

Universidade de Brasília
Instituto de Ciências Biológicas
Programa de Pós – Graduação em Biologia Animal



Região 3'UTR do Gene HLA-G em Populações Humanas do Centro - Oeste

Rafaela De Cesare Parmezan Toledo

Brasília
2011

Região 3'UTR do Gene HLA-G em Populações Humanas do Centro - Oeste

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Animal como requisito parcial para obtenção do grau de mestre em Biologia Animal.

Rafaela De Cesare Parmezan Toledo
Orientadora: Dra. Silviene Fabiana de Oliveira

Brasília
2011

Trabalho desenvolvido no Laboratório de Genética da Universidade de Brasília – UnB e no Laboratório de Biotecnologia da Universidade Católica de Brasília, com o suporte financeiro da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

“Of course the Neverlands vary a good deal. John’s, for instance, had a lagoon with flamingoes flying over it at which John was shooting, while Michael, who was very small, had a flamingo with lagoons flying over it.”

Peter Pan, J.M. Barrie

À minha família,

Dedico.

Agradecimentos

À Dra. Silviene Fabiana de Oliveira, pela orientação, amizade e paciência.

Aos membros da banca examinadora pela disponibilidade de avaliar e contribuir ao trabalho.

À Dra. Rosângela Vieira, da Universidade Católica de Brasília, pelo apoio e por tornar o laboratório da Universidade Católica disponível ao nosso trabalho.

Ao Dr. Erick da Cruz Castelli, da Universidade Federal de Goiás, por dividir seu conhecimento comigo e pela disponibilidade prestada ao me ensinar tudo o que precisei e mais.

Ao Dr. Celso Teixeira Mendes Júnior, da Universidade de São Paulo – Ribeirão Preto, pelo conhecimento compartilhado nas análises estatísticas.

Às agências de fomento CAPES e CNPq, pelo apoio financeiro.

Aos laboratórios de Biologia Molecular da UnB, e o Laboratório de Biotecnologia da UCB, pelo apoio laboratorial.

Aos membros da equipe e amigos do Laboratório de Genética da UnB, Silviene, Nazaré, Carol, Mila, Diana, Arthur, Gustavo, Neide, Dudu, Tatiana e Arcanjo, o apoio incondicional de vocês me ajudaram a chegar até aqui.

Aos membros da equipe do Laboratório de Biotecnologia da UCB, aos técnicos Ida, Willian e André por toda ajuda e gentileza prestadas a mim ao longo desse trabalho. À Larissa, Natália e Nina, pela ajuda e amizade.

Aos meus queridos amigos Kdol, Rafa, Fê-chan, Keka, Dominó, Carioca, Tis, Jonas, Sabrina, Fabrícus, Marina e todos os queridíssimos amigos do “círculo da confiança” do 2º/2004, por me proporcionarem momentos de descontração fora do laboratório e por aguentarem minhas reclamações mesmo não entendendo nada do meu projeto e seus problemas.

À minha família, especialmente aos meus pais Arthur e Regina, sem os quais nada disso seria possível. Obrigada pelo apoio, pela confiança, incentivo e por sempre acharem o máximo tudo o que eu faço, mesmo nem sempre concordando ou compreendendo.

Ao meu grande amor, Henrique, por toda a paciência e amor que existe nesse mundo.

Agradeço.

Índice

Índice de Tabelas 3

Índice de Figuras 4

Resumo 5

Abstract 7

Introdução 9

Genética de Populações 9

População brasileira e História demográfica 11

Distrito Federal 12

Kalunga 13

Perfil genético de Distrito Federal e *Kalunga* 14

HLA-G 16

Objetivos 24

Objetivos específicos 24

Material e Métodos 25

Populações 25

Aspectos éticos 25

Material biológico 25

Análises laboratoriais 25

Caracterização genética da região 3'UTR do gene HLA-G 27

Desequilíbrio de ligação 27

Análises Haplotípicas da região 3'UTR do gene HLA-G 27

Testes de seleção	28
Resultados	30
Caracterização genética da região 3'UTR do gene HLA-G	30
Teste de Desequilíbrio de ligação	33
Análises Haplotípicas da região 3'UTR do gene HLA-G	34
Testes de seleção	36
Discussão	38
Caracterização genética da região 3'UTR do gene HLA-G	37
Desequilíbrio de ligação	38
Análises Haplotípicas da região 3'UTR do gene HLA-G	39
Testes de seleção	41
Considerações finais	43
Conclusões	45
Referências	46

Índice de Tabelas

Tabela 1. Frequências gênicas, genotípicas, heterozigose observada e esperada dos marcadores genéticos da região 3'UTR do gene HLA-G em Distrito Federal e *Kalunga*.

Tabela 2. Teste de diferenciação gênica e genotípica, e Theta-P entre Distrito Federal e *Kalunga* avaliado para os marcadores genéticos da região 3'UTR do gene HLA-G.

Tabela 3. Haplótipos da região 3'UTR do HLA-G e suas frequências nas populações Distrito Federal e *Kalunga* ($2n$ = número de cromossomos avaliados) e valores de Theta-P (em parênteses IC = Intervalo de Confiança).

Tabela 4. Resultados dos testes de neutralidade realizados com os marcadores genéticos da região 3'UTR do HLA-G nas populações Distrito Federal e *Kalunga*.

Índice de Figuras

Figura 1. Mapa indicando a localização geográfica das populações: alfinete vermelho – Distrito Federal; alfinete azul - *Kalunga*.

Figura 2. Estimativas de ancestralidade nas populações Brasília e *Kalunga* para diferentes classes de marcadores moleculares. AIM (Marcadores indicativos de ancestralidade) autossômico, STR (microssatélites) autossômico, haplogrupos de DNAm e STRs do Cromossomo Y.

Figura 3. Estrutura molecular do gene HLA-G - éxons representados em preto, íntrons em branco e a região 3'UTR está representada pela área rachurada com destaque para o InDel de 14pb. Esse gene apresenta um códon de parada no éxon 6.

Figura 4. Representação do RNAm primário, dos transcritos após o processamento alternativo e das proteínas codificadas pelo HLA-G. Foi descrito que esse gene produz sete proteínas diferentes por processamento alternativo, são quatro proteínas de membrana e três proteínas solúveis. Modificado de Donadi *et al.* 2010.

Figura 5. Região 3'UTR do gene HLA-G ressaltando os sítios variáveis descritos na literatura para essa região (em negrito), o fragmento de 92pb que é descartado no processamento alternativo, e o motivo de repetição AU (retângulo) (modificado de Donadi *et al.* 2010).

Figura 6. Matriz de desequilíbrio de ligação entre os oito marcadores analisados par a par: A - Distrito Federal, B - *Kalunga*. As barras acima indicam a posição relativa dos marcadores na região 3'UTR no gene do HLA-G, enumerados de um a oito. Branco - ausência de desequilíbrio, azul - baixo desequilíbrio e vermelho - forte desequilíbrio.

Figura 7. Modelos que explicam o desequilíbrio de ligação por miscigenação. Modificado de Pfaff *et al.* (2001).

Resumo

Kimura propôs que o principal fator evolutivo na evolução humana foi a deriva genética. Porém, nos últimos anos diversos trabalhos tem indicado que a seleção natural também pode ter tido um papel importante na evolução da espécie. Nesse trabalho foi avaliado se há forças evolutivas atuando sobre a região 3'UTR do HLA-G em duas populações do Centro-Oeste brasileiro, Distrito Federal e *Kalunga*, cujas histórias demográficas são bem conhecidas. O Distrito Federal apresenta uma população que foi recentemente formada pela migração de pessoas advindas de várias regiões do Brasil, tornando-se uma região urbana geneticamente representativa do país. *Kalunga* é uma população semi-isolada rural formada basicamente por escravos fugidos e abandonados. O HLA G apresenta baixo nível de polimorfismo, distribuição e expressão restrita a tecidos/órgãos específicos e propriedades biológicas que levam a tolerância imunológica. O maior conhecimento atual sobre o HLA-G em comparação aos demais genes de classe Ib deve-se à observação de associação de alelos e padrões de expressão com abortos espontâneos recorrentes e diversas doenças. Devido à associação do gene HLA-G com várias doenças e por ele produzir respostas diferentes a cada doença, é possível pensar que esteja sob direta influência da seleção natural, sendo que no caso da região 3'UTR do HLA-G, acredita-se que esteja sob seleção balanceadora. A hipótese desse trabalho é que a região 3'UTR do gene HLA-G deve apresentar uma série de mutações, várias ainda desconhecidas, e que essa região genômica possa estar sob pressão seletiva. Para testar essa hipótese novas mutações foram buscadas e oito SNPs (*single nucleotide polymorphisms*) foram avaliados na região 3'UTR do gene HLA-G a partir do seqüenciamento total do éxon 8 de 61 indivíduos de *Kalunga* e 72 do Distrito Federal. Com os dados levantados foram avaliados parâmetros de genética de populações e a ocorrência de seleção utilizando os *Teste D de Tajima*, *Teste F de Fu & Li* e *Teste D de Fu & Li*, utilizando o DNASP v5.10.01 e o teste de neutralidade de Ewens–Watterson com o Pypop v0.7.0. Como resultado, uma mutação nova foi identificada na população *Kalunga*. As populações não mostraram diferenças significativas, para os oito *loci* avaliados, quanto as frequências gênicas e genotípicas. Ambas se adequaram ao equilíbrio de Hardy-Weinberg. Foi observado forte desequilíbrio de ligação. Foram observados 11 haplótipos diferentes no Distrito Federal e na população *Kalunga* oito. As populações não diferiram quanto as frequências haplotípicas. O teste *F de Fu & Li* e o teste de *Ewens-Watterson* foram significativos

para ambas as populações, o teste *D de Tajima* foi significativo para o Distrito Federal e o teste *D de Fu & Li* não foi significativo para nenhuma das populações. É possível perceber que apesar de todas as diferenças contidas na história demográfica das duas populações, para a região 3'UTR do HLA-G há forças evolutivas mantendo essa região semelhante nas duas populações. Pelo teste *F de Fu & Li* e pelo teste de *Ewens-Watterson* é visto que provavelmente houve atuação de seleção balanceadora desde muito tempo até hoje.

Abstract

Kimura proposed that the main evolutionary factor in human evolution is genetic drift. However, in recent years several studies have shown that natural selection may also have played an important role in the evolution of species. In this study, we assessed whether there are evolutionary forces acting on the 3'UTR region of the HLA-G in two populations of the Brazilian Center-West, Federal District and Kalunga, whose demographic histories are well known. The Federal District has a population that was recently formed by the migration of people from many different regions of Brazil, becoming an urban region genetically representative of the country. Kalunga is a semi-isolated rural population basically formed by runaway and abandoned slaves. HLA G shows a low level of polymorphism, distribution and expression restricted to tissues / organs and specific biological properties that lead to immunological tolerance. The most current knowledge on the HLA-G in comparison with other class Ib genes is due to the observation of association of alleles and patterns of expression with recurrent miscarriages and various diseases. Due to the association of HLA-G gene with various diseases and for him to produce different answers to different disease it is possible to think that this region is under the direct influence of natural selection. The case of the 3'UTR of the HLA-G, it is believed that is under balancing selection. The hypothesis of this study is that the 3'UTR of the gene HLA-G should present a series of mutations, several still unknown, and that this genomic region may be under selective pressure. To test this hypothesis were sought new mutations, one InDel and seven SNPs (single nucleotide polymorphisms) were evaluated in the 3'UTR of the HLA-G gene from the total sequencing of exon 8 of 61 individuals and 72 Kalunga of the Federal District. With the data collected were evaluated parameters of population genetics and the occurrence of selection using the Tajima D test, Fu &Li F test, and Fu & Li D test, using the DNASP v5.10.01 and the test of neutrality of Ewens-Watterson with Pypop v0.7.0. As a result, a new mutation was identified in the population Kalunga. The populations showed no significant differences for the eight markers examined for allele frequency and genotype. Both fitted the Hardy-Weinberg. We observed strong linkage disequilibrium. We observed 11 different haplotypes in the Federal District and the population Kalunga eight. The populations did not differ in haplotype frequencies. The Fu &Li F test and Ewens-Watterson test were significant for both populations, the test

was significant Tajima D for the Federal District and the Fu & Li D test was not significant for any population. It is possible to verify that despite all the differences included in demographic history between populations, to the 3'UTR of the HLA-G is evolutionary forces maintaining this region similar in both populations. The Fu & Li F test and the Ewens-Watterson test is seen that there was likely action of balancing selection for a long time until today.

Introdução

Genética de Populações

A evolução, origem e dispersão do homem moderno é objeto de estudo de várias áreas do conhecimento, como, por exemplo, paleoantropologia, arqueologia, linguística e genética. De acordo com a hipótese mais aceita pela comunidade científica, o *Homo sapiens* surgiu no continente Africano há cerca de 200 mil anos atrás. Há aproximadamente cem mil anos os continentes Europa, Ásia e África estavam ocupados por várias espécies do gênero *Homo*. Contudo, há 30 mil anos o *Homo sapiens* ocupava praticamente todos os continentes em sua forma moderna e as demais espécies do gênero *Homo* haviam desaparecido. Hoje, as evidências anatômicas, arqueológicas e genéticas mostram que o *Homo sapiens* moderno é uma espécie recente que se originou na África e se dispersou por todos os continentes (Laval *et al.* 2010).

O processo de dispersão do *Homo sapiens* foi lento e marcado pela diferenciação entre populações da espécie. Na tentativa de entender a origem e manutenção da variação encontrada entre as populações de nossa espécie, Kimura (1968) propôs a Teoria de evolução neutra, que levou a intensas discussões na sociedade científica da época e que, posteriormente, passou a ser a ideia mais aceita para explicar a evolução da espécie humana. Essa teoria diz que a variabilidade encontrada na espécie é um balanço entre o surgimento de mutações novas e o desaparecimento dessas mutações pela deriva genética. Ele percebeu que as taxas de mutação em mamíferos são altas, taxas essas teoricamente incompatíveis com a sobrevivência. Portanto, para explicar essa inconsistência Kimura propôs que grande parte das mutações e da variabilidade genética vista entre as populações humanas deve ser neutra. A teoria da evolução neutra é basicamente ditada pelo surgimento de novas mutações, mas a permanência dessas mutações em populações com tamanho efetivo finito é resultado de acasalamentos ao acaso, portanto sendo regida pelo mecanismo evolutivo conhecido como deriva genética.

