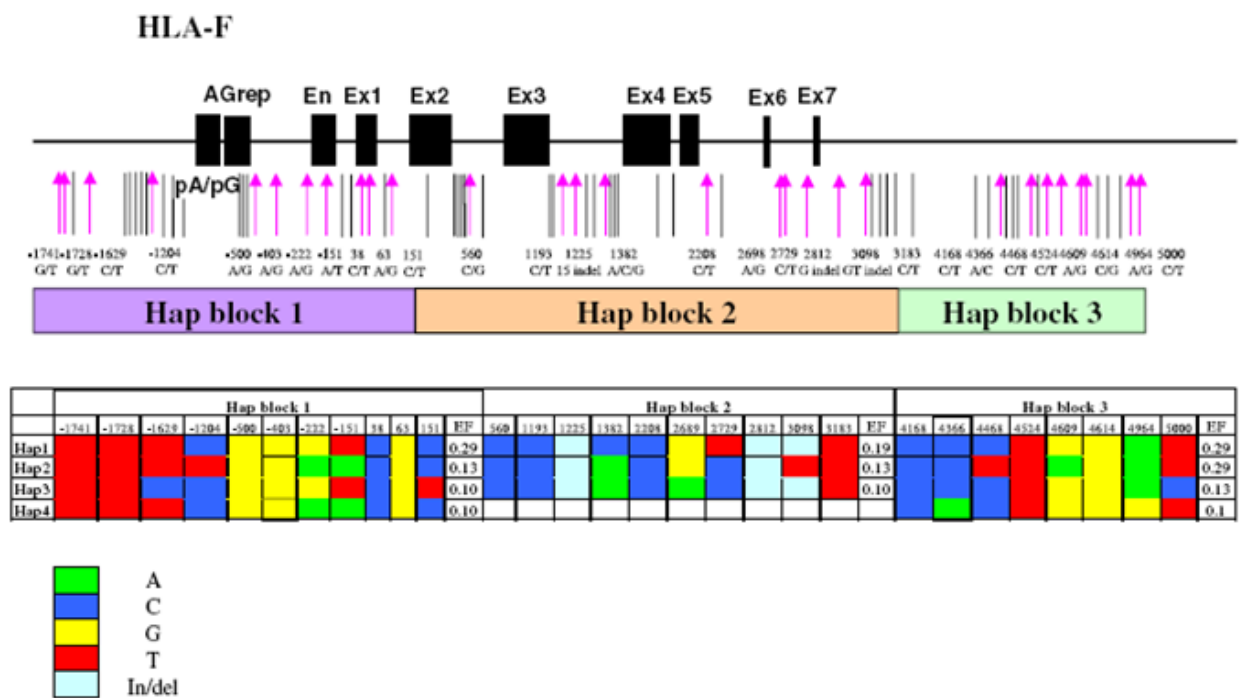


Figura 9. Representação do gene *HLA-F*. As estruturas éxon-ínton estão indicadas pelos retângulos pretos. Imediatamente abaixo, linhas verticais e flechas indicam as posições dos SNPs identificados. Abaixo das linhas e flechas estão indicadas as posições das variantes alélicas, sendo as posições designadas de acordo com a convenção aceita. Blocos haplotípicos estão representados logo abaixo das variantes alélicas. Fonte: PYO et al., 2006.

Tabela 2. Alelos codificantes de proteínas do gene *HLA-F* (exceto *HLA-F*Nulo*). Fonte: Moscoso et al., 2006. Adaptado.

	Éxon 1	Éxon 2	Éxon 3	Éxon 4
<i>HLA-F*01:01</i>	GCC (Ala)	CCG (Pro)	TAC (Tyr)	TCT (Ser)
<i>HLA-F*01:02</i>	GTC (Val)			
<i>HLA-F*01:03</i>				CCT (Pro)
<i>HLA-F*01:04</i>		CAG (Gln)		
<i>HLA-F*Nulo</i>			TAA (Parada)	

Alguns estudos também reportaram associações entre SNPs em regiões próximas ao gene *HLA-F* e patologias. Podemos usar como exemplo o estudo realizado em 2009 por PARK et al. o qual identificou um polimorfismo, associado ao desenvolvimento de distrofia corneal tipo Avellino, que está em alto desequilíbrio de ligação com o gene *HLA-F*. Os autores do estudo realizado em 551 indivíduos de origem coreana e vietnamita ainda sugeriram que o gene *HLA-F* seja um forte candidato associado ao desenvolvimento da doença. Em 2009 TSE et al. encontraram evidências de que dois polimorfismos possam estar altamente associados ao risco de desenvolvimento de carcinoma nasofaríngeo. Estes polimorfismos estão localizados a montante da região reguladora do gene *HLA-F* e estão fortemente ligados ao alelo *HLA-G*01:05N*. Em 2004 NOGUCHI et al. investigaram a expressão de genes *HLA* de classe I não clássicos em células tumorais. Verificou-se a expressão dos genes *HLA-E* e *HLA-F*, entretanto, a expressão de *HLA-F* foi considerada mais específica, estando presente no soro de quase 62% de pacientes acometidos por câncer e ausente nas amostras controle. No mesmo estudo ainda foi sugerido que anticorpos anti-*HLA-F* deveriam ser considerados como possíveis marcadores tumorais, ajudando no diagnóstico de câncer (NOGUCHI et al., 2004). LIN et al. (2011) ainda revelaram que a expressão do *HLA-F* em pacientes com câncer de pulmão de células não pequenas seria um prognóstico desfavorável. A média de sobrevida de pacientes que expressaram a proteína *HLA-F* foi de 10 meses, contrastando em 7 meses a menos quando comparados com pacientes que não apresentaram a expressão deste gene. Análises multivariadas revelaram que o *HLA-F* pode ser um prognóstico independente indicador em pacientes com câncer de pulmão de células não pequenas (LIN et al., 2011). Outro dado relevante foi apontado ainda em relação ao

gene *HLA-F* e câncer. Quando comparada com a frequência do alelo *HLA-F*01:04* na população chinesa, a frequência deste alelo foi significativamente alta (4,3%) em pacientes acometidos por hepatite B com carcinoma hepatocelular. Nenhum indivíduo tipado era homocigoto para este alelo. Também foi observado forte desequilíbrio de ligação entre SNPs da região codificadora do gene *HLA-G* e *HLA-F*, assim como entre *HLA-E* e a região 3' não traduzida (3' UTR – 3' *Untranslated Region*) de *HLA-G*, sugerindo que *HLA-G*01:05N* está associado com certos alelos de *HLA-E* e *HLA-F*. A frequência observada do alelo *HLA-F*01:01*, na população chinesa saudável de etnia Han, foi de 97,3%. Por outro lado, as frequências dos alelos *HLA-F*01:03* e *HLA-F*01:04* foram de 1,9% e 0,8%, respectivamente. Entretanto, não houve registro da presença do alelo *HLA-F*01:02* nesta população (ZHANG et al., 2012). O gene *HLA-F* também foi relacionado com a Doença de Behçet. Através de um estudo de associação de extensão genômica (GWAS – *Genome-Wide Association Study*) envolvendo 1.215 casos desta patologia e 1.278 controles, todos provenientes da Turquia. Foi possível identificar SNPs candidatos e possíveis vias (*ICSNPathway – Identify Candidate casual SNPs and Pathways*), fazendo parte do grupo de candidatos dois SNPs do gene *HLA-F* (LEE et al., 2012).

1.7. O *HLA-F* e a evolução do MHC

Através de análises de matrizes de pontos (*Dot Matrix*) usando uma sequência *HLA* de classe I composta por 1.796.938 pares de base contra ela mesma, foi possível revelar várias duplicações em grande escala. Análise de regiões homólogas e orientação gênica forneceram dados que sugerem fortemente que sucessivas duplicações segmentares deram origem ao presente sistema *HLA* de classe I. Onze genes *HLA* de classe I, entre eles o *HLA-F*, além de estarem orientados da mesma forma, também apresentam homologia de 8 a 20 kb a montante de suas sequências. Por análises filogenéticas de genes *HLA* de classe I e *MIC* (*MHC class I Chain-Related*) foi possível inferir a cinética das sete rodadas de duplicações sucessivas que originou o atual MHC. As evidências indicam que inicialmente os genes ancestrais de *HLA-F* e *MICE* ramificaram-se de um grupo principal de genes. Estes genes, que estão fisicamente próximos um ao outro, possivelmente representam genes ancestrais dos genes *HLA* de classe I e *MIC*. De acordo com o modelo proposto, dendrogramas indicam que a duplicação do gene

ancestral de *HLA-F* deu origem aos genes *MICE* e *HLA-G*. Como consequência, uma unidade básica de genes foi gerada, posteriormente, duas duplicações segmentares independentes desta unidade geraram os segmentos *HLA-A-MICF* e *MICA-HLA-B*. Após única duplicação o último segmento originou *MICB* e *HLA-C*. A partir da duplicação de parte do segmento *HLA-A-HCGIV-5* originou-se *HLA-W-HCGIV-4*. De maneira similar, quatro duplicações subsequentes levaram à organização atual da região HLA de classe I. Outras evidências também foram levadas em consideração nesta hipótese, como a ordem cronológica de geração de membros das subfamílias *Alu*, *LINE* e *LTR*, servindo como relógio molecular para definir o tempo de geração dos segmentos de interesse (SHIINA et al., 1999). KULSKI, ANZAI e INOKO (2005) sugeriram uma distinta hipótese da evolução do MHC/*HLA* também envolvendo o gene *Patr/HLA-F* (*Patr* refere-se ao MHC da espécie *Pan troglodytes*). Esta hipótese implica na herança de grupos segmentares de genes, ou duplicons, de uma espécie ancestral como um par estrutural AB. Inicialmente ERV-16 foi inserido no duplicon progenitor do MHC de classe I/*MIC* do par estrutural AB. O gene *MIC* foi ligado à categoria B mas não à categoria A. Alguns transposons e retrotransposons coevoluíram com as categorias ancestrais de duplicons A e/ou B. *AluJ1* foi inserido na categoria ancestral A e *AluJ3* na categoria B aproximadamente há 50-100 milhões de anos. Através de uma duplicação em tandem da categoria B, foi gerado o duplicon AB'B com inserção de *AluJ4* no duplicon B. Outra duplicação em tandem, porém seguida de deleção parcial, foi gerado o duplicon BAB'B, contendo *AluJ4* em ambos duplicons B mas não em B'. O grupo segmentar ancestral F que abrigara o gene *Patr/HLA-F* pode ter emergido nesta época como o primeiro produto de duplicação ortólogo fixado que não mais participou de duplicações subsequentes. Neste contexto, o gene *Patr/HLA-F* pode ser visto como o primeiro a ser fixado, ou seja, o mais antigo gene que não sofreu duplicações posteriores em um grupo de genes do MHC de primatas. Outros eventos subsequentes, como duplicações em bloco, contribuíram posteriormente para a atual disposição do sistema MHC. Também foi ressaltado que esta hipótese é mais simples e parcimoniosa que a anterior por inferir que ocorreram apenas quatro duplicações em tandem e não ocorreu transposições, ao contrário da hipótese anterior que sugeriu sete duplicações e três transposições (KULSKI, ANZAI e INOKO, 2005).

2. Objetivos

2.1. Objetivo Geral

O objetivo geral do trabalho proposto foi descrever as frequências alélicas e genotípicas do gene *HLA-F* em uma amostra composta por 201 indivíduos doadores voluntários de medula óssea, sendo estes autodenominados euro-brasileiros.

2.2. Objetivos Específicos

- Desenvolver oligonucleotídeos iniciadores específicos para a correta amplificação e sequenciamento das regiões de interesse;
- Otimizar protocolos de PCR e sequenciamento para a tipagem de indivíduos para o gene *HLA-F*;
- Determinar as variantes polimórficas dos éxons 1, 2 e 4 do gene *HLA-F*;
- Estimar as frequências genotípicas e alélicas da amostra em questão;
- Averiguar a conformidade das distribuições genotípicas para o equilíbrio de Hardy-Weinberg;
- Comparar os resultados obtidos com aqueles relatados na literatura.

3. Justificativa

Com o passar dos anos houve um aumento no número de estudos que compreendem os genes *HLA* de classe Ib, o que permitiu o esclarecimento de distintas associações relacionadas às funções do sistema imune, antes não elucidadas. Todavia, a quantidade de estudos envolvendo os genes *HLA* de classe Ib ainda está muito inferior em relação à quantidade de estudos realizados acerca dos genes *HLA* de classe Ia. A compreensão dos mecanismos que regem às interações intergênicas deste complexo e suas consequências biológicas ainda não foram completamente elucidadas. Outro ponto de discussão que deve ser ressaltado é a escassez de dados referentes às frequências alélicas do gene *HLA-F* tanto na população brasileira quanto em outros grupos.

O atual estudo tem como principal alicerce o fato de haver um número consideravelmente baixo de trabalhos relatando polimorfismos do gene *HLA-F* bem como suas respectivas frequências alélicas. Outrossim, este estudo populacional é inédito no Brasil.

4. Artigo

Os tópicos Material e Métodos, Resultados e parte da Discussão serão apresentados a seguir em forma de artigo como um capítulo desta dissertação. O artigo a seguir está formatado de acordo com as normas da revista *Tissue Antigens*.

Short Title: ***HLA-F* polymorphisms in a Brazilian population**

***HLA-F* polymorphisms in a Euro-Brazilian population from Southern Brazil**

Luis Felipe Santos Manvailer, Pryscilla Fanini Wowk, Sibelle Botogosque Mattar, José Samuel da Siva, Maria da Graça Bicalho, Valéria Maria Munhoz Sperandio Roxo

Laboratory of Immunogenetics and Histocompatibility, Genetics Department, Universidade Federal do Paraná, Curitiba, Brazil.

Correspondence to Prof. Valéria Maria Munhoz Sperandio Roxo

Laboratório de Imunogenética e Histocompatibilidade,

Departamento de Genética, Universidade Federal do Paraná

R. Cel. Francisco H. dos Santos S/N

Centro Politécnico – Jardim das Américas

CP 19071, CEP 81.530-990, Curitiba–PR, Brasil

Tel: +55 41 3361 1729, Fax: +55 41 3266 2042

e-mail: valroxo@ufpr.br

The authors do not present potential conflicts of interest.

Abstract

HLA-F is a non-classical major histocompatibility complex (MHC) gene. It codes class Ib MHC molecules with restricted distribution and less nucleotide variations than MHC class Ia genes. Of the 22 alleles registered on the IMGT database only four alleles encode for proteins that differ in their primary structure. To estimate genotype and allele frequencies, this study targeted on known protein coding regions of the *HLA-F* gene. Genotyping was performed by Sequence-Based Typing (SBT). The sample was composed by 199-unrelated bone marrow donors from the Brazilian Bone Marrow Donor Registry (REDOME), Euro-Brazilians, from Southern Brazil. About 1673 bp were analyzed. The most frequent allele was *HLA-F*01:01* (87.19%), followed by *HLA-F*01:03* (12.31%), *HLA-F*01:02* (0.25%) and *HLA-F*01:04* (0.25%). Significant linkage disequilibrium (LD) was verified between *HLA-F* and *HLA* classes I and II alleles. This is the first study regarding *HLA-F* polymorphisms in a Euro-Brazilian population contributing to the Southern Brazilian genetic characterization.

Keywords: Euro-Brazilians, *HLA-F*, linkage disequilibrium, polymorphisms, population.

Introduction

The *HLA-F* gene, along with *HLA-E* and *HLA-G* genes, is a class Ib human major histocompatibility complex (MHC) gene and is the most telomeric among the human leukocyte antigen (HLA) genes. In comparison with the *HLA* classical genes, the non-classical ones have restricted distribution, substantially less nucleotide variations, limited protein variability, distinct immune response modulation and conspicuous structural features [1]. *HLA-F* is composed by eight exons and seven introns and its expression is mainly detected in the cytoplasm of peripheral blood B cells, as well as tissues such as tonsils in adults, thymus and fetal liver [2-4]. *HLA-F* tetramers bound to *ILT2* and *ILT4*, suggesting that the *HLA-F* molecule could possibly reach cell surface [5]. Ishitani et al. were the first to show *HLA-F* cell surface expression; this study also was the first to report the

expression of all three HLA non-classical molecules simultaneously [6]. Shobu et al. documented HLA-F surface expression on decidual trophoblasts [7], although other studies found conflicting results, detecting only intracellular HLA-F expression [8, 9]. When activated, CD4⁺CD25⁻ T cells show strong HLA-F expression on cell surface and this suggests that HLA-F has a potential use as marker for an activated immune response, for instance in graft transplantation procedures [10]. This HLA Ib molecule is completely independent from TAP although the trafficking between endoplasmic reticulum and cell surface is only possible by the addition of a C-terminal valine residue in this molecule's cytoplasmic tail [11, 12]. Additional data suggest that HLA-F is expressed independently of peptide [13]. It is also suggested that HLA-F can be involved in the stabilization and transport of MHC-I open conformer form (OC) to the surface, as well as from the surface, having a role in the antigen cross-presentation [14].

The first null allele of this gene, a non-sense mutation, was detected in the third *HLA-F* exon, resulting in a truncated molecule. It was observed only in heterozygosity and the genotype frequency was 0.56% in the Japanese population [15]. A 1243C/G transversion was detected in exon 2 which resulted in a synonymous substitution found in linkage disequilibrium (LD) with *HLA-A*31* [16]. In 1990, Lury et al. detected a size reduction in the HLA-F protein cytoplasmic tail [17]. This shortening is because of a very striking polymorphism in this gene which involves the splicing of the seventh exon due to a transition that changes the splicing site between introns 6 and exon 7 [2, 18, 19]. Three previously unreported polymorphisms were detected concerning exons 1, 2 and 4, determining three HLA-F allelic variants: *HLA-F*01:02*, *HLA-F*01:04* and *HLA-F*01:03*, respectively [19]. In the position +38, located in exon 1, a C/T transition is responsible for the definition of the *HLA-F*01:02* allele. In exon 4, position +1771, a T/C transition can be detected leading to the *HLA-F*01:03* allele. A transversion (C/A) in the second exon, at +342, characterizes the *HLA-F*01:04* allele.

Significant associations have been documented with Avellino corneal dystrophy, nasopharyngeal carcinoma, gastric adenocarcinoma and Behcet's disease [20-23]. Noguchi

et al. suggested that anti-HLA-F antibodies should be investigated as possible diagnostic tumor markers [24]. The expression of this gene was also associated with poor survival in both non-small-cell lung cancer and esophageal squamous cell carcinoma patients leading the authors to propose that HLA-F could be a prognostic indicator in patients with these types of cancer [25, 26].

In this study a sample composed of 199 Euro-Brazilians was genotyped, comprising approximately 1673 bp analyzed per individual. LD with the HLA-A, HLA-B and HLA-DRB1 genes was also calculated. The results obtained herein contribute to the scarce information in the literature regarding HLA-F, apart from supplying significant LD data.