Entender a variação e a diferenciação entre populações é um dos ramos da genética de populações. A genética de populações, como ramo da ciência surgiu nas primeiras décadas do século XX como fruto do trabalho de Ronald Fisher, John Burdon

Sanderson Haldane e Sewall Wright. Ela trata das Leis de Mendel e outros princípios da genética aplicados a populações, não a organismos isolados (Hartl & Clark 2010). Nesse contexto, a genética de populações se tornou uma importante forma de avaliar aspectos importantes em diversas áreas, dentre elas genética, genômica, biologia evolutiva, medicina, biologia molecular e biologia celular. Além disso, a genética de populações inclui o estudo de vários fatores que resultam em mudanças evolutivas nas espécies ao longo do tempo. Ao definir o *pool* genético dentro do qual a evolução ocorre, os princípios de genética de populações são básicos para que se tenha uma perspectiva evolutiva ampla sobre a biologia (Hartl & Clark 2010).

O estudo de genética de populações em humanos tem sido feito a partir de dados de marcadores moleculares, que são regiões do genoma que apresentam variabilidade e possibilitam diferenciar indivíduos e/ou populações. O estudo desses marcadores auxilia no entendimento das relações evolutivas e diferenciação entre populações, assim como na reconstrução das rotas e origens do povoamento (Salzano & Callegari-Jacques 1988, Saitou 1995). A diferenciação populacional analisada sob a ótica da evolução neutra é basicamente influenciada pela sua história demográfica (deriva genética e fluxo gênico) (Wu & Zhang 2011). No entanto, apesar do grande reconhecimento dado a teoria de evolução neutra, debate sobre quais e como os fatores evolutivos moldaram a espécie humana continuam. O questionamento do papel da seleção natural na história evolutiva da espécie *Homo sapiens* voltou a cena e ultimamente vários trabalhos tem sido feitos com o intuito de melhor avaliar esse mecanismo evolutivo não neutro.

Seleção natural é um mecanismo evolutivo definido como sobrevivência e/ou reprodução diferencial dos indivíduos em decorrência de características adaptativas, onde essas tendem a ter maior frequência na população ao longo do tempo (Sabeti *et al.* 2006). Apesar da teoria neutralista explicar grande parte da variação do genoma humano, existem vertentes que dão maior importância a seleção natural uma vez que parte da variabilidade genética encontra-se relacionada a caracteres que podem apresentar valor adaptativo para a população, sendo assim objeto de sua ação. A contribuição de cada um desses diferentes mecanismos evolutivos, deriva genética e seleção natural, ainda é pauta de discussões; contudo, hoje não há dúvidas da importância dos dois mecanismos na evolução humana (Laval *et al.* 2010).

Hoje é sabido que a seleção natural atua diretamente sobre um gene ou genes com funções adaptativas, ao contrário da evolução neutra, que atua sobre todos os genes de

forma igual (Wu & Zhang 2011). Sugere-se que genes relacionados ao sistema imune são comumente alvo de seleção natural. Isso é atribuído ao fato de que muitos desses genes são responsáveis por características adaptativas. Em diversas populações tem sido observado excesso de heterozigotos para diferentes genes do sistema imune, sendo que esse fenômeno deve ser resultado de seleção balanceadora nas populações, a qual reduz a diferenciação populacional (Ferrer-Admetlla *et al.* 2008; Fumagali *et al.* 2009).

População brasileira e história demográfica

A população brasileira foi objeto de estudo de vários estudos de genética de população. Sabe-se que sua formação se deu pela miscigenação de três grupos principais: europeus, africanos e ameríndios. A miscigenação desses três grupos fez da população brasileira uma das populações mais heterogêneas do mundo, sendo descrita como a de maior variabilidade para vários marcadores moleculares (Parra *et al.* 2003).

Os europeus começaram a migrar para a região que se tornaria o Brasil no ano de 1500, como parte do evento histórico conhecido como grandes navegações. A maioria dos europeus que vieram para o Brasil partiu da Península Ibérica, principalmente de Portugal. Somente a partir do fim do século XIX, começo do século XX, o governo brasileiro incentivou a migração de europeus. Nesse período o Brasil recebeu mais de cinco milhões de imigrantes europeus e do Oriente Médio. Desses imigrantes 34% eram italianos, 29% portugueses, 14% espanhóis, 5% japoneses, 4% alemães, 2% sírios e libaneses, e 12% de outros grupos (IBGE 2000).

Os africanos que chegaram ao território brasileiro foram trazidos de forma forçada como mão de obra escrava. Acredita-se que mais de cinco milhões de africanos foram trazidos pelo tráfico para o Brasil entre 1550 e 1855 (Reis & Gomes 1996). O comércio de escravos tinha preferência por jovens do sexo masculino, e o grupo mais afetado foram os povos da África subsaariana do grupo linguístico *bantu* (Reis & Gomes 1996). Os escravos não se distribuíram igualmente pelo Brasil, sendo concentrado na região nordeste, onde havia muitas fazendas de cultivo de cana-de-açúcar. O grande número de escravos no nordeste brasileiro pode ser verificado, além dos dados históricos, pela herança genética nessa região brasileira. Apesar dos senhores de escravos terem práticas para evitar motins e rebeliões, os escravos tiveram como forma de resistência a formação de quilombos. O termo quilombo refere-se a comunidades formadas à época

da escravidão por escravos fugidos e/ou abandonados, geralmente em locais de difícil acesso, que tornaram-se isoladas para evitar a recaptura. Ainda hoje é possível identificar comunidades que foram formadas nesse contexto. Atualmente comunidades derivadas de quilombos podem ser reconhecidas como remanescentes de quilombos com auxílio da Fundação Cultural Palmares, a partir de estudos históricos e antropológicos, o que garante a posse da terra e outros direitos àqueles que ainda vivem nessas comunidades (Decreto nº 4887, do dia 23 de Novembro de 2003).

Outro grupo formador da população brasileira foram os ameríndios, que formam o grupo nativo das Américas. Estudos recentes mostram que a ocupação do continente Americano como um todo ocorreu há pelo menos 12 mil anos (Wang *et al.* 2007). O tamanho da população ameríndia à época da chegada dos portugueses é estimado de 1 a 10 milhões de indivíduos (FUNAI 2011), sendo que no território brasileiro havia, principalmente, Tupi-Guaranis, grupos que compartilham semelhanças culturais e linguísticas (Fausto 2001). O contato do europeu com a população nativa nem sempre foi pacífico. Esse contato gerou miscigenação, mas também gerou uma redução drástica no tamanho da população indígena brasileira. Hoje, eles são estimados em 460 mil habitantes vivendo em todo o país distribuído em terras indígenas e zonas urbanas (FUNAI 2011). Diversas etnias foram extintas, mas a miscigenação dos três grupos principais grupos parentais (africanos, europeus e ameríndios) hoje é percebida na aparência do povo brasileiro, em sua cultura e em estudos genéticos.

Dentro da mesma região geográfica brasileira, observam-se diferenças entre as populações humanas que a compõem, apesar da similaridade do contexto histórico de povoamento brasileiro. O presente estudo visa avaliar duas populações, Distrito Federal (população urbana) e *Kalunga* (população remanescente de quilombo rural). Apesar de fazerem parte do contexto histórico do mesmo, as duas populações foram escolhidas por apresentarem histórias de formação peculiares.

Distrito Federal

A população do Distrito Federal (DF) é considerada como um parâmetro nacional de população urbana brasileira. A história do DF, localizado no Centro Oeste brasileiro (Figura 1), faz com que sua população seja peculiar. O DF começou a ser construído em 1956 para ser a nova capital federal do Brasil, sendo que durante sua construção houve uma intensa migração de várias pessoas de todas as regiões brasileiras a procura de

trabalho e oportunidade. Podemos destacar migrações de regiões como nordeste, sudeste e sul, sendo que essas regiões já apresentavam intensa miscigenação dos grupos parentais (Godinho 2008, Gontijo 2008). Dessa forma é possível inferir que o Distrito Federal é uma população geneticamente representativa do Brasil (Barcelos 2006). Sua população foi estimada, em 2007, como sendo de 2.455.903 habitantes distribuídos em 29 regiões administrativas (CODEPLAN 2009).

Kalunga

Kalunga constitui histórica e numericamente um dos maiores quilombos do Brasil e o principal remanescente de quilombo da região Centro-Oeste. Com aproximadamente 6000 habitantes, esta comunidade se localiza no nordeste do Estado de Goiás (Centro-Oeste brasileiro – Figura 1) e ocupa uma área de aproximadamente 258 mil hectares próxima às cidades de Monte Alegre e Cavalcante. *Kalunga* foi reconhecido pelo Estado Brasileiro, em 2006, como um remanescente de quilombo (Fundação Cultural Palmares, 2009).

O quilombo foi formado na época da exploração de ouro na atual cidade de Monte Alegre (GO), local conhecido como “Minas do Tocantins”. Segundo Soares (1995), teria sido formado no final do século XVIII por escravos fugidos. Entretanto, de acordo com relatos dos moradores, ao menos parte de seus ancestrais teriam sido abandonados pelos bandeirantes quando o ouro das minas acabou. Temendo o retorno dos bandeirantes, os escravos abandonados teriam se juntado ao quilombo. Hoje, o remanescente desse antigo quilombo é formado por nove subcomunidades, dentre elas Engenho I e II, Vão de Moleque e Vão de Almas.

Kalunga ficou isolada da área urbana durante gerações. Mais recentemente, tem mantido contato, principalmente com o município de Cavalcante (Goiás), com o objetivo de comércio e cidadania (saúde e educação, prioritariamente), e através de programas de pesquisa e extensão universitária.



Figura 1. Mapa indicando a localização geográfica das populações: Distrito Federal - alfinete vermelho; *Kalunga* - alfinete azul.

Perfil genético de Distrito Federal e Kalunga

O conhecimento do perfil genético e demográfico de populações é um passo necessário para diferenciar a história evolutiva neutra e a história de seleção natural. Nesse contexto, as populações Kalunga e Distrito Federal são especialmente interessantes para avaliação do impacto da seleção natural pois, foram bem caracterizadas com a utilização de marcadores neutros.

De acordo com dados da literatura, é visto que ambas as populações são diferentes de acordo com sua constituição genética e ancestralidade considerando marcadores genéticos seletivamente neutros (Ribeiro 2009, Gontijo 2008, Godinho 2008, Pedrosa 2006, Barcelos 2006). A Figura 2 é uma compilação de dados de estimativas de contribuição parental para marcadores autossômicos, DNA mitocondrial e cromossomo Y. Dentre os marcadores autossômicos há marcadores indicativos de ancestralidade (AIM's – do inglês *Ancestry Informative Markers*) e marcadores microsátélites (STR's – do inglês *Short Tandem Repeats*). Esses marcadores contam a história demográfica de cada população de uma forma distinta em decorrência de suas diferentes localizações

genômicas. Por exemplo, o DNAm_t conta a história de linhagem matrilinea, enquanto o cromossomo Y conta a história da linhagem masculina.

Os marcadores AIMs e STRs mostram a contribuição dos três grupos parentais nas duas populações, como era de se esperar, de acordo com os dados históricos. Esses marcadores mostram que o Distrito Federal teve em sua formação uma contribuição maior da parental europeia (Figura 2) (Godinho 2008). No entanto, as demais contribuições somam quase 40%, isso se deve ao fato de que o DF foi formado por pessoas que vieram de regiões onde já havia miscigenação prévia. Em *Kalunga* esses mesmos marcadores mostram que a parental de maior contribuição é a africana, como esperado (Figura 2), porém com uma participação significativa da parental europeia.

Os marcadores do DNAm_t contam a história da linhagem materna dessas duas populações. Pode ser observado que no DF há uma contribuição semelhante de cada grupo parental, sendo que a porção não-Ameríndia indicada no gráfico foi interpretada como um componente europeu (Barcelos 2006 – Figura 2). Enquanto em *Kalunga* não há nenhum registro de haplogrupos europeus.

O cromossomo Y mostra a história da linhagem paterna em ambas as populações. Através dos haplogrupos observados vemos que homem europeu teve reprodução diferencial com as mulheres de todos os grupos parentais (Barcelos 2006; Ribeiro 2009). Isso é percebido em ambas as populações, onde a contribuição de haplogrupos Y europeus é acima de 60%, inclusive em *Kalunga* um remanescente de quilombo (Figura 2).

Portanto, com a comparação desses marcadores é possível perceber que as duas populações possuem histórias diferentes que se refletem na constituição genética. Enquanto uma é fruto de um intenso processo de miscigenação a outra se manteve isolada até um período recente. Isso torna cada uma delas únicas e diferentes quanto a sua composição genética e ancestralidade.

Os marcadores mostrados até aqui são bastante tradicionais em estudos de genética de populações e mostra bem a teoria neutralista proposta por Kimura. Contrapondo as ideias de Kimura, tem se buscado mais informações sobre a seleção natural na história da espécie *Homo sapiens*. Para esse tipo de questão tem-se buscado por marcadores que estejam relacionados a características adaptativas. Como já dito, os genes relacionados ao sistema imune estão comumente sob seleção, sendo assim um bom modelo. Por isso, o gene escolhido para esse estudo foi o gene HLA-G.

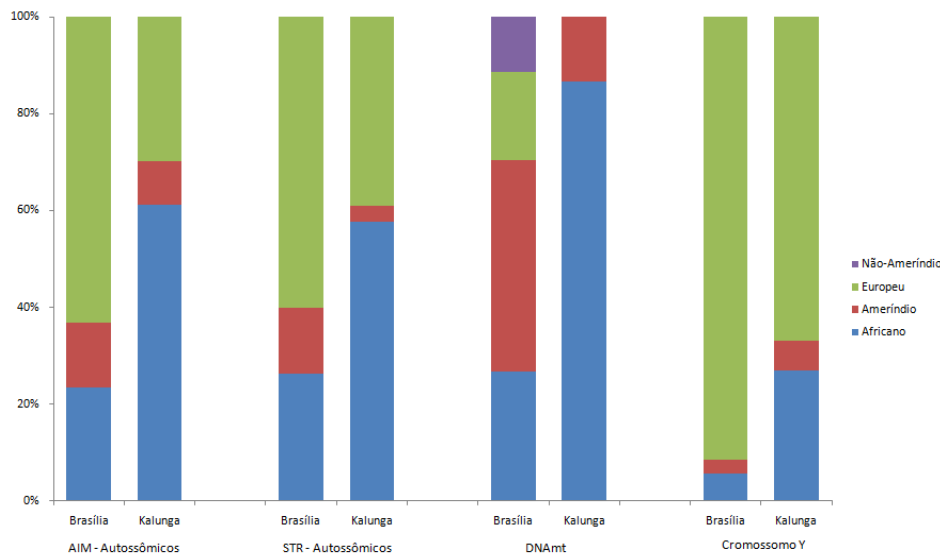


Figura 2. Estimativas de ancestralidade nas populações do Distrito Federal e *Kalunga* para diferentes classes de marcadores moleculares. AIMS (Marcadores indicativos de ancestralidade) autossômico, STRs (microssatélites) autossômico, haplogrupos de DNAmT e STRs do cromossomo Y.