Material and Methods

Sample

Genomic DNA, extracted from peripheral blood leukocytes by the *salting out* method, was obtained from 199 adult unrelated bone marrow donors belonging to the Brazilian Bone Marrow Donor Registry (REDOME), who live in Curitiba, the capital city of the southern state of Paraná (Figure 1). All participants provided their personal data and informed consent. These studies were approved by the Ethics Committee of UFPR's Clinical Hospital (HCUFPR/CEP-HC n037 ext. 019/2001-07). The ethnic classification was based on self-perception of respondents, as well as the information of the ethnic origin and composition of families involved. According to the 2010 IBGE (Instituto Brasileiro de Geografia e Estatística) Census, 78.9% of Curitiba residents are White people, 16.8% Pardo (Multiracial) people, 2.9% Black people (2.9%), 1.4% Asian people and 0.2% Amerindian people (0.2%). The population of Curitiba is mainly composed by European-Caucasoid descendants, resulted from the Paraná State government program of colonization established during XIX and XX century [27]. Data from Curitiba population showed a predominance of a Caucasian sample, as well as observed at North of Paraná and Rio Grande do Sul. The HLA allelic and haplotypes frequencies were related between these populations with a similar immigration history [28-30].

HLA-F Genotyping

Two pairs of primers were developed through the Primer3 Plus software [31] and dnastar package version 4.0 (DNASTAR, Madison, WI) in order to amplify the regions of interest. Sixteen SNPs located in three exons [1, 2, 4] and four introns [1, 2, 4, 5] comprising approximately 1673 bp. The first pair of primers (5' CGCAGTCCCAGGTTCTAAA 3') and (5' AGATGTAATCCTTGCCGTCGT 3') amplified a region comprising part of the promoter and the beginning of the third exon, resulting in 884 bp polymerase chain reaction (PCR) products. The second pair of primers (5' GGAGTGCAAAGTGCCTGAAT 3') and (5' AGGTCCCATCACCACCATAA 3') amplified part of intron 3 extending to part of intron 5, resulting in 789 bp PCR products. The Primer-BLAST (Basic Local Alignment Search Tool) [32] program was used to check the specificity of the probes developed. PCR was carried out according to the following conditions: final volume of 25 μ L containing 1 \times PCR buffer (20 mM Tris-HCl pH 8.4, 50 mM KCl), 37.5 mM MgCl₂, 0.45 mM of each dNTPs, 7.5 pMols of each primer, 2.25 U of Taq DNA polymerase Platinum (Invitrogen, Carlsbad, CA) and 100 ng of genomic DNA. The cycling program was: 94°C for 5 min, 94°C for 30 s, 60°C (for first pair of primers) and 63°C (second pair of primers) for 30 s, 72°C for 3 min, the last three steps were repeated 35 times, with a final extension step at 72°C for 5 min (GeneAmp PCR System 9700, Applied Biosystems, Foster City, California, USA). The PCR products obtained were purified with Exonuclease I and Shrimp Alkaline Phosphatase (USB Corporation, Staufien, Germany), followed by a sequencing reaction carried out with ABI Prism Large Dye Terminator Cycle Sequencing Kit v.3.1 (Applied Biosystems, Foster City, CA). The sequencing products were eluted via electrophoresis using an ABI3130 Genetic Analyzer (Applied Biosystems). The primers employed for direct DNA sequencing were the same as those used for PCR amplification. Genotyping was performed by Sequence-Based Typing (SBT). DNA sequences were analyzed using BioEdit version 7.0.9.0. Genotype and allele frequencies were obtained by direct counting.

HLA-A, HLA-B and HLA-DRB1 Genotyping

HLA low-resolution typing for the HLA-A, -B and -DRB1 was carried out with polymerase chain reaction-sequence specific oligonucleotide probes (PCR-SSOP) using the LABType® SSOP Kits (One Lambda, Canoga Park, CA) according to the manufacturer's instructions.

Haplotype inference and Linkage Disequilibrium (LD)

Hardy–Weinberg equilibrium, haplotype inference and LD parameters were calculated using the Software Arlequin 3.5 [33]. The Hardy–Weinberg Equilibrium was tested with exact test using default parameters, ELB algorithm were used to phase reconstruction of haplotypes and D, D' and r2 parameters were calculated for LD.

Results

Through the analysis of the DNA sequences, 199 individuals were genotyped concerning the four HLA-F protein coding alleles. 152 of 199 individuals were homozygotes for the *HLA-F*01:01* allele whereas 42 were heterozygotes *HLA-F*01:01/01:03* and 3 presented the transition T/C in position +1771 in both phases, characterizing the homozygous *HLA-F*01:03/01:03* genotype. The alleles *HLA-F*01:02* and *HLA-F*01:04* were also detected in this sample in only two individuals whose alleles were in heterozygosity with *HLA-F*01:01* and *HLA-F*01:03*, respectively. All four HLA-F protein coding alleles were detected in the Euro-Brazilian sample of this study. *HLA-F*01:01* was the most frequent allele (87.19%), followed by *HLA-F*01:03* (12.31%), *HLA-F*01:02* (0.25%) and *HLA-F*01:04* (0.25%). The alleles and genotypes frequencies are summarized in Table 1. From the Euro-Brazilians allele frequencies obtained, it was possible to compare those frequencies with previous data from Chinese populations (Table 2).

All *loci* studied were in Hardy-Weinberg Equilibrium ($P > 0.05$). The phase reconstruction, two by two, between *HLA-F* and *HLA-A*, *HLA-B* and *HLA-DRB1* are depicted in Table 3. A graphical overview of LD between *HLA-F* and *HLA-A*, *HLA-B* and *HLA-DRB1* is

represented in Figure 2. The *HLA-F*01:01* allele is strongly linked to *HLA-A*02* while *HLA-F*01:03* is strongly associated to *HLA-A*03* and in repulsion with *HLA-A*02*. The *HLA-F*01:02* is in strong LD with *HLA-A*34*, *HLA-B*35* and *HLA-DRB1*08*; also in strong LD are *HLA-B*14* with *HLA-F*01:04*.

Discussion

Although the *HLA-F* gene has been increasingly studied in the last decade, much still remains unclear about this gene. The use of data such as composition and nucleotide variations, genotype frequencies, allele frequencies and LD in different populations are of utmost importance to understand the demographic and evolutionary history of populations. Such inferences, together with gene expression studies, provide new paths to the understanding of the exact role of the *HLA-F* gene and its product in the immunologic scenario.

Concerning nucleotide variations, these should not be analyzed only under the molecular perspective, as basis of evolutionary inferences, but should also be considered in association studies of pathologic conditions. Zhang et al. verified a significant difference in the *HLA-F*01:04* allele frequency in patients affected by hepatitis B and hepatocellular carcinoma (4.3%) compared to healthy Han Chinese individuals (0.8%); odds ratio (OR) [95% confidence interval (CI)] = 5.87 (1.28 ± 26.78) [34]. In this study population, concerning Euro-Brazilians, a distinct frequency was obtained for *HLA-F*01:04* (0.25%) while in other four Chinese populations this allele was not detected [35]. The allele frequencies of Euro-Brazilian and Chinese populations are depicted in Table 2. Considering these frequencies, Euro-Brazilian population differs from all the aforementioned Chinese populations, using Exact Test of population differentiation implemented at Arlequin 3.5 [33] and significance level at 0.05. At the moment there are no other population studies focusing on the *HLA-F* gene, thus, it is only possible to compare the present results with those found in the aforementioned populations.

The differences found in the allele frequencies are expected considering the hypothesis of the colonization of the American continent through the Bering Strait followed by subsequent migrations to the Southern regions [36]. Nearby populations are expected to be genetically more similar than distant ones. However, this scenario can be different in Brazil due to the high degree of admixture arising from the migratory waves which specially took place in the last century, resulting in high gene flow mainly from German, Spanish, Italian, Ukrainian and Japanese immigrants to the southern regions of Brazil and especially to the state of Paraná.

Despite the absence of such data, the genotype and allele frequencies found in this study are expected to be similar to the respective values of European populations, especially with those whose contribution was important to the ethnic composition of this state.

The verification of LD between pairs of *HLA-F* and *HLA-A*, *HLA-B* and *HLA-DRB1* alleles provides additional interesting information which might benefit future association studies. It is possible to have a sample-size effect associated to the LD involving both *HLA-F*01:02* and *HLA-F*01:04* alleles, as they were detected only once in separate individuals.

This is the first study to characterize nucleotide composition of the *HLA-F* gene and to estimate allele and genotype frequencies in a Euro-Brazilian population sample, contributing to the characterization and the study of the *HLA-F* gene. The information presented here supply data for further comparisons which might be useful not only as basis for demographic and evolutionary inferences but also for association studies that might relate *HLA-F* as a genetic marker.

References

1. Donadi EA, Castelli EC, Arnaiz-Villena A, Roger M, Rey D, Moreau P. Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. *Cell Mol Life Sci* 2011; **68**: 369–95.
2. Geraghty DE, Wei XH, Orr HT, Koller BH. Human leukocyte antigen F (HLA-F). An expressed HLA gene composed of a class I coding sequence linked to a novel transcribed repetitive element. *J Exp Med* 1990; **171**: 1–18.