HLA-G

Como modelo para verificar se há a existência de seleção natural nas populações do Distrito Federal e *Kalunga* foi selecionada a região 3'UTR do gene HLA-G para análise. Esse gene faz parte do complexo de histocompatibilidade principal (MHC) em humanos. O MHC é um grupo de moléculas do sistema imunológico cuja principal função é apresentar fragmentos peptídicos de agentes patogênicos na superfície celular para induzir a resposta imune específica (Janeway *et al.* 2002). Esse sistema é codificado por aproximadamente 200 genes localizados no braço curto do cromossomo 6 em humanos (6p21.3), também conhecido como Antígeno Leucocitário Humano (HLA, do inglês – *Human Leukocyte Antigen*) (Hviid 2006a, Janeway *et al.* 2002). Os genes desse complexo são subdivididos em HLA clássicos de classe Ia (HLA-A, -B e -C) e II (HLA-DR, -DQ e -DP) e HLA não-clássicos de classe Ib (HLA-E, -F e -G). Os genes HLA de classe Ia e II são considerados os genes de maior variação do genoma e seu papel na apresentação de antígenos já é bem compreendido (Janeway *et al.* 2002). Na década de 1980, começou a discussão sobre particularidades presentes nos genes do grupo HLA não-clássicos classe Ib (Apps *et al.* 2008, Hviid 2006a). Dentre as

diferenças, podemos ressaltar o baixo nível de polimorfismo, expressão restrita a tecidos e/ou órgãos e função.

Dentre os genes HLA não-clássicos, o HLA-G, assim como os demais HLAs de classe I, tem uma organização molecular constituída por uma região promotora 5', que se estende por 1,4kb desde o ATG, oito éxons e sete íntrons (Solier *et al.* 2001, Castro *et al.* 2000). No entanto, esse gene apresenta algumas características únicas como: um códon de parada no éxon 6; o éxon 8, último éxon, é a região 3'UTR (do inglês – *untranslated region*, Figura 3), região não traduzida do RNAm maduro (Solier *et al.* 2001, Castro *et al.* 2000), e um padrão particular de expressão por processamento alternativo.

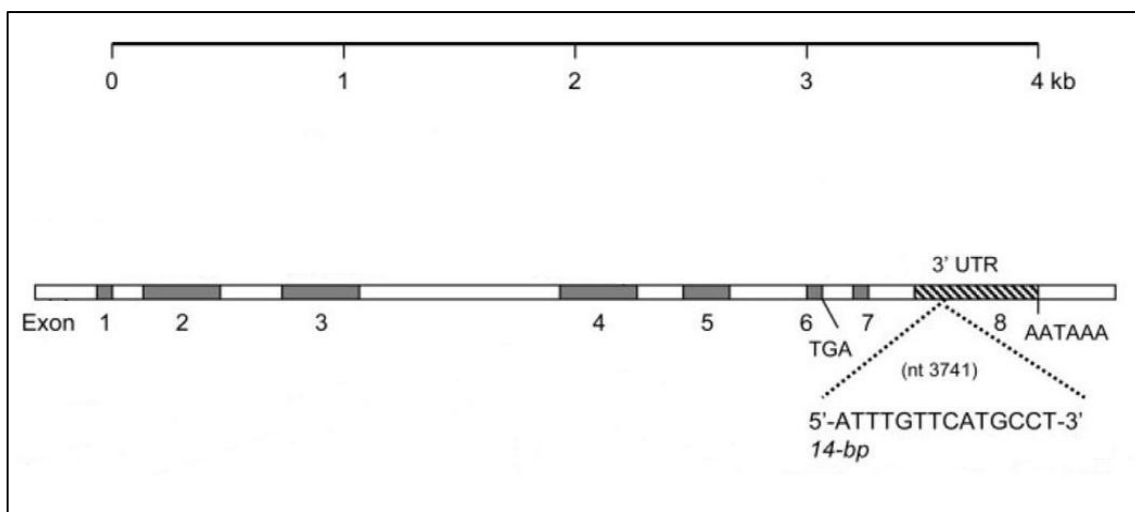


Figura 3. Estrutura molecular do gene HLA-G - éxons representados em preto, íntrons em branco e a região 3'UTR está representada pela área rachurada com destaque para o InDel de 14pb. Esse gene apresenta um códon de parada no éxon 6.

Até o momento, foram reconhecidos oficialmente 44 alelos pelo IMGT, sendo que dois alelos são nulos (IMGT – *International Immunogenetics Information Systems*, <http://www.ebi.ac.uk/imgt/hla/>, acesso em Janeiro de 2011). A maioria dos alelos do gene HLA-G apresentam variações na sequência de regiões codificadoras, sendo que a maior parte deles são substituições sinônimas. Apesar da quantidade de alelos, o HLA-G produz apenas sete tipos de proteínas diferentes geradas por processamento alternativo do transcrito primário, sendo quatro delas proteínas de membrana e três isoformas solúveis (Donadi *et al.* 2010, Figura 4). As proteínas codificadas por esse gene possuem características únicas como: baixa variabilidade, várias proteínas de membrana e isoformas solúveis geradas por processamento alternativo do transcrito primário, cauda citoplasmática reduzida, modulação do sistema imune e expressão restrita a alguns tecidos (Donadi *et al.* 2010).

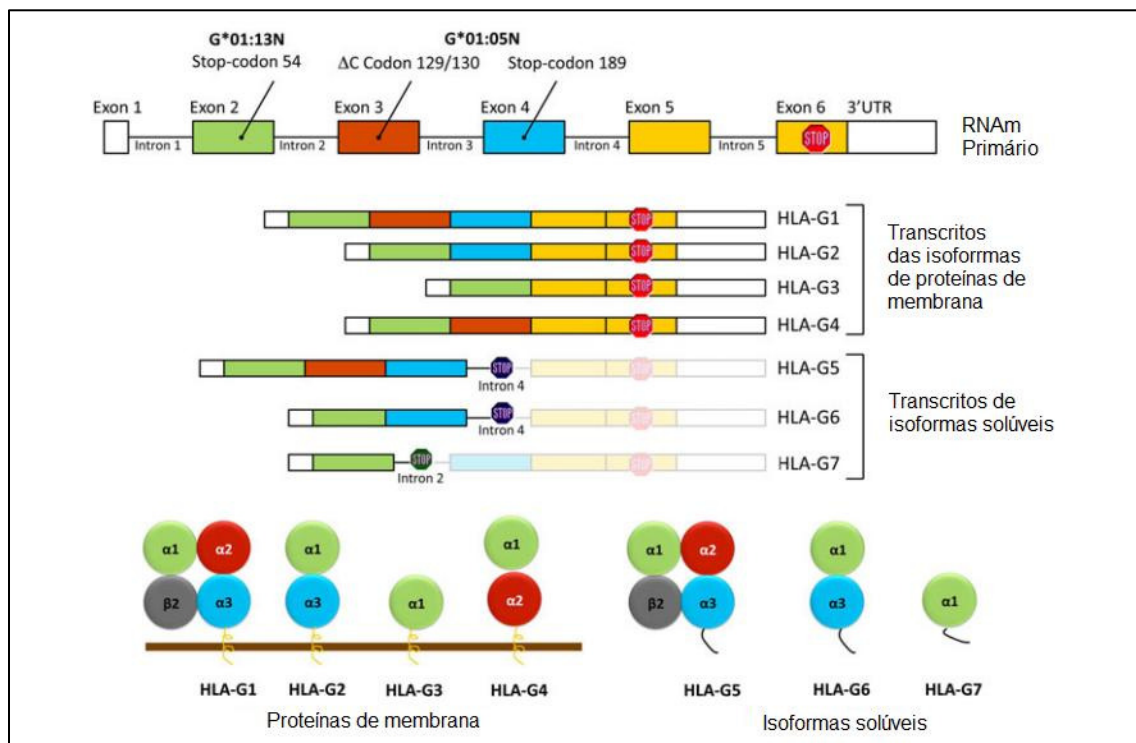


Figura 4. Representação do RNAm primário, dos transcritos após o processamento alternativo e das proteínas codificadas pelo HLA-G. Foi descrito que esse gene produz sete proteínas diferentes por processamento alternativo, são quatro proteínas de membrana e três proteínas solúveis. Modificado de Donadi *et al.* 2010.

Os estudos envolvendo o HLA-G iniciaram após o sequenciamento do gene por Geraghty *et al.* em 1987. Os primeiros estudos descreveram que o HLA-G apresenta

uma grande expressão nas células do trofoblasto (Carosella *et al.* 2008), porção embrionária da placenta pela qual o feto obtém oxigênio e alimento a partir da mãe. Essas células são necessárias para a implantação do embrião no útero materno (Gilbert 2000). Durante a gravidez, o sistema imune da mãe está em contato direto com células e tecidos do feto, que é semialogênico (parcialmente diferente do material genético da mãe). Percebeu-se que devem existir sistemas de modulação do sistema imune para que a mãe não rejeite o próprio feto (Hviid 2006b; Gilbert 2000). Hoje, é sabido que o HLA-G está envolvido na inibição dos linfócitos T citotóxicos e das células NK (*Natural Killer*) permitindo, dessa forma, que o feto se desenvolva de maneira normal num ambiente geneticamente estranho (LeMout *et al.* 2004). A baixa variabilidade do HLA-G assegura a sobrevivência do feto e representa uma evidência de ação de seleção contra mutações que poderiam resultar em proteínas altamente antigênicas (Hviid *et al.* 2006b, Hviid *et al.* 2003, Ober *et al.* 2003).

Hoje, sabe-se também que além das células embrionárias, o HLA-G é expresso em baixas quantidades em tecidos específicos em adultos, como por exemplo, no timo, córnea, células de Langerhans, eritrócitos e células precursoras do endotélio (Carosella *et al.* 2008b). Entretanto, a sua expressão pode ser induzida em casos específicos como em alguns cânceres, transplante de órgãos, infecções virais crônicas, esclerose múltipla e algumas doenças inflamatórias (Carosella *et al.* 2008b, Donadi *et al.* 2010). Dessa forma, diversos trabalhos demonstram que a modulação da expressão do HLA-G pode beneficiar os resultados de transplantes de órgãos, doenças auto-imunes e gestações. Contudo, a presença dos antígenos do HLA-G foi associada a consequências negativas em tratamentos de doenças como câncer e infecções virais, pois a tolerância dada por essas moléculas a células mutadas e/ou infectadas evitam respostas do sistema imune inato e adaptativo (Rizzo *et al.* 2008b).

A expressão tanto induzida quanto a constitutiva dos transcritos do HLA-G é principalmente controlada pela região promotora 5' que exibe características estruturais únicas dentre os genes do HLA (Solier *et al.* 2001). Em contraste com o HLA clássico de classe I, tem sido encontrada pouca variação em sua região codificadora. Já a região promotora do HLA-G é extremamente variável (Tan *et al.* 2005). A variabilidade presente nas regiões codificadora e não-codificadora podem afetar todas as propriedades do gene. Como mostrado na Figura 4, o éxon 7 está sempre ausente nos RNAm maduros e, devido ao códon de parada no éxon 6, o éxon 8 não é traduzido sendo este a

região 3'UTR do RNAm maduro. Essa região contém vários elementos de regulação, incluindo sinais de poliadenilação e motivos de repetição AU, assim como sinais que regulam a expressão de RNAm tanto no espaço quanto no tempo (Figura 5).

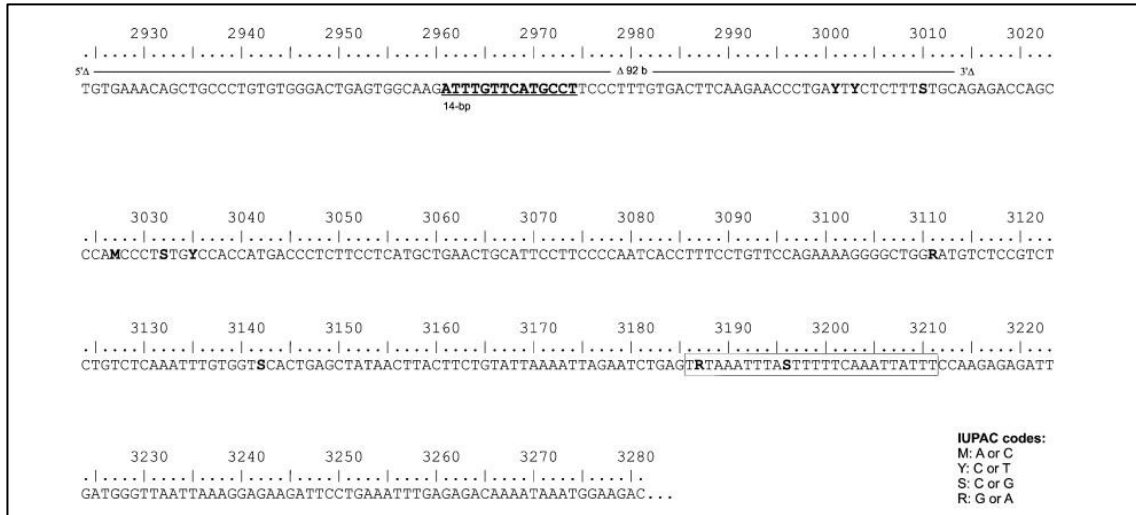


Figura 5. Região 3'UTR do gene HLA-G ressaltando os sítios variáveis descritos na literatura para essa região (em negrito), o fragmento de 92pb que é descartado no processamento alternativo, e o motivo de repetição AU (retângulo) (modificado de Donadi *et al.* 2010).

A região 3'UTR é transcrita mas não é traduzida. O que se sabe é que os transcritos primários devem ser processados e ligados a proteínas antes de serem exportados ao citoplasma (Aguilera 2005). Nesse processo a proteína Cap é ligada, os íntrons são retirados pelo spliceossomo e a porção 3' final é clivada e poliadenilada. O complexo ligado a proteína Cap se liga as estruturas Cap5' monometiladas e as proteínas poli(A) à cauda 3' para exportar as partículas ribonucleoprotéicas mensageiras necessárias para o transporte e tradução do RNAm (Aguilera 2005, Kuersten & Godwin 2003). As proteínas que se ligam ao RNAm podem influenciar na tradução, localização e degradação. Dessa forma, qualquer variação na região 3'UTR de qualquer gene pode influenciar as propriedades ligantes. A disponibilidade de RNAm é constantemente balanceada pela retenção e decaimento no momento em que transcritos deletérios ou não-funcionais são eliminados antes da tradução (Aguilera 2005, Kuersten & Godwin 2003).

A região 3'UTR tem função regulatória, como, por exemplo, controle da expressão e influência na estabilidade dos níveis de transcrição do RNA mensageiro (Tan *et al.* 2007; Harrison *et al.* 1993). No entanto, ainda não foram descritas funções para todas as variações observadas, tendo-se conhecimento sobre as funções de três variações: InDel de 14 pb, e os SNPs +3142 C/G e o +3187 A/G (Figura 5).