3. Houlihan JM, Biro PA, Fergar-Payne A, Simpson KL, Holmes CH. Evidence for the expression of non-HLA-A,-B,-C class I genes in the human fetal liver. *J Immunol* 1992; **149**: 668–75.
4. Wainwright SD, Biro PA, Holmes CH. HLA-F is a predominantly empty, intracellular, TAP-associated MHC class Ib protein with a restricted expression pattern. *J Immunol* 2000; **164**: 319–28.
5. Lepin EJ, Bastin JM, Allan DS et al. Functional characterization of HLA-F and binding of HLA-F tetramers to ILT2 and ILT4 receptors. *Eur J Immunol* 2000; **30**: 3552–61.6.
6. Ishitani A, Sageshima N, Lee N et al. Protein expression and peptide binding suggest unique and interacting functional roles for HLA-E, F, and G in maternal-placental immune recognition. *J Immunol* 2003; **171**: 1376–84.
7. Shobu T, Sageshima N, Tokui H et al. The surface expression of HLA-F on decidual trophoblasts increases from mid to term gestation. *J Reprod Immunol* 2006; **72**: 18–32.
8. Nagamatsu T, Fujii T, Matsumoto J, Yamashita T, Kozuma S, Taketani Y. Human leukocyte antigen F protein is expressed in the extra-villous trophoblasts but not on the cell surface of them. *Am J Reprod Immunol* 2006; **56**: 172–7.
9. Apps R, Gardner L, Traherne J, Male V, Moffett A. Natural-killer cell ligands at the maternal-fetal interface: UL-16 binding proteins, MHC class-I chain related molecules, HLA-F and CD48. *Hum Reprod* 2008; **23**: 2535–48.
10. Lee N, Ishitani A, Geraghty DE. HLA-F is a surface marker on activated lymphocytes. *Eur J Immunol* 2010; **40**: 2308–18.
11. Lee N, Geraghty DE. HLA-F surface expression on B cell and monocyte cell lines is partially independent from tapasin and completely independent from TAP. *J Immunol* 2003; **171**: 5264–71.
12. Boyle LH, Gillingham AK, Munro S, Trowsdale J. Selective export of HLA-F by its cytoplasmic tail. *J Immunol* 2006; **176**: 6464–72.
13. Goodridge JP, Burian A, Lee N, Geraghty DE. HLA-F complex without peptide binds to MHC class I protein in the open conformer form. *J Immunol* 2010; **184**: 6199–208.
14. Goodridge JP, Lee N, Burian A et al. HLA-F and MHC-I open conformers cooperate in a MHC-I antigen cross-presentation pathway. *J Immunol* 2013; **191**: 1567–77.
15. Uchigiri C, Mizuno S, Wada K et al. An identification of the *HLA-F* null allele in Japanese. *Immunogenetics* 1997; **45**: 466–7.
16. Kunishima S, Nagae M, Mizuno S, Kamiya T, Ozawa K. A new polymorphism in the *HLA-F* gene (67Ala[GCC] to Ala[GCG]). *Immunogenetics* 1999; **49**: 147–8.
17. Lury D, Epstein H, Holmes N. The human class I MHC gene *HLA-F* is expressed in lymphocytes. *Int Immunol* 1990; **2**: 531–7.

18. He X, Xu L, Liu Y, Zeng Y. Identification of a novel *HLA-F* allele – HLA-F*010102. *Tissue Antigens* 2004; **63**: 181–3.
19. Pyo CW, Williams LM, Moore Y et al. *HLA-E*, *HLA-F*, and *HLA-G* polymorphism: genomic sequence defines haplotype structure and variation spanning the nonclassical class I genes. *Immunogenetics* 2006; **58**: 241–51.
20. Park M, Kim DJ, Kim KJ et al. Genetic associations of common deletion polymorphisms in families with Avellino corneal dystrophy. *Biochem Biophys Res Commun* 2009; **387**: 688–93.
21. Tse KP, Su WH, Chang KP et al. Genome-wide association study reveals multiple nasopharyngeal carcinoma-associated loci within the HLA region at chromosome 6p21.3. *Am J Hum Genet* 2009; **85**: 194–203.
22. Ishigami S, Arigami T, Setoyama T et al. Clinical-pathological implication of human leukocyte antigen-F-positive gastric adenocarcinoma. *J Surg Res* 2013; **184**(2): 802–6.
23. Lee YH, Choi SJ, Ji JD, Song GG. Genome-wide pathway analysis of a genome-wide association study on psoriasis and Behcet's disease. *Mol Biol Rep* 2012; **39**: 5953–9.
24. Noguchi K, Isogai M, Kuwada E, Noguchi A, Goto S, Egawa K. Detection of anti-*HLA-F* antibodies in sera from cancer patients. *Anticancer Res* 2004; **24**: 3387–92.
25. Lin A, Zhang X, Ruan YY, Wang Q, Zhou WJ, Yan WH. *HLA-F* expression is a prognostic factor in patients with non-small-cell lung cancer. *Lung Cancer* 2011; **74**: 504–9.
26. Zhang X, Lin A, Zhang JG et al. Alteration of *HLA-F* and *HLA I* antigen expression in the tumor is associated with survival in patients with esophageal squamous cell carcinoma. *Int J Cancer* 2013; **132**: 82–9.
27. Wowk PF, Mattar SB, Poerner F, Tomaz de Rezende GY, da Graça Bicalho M. Two novel alleles, *HLA-B*40:125* and *B*40:129*, in the Brazilian population. *Tissue Antigens* 2012; **79**: 137–8.
28. Ruiz TM, da Costa SM, Ribas F, Luz PR, Lima SS, da Graça BM. Human leukocyte antigen allelic groups and haplotypes in a brazilian sample of volunteer donors for bone marrow transplant in Curitiba, Paraná, Brazil. *Transplant Proc* 2005; **37**: 2293–6.
29. Bardi MS, Jarduli LR, Jorge AJ et al. *HLA-A*, *B* and *DRB1* allele and haplotype frequencies in volunteer bone marrow donors from the north of Parana State. *Rev Bras Hematol Hemoter* 2012; **34**: 25–30.
30. Bortolotto AS, Petry MG, da Silveira JG et al. *HLA-A*, *-B*, and *-DRB1* allelic and haplotypic diversity in a sample of bone marrow volunteer donors from Rio Grande do Sul State, Brazil. *Hum Immunol* 2012; **73**: 180–5.
31. Untergasse A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* 2007; **35** (Web Server Issue): W71–4.

32. Ye J, Coulouris G, Zaretskava I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 2012; **13**: 134.
33. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 2010; **10**: 546–67.
34. Zhang J, Pan L, Chen L, Feng X, Zhou L, Zheng S. Non-classical MHC-lota genes in chronic hepatitis B and hepatocellular carcinoma. *Immunogenetics* 2012; **64**: 251–8.
35. Pan FH, Liu XX, Tian W. Characterization of the *HLA-F* polymorphism in four distinct populations in Mainland China. *Int J Immunogenet* 2013: 369–76.
36. Wang S, Lewis CM, Jakobsson M et al. Genetic variation and population structure in native Americans. *PLoS Genet* 2007; **3**: e185.

Figure 1. Geographic location of Curitiba, the capital city of Paraná state. The maps are not drawn to scale. The geographical coordinates of Curitiba are $25^{\circ}25'42''$ S, $49^{\circ}16'23''$ W.



Table 1. *HLA-F* allele and genotype frequencies in the Euro-Brazilian population.

Alleles	Frequencies	Genotypes (<i>HLA-F</i> *)	AF (RF)
<i>HLA-F</i> *01:01	0.8719	<i>01:01/01:01</i>	152 (0.7638)
<i>HLA-F</i> *01:02	0.0025	<i>01:01/01:03</i>	42 (0.2110)
<i>HLA-F</i> *01:03	0.1231	<i>01:03/01:03</i>	3 (0.0151)
<i>HLA-F</i> *01:04	0.0025	<i>01:03/01:04</i>	1 (0.0050)
		<i>01:01/01:02</i>	1 (0.0050)

AF (absolute frequencies), RF (relative frequencies)

Table 2. Comparison among Euro-Brazilian and Chinese *HLA-F* allele frequencies.

Alleles	Euro-Brazilian (n= 199)	Han Chinese* (n= 520)	HNH** (n= 319)	IMH** (n= 104)	IMM** (n= 160)	GDH** (n= 107)
<i>HLA-F</i> *01:01	0.8719	0.9730	0.9890	0.9423	0.9469	0.9766
<i>HLA-F</i> *01:02	0.0025	0.0000	0.0000	0.0000	0.0000	0.0000
<i>HLA-F</i> *01:03	0.1231	0.0190	0.0110	0.0577	0.0469	0.0234
<i>HLA-F</i> *01:04	0.0025	0.0080	0.0000	0.0000	0.0063	0.0000

*According to Zhang J, Pan L, Chen L, Feng X, Zhou L, Zheng S., 2012 (34).

**According to Pan FH, Liu XX, Tian W., 2013 (35).

HNH, Hunan Han population; IMH, Inner Mongolia Han population; IMM, Inner Mongolia Mongol population; GDH, Guandong Han population.

Table 3. Haplotype frequencies of *HLA-F-A*, *-B* and *-DRB1* in Euro-Brazilians.