Os alelos do HLA-G que apresentam a inserção de 14 pb (5'-ATTTGTTTCATGCCT-3') na região 3'UTR (Harrison *et al.* 1993) foram associados a uma baixa produção de mRNA para a maioria de proteínas de membrana e isoformas solúveis em amostras de trofoblastos (Hviid *et al.* 2003). Por outro lado, dos alelos com a inserção, parte dos transcritos primários são processados e tem a remoção dos primeiros 92 pb do éxon 8 (Hviid *et al.* 2003). Esses transcritos do HLA-G foram descritos como sendo mais estáveis, apesar de serem menores que os mRNAs completos (Rousseau *et al.* 2003).

A região 3'UTR é também alvo de microRNAs – pequenos RNAs não codantes com aproximadamente 22 nucleotídeos. Os microRNAs podem diminuir a expressão impedindo a tradução, degradando o RNA, ou ainda pode realizar ambos os processos (Veit & Chies 2009; Kuersten & Goodwin 2003). O SNP +3142C/G está localizado no éxon 8 a 168 pb do local do InDel de 14 pb (Figura 5), dentro de uma região de 22 pb que é reconhecida por três microRNAs (miRNA) codificados por genes distintos. O alelo +3142G favorece a ligação dos miRNAs em tal região, regulando negativamente a expressão do HLA-G pela supressão da transcrição e pela degradação do mRNA HLA-G (Tan *et al.* 2007).

O outro SNP do qual há estudos demonstrando funcionalidade é o +3187 A/G (Figura 5). Um trabalho demonstrou que a presença de adenina na posição +3187 está associada com pré-eclampsia em uma população Canadense (Yie *et al.* 2008). Esse SNP está próximo a um motivo de repetição de AU que media a degradação de RNAm. O mesmo estudo demonstrou que esse SNP está associado com um decréscimo de RNAm *in vitro* e que a presença da adenina também provoca decréscimo na expressão de HLA-G (Yie *et al.* 2008).

O maior conhecimento atual sobre o HLA-G em comparação aos demais genes de classe Ib deve-se à observação de associação de alelos e padrões de expressão com abortos espontâneos recorrentes (Yan *et al.* 2006, Ober *et al.* 2003, Aldrich *et al.* 2001) e diversas doenças, entre elas, pênfigo vulgar (Gazit *et al.* 2006), síndrome do intestino

irritado (Glas *et al.* 2007), sarcoidose (Hviid *et al.* 2006b), pré-eclampsia (Moureau *et al.* 2008, Vianna *et al.* 2007), asma e doenças pulmonares (Nicolae *et al.* 2005, Rizzo *et al.* 2005), cardiomiopatia dilatada idiopática (Lin *et al.* 2007b), retocolite ulcerativa (Torre *et al.* 2004), psoríase (Aractingi *et al.* 2001), esclerose múltipla (Mitsdoeffer *et al.* 2005), e diversos tumores (Castelli *et al.* 2008, Lin *et al.* 2007a, LeMaoutt *et al.* 2003, Wiendl *et al.* 2002). Também foi observada associação entre alelos e expressão do HLA-G e resultado de transplantes de coração e rins (Crispim *et al.* 2008, Creput *et al.* 2003a, Creput *et al.* 2003b, Lila *et al.* 2000).

Devido à associação entre alelos e/ou genótipos do gene HLA-G com várias doenças e pelas diferentes respostas observadas, é possível inferir que esse gene esteja sob influência direta da seleção natural. Já foi descrito que a região 3'UTR, região abordada no presente estudo, está sob seleção balanceadora (Tan *et al.* 2005). Como já comentado, a seleção balanceadora é aquela cujos genótipos heterozigotos são os de maior valor adaptativo, sendo que ao longo do tempo a frequência deles tende a aumentar (Futuyma 2002).

Esse gene também tem sido alvo de estudos de epidemiologia genética, uma vertente da aplicação da genética de populações cujo objetivo é detectar a origem e distribuição de genes ou alelos que estejam diretamente ou indiretamente relacionados a doenças em uma determinada população (Arnaiz-Villena *et al.* 2002). No caso do HLA-G, o principal foco dos trabalhos realizados até o momento envolve a associação do polimorfismo inserção/deleção de 14 bp com doenças (Hviid 2006, Apps *et al.* 2008, Carosella *et al.* 2008a). Dentro desse contexto, nosso grupo de pesquisa avaliou a ocorrência de dois polimorfismos do gene HLA-G (HLA-G 14pb e HLA-G G*0105N) em mulheres da comunidade *Kalunga* buscando verificar se há associação entre alelos de HLA-G e incidência de abortos espontâneos (Diniz 2008). Essa comunidade apresenta uma alta taxa de abortos espontâneos noticiados - 8,32% das gestações totais (Diniz 2008), o que é similar, porém mais alta, do que o observado por nossa equipe em outros remanescentes de quilombos brasileiros (Oliveira *et al.* 2004). Não foi observada associação entre esses dois polimorfismos e ocorrência de abortos espontâneos nessa comunidade (Diniz 2008).

Por outro lado, até o momento há poucos dados disponíveis sobre frequências populacionais dos alelos de HLA-G, tanto no Brasil como em outros países, onde podemos destacar trabalhos com populações chinesas (Yan *et al.* 2008, Chen *et al.*

2008). No Brasil existem estimativas de frequência e levantamentos de polimorfismos para populações Ameríndias amazônicas (Mendes-Junior *et al.* 2007b) e para duas populações urbanas do sudeste brasileiro (Castelli *et al.* 2009a, Castelli *et al.* 2007).

Porém esse volume de dados é insuficiente para a avaliação do comportamento da distribuição de frequências de alelos desse gene, ainda mais considerando o histórico de formação da população brasileira. Essa população brasileira é resultado de vários anos de história de intercruzamentos com diferentes intensidades de contribuição dos grupos parentais ao longo do país. O Brasil tem um histórico de intensa migração e miscigenação em sua população, por isso é considerado como uma das populações mais heterogêneas do mundo (Parra *et al.* 2003), e demonstrou a maior variabilidade genética para o gene HLA-G (Castelli *et al.* 2007, Castelli *et al.* 2008, Castelli *et al.* 2010). Nesse trabalho foi avaliada a variabilidade genética da região 3'UTR do HLA-G em duas populações do Centro-Oeste brasileiro, uma urbana trihíbrida e uma rural afrodescendente, região geográfica ainda não analisada para padrão de distribuição dos alelos da região 3'UTR do HLA-G.

Dado o perfil genético da população brasileira associado ao pouco conhecimento disponível para essa população quanto ao gene HLA-G, podemos pensar que uma série de mutações no éxon 8 desse gene ocorrem nessa população e ainda não foram descritas. Parte dessas mutações pode estar relacionada com patologias, o que colocaria essa região genômica sob pressão seletiva e levaria a diferenciação populacional na ausência de fatores que anulem a seleção, como acesso aos serviços de saúde.

Objetivos

A variação do genoma é uma das ferramentas mais úteis na reconstrução de história demográfica de populações humanas assim como no entendimento das bases moleculares de doenças e fenótipos complexos. Tendo isso em mente, o objetivo desse trabalho foi buscar novas mutações e analisar polimorfismos anteriormente descritos na região 3' UTR do HLA-G e traçar um panorama de sua distribuição em populações brasileiras de diferentes composições genéticas (trihíbrida urbana e predominantemente afrodescendente rural). A hipótese aqui colocada é que a região 3' UTR do gene HLA-G deve apresentar uma série de mutações, diversas já descritas e outras ainda desconhecidas, devido a história demográfica das populações e que essa região genômica está sob pressão seletiva.

Objetivos específicos:

- a) Buscar novas mutações e polimorfismos na região 3'UTR do gene HLA-G a partir do sequenciamento total do éxon 8 de 134 indivíduos de duas populações brasileiras, Distrito Federal e *Kalunga*;
- b) Avaliar a variação na região 3'UTR, assim como a distribuição e frequência de mutações e polimorfismos, se observados, nas populações com ênfase na comparação entre populações urbanas e rurais;
- c) Avaliar a ocorrência de seleção nesta região genômica utilizando os dados obtidos dessas duas populações.

Material e Métodos

Populações

O universo amostral deste trabalho foi constituído por 72 amostras da população do Distrito Federal e 61 amostras provenientes da população de *Kalunga*. As amostras do Distrito Federal foram coletadas em 2002 e 2008 e as amostras de *Kalunga* em 2001/2002 e 2007. Em todos os casos, as amostras selecionadas foram de pessoas não aparentadas entre elas (grau de parentesco máximo aceitável = 1/16).

Aspectos éticos

O protocolo e o termo de consentimento livre e esclarecido para a coleta e análise das amostras das populações de Distrito Federal e *Kalunga* foram aprovados pelo Conselho Nacional de Ética em Pesquisa (CONEP - registros nº 2938 e 4515, respectivamente). Todos os voluntários dessa pesquisa concordaram com sua participação na pesquisa.

Material biológico

As amostras de DNA utilizadas nesse trabalho já estavam disponíveis para uso, pois pertencem ao banco de DNA do Laboratório de Genética da Universidade de Brasília. As amostras de sangue venoso foram coletadas utilizando o sistema de coleta à vácuo com tubos contendo EDTA como anticoagulante. O DNA foi extraído das amostras com o kit comercial DNA *Illustra Blood Genomic*® da GE segundo instruções do fabricante e quantificado em NanoDrop®.

Análises laboratoriais

Toda a região 3'UTR do gene HLA-G (Figura 5) foi amplificada por PCR (reação em cadeia da polimerase), conforme descrito em Castelli *et al.* (2009b), utilizando os iniciadores HLAG8F – TGTGAAACAGCTGCCCTGTGT e HLAG8R – GTCTTCCATTTATTTTGTCTCT. A reação foi realizada com um volume final de 25 µL contendo 1,2X de tampão de PCR (0,2 M Tris-HCl pH 8,5; 0,5 M KCl), 0,2 mM de cada dNTP, 1,5 mM de MgCl₂, 5 pmol de cada iniciador, 0,5 unidade de *Platinum Taq DNA-polymerase* (Invitrogen, Carlsbad, CA) e 200 ng de DNA genômico. O ciclo inicial de desnaturação foi de 94°C por 3 minutos, seguido por 31 ciclos de 94°C por 30

segundos, 58°C por 25 segundos e 72°C por 25 segundos, seguidos por um passo de extensão final a 72°C por 5 minutos.

Os fragmentos amplificados por PCR (345 pb para o alelo com deleção e 359 pb para o alelo com inserção) foram visualizados em gel vertical de poliacrilamida não desnaturante com glicerol a 6% corado com nitrato de prata 1%. A eletroforese visou confirmar a amplificação do fragmento e genotipagem do InDel de 14 pb, que posteriormente foi confirmada por sequenciamento. As demais variações foram verificadas por sequenciamento.

Para realizar a reação de sequenciamento, o material amplificado foi purificado utilizando as enzimas EXO e Sap (*Exo-nuclease* e *Shrimp alkaline phosphatase* - Affymetrix®). A reação com volume final de 6µl foi constituída por 1µl de EXO/Sap (proporção de 1 EXO : 9 Sap) e 5µl de produto de PCR. Essa etapa foi necessária para retirar o excesso de iniciadores e dNTPs não incorporados durante a reação de PCR.

O material purificado foi utilizado na reação de sequenciamento. Essa reação, com um volume final de 10,0µl, foi composta por 1,7µl de tampão e 0,65µl de fluoróforo do *BigDye™ Terminator kit* (Applied Biosystems, Foster city, CA USA), 32 pmol de iniciador *reverse* - visando evitar possíveis sobreposições em amostras que são heterozigotas para a deleção de 14 pb - e 2,0µl de produto de PCR purificado. A reação ocorreu seguindo o ciclo padrão recomendado pelo fabricante de 25 ciclos a 96°C por 10 segundos, 50°C por 5 segundos e 60°C por 4 minutos.

Ao fim dos ciclos, os produtos das reações foram submetidos ao processo de precipitação. A cada produto de reação, foram acrescentados 1µl de EDTA 125mM, 1µl de acetato de sódio a 3M e 25µl de etanol 100%. As placas contendo as soluções foram vedadas, invertidas e submetidas a descanso por 15 minutos a temperatura ambiente. Após o descanso, o material foi centrifugado por 45 minutos a 1650G e o sobrenadante, descartado. O *pellet* foi submetido a uma lavagem com 35µl de etanol a 70% e centrifugado por 15 minutos à temperatura de 4°C. Por fim, o sobrenadante foi descartado e o *pellet*, secado a vácuo. A preparação final foi a desnaturação do material utilizando 10µl de formamida ultrapura, levada ao termociclador por 5min à temperatura de 94°C. Só então as amostras foram levadas ao sequenciador automático *ABI3130 Genetic Analyzer* (Applied Biosystems, Foster city, CA USA).

O alinhamento das sequências foi feito utilizando o programa Chromas LITE v2.01 (Software gratuito disponível para *download* em

http://www.technelysium.com.au/chromas_lite.html). Toda a variação encontrada na região 3'UTR foi individualmente anotada utilizando-se o software SNPex desenvolvido pelo Dr. Erick da Cruz Castelli (programa gratuito hospedado em www.fmrp.usp.br/immunogen/snpex), adaptado para inferir uma possível ocorrência de haplótipos já conhecidos nessa região.

Caracterização genética da região 3'UTR do gene HLA-G

Para a caracterização genética e comparativa das duas populações quanto a região 3'UTR foi realizada uma série de análises, dentre elas frequência gênica e genotípica, heterozigose observada e esperada, avaliação da aderência dos dados genotípicos ao Equilíbrio de Hardy-Weinberg e diferenciação gênica e genotípica por *locus*. Foi também realizado o teste Theta-p. A estimativa de frequência foi realizada por contagem direta. A aderência ao Equilíbrio de Hardy-Weinberg foi avaliada utilizando o teste exato de Fisher. As diferenciações gênica e genotípica foram avaliadas utilizando a metodologia da cadeia de Markov utilizando os parâmetros *default*. Essas análises foram realizadas utilizando o programa GENEPOP v3.4 (Raymond & Rousset 1995). Para o cálculo de Theta-P e da heterozigose foi utilizado o programa *Genetic Data Analysis - GDA* (Lewis & Zaykin 1997).

Desequilíbrio de ligação

O desequilíbrio de ligação é definido por uma herança não independente entre dois ou mais marcadores. Isso pode ocorrer por diversos fatores, entre eles a proximidade física entre os marcadores avaliados, miscigenação e estruturação populacional (Pfaff *et al.* 2001). Para verificar a ocorrência de desequilíbrio entre as variações que apresentaram frequências polimórficas da região 3'UTR nas populações em análise, foram utilizados os parâmetros *default* do programa HAPLOVIEW v4.2 (Barrett *et al.* 2005).

Análises Haplotípicas da região 3'UTR do gene HLA-G

O cálculo da probabilidade de fase gamética da região 3'UTR para cada cromossomo analisado foi realizado com o programa PHASE (Stephens *et al.* 2001). Para a entrada dos dados no programa a posição relativa das variações observadas em frequência polimórfica em ambas as populações na região 3'UTR do gene HLA-G ao

longo da região foi informada. Na Figura 5 foi apresentada a posição relativa entre os marcadores.

Para a estimativa da probabilidade foram utilizados os seguintes parâmetros: cinco corridas independentes utilizando diferentes valores de partida, número de interações=1000, intervalo de rarefação=1 e *burn in value*=100. As amostras utilizadas para inferências de frequências haplotípicas, heterozigose esperada e observada e teste de diferenciação populacional foram aquelas que resultaram de forma idêntica em todas as corridas e com probabilidade haplotípica acima de 98%.

Testes de Seleção

Para avaliação da seleção natural foram realizados os seguintes testes: *Teste D de Tajima*, *Teste F de Fu & Li* e *Teste D de Fu & Li*, todos utilizando os parâmetros *default* do programa DNASP v5.10.01 (Librado & Rozas 2009). Além desses, foi feito o teste de neutralidade de Ewens–Watterson utilizando os parâmetros *default* do programa Pypop v0.7.0 (Lancaster *et al.* 2003).

O *teste D de Tajima* leva em consideração a diversidade da sequência (Tajima 1989), sendo que a base deste é que os sítios segregantes e o número médio de diferenças par a par diferem fundamentalmente na presença de seleção. A fórmula do cálculo se baseia na diferença de nucleotídeos menos a razão entre sítios segregantes e tamanho da amostra numa janela de 250 mil anos (Sabeti *et al.* 2006). Como resultado O teste possibilita verificar ocorrência de seleção e indicar o tipo de seleção que estaria ocorrendo.

Os *testes D e F de Fu & Li* baseia-se no padrão de frequências, sendo que o teste *D* avalia excesso de mutações jovens e o *F* excesso de mutações antigas. Esse teste também avalia assinaturas em uma janela histórica de 250 mil anos. Dessa forma, a comparação do número de mutações antigas e novas ao número de mutações esperadas sob seleção neutra é um teste robusto para indicar seleção (Fu & Li 1993).

O *Teste de neutralidade de Ewens-Watterson*, o *F* normalizado, foi desenvolvido a partir do trabalho de Ewens *et al.* (1972) e de Watterson (1978). O teste adequado aos dados desse trabalho compara a homozigose observada em uma amostra com aquela esperada a partir da equação proposta por Ewens *et al.* (1972) que proporciona uma base de comparação de frequências alélicas observadas em amostras com aquelas

esperadas sob o modelo de alelos infinitos em neutralidade. Nesse modelo é possível avaliar assinaturas de seleção em janelas de evolução recente.

Resultados

Caracterização genética da região 3'UTR do gene HLA-G

No presente trabalho foram observadas oito variações, descritas previamente (Castelli *et al.* 2009), com frequências polimórficas na região 3'UTR. Das variações observadas, uma é um InDel de 14 pb (número de acesso: rs1704, <http://www.ncbi.nlm.nih.gov/>) e os outros sete são variações de ponto (SNP – do inglês: *Single Nucleotide Polimorphism*): +3003 C/T (rs1707); +3010 G/C (rs1710); +3027 A/C (rs17179101); +3035 C/T (rs17179108); +3142 G/C (rs1063320); +3187 G/A (rs9380142); +3196 G/C (rs1610696). Além desses, foi identificada uma mutação de ponto, não descrita previamente, em um indivíduo da população *Kalunga* na posição +3052 C/T.

A Tabela 1 mostra frequências gênicas e genotípicas e o teste de aderência ao Equilíbrio de Hardy-Weinberg (EHW) e heterozigose esperada e observada dos marcadores genéticos nas duas populações. Não foi observado nenhum marcador fora do equilíbrio. No Distrito Federal, cinco dos oito marcadores apresentaram uma heterozigose observada ligeiramente maior do que a esperada. Em *Kalunga* o mesmo ocorreu com seis marcadores. O resultado do teste de diferenciação gênica e genotípica entre as duas populações está apresentado na Tabela 2, onde se observa diferenciação gênica e genotípica entre os marcadores +3003 C/T e +3187 A/G. Na mesma tabela também estão apresentados os resultados do teste Theta-P. O intervalo de confiança apresentado inclui o zero, portanto o valor não é significativo estatisticamente.

Tabela 1. Frequências gênicas, genótípicas, heterozigose observada e esperada dos marcadores genéticos da região 3'UTR do gene HLA-G em Distrito Federal e *Kalunga*.

Marcadores genéticos		Distrito Federal		<i>Kalunga</i>	
Del/Ins 14pb	Alelo	Del	0,6250	Del	0,5984
		Ins	0,3750	Ins	0,4016
	Genótipo	Del/Del	0,4166	Del/Del	0,3443
		Del/Ins	0,4166	Del/Ins	0,5082
		Ins/Ins	0,1668	Ins/Ins	0,1475
	pEHW		0,3268		0,7919
	He		0,4720		0,4846
	Ho		0,4166		0,5082
3003 T/C	Alelo	T	0,8944	T	0,7931
		C	0,1056	C	0,2069
	Genótipo	T/T	0,8169	T/T	0,6034
		T/C	0,1549	T/C	0,3793
		C/C	0,0282	C/C	0,0173
	pEHW		0,1584		0,4252
	He		0,1775		0,3172
	Ho		0,1549		0,3793
3010 G/C	Alelo	G	0,4627	G	0,4426
		C	0,5373	C	0,5574
	Genótipo	G/G	0,2238	G/G	0,1803
		G/C	0,4777	G/C	0,5246
		C/C	0,2985	C/C	0,2951
	pEHW		0,8065		0,7956
	He		0,4722		0,4975
	Ho		0,4776		0,5245
3027 C/A	Alelo	C	0,9375	C	0,9426
		A	0,0625	A	0,0574
	Genótipo	C/C	0,8750	C/C	0,9018
		C/A	0,1250	C/A	0,0819
		A/A	0,0000	A/A	0,0163
	pEHW		1,0000		0,1663
	He		0,1180		0,1090
	Ho		0,1250		0,0819

continuação tabela 1

Marcadores genéticos		Distrito Federal		<i>Kalunga</i>	
3035 C/T	Alelo	C	0,8542	C	0,8167
		T	0,1458	T	0,1833
	Genótipo	C/C	0,7222	C/C	0,6667
		C/T	0,2639	C/T	0,3000
		T/T	0,0139	T/T	0,0333
	pEHW		1,0000		1,0000
	He		0,2508		0,2988
	Ho		0,2638		0,3000
3142 G/C	Alelo	G	0,5500	G	0,5574
		C	0,4500	C	0,4426
	Genótipo	G/G	0,2714	G/G	0,2951
		G/C	0,5571	G/C	0,5246
		C/C	0,1715	C/C	0,1803
	pEHW		0,3439		0,7956
	He		0,4920		0,4975
	Ho		0,5571		0,5245
3187 A/G	Alelo	A	0,7324	A	0,8525
		G	0,2676	G	0,1475
	Genótipo	A/A	0,5070	A/A	0,0328
		A/G	0,4507	A/G	0,2295
		G/G	0,0423	G/G	0,7377
	pEHW		0,3607		0,6006
	He		0,3909		0,2536
	Ho		0,4507		0,2295
3196 C/G	Alelo	C	0,7708	C	0,2049
		G	0,2292	G	0,7951
	Genótipo	C/C	0,6111	C/C	0,0163
		C/G	0,3195	C/G	0,3770
		G/G	0,0694	G/G	0,0163
	pEHW		0,5011		0,4290
	He		0,3557		0,3285
	Ho		0,3194		0,3770

Tabela 2. Teste de diferenciação gênica e genotípica, e Theta-P entre Distrito Federal e *Kalunga* avaliado para os marcadores genéticos da região 3'UTR do gene HLA-G.

Marcadores	Gênica		Genotípica		Theta-P (IC)
	p	dp	p	dp	
InDel 14pb	0,7055	0,0023	0,7124	0,0016	-0,0064
3003 C/T*	0,0355	0,0008	0,0335	0,0010	0,0322
3010 C/G	0,6232	0,0029	0,8025	0,0011	-0,0047
3027 A/C	1,0000	0,0000	1,0000	0,0000	-0,0079
3035 C/T	0,4059	0,0025	0,5004	0,0023	-0,0010
3142 C/G	0,9017	0,0008	1,0000	0,0000	-0,0067
3187 A/G*	0,0151	0,0005	0,0194	0,0007	0,0375
3196 C/G	0,6581	0,0021	0,6590	0,0017	-0,0058
Total	-	-	-	-	0,0033(0,0173 a -0,0057)

* = Valores significativos.

Teste de desequilíbrio de ligação

A Figura 6 indica o resultado do teste de desequilíbrio de ligação realizado para as duas populações considerando os oito marcadores par a par, sendo que A representa Distrito Federal e B *Kalunga*. Em um total de 28 comparações, no Distrito Federal foram observados 17 pares de marcadores em forte desequilíbrio, oito pares em baixo desequilíbrio e três pares com ausência de desequilíbrio. Em *Kalunga* foram encontrados 16 pares em forte desequilíbrio de ligação, 11 pares em baixo desequilíbrio e um par com ausência de desequilíbrio. A ausência de desequilíbrio foi observada entre os pares +3003 e +3027, +3003 e +3035; e entre o +3027 e +3196 para Distrito Federal e entre +3003 e +3196 para *Kalunga*.

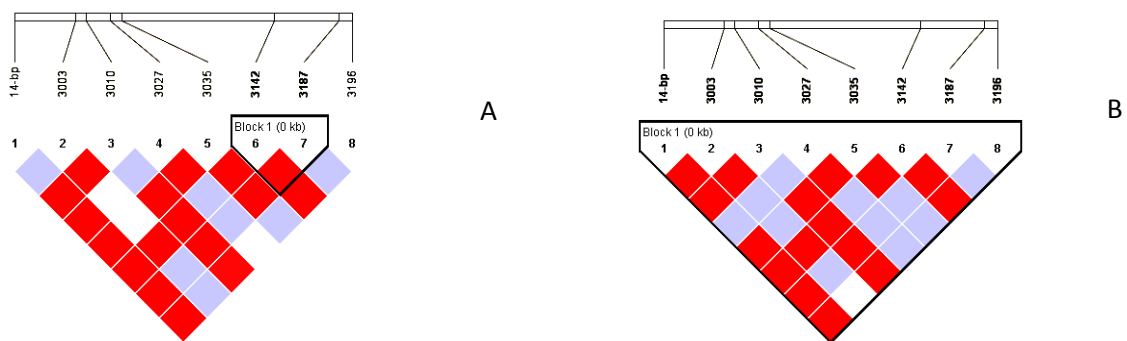


Figura 6. Matriz de desequilíbrio de ligação entre os oito marcadores analisados par a par: A - Distrito Federal, B - *Kalunga*. As barras acima indicam a posição relativa dos marcadores na região 3'UTR no gene do HLA-G, enumerados de um a oito. Branco - ausência de desequilíbrio, azul - baixo desequilíbrio e vermelho - forte desequilíbrio.

Análises Haplotípicas da região 3'UTR do gene HLA-G

Três amostras do Distrito Federal não se adequaram às condições de definição haplotípica e foram retiradas das análises subsequentes. A análise utilizando o programa PHASE apontou a provável ocorrência de onze haplótipos diferentes, denominados de UTR1 a UTR11. A nomenclatura utilizada foi a proposta por Castelli *et al.* (2009). O indivíduo *Kalunga* que apresentou a mutação não descrita foi definido como sendo heterozigoto UTR3/UTR4.

Na Tabela 3 estão apresentadas as frequências haplotípicas observadas em ambas as populações. No Distrito Federal, o haplótipo mais frequente foi o UTR1, que somado aos haplótipos que tiveram frequências iguais como segundo mais frequente, UTR 2 e UTR 3, representaram 55,3% dos haplótipos encontrados. Em *Kalunga* o haplótipo mais frequente foi o UTR4, seguido do UTR3, que juntos representaram 37,72% dos haplótipos da população. Além disso, é importante apontar que nessa população não foi encontrado nenhum indivíduo com os haplótipos UTR9, UTR10 e UTR11. O teste de diferenciação gênica e genotípica realizado para os haplótipos teve como resultado $p=0,0472$ ($dp=0,0030$) e $p=0,0599$ ($dp=0,0044$) respectivamente. A heterozigose observada e esperada para o Distrito Federal foi de $He=0,8626$ e $Ho=0,8059$, e em *Kalunga* foi de $He=0,8650$ e $Ho=0,8596$. Também na Tabela 3, estão apresentados os resultados do teste Theta-P para os haplótipos das duas populações. O Theta-P apresentou valores baixos. Além disso, o intervalo de confiança apresentado não incluiu o zero, o que tornou o resultado significativo.

Tabela 3. Haplótipos da região 3'UTR do HLA-G e suas frequências nas populações Distrito Federal e *Kalunga* ($2n$ = número de cromossomos avaliados) e valores de Theta-P (em parênteses IC = Intervalo de Confiança).

Haplótipos	Frequências		Haplótipos da região 3'UTR do HLA-G								Theta-P (IC)
	DF $2n = 252$	Kalunga $2n = 228$	14 pb	3003	3010	3027	3035	3142	3187	3196	
UTR 1	0,2652	0,1404	Del	T	G	C	C	C	G	C	0,0361
UTR 2	0,1439	0,1228	Ins	T	C	C	C	G	A	G	-0,0082
UTR 3	0,1439	0,1667	Del	T	C	C	C	G	A	C	-0,0058
UTR 4	0,1136	0,2105	Del	C	G	C	C	C	A	C	0,0282
UTR 5	0,0758	0,1316	Ins	T	C	C	T	G	A	C	0,0044
UTR 6	0,0530	0,0789	Del	T	G	C	C	C	A	C	-0,0039
UTR 7	0,0682	0,0526	Ins	T	C	A	T	G	A	C	-0,0070
UTR 8	0,0909	0,0965	Ins	T	G	C	C	G	A	G	-0,0071
UTR 9	0,0227	-	Del	T	G	C	C	G	A	C	0,0128
UTR10	0,0152	-	Ins	T	G	C	T	G	A	C	0,0058
UTR11	0,0076	-	Ins	T	C	C	C	G	A	C	-0,0012
Total	-	-	-	-	-	-	-	-	-	-	0,0085 (0,0085 a 0,0085)

Testes de Seleção

Os testes para avaliação de seleção natural forneceram como resultados os dados apresentados na Tabela 4. Distrito Federal apresentou a maior diversidade nucleotídica, assim como o maior número de haplótipos. A maior diversidade haplotípica foi encontrada em *Kalunga*. Distrito Federal apresentou resultado significativo para o *Teste D de Tajima*, o *Teste F de Fu & Li* e para o teste de *Ewens-Watterson*. *Kalunga* teve resultado significativo para o *Teste F de Fu & Li* e para o teste de *Ewens-Watterson*. O *Teste D de Fu & Li* não foi significativo para nenhuma das duas populações.