Abs Freq.	Rel Freq.	Haplotype <i>HLA-A/HLA-F</i>	Abs Freq.	Rel Freq.	Haplotype <i>HLA-B/HLA-F</i>	Abs Freq.	Rel Freq.	Haplotype <i>HLA-DRB1/HLA-F</i>
133	33.42%	*02/*01:01	43	10.80%	*35/*01:01	57	14.32%	*11/*01:01
43	10.80%	*01/*01:01	42	10.55%	*51/*01:01	46	11.56%	*04/*01:01
27	6.78%	*03/*01:03	38	9.55%	*15/*01:01	45	11.31%	*13/*01:01
27	6.78%	*11/*01:01	37	9.30%	*44/*01:01	40	10.05%	*07/*01:01
25	6.28%	*31/*01:01	26	6.53%	*18/*01:01	33	8.29%	*15/*01:01
20	5.03%	*24/*01:01	15	3.77%	*39/*01:01	33	8.29%	*03/*01:01
19	4.77%	*32/*01:01	15	3.77%	*08/*01:01	23	5.78%	*16/*01:01
18	4.52%	*68/*01:01	13	3.27%	*40/*01:01	22	5.53%	*01/*01:01
14	3.52%	*23/*01:03	13	3.27%	*07/*01:03	17	4.27%	*08/*01:01
13	3.27%	*33/*01:01	13	3.27%	*07/*01:01	15	3.77%	*01/*01:03
11	2.76%	*26/*01:01	12	3.02%	*14/*01:01	12	3.02%	*07/*01:03
11	2.76%	*29/*01:01	11	2.76%	*38/*01:01	12	3.02%	*14/*01:01
9	2.26%	*30/*01:01	10	2.51%	*27/*01:01	8	2.01%	*13/*01:03
7	1.76%	*66/*01:01	9	2.26%	*57/*01:01	8	2.01%	*04/*01:03
6	1.51%	*24/*01:03	9	2.26%	*58/*01:01	8	2.01%	*12/*01:01
6	1.51%	*25/*01:01	9	2.26%	*49/*01:01	6	1.51%	*10/*01:01
4	1.01%	*03/*01:01	8	2.01%	*44/*01:03	5	1.26%	*09/*01:01
1	0.25%	*02/*01:04	7	1.76%	*14/*01:03	2	0.50%	*14/*01:03
1	0.25%	*30/*01:03	7	1.76%	*50/*01:01	2	0.50%	*16/*01:03
1	0.25%	*11/*01:03	6	1.51%	*55/*01:01	1	0.25%	*13/*01:04
1	0.25%	*74/*01:01	6	1.51%	*53/*01:01	1	0.25%	*11/*01:03
1	0.25%	*34/*01:02	6	1.51%	*52/*01:01	1	0.25%	*08/*01:02
			5	1.26%	*48/*01:01	1	0.25%	*15/*01:03
			4	1.01%	*45/*01:03			
			4	1.01%	*37/*01:01			
			4	1.01%	*41/*01:01			
			3	0.75%	*50/*01:03			
			3	0.75%	*57/*01:03			
			3	0.75%	*35/*01:03			
			2	0.50%	*13/*01:03			
			2	0.50%	*49/*01:03			
			2	0.50%	*13/*01:01			
			2	0.50%	*38/*01:03			
			2	0.50%	*73/*01:01			
			1	0.25%	*14/*01:04			
			1	0.25%	*56/*01:01			
			1	0.25%	*39/*01:03			
			1	0.25%	*58/*01:03			
			1	0.25%	*35/*01:02			
			1	0.25%	*42/*01:01			
			1	0.25%	*45/*01:01			

Note: ELB Algorithm was used for phase reconstruction.

Abs Freq. (absolute frequencies), Rel Freq. (relative frequencies).

5. Discussão

Apesar de nos últimos anos ter aumentado consideravelmente o número de trabalhos envolvendo o gene *HLA-F*, muito ainda tem a ser revelado sobre este gene. A utilização de dados como composição e variações nucleotídicas, assim como, as respectivas frequências em diferentes populações, é de extrema relevância no entendimento da história demográfica e evolutiva das populações. Tais inferências, agregadas a estudos de expressão gênica, podem abrir novos caminhos que levem a favor do entendimento da função do gene em questão.

Ainda quanto às variações nucleotídicas, estas não apenas devem ser analisadas sob a ótica de relógio molecular, como base para inferências evolutivas, mas também devem ser extendidas a investigações de associação a condições patológicas. Zhang et al., 2012 verificaram uma diferença significativa na frequência do alelo *HLA-F*01:04* em indivíduos acometidos por hepatite crônica e carcinoma hepatocelular (4,3%) em relação à frequência obtida na população chinesa saudável (0,8%). No presente trabalho, que teve a amostra composta por euro-brasileiros, foram obtidas frequências distintas para o mesmo alelo (0,25%). Contudo, outro dado mostrou-se semelhante em ambas as populações, o alelo *HLA-F*01:04* somente foi detectado em heterozigose. A tabela 3 ainda ressalta as diferenças encontradas nas frequências alélicas entre amostras das populações euro-brasileiras e chinesas de etnia Han.

Tabela 3. Comparação das frequências alélicas entre amostras de diferentes populações ($F_{ST} = 0.09400$; $P < 0.0001$)

Alleles	Frequency (Euro-Brazilians)	Frequency (Han Chinese)
<i>HLA-F*01:01</i>	0.8719	0.9730
<i>HLA-F*01:02</i>	0.0025	0.0000
<i>HLA-F*01:03</i>	0.1231	0.0190
<i>HLA-F*01:04</i>	0.0025	0.0080

A divergência encontrada nas frequências alélicas entre as populações corroboram com a hipótese de MIDDLETON, MANAVALAN e FERNANDEZ-VINA, 2004 a qual afirma que populações com origens em comum tendem a compartilhar alelos característicos e a apresentarem frequências alélicas similares entre genes do

sistema HLA. Levando em consideração outra hipótese, fortemente fundamentada em dados genéticos e linguísticos, a qual sugere que a colonização do continente americano ocorreu em um único evento principal pelo estreito de Bering há milhares de anos com subseqüentes migrações rumo ao sudoeste e então sudeste do continente (WANG et al., 2007), espera-se que as populações localizadas nestas regiões apresentem mais semelhanças genéticas entre si do que entre populações distantes, como as européias, por exemplo. Entretanto, esse panorama pode ser diferente quando levado em consideração o fato da população brasileira ter como característica seu alto grau de miscigenação. Este atributo é oriundo das ondas migratórias decorrentes de séculos anteriores resultando em uma alta taxa de fluxo gênico, provavelmente diminuindo de maneira considerável as distâncias genéticas entre as populações. As ondas migratórias no estado do Paraná eram compostas principalmente por imigrantes alemães, espanhóis, italianos, ucranianos e japoneses. De acordo com o censo realizado pelo IBGE em 2000 a população de Curitiba está composta por 77,4% de brancos, 18,2% de pardos, 2,9% de negros e 1,4% de amarelos e indígenas. Com base nessas evidências, espera-se, portanto, que as proporções genotípicas e alélicas encontradas na amostra da população deste estudo não estejam distantes das respectivas proporções genotípicas de populações européias, principalmente das que representaram maior percentual de imigração no estado.

Ao longo da evolução dos vertebrados, os genes MHC que codificam para proteínas que apresentam peptídeos não-próprios para as células T representou uma vantagem seletiva decisiva para o sucesso evolutivo deste grupo. Nos humanos estes genes, *HLA*, além de participarem da resposta imune contra doenças infecciosas (DOHERTY e ZINKERNAGEL, 1975) também desempenham um importante papel na interação materno-fetal (HEDRICK e THOMSON, 1988). Geralmente os genes que fazem parte do sistema HLA apresentam grande diversidade alélica nas populações e alta taxa de heterozigosidade. Com isso surgiu a hipótese da vantagem dos heterozigotos, a qual sugere que esta seria resultante do reconhecimento de uma maior variedade de peptídeos provenientes de agentes patogênicos (DOHERTY e ZINKERNAGEL, 1975). Todavia, esta superioridade seria alcançada apenas na condição de coinfeção e os perfis de susceptibilidade dos patógenos devem ser opostos ou recíprocos (DOHERTY e ZINKERNAGEL, 1975; HUGHES e NEI, 1992; POTTS e SLEV, 1995; McCLELLAND, PENN e POTTS,

2003). Os resultados desta dissertação foram condizentes com a diversidade alélica esperada para o gene *HLA-F*, entretanto, quando os alelos são analisados se estão em condições de hetero ou homozigose, houve uma considerável diferença, prevalecendo a existência dos alelos desta amostra em homozigose. Ainda que a proteína HLA-F possa ser expressa sem a presença de peptídeo (GOODRIDGE et al., 2010) e sua função pode ser independente da apresentação de peptídeo aos linfócitos T (BOYLE et al., 2006), ainda há a possibilidade desta molécula participar do monitoramento de infecções no aparelho de Golgi (BOYLE et al., 2006). Entretanto, as frequências genotípicas e alélicas deste estudo devem ser refinadas, em um futuro trabalho, com o delineamento de novos pares de oligonucleotídeos iniciadores que permitam a análise de maior quantidade de SNPs e, portanto, resultando na determinação dos alelos em maior resolução.

Nas regiões analisadas também foi confirmada uma combinação de SNPs, em duas amostras, que ainda não foi descrita na literatura e, portanto, não foi possível definir os alelos desses indivíduos. Levando em consideração a escassez de estudos de caracterização populacional para o gene *HLA-F*, foi levantada a hipótese da presença de um novo alelo. Em relação ao possível novo alelo encontrado neste trabalho, a priori tende a ser resultado de conversão interalélica, envolvendo o segundo íntron e o quarto éxon do gene em questão. Esta é apenas uma inferência inicial baseada na combinação dos polimorfismos envolvidos e em relatos frequentes deste mecanismo para genes do sistema HLA (MADRIGAL et al., 1992; ZANGENBERG et al., 1995; PARHAM et al., 1997), todavia, análises mais aprofundadas serão realizadas para a confirmação do surgimento deste alelo e dos possíveis mecanismos envolvidos.

Como perspectiva futura, há a intenção do desenvolvimento de oligonucleotídeos iniciadores para aumentar a resolução das tipagens realizadas nesta amostra e posterior aplicação de testes de neutralidade com a finalidade de realizar inferências evolutivas com a menor margem de erro possível, pois será inserido na análise um maior número de variações nucleotídicas, como outros SNPs não analisados neste estudo e também *indels*.

Este estudo é o primeiro a caracterizar a composição nucleotídica das regiões analisadas do gene *HLA-F*, bem como o pioneiro no relato das frequências genotípicas e alélicas para este gene em uma amostra da população brasileira, contribuindo com a caracterização genotípica de euro-brasileiros, além de fornecer

dados para posteriores comparações que possam servir como base de inferências que ajudem a desvendar o papel do gene *HLA-F*. Outros estudos relacionando as frequências genotípicas e alélicas do gene *HLA-F* em diferentes populações, associada a uma varredura genética detalhada deste gene, devem revelar novos dados que possam contribuir para a elaboração de novas hipóteses que levem a uma melhor compreensão da função deste gene.

6. Referências Bibliográficas

<http://www.dnastar.com> - acessado em setembro de 2010.

<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi> - acessado em setembro de 2010.

http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHomeAd - acessado em março de 2011.

ADAMASHVILI, I.M.; KELLEY, R.E.; PRESSLY, T.; MCDONALD, J.C. Soluble HLA: patterns of expression in normal subjects, autoimmune diseases, and transplant recipients. **Rheumatol. Int.**, 25: 491–500, 2005.

ALAM, S.M.; TRAVERS, P.J.; WUNG, J.L.; NASHOLDS, W.; REDPATH, S.; JAMESON, S.C.; GASCOIGNE, N.R. T cell receptor affinity and thymocyte positive selection. London: **Nature**, 381: 616-620, 1996.

APARICIO, P.; ALONSO J.M.; TORIBIO M.L.; MARCOS M.A.R.; PEZZI L.; Martinez-A., C. Isolation and characterisation of (gamma, delta) CD4⁺ T cell clones derived from human fetal liver cells. **J. Exp. Med.** 170: 1009, 1989.

APPS, R.; GARDNER, L.; SHARKEY, A.M.; HOLMES, N.; MOFFETT A. A homodimeric complex of HLA-G on normal trophoblast modulates antigen-presenting cells via LILRB1. **Eur. J. Immunol.**, 37: 1924–1937, 2007.

APPS, R.; GARDNER, L.; SHARKEY, A.M.; HIBY, S.E.; MOFFETT, A. Conformation of human leucocyte antigen-C molecules at the surface of human trophoblast cells. **Immunology**, 124(3): 322-8, 2008(a).

APPS, R.; GARDNER, L.; TRATHERNE, J.; MALE, V.; MOFFETT, A. Natural-killer cell ligands at the maternal–fetal interface: UL-16 binding proteins, MHC class-I chain related molecules, HLA-F and CD48. **Human Reproduction**, 23(11): 2535–2548, 2008(b).

ARNAIZ-VILLENA, A.; DIMITROSKI, K.; PACHO, A.; MOSCOSO, J.; GÓMEZ-CASADO, E.; SILVERA-REDONDO, C.; VARELA, P.; BLAGOEVSKA, M.; ZDRAVKOVSKA, V.; MARTÍNEZ-LASO, J. HLA genes in Macedonians and the sub-Saharan origin of Greeks. **Tissue Antigens**, 57: 118-127, 2001.

AYED, K.; AYED-JENDOUBI, S.; SFAR, I.; LABONNE, M.P.; GEBUHRER, L. HLA class-I and HLA class-II phenotypic, gene and haplotypic frequencies in Tunisians by using molecular typing data. **Tissue Antigens**, 64: 520-532, 2004.

BALDWIN, A.S.; Jr.; SHARP, P.A. Binding of a nuclear factor to a regulator sequence in the promoter of the mouse H-2Kb class I major histocompatibility gene. **Mol. Cell. Biol.**, 7: 305, 1987.

BODMER, J.G.; MARSH, S.G.E.; ALBERT, E.D.; BODMER, W.F.; BONTROP, R.E.; CHANNON, D.; DUPONT, B.; ERLICH, H.A.; MACH, W.R.; PARHAM, P.; SASAZUKI, T.; SCHNEUDEUR, G.H.T.H.; STROMINGER, J.L.; SVEJGAARD, A.; TERASAKI, P.I. Nomenclature for factors of the HLA system. **Tissue Antigens.**, 46: 1-18, 1995.

BOYLE, L.H.; GILLINGHAM, A.K.; MUNRO, S.; TROWSDALE, J. Selective Export of HLA-F by Its Cytoplasmic Tail. **Journal of Immunology**, 176: 6464-6472, 2006.

BREUNING, M.H.; VAN DEN BERG-LOONEN, E.M.; BERNINI, L.F.; BIJLSMA, J.B.; VAN LOGHEM, E.; MEERA KHAN, P.; NIJENHUIS, L.E. Localization of HLA on the short arm of chromosome 6. **Hum Genet**, 37: 131-139, 1977.

BRILEY, G.P., HISSONG, M.A.; CHIU, M.L.; LEE, D.C. The carboxylterminal valine residues of proTGF α are required for its efficient maturation and intracellular routing. **Mol. Biol. Cell**, 8: 1619–1631, 1997.

CAMPBELL, R.D.; TROWSDALE, J. Map of the human MHC. **Immunol. Today** 14: 349-352, 1993.

CAVALLI-SFORZA, L.L. **Genes, Povos e Línguas**. São Paulo: Companhia das Letras, 2003.

DOHERTY P.C.; ZINKERNAGEL R.M. A biological role for the major histocompatibility antigens. **Lancet**, 1: 1406-1409, 1975.

EXCOFFIER, L.; LAVAL, G.; SCHNEIDER, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. **Evolutionary Bioinformatics Online** 1:47-50, 1995.

GATTI, E.; PIERRE, P. Understanding the cell biology of antigen presentation: the dendritic cell contribution. **Current Opinion in Cell Biology** 15: 468–473, 2003.

GERAGHTY, D.E.; XIAOHUA, W.; ORR, H.T.; KOLLER, B.H. Human leukocyte antigen F (HLA-F), an expressed HLA gene composed of a class I coding sequence linked to a novel transcribed repetitive element. **J. Exp. Med.**, 171: 1-18, 1990.

GERAGHTY, D.E.; KOLLER, B.H.; HANSEN, J.A.; ORR, H.T. The HLA class I gene family includes at least six genes and twelve pseudogenes and gene fragments. **Journal of Immunology**, 149: 1934-1946, 1992.

GERAGHTY, D.E. Structure of the HLA class I region and expression of its resident genes. **Immunology**, 5: 3-7, 1993.

GOBIN, S.J.; KEIJSERS, V.; VAN ZUTPHEN, M.; VAN DEN ELSEN, P.J. The role of enhancer A in the locus-specific transactivation of classical and nonclassical HLA class I genes by nuclear factor kappa B. **Journal of Immunology**, 161: 2276–2283, 1998.

GOBIN, S.J.; VAN DEN ELSEN, P.J. Transcriptional regulation of the MHC class Ib genes HLA-E, HLA-F, and HLA-G. **Hum. Immunol.**, 61: 1102–1107, 2000.

GOODRIDGE J.P.; BURIAN A.; LEE, N.; GERAGHTY, D.E. HLA-F complex without peptide binds to MHC Class I protein in the open conformer form. **J. Immunol.**, 184(11): 6199-208, 2010.

GUSSOW, D.; REIN, R.S.; MEIJER, L.; DE HOOG, W.; SEEMANN, G.H.A.; HOCHSTENBACH, F.M.; PLOEGH, H.L. Isolation, expression, and the primary structure of the HLACw1 and HLA-Cw2 genes: evolutionary aspects. **Immunogenetics**, 25: 313, 1987.

HAUPTMANN, G.; BAHRAM, S. Genetics of the central MHC. **Current Opinion on Immunology**, 16(5): 668-672, 2004.

HE, X.; XU, L.; LIU, Y.; ZENG, Y. Identification of a novel HLA-F allele – HLA-F*010102. **Tissue Antigens**, 63: 181-183, 2004.

HEDRICK P.W.; THOMSON G. Maternal-fetal interactions and the maintenance of HLA polymorphism. **Genetics**, 119: 205-212, 1988.

HORTON, R.; WILMING, L.; RAND, V.; LOVERING, R.C.; BRUFORD, E.A.; KHODIYAR, V.K.; LUSH, M.J.; POVEY, S.; TALBOT, C.C. Jr; WRIGHT, M.W.; WAIN, H.M.; TROWSDALE, J.; ZIEGLER, A.; BECK, S. Gene map of the extended human major Histocompatibility complex. **Nature Rev. Genet.**, 5: 889–899, 2004.

HOULIHAN, J.M.; BIRO, P.A.; FERGAR-PAYNE, A.; SIMPSON, K.L.; HOLMES, C.H. Evidence for the expression of non-HLA-A,-B,-C class I genes in the human fetal liver. **The Journal of Immunology**, 149(2): 668-675, 1992.

HUGHES, A.L.; NEI, M. Models of host-parasite interaction and MHC polymorphism. **Genetics**, 132: 863-864, 1992.

IODICE, L., SARNATARO, S.; BONATTI, S. The carboxyl-terminal valine is required for transport of glycoprotein CD8 α from the endoplasmic reticulum to the intermediate compartment. **J. Biol. Chem.**, 276: 28920–28926, 2001.

ISHITANI, A.; SAGESHIMA, N.; LEE, N.; DOROFEEVA, N.; HATAKE, K.; MARQUARDT, H.; GERAGHTY, D.E. Protein expression and peptide binding suggest unique and interacting functional roles for HLA-E, F, and G in maternal-placental immune recognition. **J. Immunol.**, 171: 1376-1384, 2003.

ISRAEL, A.; KIMURA, A.; KIERAN, M.; YANO, O.; KANELLOPOULOS, J.; LE BAIL, O.; KOURILSKY, P. A common positive trans-acting factor binds to enhancer sequences in the promoters of mouse H-2 and N2-microglobulin genes. **Proc. Natl. Acad. Sci. USA**, 84: 2653, 1987.

KIMURA, A.; ISRAEL, A.; LE BAIL, O.; KOURILSKY, P. Detailed analysis of the mouse H-2Kb promoter: enhancer-like sequences and their role in the regulation of class I gene expression. **Cell**, 44: 261, 1986.