Tabela 4. Resultados dos testes de neutralidade realizados com os marcadores genéticos da região 3'UTR do HLA-G nas populações Distrito Federal e *Kalunga*.

População	π	k	Π	Testes de neutralidade			
				D (p)	F (p)	D ¹ (p)	F
Distrito Federal	0,0089	11	0,8300	2,1419* p < 0,05	1,8458* p < 0,05	1,2431 p > 0,10	-1,4140* p = 0,0086
<i>Kalunga</i>	0,0087	8	0,8410	2,0457 0,10 > p > 0,5	1,8079* p < 0,05	1,2501 p > 0,10	-1,6354* p = 0,0008

Tamanho = Tamanho do fragmento; π = Diversidade Nucleotídica; k = Número de haplótipos; Π = Diversidade haplotípica; D = *Teste D de Tajima*; F = *Teste F de Fu & Li*; D¹ = *Teste D de Fu & Li*; F = *Ewens-watterson* F normalizado, * = Valores significativos.

Discussão

Caracterização genética da região 3'UTR do gene HLA-G

As frequências alélicas observadas em cada marcador no Distrito Federal são, de modo geral, semelhantes às apresentadas por *Kalunga*. Todos os marcadores estão em equilíbrio de Hardy-Weinberg, fato também observado em Castelli *et al.* (2009) em uma população urbana brasileira. Esse é o único artigo, disponível até o momento, que avaliou os mesmos parâmetros nos mesmos marcadores avaliados neste trabalho. A frequência do alelo deletado do InDel de 14 pb foi de aproximadamente 60% em ambas as populações. A similaridade da frequência foi confirmada pelo teste de diferenciação gênica e genotípica, não significativo para as frequências do InDel. Ainda, é interessante destacar que populações de diferentes constituições genéticas derivadas de diferentes ancestralidades apresentaram valores similares na frequência do alelo deletado, como, por exemplo, uma população brasileira (Castelli *et al.* 2009), uma dinamarquesa (Hylenius *et al.* 2004), uma população chinesa Han (Chen *et al.* 2008) e em tribos amazônicas (Mendes-Junior *et al.* 2007). Como dito, esse marcador está associado a baixa expressão gênica e à estabilidade do RNAm (Rizzo *et al.* 2008).

Dos oito marcadores analisados, as frequências alélicas de dois deles mostraram diferenças estatisticamente significativas entre as populações estudadas. No SNP +3003 T/C Distrito Federal apresentou uma frequência de 89% do alelo T, enquanto *Kalunga* apresentou 79%. Esse SNP não teve até o momento funcionalidade descrita, porém sabe-se que quando há o processamento alternativo induzido pelo alelo com a inserção de 14pb, o SNP +3003 está contido nos 92 pb que são descartados (Hviid *et al.* 2003). Além disso, os valores observados em uma população brasileira do estado de São Paulo (Castelli *et al.*, 2009) são valores intermediários aos encontrados no Distrito Federal e *Kalunga*. No SNP +3187 A/G, foi observada em *Kalunga* a frequência de 85% do alelo A e no Distrito Federal 73%, diferença essa que teve como reflexo o valor significativo no teste de diferenciação gênica e genotípica. Frequência semelhante ao Distrito Federal para esse SNP é observada em uma população brasileira do estado de São Paulo (Castelli *et al.* 2009). O alelo +3187*A é relatado como associado a uma baixa produção de RNAm (Yie *et al.* 2008).

A heterozigose observada, de forma geral, foi levemente maior que a esperada. Contudo, essa diferença não foi significativa. Por isso, baseado nesses dados não é

possível afirmar se o excesso de heterozigotos em ambas as populações é sinal de ação de mecanismos evolutivos mantendo esse padrão em populações distintas.

O teste Theta-P é similar ao *Fst* e são testes amplamente utilizados em genética de populações para marcadores seletivamente neutros. Aplicado a marcadores neutros esse teste fornece um resultado de diferenciação populacional baseado numa matriz de distância (Weir & Cockerham 1984). Contudo, os marcadores utilizados nesse trabalho foram descritos como marcadores que estão sob seleção, portanto não neutros. O teste de Theta-P aplicado aos marcadores não neutros possibilita inferência sobre seleção natural (Barreiro *et al.* 2008). O intervalo de confiança (IC) é um resultado similar ao valor de *p* e indica se há seleção atuando sobre os marcadores analisados. A significância dos valores do IC é dada quando o intervalo resultante não inclui o valor zero (Weir & Cockerham 1984). Os valores de Theta-P mostram que tipo de seleção é mais provável de estar atuando. Valores baixos, independente do sinal do resultado (positivo ou negativo), são indicativos de que os marcadores estão passando por seleção balanceadora, enquanto valores mais altos indicam seleção positiva (Barreiro *et al.* 2008). Os resultados obtidos mostram que não há seleção atuando sobre os marcadores separadamente.

Desequilíbrio de ligação

O teste do desequilíbrio de ligação indica se a segregação entre os marcadores avaliados está ocorrendo de forma independente ou não. O desequilíbrio observado em ambas as populações pode ser resultado de dois mecanismos. Uma explicação é a proximidade física entre os marcadores, pois a proximidade física diminui a probabilidade de recombinação entre os marcadores. A região 3'UTR do gene HLA-G tem como extensão total máxima 359pb, porém do começo do Indel de 14pb ao último SNP avaliado da região a distância diminui para no máximo 263pb, portanto uma região muito pequena.

Outra possibilidade é relacionada a história de miscigenação das duas populações. Pfaff *et al.* (2001) propôs que pode ocorrer desequilíbrio de ligação entre marcadores em consequência de miscigenação (Figura 7). No modelo A (Figura 7) o desequilíbrio ocorre no momento da miscigenação e tende a desaparecer ao longo do tempo. Já o modelo B (Figura 7) mostra que o desequilíbrio ocorre no momento da miscigenação e é mantido na população ao longo do tempo pelo constante fluxo gênico (Pfaff *et al.* 2001).

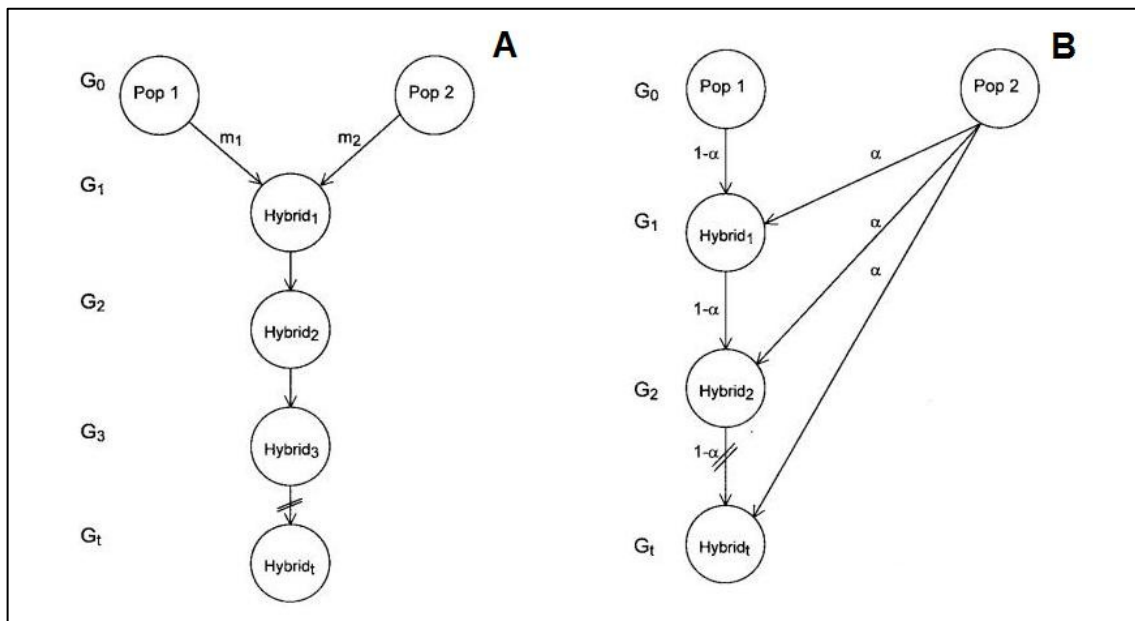


Figura 7. Modelos que explicam o desequilíbrio de ligação por miscigenação. Modificado de Pfaff *et al.* (2001).

Apesar das duas populações se adequarem aos modelos propostos por Pfaff *et al.* (2001), acredita-se que o desequilíbrio em ambas as populações seja resultado da proximidade física entre os marcadores, sendo que o desequilíbrio de ligação dessa região do HLA-G já foi registrado para outras populações (Donadi *et al.* 2010, Castelli *et al.* 2009; Nicolae *et al.* 2005). O desequilíbrio acentuado dessa região mostra que a análise mais adequada para regiões como essa é na forma de haplótipos.

Análises Haplótípicas da região 3'UTR do gene HLA-G

O Distrito Federal apresentou frequências haplotípicas semelhantes às observadas por Castelli *et al.* (2009) na população de Ribeirão Preto (Estado de São Paulo). As duas áreas são urbanas e em ambas as populações os haplótipos mais frequentes foram o

UTR1. No Distrito Federal esse haplótipo representa 26,52%, enquanto em Ribeirão Preto 25,8%. As populações diferiram quanto ao segundo haplótipo mais frequente: enquanto no Distrito Federal observou-se que os haplótipos UTR2 e UTR3 apresentaram a mesma frequência (14,39%), em Ribeirão Preto o segundo mais frequente foi o UTR2 representando 24,2% dos haplótipos encontrados (Castelli *et al.* 2009). Em *Kalunga*, os haplótipos mais frequentes foram UTR3 e UTR4. Juntos os dois representaram 37,72% dos haplótipos encontrados. A diferença quanto aos principais haplótipos nas duas populações poderia ser explicada pelo histórico de formação e a ancestralidade das duas populações, sendo assim parcialmente resultado de efeito fundador. Os testes de diferenciação gênica e genotípica, considerando um valor α de 0,05, resultaram em valores limítrofes (respectivamente, $p=0,047$ ($dp=0,003$) e $p=0,059$ ($dp=0,004$)). Quando se considera o desvio padrão, verifica-se que o intervalo obtido inclui o valor α . Portanto, apesar das diferenças observadas nas frequências haplotípicas e nos haplótipos mais comuns em cada população, não há diferenças quanto à distribuição de frequências nas duas populações.

Outro aspecto a ser abordado é a funcionalidade dos SNPs constituintes de cada haplótipo e a relação função/frequência. Foi observada a ocorrência de um padrão na constituição haplotípica: haplótipos com o alelo inserção do InDel de 14 pb necessariamente apresentam, também, em sua constituição os alelos +3142*G e +3187*A. Isso caracteriza os haplótipos UTR2, UTR5, UTR7, UTR8, UTR10 e UTR11. Todos esses alelos estão associados a uma baixa expressão gênica (Tan *et al.* 2009; Veit & Chies 2007), dessa forma, em teoria, esses haplótipos devem ser característico de baixa expressão (Castelli *et al.* 2009). Considerando a questão da expressão gênica, observamos que a soma das frequências dos supostos haplótipos de baixa expressão fornecem valores similares nas duas populações. Assim, mesmo com um número de haplótipos diferentes, verifica-se que a região 3'UTR do gene HLA-G são similares quanto a funcionalidade quando compara-se Distrito Federal e *Kalunga*.

A heterozigose esperada foi ligeiramente maior que a observada para os haplótipos nas duas populações. Dessa forma, temos que a heterozigose, nesse estudo, não foi um bom parâmetro para verificar seleção. Já o teste Theta-P foi significativo, pois o intervalo de confiança não incluiu o zero. Esse resultado indicou que há ação da seleção sobre os haplótipos, e não sobre os marcadores separadamente. Os valores de

Theta-P foram bastante baixos, o que indicou ação de seleção balanceadora (Barreiro *et al.* 2008).

Testes de Seleção

A região 3'UTR do HLA-G apresenta uma diversidade nucleotídica 1,4 vezes maior que a região promotora do mesmo gene (Tan *et al.* 2005). Nesse trabalho foi observada uma diversidade nucleotídica, em ambas as populações, onze vezes maior que a média de 0,00075 estimada para o genoma humano (Sachidanandam *et al.* 2001). Além disso, foi verificado que o Distrito Federal apresenta um número maior de haplótipos comparado à *Kalunga*, provavelmente resultado de efeito fundador. A população do Distrito Federal foi inicialmente constituída a partir da migração de pessoas de várias partes do Brasil, sendo que as populações de origem já apresentavam miscigenação. Já *Kalunga* teve em sua história a migração preferencial de um grupo, africanos e/ou afroderivados, em sua formação. Contudo, ainda que o número de haplótipos no Distrito Federal seja maior, a diversidade haplotípica foi menor. Isso se deve ao fato de que o cálculo da diversidade haplotípica leva em consideração o tamanho da amostra e a distribuição uniforme das frequências (Librado & Rozas 2009). Assim, apesar do menor número, *Kalunga* apresentou frequências mais uniformes do que o Distrito Federal.

Buscando verificar se a região 3'UTR do gene HLA-G está sendo ou foi alvo de seleção, foram realizados diversos testes de seleção. O teste *D de Tajima* (Tajima 1989) mostra se há seleção e o valor de D mostra que tipo de seleção estaria ocorrendo. Nos resultados aqui obtidos, os valores estimados de D foram positivos, o que pode ser reflexo de dois fatores. Nas populações vemos que ambos os valores de D foram positivos e maiores que dois. Esse resultado de D sugere algum tipo de seleção balanceadora, na qual os genótipos heterozigotos são favorecidos, ou algum tipo de seleção diversificadora, na qual os genótipos que contem alelos menos frequentes são favorecidos. Outra explicação para a ocorrência de valores de D positivos é se a população amostrada foi formada a partir de uma mistura recente entre duas populações distintas (Hartl & Clark 2010). No presente trabalho, apesar do valor de D ser positivo e maior que dois em ambas as populações, somente Distrito Federal apresentou um valor de p significativo para o *Teste D de Tajima*. Dados os parâmetros desse teste específico e o tempo avaliado pelo mesmo, não podemos julgar se a significância pode ser fruto de

seleção ou resultado da recente formação da população com o resultado desse teste. No caso de *Kalunga*, o valor de p não foi significativo, mas foi um valor limítrofe. Sugere-se, pelos valores obtidos, que possivelmente um número amostral maior seja suficiente para que o teste se torne significativo.