KING, A.; ALLAN, D.S.; BOWEN, M; POWIS, S.J.; JOSEPH, S.; VERMA, S., HIBY, S.E.; McMICHAEL, A.J.; LOKE, Y.W.; BRAUD, V.M. HLA-E is expressed on trophoblast and interacts with CD94/NKG2 receptors on decidual NK cells. **Eur J Immunol.**, 30: 1623–1631, 2000.

KLEIN, J. Biology of the Mouse Histocompatibility Complex. **Springer-Verlag New York Inc.**, 620, 1975.

KLEIN, J.; SATO, A. The HLA system. First of two parts. Advances in immunology. **New England Journal of Medicine**, 343(10): 702-709, 2000.

KULSKI, J.K.; ANZAI, T.; INOKO, H. ERVK9, transposons and the evolution of MHC class I duplicons within the alpha-block of the human and chimpanzee. **Cytogenet Genome Res**, 110: 181–192, 2005.

KUNISHIMA, S.; NAGAE, M.; MIZUNO, S.; KAMIYA, T.; OZAWA, K. A new polymorphism in the *HLA-F* gene (67Ala[GCC] to Ala[GCG]). **Immunogenetics**, 49: 147-148, 1999.

LAHIRI, D.K.; NURNBERGER, J.I. JR. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. **Nucleic Studies**, Oxford, 19: 5444, 1991.

LAWLOR, D.A.; ZEMMOUR, J.; ENNIS, P.D.; PARHAM, P. Evolution of class-I MHC genes and proteins: from natural selection to thymic selection. **Annu. Rev. Immunol.**, 8: 23-63, 1990.

Le BOUTEILLER P. HLA-G: on the track of immunological functions. **Eur. J. Immunogenet.**, 24: 397-408, 1997.

LEE, N.; GERAGHTY, D. HLA-F Surface Expression on B Cell and Monocyte Cell Lines Is Partially Independent from Tapasin and Completely Independent from TAP. **Journal of Immunology**, 171: 5264-5271, 2003.

LEE, N.; ISHITANI, A.; GERAGHTY, D.E. HLA-F is a surface marker on activated lymphocytes. **Eur. J. Immunol.**, 40: 2308-2318, 2010.

LEE, Y.H.; CHOI, S.J.; JI, J.D.; SONG, G.G. Genome-wide pathway analysis of a genome-wide association study on psoriasis and Behcet's disease. **Mol Biol Rep**, 39: 5953-9, 2012.

LEPIN, E.J.; BASTIN, J.M.; ALLAN, D.S.; RONCADOR, G.; BRAUD, V.M.; MASON, D.Y.; VAN DER MERWE, P.A.; McMICHAEL, A.J.; BELL, J.I.; POWIS, S.H.; O'CALLAGHAN, C.A. Functional characterization of HLA-F and binding of HLA-F tetramers to ILT2 and ILT4 receptors. **Eur. J. Immunol.**, 30: 3552, 2000.

LIN, A.; ZHANG, X.; YAN-YUN, R.; QING, W.; WEN-JUN, Z.; WEI-HUA, Y. HLA-F expression is a prognostic factor in patients with non-small-cell-lung cancer. **Lung Cancer**, 74(3):504-9, 2011.

LOPEZ DE CASTRO, J.A.; BARBOSA, J.A.; KRANGEL, M.S.; BIRO, P.A.; STROMINGER, J.L. Structural analysis of the functional sites of class I HLA antigens. **Immunol. Rev.**, 85: 149, 1985.

LUO M., EMBREE J., RAMDAHIN S., NDINYA-ACHOLA J., NJENGA S., BWAYO J.B., PAN S., MAO X., CHEANG M., STUART T., BRUNHAM R.C., PLUMMER F.A. HLA-A and HLA-B in Kenya, Africa: Allele frequencies and identification of HLAB*1567 and HLA-B*4426. **Tissue Antigens** 59: 370-380, 2002.

LURY, D.; EPSTEIN, H.; HOLMES, N. The human class I MHC gene HLA-F is expressed in lymphocytes. **Int Immunol.**, 2: 531-7, 1990.

MADRIGAL, J.A.; BELICH, M.P.; HILDEBRAND, W.H.; BENJAMIN, R.J.; LITTLE, A.; ZEMMOUR, J.; ENNIS, P.D.; WARD, F.E.; PETZL-ERLER, M.L.; DUTOIT, E.D.; PARHAM, P. Distinctive HLA-A,B antigens of Black populations formed by interallelic conversion. **Journal of Immunology**, 149: 3411-3416, 1992.

MALISSEN, M.; MALISSEN, B.; JORDAN, B.R. Exon/intron organization and complete nucleotide sequence of an HLA gene. **Proc. Natl. Acad. Sci. U.S.A.**, 79: 893-897, 1982.

MATSUZAKA, Y.; MAKINO, S.; NAKAJIMA, K.; TOMIZAWA, M.; OKA, A.; BAHRAM, S.; KULSI, J.K.; TAMIYA, G.; INOKO, H. New polymorphic microsatellite markers in the human MHC class III region. **Tissue Antigens**, 57(5): 397-404, 2001.

McCLELLAND, E.E.; PENN, D.J.; POTTS, W.K. Major histocompatibility complex heterozygote superiority during coinfection. **Infection and Immunity**, 71(4): 2079-2086, 2003.

McDEVITT, H.O. Discovering the role of the major histocompatibility complex in the immune response. **Annual review of immunology**, 18: 1-17, 2000.

MIDDLETON D., MANAVALAN J.S., FERNANDEZ-VINA M.A., Special issue: HLA alleles and other immunogenetic polymorphism frequencies from worldwide populations. **Human Immunology** 65: 855-1272, 2004.

MOSCOSO, J.; IGNACIO SERRANO-VELA, J.; ARNAIZ-VILLENA, A. MHC-F polymorphism and evolution **Tissue Antigens**, 69: 136-139, 2007.

MUNGALL, A.J.; PALMER, S.A.; SIMS, S.K.; EDWARDS, C.A.; ASHURST, J.L.; WILMING, L.; JONES, M.C.; HORTON, R.; HUNT, S.E.; SCOTT, C.E.; GILBERT, J.G.R.; CLAMP, M.E.; BETHEL, G.; MILNE, S.; AINSCOUGH, R.; ALMEIDA, J.P.; AMBROSE, K.D.; ANDREWS, T.D.; ASHWELL, R.I.S.; BABBAGE, A.K.; BAGGULEY, C.L.; BAILEY, J.; BANERJEE, R.; BARKER, D.J.; BARLOW, K.F.; BATES, K.; BEARE, D.M.; BEASLEY, H.; BEASLEY, O.; BIRD, C.P.; BLAKEY, S.; BRAY-ALLEN, S.; BROOK, J.; BROWN, A.J.; BROWN, J.Y.; BURFORD, D.C.; BURRILL, W.; BURTON, J.; CARDER, C.; CARTER, N.P.; CHAPMAN, J.C.; CLARK, S.Y.; CLARK, G.; CLEE, C.M.; CLEGG, S.; COBLEY, V.; COLLIER, R.E.; COLLINS, J.E.; COLMAN, L.K.; CORBY, N.R.; COVILLE, G.J.; CULLEY, K.M.; DHAMI, P.; DAVIES, J.; DUNN, M.; EARTHROWL, M.E.; ELLINGTON, A.E.; EVANS, K.A.;

FAULKNER, L.; FRANCIS, M.D.; FRANKISH, A.; FRANKLAND, J.; FRENCH, L.; GARNER, P.; GARNETT, J.; GHORI, M.J.R.; GILBY, L.M.; GILLSON, C.J.; GLITHERO, R.J.; GRAFHAM, D.V.; GRANT, M.; GRIBBLE, S.; GRIFFITHS, C.; GRIFFITHS, M.; HALL, R.; HALLS, K.S.; HAMMOND, S.; HARLEY, J.L.; HART, E.A.; HEATH, P.D.; HEATHCOTT, R.; HOLMES, S.J.; HOWDEN, P.J.; HOWE, K.L.; HOWELL, G.R.; HUCKLE, E.; HUMPHRAY, S.J.; HUMPHRIES, M.D.; HUNT, A.R.; JOHNSON, C.M.; JOY, A.A.; KAY, M.; KEENAN, S.J.; KIMBERLEY, A.M.; KING, A.; LAIRD, G.K.; LANGFORD, C.; LAWLOR, S.; LEONGAMORNLETT, D.A.; LEVERSHA, M.; LLOYD, C. R.; LLOYD, D.M.; LOVELAND, J.E.; LOVELL, J.; MARTIN, S.; MASHREGHI-MOHAMMADI, M.; MASLEN, G.L.; MATTHEWS, L.; MCCANN, O.T.; MCLAREN, S.J.; MCLAY, K.; MCMURRAY, A.; MOORE, M.J.F.; MULLIKIN, J.C.; NIBLETT, D.; NICKERSON, T.; NOVIK, K.L.; OLIVER, K.; OVERTON-LARTY, E. K.; PARKER, A.; PATEL, R.; PEARCE, A.V.; PECK, A.I.; PHILLIMORE, B.; PHILLIPS, S.; PLUMB, R.W.; PORTER, K.M.; RAMSEY, Y.; RANBY, S.A.; RICE, C.M.; ROSS, M.T.; SEARLE, S.M.; SEHRA, H.K.; SHERIDAN, E.; SKUCE, C.D.; SMITH, S.; SMITH, M.; SPRAGGON, L.; SQUARES, S.L.; STEWARD, C.A.; SYCAMORE, N.; TAMLYN-HALL, G.; TESTER, J.; THEAKER, A.J.; THOMAS, D.W.; THORPE, A.; TRACEY, A.; TROMANS, A.; TUBBY, B.; WALL, M.; WALLIS, J.M.; WEST, A.P.; WHITE, S.S.; WHITEHEAD, S.L.; WHITTAKER, H.; WILD, A.; WILLEY, D.J.; WILMER, T.E.; WOOD, J.M.; WRAY, P.W.; WYATT, J.C.; YOUNG, L.; YOUNGER, R.M.; BENTLEY, D.R.; COULSON, A.; DURBIN, R.; HUBBARD, T.; SULSTON, J.E.; DUNHAM, I.; ROGERS, J.; BECK, S. The DNA sequence and analysis of human chromosome 6. **Nature**, 425: 805-811, 2003.

NAGAMATSU, T.; TOMOYUKI, F.; MATSUMOTO, J.; YAMASHITA, T.; KOZUMA, S.; TAKETANI, Y. Human leukocyte antigen F protein is expressed in the extravillous trophoblasts but not on the cell surface of them. **Am. J. Reprod. Immunol.**, 56: 172-177, 2006.

NOGUCHI, K.; ISOGAI, M.; KUWADA, E.; NOGUCHI, A.; GOTO, S.; EGAWA, K. Detection of anti-HLA-F antibodies in sera from cancer patients. **Anticancer Research**, 24: 3387-3392, 2004.

PARHAM, P.; ARNETT, K.L.; ADAMS, E.J.; LITTLE, A.; TEES, K.; BARBER, L.D.; MARSH, S.G.E.; OHTA, T.; MARKOW, T.; PETZL-ERLER, M.L. Episodic evolution and turnover of HLA-B in the indigenous human populations of the Americas. **Tissue Antigens**, 50: 219-232, 1997.

PARK, M.; KIM, D.J.; KIM, K.J.; HONG, C.B.; KIM, Y.J.; CHEONG, H.S.; SHIN, H.D.; LEE, E.J.; KIM, H.N.; CHUNG, H. W.; KIM, E.K.; LEE, J.Y.; KIM, H.L. Genetic associations of common deletion polymorphisms in families with Avellino corneal dystrophy. **Biochem. Biophys. Res. Commun.**, 387(4): 688-93, 2009.

PIANCATELLI, D.; CANOSSI, A.; AURELI, A.; OUMHANI, K.; DEL BEATO, T.; DI ROCCO, M.; LIBERATORE, G.; TESSITORE, A.; WITTER, K.; EL AOUAD, R.; ADORNO, D. Human leukocyte antigen-A, -B, and -Cw polymorphism in a Berber population from North Morocco using sequence-based typing. **Tissue Antigens**, 63: 158-172, 2004.

PIMTANOTHAI, N.; HURLEY, C.K.; LEKE, R.; KLITZ, W.; JOHNSON, A.H. HLA-DR and -DQ polymorphism in Cameroon. **Tissue Antigens**, 58: 1-8, 2001.

POTTS, W.K.; SLEV, P.R. Pathogen-based models favoring MHC genetic diversity. **Immunol. Rev.**, 143: 181-197, 1995.

PYO, C-H.; WILLIAMS, L.M.; YUKI, M.; HYODO, H.; LI, S.S.; ZHAO, L.P.; GERAGHTY, D.E.; HLA-E, HLA-F and HLA-G polymorphism: genomic sequence defines haplotype structure and variation spanning the nonclassical class I genes. **Immunogenetics**, 58: 241–251, 2006.

ROUSSEAU, P.; MASTERNAK, K.; KRAWCZYK, M.; REITH, W.; DAUSSET, J.; CAROSELLA, E.D.; MOREAU, P. In vivo, RFX5 binds differently to the human leucocyte antigen-E, -F, and -G gene promoters and participates in HLA class I protein expression in a cell type-dependent manner. **Immunology**, 111: 53–65, 2004.

SANCHEZ-VELASCO, P.; GOMEZ-CASADO, E.; MARTINEZ-LAZO, J.; MOSCOSO, J.; ZAMORA, J.; LOWY, E.; SILVERA, C.; CEMBORAIN, A.; LEYVA-COBIÁN, F.; ARNAIZ-VILLENA, A. HLA alleles in isolated populations from North Spain: origin of the Basques and ancient Iberians. **Tissue Antigens**, 61: 384-392, 2003.

SHIINA, T.; TAMIYA, G.; OKA, A.; TAKISHIMA, N.; YAMAGATA, T.; KIKKAWA, E.; IWATA, K.; TOMIZAWA, M.; OKUAKI, N.; KUWANO, Y.; WATANABE, K.; FUKUZUMI, Y.; ITAKURA, S.; SUGAWARA, C.; ONO, A.; YAMAZAKI, M.; TASHIRO, H.; ANDO, A.; IKEMURA, T.; SOEDA, E.; KIMURA, M.; BAHRAM, S.; INOKO, H. Molecular dynamics of MHC genesis unraveled by sequence analysis of the 1,796,938-bp HLA class I region. **Proc Natl Acad Sci U.S.A.**, 96(23): 13282-7, 1999.

SHIINA, T.; INOKO, H.; KULSKI, J.K. An update of the HLA genomic region, locus information and disease associations. **Tissue Antigens**, 64: 631-649, 2004.

SHOBU, T.; SAGESHIMA, N.; TOKUI, H.; OMURA, M.; SAITO, K.; NAGATSUKA, Y.; NAKANISHI, M.; HAYASHI, Y.; HATAKE, K.; ISHITANI, A. The surface expression of HLA-F on decidual trophoblasts increases from mid to term gestation. **Journal of Reproductive Immunology**, 72: 18-32, 2006.

SNELL, G.D. Methods for the study of histocompatibility genes. **J Genet**, 49: 87-108, 1948.

STROYNOWSKI, I. Molecules related to class I major histocompatibility complex antigens. **Annu. Rev. Immunol.**, 8: 501, 1990.

SULLIVAN, L.C.; HOARE, H.L.; McCLUSKEY, J.; ROSSJOHN, J.; BROOKS, A.G. A structural perspective on MHC class Ib molecules in adaptive immunity. **Trends Immunol.**, 27: 413-420, 2006.

TABAYOYONG, W.B.; ZAVAZAVA, N. Soluble HLA revisited. **Leuk. Res.**, 31: 121–125, 2007.

The MHC sequencing consortium, Complete sequence and gene map of a human major histocompatibility complex. **Nature**, 401: 921–3, 1999.

TRANSY, C.; NASH, S.R.; DAVID-WATINE, B.; COCHET, M.; HUNT III, S.W.; HOOD, L.E.; KOURILSKY, P. A low polymorphic mouse H-2 class I gene from the Tla complex is expressed in a broad variety of cell types. **J Exp. Med.**, 166: 341, 1987.

TSE, K.P.; WEN-HUI, S.; KAI-PING, C.; NGAN-MING, T.; CHIA-JUNG, Y.; PETRUS, T.; LEE-CHU, S.; CHUEN, H.; MIN-LEE, Y.; SHENG-PO, H.; HONG-YI, L.; MING-HSI, W.; LI-PING, L.; LIH-CHYANG, C.; SHEUE-RONG, L.; TIMOTHY, J.J.; YU-SUN, C.; YIN, S.Y. Genome-wide Association Study Reveals Multiple Nasopharyngeal Carcinoma-Associated Loci within the HLA Region at Chromosome 6p21.3. **The American Journal of Human Genetics**, 85: 194-203, 2009.

UCHIGIRI, C.; MIZUNO, S.; WADA, K.; TSUTSUMI, M.; KATO, T.; KAMIYA, T.; OZAWA, K. An Identification of the HLA-F null allele in Japanese. **Immunogenetics**, 45: 466-567, 1997.

URENA, J.M.; MERLOS-SUAREZ, A.; BASELGA, J.; ARRIBAS, J. The cytoplasmic carboxy-terminal amino acid determines the subcellular localization of proTGF-(α) and membrane type matrix metalloprotease (MT1-MMP). **J. Cell Sci.**, 112(6): 773–784, 1999.

WAINWRIGHT, S.D.; BIRO, P.A.; HOLMES, C.H. HLA-F is a predominantly empty, intracellular, TAP-associated MHC class Ib protein with a restricted expression pattern. **J. Immunol.**, 164: 319, 2000.

WANG, S.; LEWIS Jr., C.M.; JAKOBSSON, M.; RAMACHANDRAN, S.; RAY, N.; BEDOYA, G.; ROJAS, W.; PARRA, M.V.; MOLINA, J.A.; GALLO, C.; MAZZOTTI, G.; POLETTI, G.; HILL, K.; HURTADO, A.M.; LABUDA, D.; KLITZ, W.; BARRANTES, R.; BORTOLINI, M.C.; SALZANO, F.M.; PETZL-ERLER, M.L.; TSUNETO, L.T.; LLOP, E.; ROTHHAMMER, F.; EXCOFFIER, L.; FELDMAN, M.W.; ROSENBERG, N.A.; RUIZ-LINARES, A. Genetic Variation and Population Structure in Native Americans. **PLOS Genetics**, 3: 185, 2007.

ZANGENBERG, G.; HUANG, M.M.; ARNHEIM, N.; ERLICH, H. New HLA-DPB1 alleles generated by interallelic gene conversion detected by analysis of sperm. **Nat Genet.**, 10(4): 407-14, 1995.

ZAVAZAVA, N. Soluble HLA class I molecules: biological significance and clinical implications. **Molecular Medicine Today**, 4: 116–121, 1998.

ZHANG, J; PAN, L.; CHEN, L.; FENG, X.; ZHOU, L.; ZHENG, S. Non-classical MHC-I genes in chronic hepatitis B and hepatocellular carcinoma. **Immunogenetics**, 64: 251-8, 2012.