Como já comentado, os testes D e F de Fu & Li se baseiam no padrão de frequências, sendo que o teste D avalia excesso de mutações jovens e o F excesso de mutações antigas. Tanto no Distrito Federal quanto em *Kalunga* observou-se valores positivos nos testes D e F . Os testes desenvolvidos por Fu & Li (1993), $Fu \& Li D$ e $Fu \& Li F$, tem uma abordagem um pouco diferenciada do teste D de Tajima. Os testes D e F consideram a distribuição das mutações na genealogia de uma amostra aleatória de genes da população. Como esse teste também avalia assinaturas em uma janela histórica de 250 mil anos, torna-se possível estimar a idade das mutações. Mutações antigas tendem a ser encontradas nos ramos internos da genealogia, enquanto mutações novas tendem a ser encontradas em ramos mais externos. Na presença de seleção purificadora, haverá uma tendência a mutações em ramos externos, isso porque alelos deletérios estão presentes em baixa frequência. Além disso, haverá excesso de mutações recentes se um alelo vantajoso foi fixado há pouco tempo na população. Dessa forma, se espera que a maioria das mutações sejam recentes. Por outro lado, se seleção balanceadora estiver exercendo pressão nos marcadores, então haverá indicação de que alguns alelos são muito antigos, ou ainda, que há uma deficiência de mutações em braços externos. Dessa forma, a comparação do número de mutações antigas e novas ao número de mutações esperadas sob seleção neutra é um teste robusto para indicar seleção (Fu & Li 1993). O teste $Fu \& Li D$ em ambas as populações se mostrou não significativo, se adequando à hipótese nula do teste, a de evolução neutra. Apesar disso, o valor de D positivo indica que há uma deficiência de mutações recentes. Já o teste F apresenta valores significativos para ambas as populações, mostrando um excesso de mutações antigas. Esses dados confirmam a indicação de que o HLA-G é um gene que tem uma longa história evolutiva (Donadi *et al.* 2010).

O Teste de neutralidade de Ewens-Watterson possibilita avaliar assinaturas de seleção em janelas de evolução recente. O p apresentado por ambas as populações é significativo, dessa forma, ambas as populações não aderem à hipótese nula de neutralidade. Além disso, as populações indicaram valores de F normalizado negativos. Para esse teste o F normalizado negativo indica que a homozigose observada é menor

que a homozigose esperada, resultado que indica seleção balanceadora (Hartl & Clark 2010, Mendes-Junior *et al.* 2007).

A análise dos resultados como um todo é a melhor forma para avaliar as populações. Os resultados positivos do teste D de Tajima mostraram uma possibilidade de seleção balanceadora. Esses resultados, aliados aos resultados significativos do *Teste F de Fu & Li*, que indica a presença de mutações antigas, e aos resultados negativos do *Teste de Ewans-Watterson*, que também sinalizam seleção balanceadora, mostraram que a região 3'UTR do HLA-G apresenta indicações de que houve uma seleção balanceadora na região, que possivelmente está atuando há um tempo bastante longo e não permitiu a diferenciação das duas populações mesmo com uma história demográfica diferente (Donadi *et al.* 2010, Mendes-Junior *et al.* 2007, Tan *et al.* 2005). Contudo, a região 3'UTR, como já dito, é melhor avaliada na forma de haplótipos como demonstrado pelo teste de desequilíbrio de ligação. Dessa maneira, a interpretação da seleção balanceadora sobre os haplótipos deve ser feita quanto a funcionalidade dos haplótipos. A seleção balanceadora aumenta a frequência dos heterozigotos, portanto, ela deve ocorrer aumentando a frequência de haplótipos de funcionalidade diferente. Dessa forma, a seleção balanceadora deve aumentar a frequência de haplótipos de expressão normal e haplótipos de baixa expressão, agrupando: expressão normal – UTR1; expressão superbaixa – UTR2, UTR5, UTR7, UTR8, UTR10 e UTR11; expressão baixa – UTR3, UTR4, UTR6 e UTR9.

Considerações Finais

De acordo com os resultados aqui mostrados, as população do Distrito Federal e *Kalunga* não apresentam diferenças estatisticamente significativas para a distribuição alélica e haplotípica da região 3'UTR do gene HLA-G. Trabalhos anteriores com essa região sugeriram a ocorrência de seleção, porém nenhum deles tinha a informação da constituição da população com marcadores reconhecidamente neutros, dados estes disponíveis para Distrito Federal e *Kalunga*. Sabe-se de trabalhos anteriores realizados com as mesmas populações e mesmas amostras que essas populações são distintas, isso é, apresentam constituição genética bastante diferente. Essa observação leva ao questionamento de porque diferentes regiões genômicas se comportam de maneira tão distinta. Somando a esse questionamento estão os testes de seleção, onde parte destes

sugere a ocorrência de seleção para essa região ou a constituição recente das populações.

Apesar do conhecimento de que ambas populações apresentam história recente, há de se comentar que distintas populações de distintas constituições apresentam similaridade de frequência levando a sugestão que o que estamos observando é indicativo de seleção. Sendo assim, ao contrário da história demográfica e do perfil genético de marcadores neutros, a região 3'UTR do gene HLA-G do Distrito Federal e *Kalunga* são similares. Enquanto a seleção neutra atua uniformemente sobre todo o genoma, a seleção natural atua sobre pontos específicos do genoma. Dessa forma, os genes seletivamente neutros mostram que o acaso levou a uma grande diferenciação populacional. Contudo, a seleção natural mantém genes de caráter adaptativo com frequências similares nas populações.

Conclusões

A análise dos oito marcadores da região 3'UTR do gene HLA-G em duas comunidades brasileiras, Distrito Federal e *Kalunga*, e a comparação com outras permitiu propor as seguintes conclusões:

- ✓ Foi encontrada uma mutação nova, ainda não descrita, em um indivíduo da população *Kalunga*;
- ✓ As populações não diferiram significativamente quanto às frequências gênicas e genotípicas;
- ✓ As populações não apresentaram diferenciação genética para os oito *loci* de acordo com a estimativa de Theta-P;
- ✓ Apesar de ambas populações apresentarem-se em equilíbrio de Hardy-Weinberg para os oito *loci* analisados, ambas populações apresentaram uma heterozigose observada ligeiramente maior que a esperada, sugerindo a atuação de forças evolutivas para manutenção dos heterozigotos;
- ✓ Os oito marcadores da região 3'UTR apresentaram-se em desequilíbrio de ligação, indicando a necessidade do estudo na forma de haplótipos;
- ✓ As populações não apresentaram diferenças estatisticamente significantes para as frequências haplotípicas, além de apresentarem heterozigose esperada maior que a observada;
- ✓ As populações não mostraram significância estatística quanto a diferenciação populacional dos haplótipos;
- ✓ As populações tem alta diversidade nucleotídica, sendo que *Kalunga* apresentou diversidade haplotípica maior que Distrito Federal;
- ✓ Brasília mostrou resultado significativo para o teste *D de Tajima*, ambas as populações apresentaram resultados significantes para o teste *F de Fu & Li* e para o teste de *Ewens-Watterson*, e nenhuma teve resultados significantes para o teste *D de Fu & Li* sugerindo a ocorrência de seleção para essa região genômica.

Referências

- Aguilera A (2005). Cotranscriptional mRNP assembly: from the DNA to the nuclear pore. *Curr Opin Cell Biol* 17:242–250.
- Aldrich CL, Stephenson MD, Karrison T, Odem RR, Branch DW, Scott JR et al. (2001). HLA-G genotypes and pregnancy outcome in couples with unexplained recurrent miscarriage. *Mol Hum Reprod* 7(12): 1167-72.
- Apps R, Gardner L, Moffett A (2008). A critical look at HLA-G. *Trends in Immunology*. 29(7)313-321.
- Amorim CEG (2009). Genética de populações e reinterpretação da história demográfica de remanescentes de quilombo: uma comparação entre três populações do nordeste brasileiro. Dissertação de mestrado. Programa de Pós-Graduação em Biologia Animal. Universidade de Brasília. Brasília – DF.
- Aractingi S, Briand N, Le Danff C, Viguier M, Bachelez H, Michel L, Dubertret L and Carosella E. (2001). HLA-G and NK receptor are expressed in psoriatic skin: a possible pathway for regulating infiltrating T cells? *Am J Pathol* 159(1): 71-7.
- Arantes LC (1999). *Análise da variação genética em treze locos STR na população do Distrito Federal*. Dissertação de Mestrado. UNB.
- Arnaiz-Vilhena A, Gomez-Casado E, Martinez-Laso J (2002). Population genetics relationships between Mediterranean populations determined by HLA allele distribution and a historic perspective. *Tissue Antigens*. 60:111-121.
- Barcelos RSS (2006) *Contribuição genética de duas populações urbanas da região Centro-Oeste brasileiro estimada por marcadores uniparentais*. Tese de doutorado. Programa de Pós-Graduação em Biologia Animal. Instituto de Ciências Biológicas. Universidade de Brasília.
- Barreiro LB, Laval G, Quach H, Patin E, Quintana-Murci L: Natural selection has driven population differentiation in modern humans. *Nat Genet* 2008, 40(3):340-345.
- Barrett JC, Fry B, Maller J, Daly MJ. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 21(2):263-265.
- Bernatchez L, Landry C. (2003) MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evol Biol* 16:363-77
- Carosella ED, Philippe M, LeMaout J, Rouas-Freiss N (2008a). HLA-G: from biology to clinical benefits. *Trends in Immunology*. 29(3):125-132.

- Carosella E, Favier B, Rouas-Freiss N, Moreau P, LeMaoult J. (2008b). Beyond the increasing complexity of the immunomodulatory HLA-G molecule. *Blood*. 111(10):
- Castelli EC, Mendes-Junior CT, Deghaide NHS, Albuquerque RS, Muniz YCN, Simões RT, Carosella ED, Moreau P, Donadi EA. (2009). The genetic structure of the 3' untranslated region of the HLA-G gene: polymorphisms and haplotypes. *Genes and Immunity*. 1: 1 – 8.
- Castelli EC, Mendes-Junior CT, Donadi EA (2007). HLA-G alleles and HLA-G 14bp polymorphisms in a Brazilian population. *Tissue Antigens*. 70:62-68.
- Castelli EC, Mendes-Junior CT, Viana de Camargo JL, Donadi EA (2008). HLA-G polymorphism and transitional cell carcinoma of the bladder in a Brazilian population. *Tissue Antigens* 72(2): 149-57.1
- Castelli EC, Moreau P, Chiromatzo AO, Mendes-Junior CT, Veiga LCS, Yaghi L, Giuliatti S, Carosella ED, Donadi EA (2009b). In silico analysis of microRNAs targeting the HLA-G 3' untranslated region alleles and haplotypes. Artigo aceito para publicação na *Genes Human Immunology*.
- Castro MJ, Morales P, Martinez-Laso J, Allende L, Rojo-Amigo R, Gonzalez-Hevilla M, Varela P, Moscoso J, Garcia-Berciano M, Arnaiz-Villena A (2000). Lack of MHC-G4 and Soluble (G5-G6) isoforms in higher primates, *Pongidae*. *Human Immunology* 61:1161-1168.
- Chen XY, Yan WH, Lin A, Xu HH, Zhang JG, Wang XX (2008). The 14bp deletion polymorphisms in HLA-G gene play an important role in the expression of soluble HLA-G in plasma. *Tissue Antigens* 72:335-341.
- Companhia do Desenvolvimento do Planalto Central (CODEPLAN). Disponível em <http://www.codeplan.df.gov.br>. Sítio da internet acessado em 2009.
- Creput C, Durrbach A, Charpentier B, Carosella ED, Rouas-Freiss N. (2003a) HLA-G: immunoregulatory molecule involved in allograft acceptance. *Nephrologie* 24(8): 451-6.
- Creput C, Durrbach A, Menier C, Guettier C, Samuel D, Dausset J, Charpentier B, Carosella E, Rouas-Freiss N. (2003b) Human leukocyte antigen-G (HLA-G) expression in biliary epithelial cells is associated with allograft acceptance in liver-kidney transplantation. *J Hepatol* 39(4): 587-94.
- Crispim JC, Duarte RA, Soares CP, Costa R, Silva JS, Mendes-Junior CT, Wastowskia I, Faggionie L, Saberf L e Donadi E. (2008) Human leukocyte antigen-G

- expression after kidney transplantation is associated with a reduced incidence of rejection. *Transpl Immunol*; 18(4): 361-7.
- Diniz, MECG (2008). *Fecundidade e Genética em Kalunga: busca de associação entre dados demográficos e marcadores moleculares num remanescente de quilombo brasileiro*. Dissertação de Mestrado. Programa de Pós-Graduação em Ciências da Saúde da Faculdade de Ciências da Saúde. Universidade de Brasília. Brasília – DF.
- Donadi EA, Castelli EC, Arnaiz-Villena A, Roger M, Rey D, Moureau P (2010). Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. *Celular and Molecular Life Science*.
- Excoffier, L. G. Laval, and S. Schneider (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *EvolutionaryBioinformatics Online* 1:47-50.
- Ewens WJ. The sampling theory of selectively neutral alleles. *Theor Popul Biol* 1972: 3: 87–112.
- Fausto, B. (2001) *História Concisa do Brasil*. São Paulo, SP.
- Ferrer-Admetlla A, Bosch E, Sikora M, Marques-Bonet T, Ramirez-Soriano A, Muntasell A, Navarro A, Lazarus R, Calafell F, Bertranpetit J (2008). Balancing selection is the main force shaping the evolution of innate immunity genes. *J Immunol*, 181(2):1315-1322.
- Fu, Y.X. and Li, W.H. (1993) Statistical tests of neutrality of mutations. *Genetics*, 133, 693–709.
- Fumagalli M, Cagliani R, Pozzoli U, Riva S, Comi GP, Menozzi G, Bresolin N, Sironi M (2009). Widespread balancing selection and pathogen-driven selection at blood group antigen genes. *Genome Res*, 19(2):199-212.
- Fundação Cultural Palmares. Disponível em [http:// www.palmares.gov.br/](http://www.palmares.gov.br/). Sítio da internet acessado em fevereiro 2011.
- Fundação Nacional do índio. Disponível em <http://www.funai.gov.br/>. Sítio da internet acessado em fevereiro 2001.
- Futuyma, DJ. *Biologia Evolutiva*. 2ed. Ribeirão Preto – FUNPEC-RP. 2002.
- Gazit E, Slomov Y, Goldberg I, Brenner S, Loewenthal R (2004). HLA-G is associated with pemphigus vulgaris in Jewish patients. *Hum Immunol* 65(1): 39-46.

- Geraghty DE, Koller BH, Orr HT (1987) A human major histocompatibility complex class I gene that encodes a protein with a shortened cytoplasmic segment. *Proc Natl Acad Sci USA* 84:9145–9149.
- Gilbert SF. *Developmental biology*. 6ed. Sinauer Associates: Sunderland. 2000.
- Glas J, Torok HP, Tonenchi L, Wetzke M, Beynon V, Teshome MY, Cotofana S, Schiemann U, Griga T, Klein W, Epplen J, Folwaczny C, Folwaczny M, Mussack T e Weiss E. (2007). The 14-bp deletion polymorphism in the HLA-G gene displays significant differences between ulcerative colitis and Crohn's disease and is associated with ileocecal resection in Crohn's disease. *Int Immunol*; 19(5): 621-6.
- Godinho NM (2008). O impacto das migrações na constituição genética de populações Latino-Americanas. Tese de doutorado. Programa de Pós-Graduação em Biologia Animal. Universidade de Brasília. Brasília – DF.
- Gontijo CC (2008). Composição genética de duas populações Afro-derivadas brasileiras inferidas a partir de marcadores indicativos de ancestralidade. Dissertação de mestrado. Programa de Pós-Graduação em Biologia Animal. Universidade de Brasília. Brasília – DF.
- Harrison GA, Humphrey KE, Jakobsen IB, Cooper DW (1993). A 14bp deletion polymorphism in the HLA-G gene. *Human Molecular Genetics*. 2:2200.
- Hartl DL, Clark AG (2010). Princípios de genética de populações. 4ed. Porto Alegre – Artmed.
- Hviid TV, Hylenius S, Rorbye C, Nielsen LG (2003). HLA-G allelic variants are associated with differences in the HLA-G mRNA isoform profile and HLA-G mRNA levels. *Immunogenetics*. 55:63-79.
- Hviid TVF (2006a). HLA-G in human reproduction: aspects of genetics, function and pregnancy complications. *Human Reproduction Update*. 12(3):209-232.
- Hviid TV, Milman N, Hylenius S, Jakobsen K, Jensen MS, Larsen LG (2006b). HLA-G polymorphisms and HLA-G expression in sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* 23(1): 30-7.
- Hylenius S, Andersen AMN, Melbye M, Hviid TVF (2004). Association of HLA-G genotype and risk of pre-eclampsia: a case control study using family triads. *Molecular Human Reproduction*. Vol10, N°4, pp237-246.
- IBGE (2000) Brasil: 500 anos de povoamento. Rio de Janeiro: IBGE. 231 p.

- IMGT – *International Immunogenetics Information Systems*. Disponível no endereço <http://www.ebi.ac.uk/imgt/hla/>. Sítio da internet acessado em novembro/2009.
- Janeway CA, Travers P, Walport M, Shlomchik M (2002). *Imunobiologia – O sistema imune na saúde e na doença*. 5ed. Porto Alegre – Artmed.
- Kimura M (1968). Evolutionary rate at the molecular level. *Nature* 217: 624–626.
- Kuersten S, Goodwin EB (2003). The power of the 3'UTR: translational control and development. *Nat Rev Genet* 4:626–637.
- Lancaster, A, MP Nelson, D Meyer, G Thomson, RM Single. 2003. PyPop: a software framework for population genomics: analyzing large-scale multi-locus genotype data. *Pac Symp Biocomput*: 514-525.
- Lajoie J, Jeanneau A, Faucher MC, Moreau P, Roger M (2008). Characterization of five novel HLA-G alleles with coding DNA base changes. *Tissue Antigens* 72(5): 502-4.
- Laval G, Patin E, Barreiro LB, Quintana-Murci L (2010). Formulating a historical and demographic model of recent human evolution based on resequencing data from noncoding regions. *Plos One*, Vol 5, issue 4.
- LeMaoult J, Le Discorde M, Rouas-Freiss N, Moreau P, Menier C, McCluskey J *et al.* (2003). Biology and functions of human leukocyte antigen-G in health and sickness. *Tissue Antigens* 62(4): 273-84.
- Lewis, P. O.; Zaykin, D. Genetic data analysis: software for the analysis of discrete genetic data. Version 1.0. 1997.
- Librado, P, J Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-1452.
- Lila N, Carpentier A, Amrein C, Khalil-Daher I, Dausset J, Carosella ED (2000). Implication of HLA-G molecule in heart-graft acceptance. *Lancet* 355(9221): 2138.
- Lin A, Yan WH, Xu HH, Gan MF, Cai JF, Zhu M *et al.* (2007a) HLA-G expression in human ovarian carcinoma counteracts NK cell function. *Ann Oncol* 18(11): 1804-9.
- Lin A, Yan WH, Xu HH, Tang LJ, Chen XF, Zhu M *et al.* (2007b). 14 bp deletion polymorphism in the HLA-G gene is a risk factor for idiopathic dilated cardiomyopathy in a Chinese Han population. *TissueAntigens* 70(5): 427-31.
- Mendes-Junior C.T., Castelli E.C., Simões A.L., Donadi E.A. (2007a). Absence of the HLA-G*0105N allele in Amerindian populations from the Brazilian Amazon region: a possible role of natural selection. *TissueAntigens*. 70: 330-324.

- Mendes-Junior C.T., Castelli E.C., Simões A.L., Donadi E.A. (2007b). HLA-G 14-bp polymorphism at exon 8 in Amerindian populations from Brazilian Amazon. *Tissue Antigens* 69:255-260.
- Mitsdoerffer M, Schreiner B, Kieseier BC, Neuhaus O, Dichgans J, Hartung HP *et al.* (2005). Monocyte-derived HLA-G acts as a strong inhibitor of autologous CD4 T cell activation and is upregulated by interferon-beta in vitro and in vivo: rationale for the therapy of multiple sclerosis. *J Neuroimmunol* 159(1-2): 155-64.
- Moreau P, Contu L, Alba F, Lai S, Simoes R, Orru S *et al.* (2008). HLA-G gene polymorphism in human placentas: possible association of G*0106 allele with preeclampsia and miscarriage. *Biol Reprod* 79(3): 459-67.
- Nicolae D, Cox NJ, Lester LA, Schneider D, Tan Z, Billstrand C *et al.* (2005). Fine mapping and positional candidate studies identify HLA-G as an asthma susceptibility gene on chromosome 6p21. *Am J Hum Genet* 76(2): 349-57.
- Ober C, Aldrich CL (1997). HLA-G polymorphisms: neutral evolution or novel function? *Journal of Reproduction Immunology*. 36:1-21.
- Ober C, Aldrich CL, Chervoneva I, Billstrand C, Rahimov F, Gray HL *et al.* (2003). Variation in the HLA-G promoter region influences miscarriage rates. *Am J Hum Genet* 72(6): 1425-35.
- Oliveira SF, Nogales AM, Novion HI (2004). Estimativas de fecundidad em poblaciones afro-descendientes rurales brasileñas. In: J. E. Egocheaga. (Org.). *Biologia de Poblaciones Humanas: Diversidad, Tiempo, Espacio*. 1 ed. Oviedo (Espanha): Universidad de Oviedo, v. 1, p. 885-895.
- Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD (2003). Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci USA* 100(1): 177-82.
- Pedrosa MAF (2006). Composição genética de quatro populações remanescentes de quilombos do Brasil com base em microssatélites e marcadores de ancestralidade. Dissertação de mestrado. Programa de Pós-Graduação em Biologia Molecular. Universidade de Brasília. Brasília – DF.
- Pfaff, C. L.; Parra, E. J.; Bonilla, C.; Hiester, K.; Mckeigue, P. M.; Kamboh, M. I.; Hutchinson, R. G.; Ferrell, R. E.; Boerwinkle, E.; Shriver, M. D. Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am J Hum Genet*, 68:198-207, 2001.
- Raymond, M. and Rousset, F., Genepop Version 3.3: A Population Genetics Software

- for Exact Tests and Ecumenicism, *J. Hered.*, 1995, vol. 86, pp. 248–249.
- Rebmann V, van der Ven K, Passler M, Pfeiffer K, Krebs D, Grosse-Wilde H (2001). Association of soluble HLA-G plasma levels with HLA-G alleles. *Tissue Antigens* 57:15–21.
- Reis JJ, Gomes FS (1996). Introdução – Uma história de liberdade. In: *Liberdade por um fio*. REIS, J.J.; GOMES, F.S (orgs.) – São Paulo, Cia das Letras: 9-25.
- Ribeiro GGBL (2009). Análise genética de marcadores do tipo STR e Indel em cromossomos sexuais humanos em populações de remanescentes de quilombo. Tese de doutorado. Programa de Pós-Graduação em Biologia Animal. Universidade de Brasília. Brasília - DF.
- Rizzo R, Campioni D, Lanza F, Baricordi OR (2008b). Cotransplantation of mesenchymal cells and a higher relapse rate: a role for HLA-G molecules? *Leukemia*. Letter to the editor.
- Rizzo R, Mapp CE, Melchiorri L, Maestrelli P, Visentin A, Ferretti S *et al.* (2005). Defective production of soluble HLA-G molecules by peripheral blood monocytes in patients with asthma. *J Allergy Clin Immunol* 115(3): 508-13.
- Rizzo R, Melchiorri L, Simone L, Stignani M, Marzola A, Gullini S, Baricordi OR (2008a). Different Production of Soluble HLA-G Antigens by Peripheral Blood Mononuclear Cells in Ulcerative Colitis and Crohn's Disease: A Noninvasive Diagnostic Tool? *Inflammatory Bowel Disease* 14:100-105.
- Rousseau P, Le Discorde M, Mouillot G, Marcou C, Carosella ED, Moreau P (2003). The 14bp deletion-insertion polymorphism in the 3'UT region of the HLA-G gene influences HLA-G mRNA stability. *Human Immunology*. 64:1005-1010.
- Sabeti PC, Schaffner SF, Fry B, Lohmuller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D, Lander ES (2006). Positive natural selection in human lineage. *Science* 312:1614.
- Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L. et al. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 409,928–933.
- Saitou, N. A genetic affinity analysis of human populations. *Hum. Evol.* 10(1): 17-33. 1995.
- Salzano, FM; Callegari-Jacques, SM. South American Indians: a case study in

- evolution. Research Monographs on Human Population Biology. Number 6. Clarendon Press. Oxford. 1988.
- Soares AA (1995). Kalunga: O direito de Existir. (Questões Antropológicas e Jurídicas Sobre Remanescentes de Quilombo). Goiânia. Min / Fundação Cultural Palmares. 253 p.
- Solier C, Mallet V, Lenfant F, Bertrand A, Huchénq A, Le Bouteiller P (2001). HLA-G unique promoter region: functional implications. *Immunogenetics*. 53(8):617-625.
- Suarez MB, Morales P, Castro MJ (1997). A new HLA-G allele (HLA-G*0105N) and its distribution in the Spanish population. *Immunogenetics*. 45:464-465.
- Suarez MB, Morales P, Castro MJ (1997). A new HLA-G allele (HLA-G*0105N) and its distribution in the Spanish population. *Immunogenetics*. 45:464-465.
- Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001; 68: 978–989.
- Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123, 585–595.
- Tan CY, Chong YS, Loganath A, Chan YH, Ravichandran J, Lee CG, Chong SS (2009). Possible gene-gene interaction of KIR2DL4 with its cognate ligand HLA-G in modulating risk for preeclampsia. *Reprod Sci* 16:1135–1143.
- Tan Z, Randall G, Fan J, Camoretti-Mercado B, Brockman-Schnieder R, Pan L, Solway J, Gern JE, Lemanske RF, Nicolae D, Ober C (2007). Allele-Specific targeting of microRNAs to HLA-G and risk of asthma. *Am J Hum Genet* 81:829-834.
- Tan Z, Shon MA, Ober C (2005) Evidence of balancing selection at the HLA-G promoter region. *Human Molecular Genetics*. 14(23):3619-3628.
- Torres MI, Le Discorde M, Lorite P, Rios A, Gassull MA, Gil A *et al.* (2004). Expression of HLA-G in inflammatory bowel disease provides a potential way to distinguish between ulcerative colitis and Crohn's disease. *IntImmunol* 16(4): 579-83.
- Veit TD, Chies JA (2009) Tolerance versus immune response microRNAs as important elements in the regulation of the HLA-G gene expression. *Transpl Immunol* 20:229–231.
- Veit TD, Vianna P, Scheibel I, Brenol CV, Brenol JC, Xavier RM, Delgado-Canedo A, Gutierrez JE, Brandalize AP, Schuler-Faccini L, Chies JA (2008) Association of

- the HLA-G 14-bp insertion/deletion polymorphism with juvenile idiopathic arthritis and rheumatoid arthritis. *Tissue Antigens* 71:440–446.
- Vianna P, Dalmaz CA, Veit TD, Tedoldi C, Roisenberg I, Chies JA (2007). Immunogenetics of pregnancy: role of a 14-bp deletion in the maternal HLA-G gene in primiparous pre-eclamptic Brazilian women. *Hum Immunol* 68(8): 668-74.
- Wang S, Lewis Jr CM, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C, Mazzotti G, Poletti G, Hill K, Hurtado AM, Labuda D, Klitz W, Barrantes R, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Llop E, Rothhammer F, Excoffier L, Feldman MW, Rosenberg NA, Ruiz-Linares A (2007). Genetic variation and population structure in native Americans. *Plos Genetics* vol3:issue11.
- Watterson GA. The homozygosity test of neutrality. *Genetics* 1978; 88: 405–17.
- Weir, B. S.; Cockerham, C. C. Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6):1358-1370, 1984.
- Wiendl H, Mitsdoerffer M, Hofmeister V, Wischhusen J, Bornemann A, Meyermann R et al. (2002). A functional role of HLA-G expression in human gliomas: an alternative strategy of immune escape. *J Immunol* 168(9): 4772-80.
- Wu DD, Zhang YP (2011). Different level of population differentiation among human genes. *BMC Evolutionary Biology* 11:16.
- Yan WH, Fan LA, Yang JQ, Xu LD, Ge Y, Yao FJ (2006). HLA-G polymorphism in a Chinese Han population with recurrent spontaneous abortion. *Int J Immunogenet* 33(1): 55-8.
- Yan WH, Lin A, Li M, Xu HH, Zhang ZP, Wang XX (2008). Analysis of the 14 bp insertion and deletion polymorphism in human leukocyte antigen-G in two Chinese ethnic populations. *Tissue Antigens*. 71:2.
- Yie SM, Li LH, Xiao R, Librach CL (2008). A single base-pair mutation in the 30 untranslated region of HLA-G mRNA is associated with pre-eclampsia. *Mol Hum Reprod* 14:649–653.