UNIVERSIDADE DE SÃO PAULO FACULDADE DE MEDICINA DE RIBEIRÃO PRETO DEPARTAMENTO DE GENÉTICA

DINÂMICA DE POLIMORFISMOS GENÉTICOS LIGADOS AO GENE DA HEMOFILIA A (*F8*) NA POPULAÇÃO BRASILEIRA

JULIANA DOBLAS MASSARO

Orientador: Prof. Dr. Aguinaldo Luiz Simões

RIBEIRÃO PRETO 2009

Universidade de São Paulo Faculdade de Medicina de Ribeirão Preto Departamento de Genética

DINÂMICA DE POLIMORFISMOS GENÉTICOS LIGADOS AO GENE DA HEMOFILIA A (*F8*) NA POPULAÇÃO BRASILEIRA

JULIANA DOBLAS MASSARO

Tese apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, como requisito parcial para a obtenção do título de Doutora em Ciências – Área de concentração: Genética.

Orientador: Prof. Dr. Aguinaldo Luiz Simões

RIBEIRÃO PRETO 2009 AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

FICHA CATALOGRÁFICA

Massaro, Juliana Doblas.

Dinâmica de polimorfismos genéticos ligados ao gene da Hemofilia A (*F8*) na população Brasileira / Juliana Doblas Massaro; orientador Aguinaldo Luiz Simões. – Ribeirão Preto, 2009.

156 f. il. 30cm.

Tese (Doutorado – Programa de Pós-Graduação em Genética.) – Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo.

1. População brasileira, 2. Fator VIII, 3. Hemofilia A,

4. diagnóstico indireto, 5. microssatélites.

FOLHA DE APROVAÇÃO

Juliana Doblas Massaro

Dinâmica de polimorfismos genéticos ligados ao gene da Hemofilia A (*F8*) na população Brasileira.

Tese apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, como requisito parcial para a obtenção do título de Doutora em Ciências – Área de concentração: Genética.

Aprovado em:

Prof. Dr	
Instituição:	Assinatura:
Prof. Dr	
Instituição:	Assinatura:
Prof. Dr	
Instituição:	Assinatura:
Prof. Dr	
Instituição:	_Assinatura:
Prof. Dr	
Instituição:	Assinatura:

Banca Examinadora



"Cascata da Coagulação" **Associação Portuguesa dos Hemofílicos** www.aphemofilicos.pt/hemofilia_o_que_e.asp.

"Aprendi com a primavera; a deixar-me cortar e voltar sempre inteira".

(Cecília Meireles)

Aos meus pais, **Roberto e Cida**, os responsáveis pelo que há de melhor em mim. Ao **Sérgio**, meu marido, pelo amor e ajuda incondicional. A **Laura**, minha filha, que sempre me proporcionou sensações de emoção inacreditáveis.

Dedico

AGRADECIMENTOS

A **Deus**, que me acompanhou em todos os momentos nesta caminhada repleta de surpresas.

Ao Prof. Dr. **Aguinaldo Luiz Simões**, pela oportunidade de realizar mais este trabalho, pela orientação, pelas ótimas sugestões, pelas explicações, discussões, confiança e amizade.

Aos **professores membros da banca examinadora**, por se mostrarem disponíveis em avaliar este trabalho.

A todos os **professores do Departamento de Genética** da FMRP, pelos ensinamentos durante todos esses anos.

Aos **funcionários do Departamento de Genética** pela acolhida e ajuda. Em especial à **Maria Aparecida**, a **Susie** e a **Sílvia**, secretárias do departamento, pela paciência, pelo apoio nos momentos difíceis e pelo carinho.

As amigas **Ana Lúcia** e **Maria do Carmo**, pela grande amizade, conforto e companheirismo, e em especial pelo impagável e incansável auxílio técnico.

Aos amigos **Celso, Cláudia e Marcelo**, pelo auxílio nas análises estatísticas, pelas discussões, revisão e apoio.

A amiga **Yara**, pelo auxílio nas análises estatísticas, formatação da tese e amizade.

A todos os amigos do Departamento de Genética: Edna (Ananda), Cláudia Caixeta, Edilene, Natália, Leonardo, Marcela, Juliana, Eddy, Rosana, Profa. Dra. Eucléia, Prof. Dr. Mestriner, Prof. Dr. Ana Lílian, Bete, Camila, Amanda, Aline, Geusa, Ronai, Eliana, Daniela, Ana Paula e Ludmila pelo excelente ambiente de trabalho, pela prazerosa companhia e grande amizade.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa concedida para o desenvolvimento desta tese.

Em especial.....

Aos meus pais, **Roberto e Cida**, pela ajuda, apoio, exemplo e amor incondicionais que sempre dedicaram a mim e principalmente pela confiança que sempre depositaram.

Ao meu companheiro **Sérgio** e a minha **filha Laura**, que foram um grande incentivo nesta caminhada e sempre serão.

Ao meu irmão **Hamilton**, a minha cunhada **Paula** e aos meus sobrinhos (**Gustavo e Guilherme**), pela alegria e apoio mesmo à distância.

A toda minha família. AMO TODOS VOCÊS!!!!

RESUMO

Massaro, J. D. **Dinâmica de polimorfismos genéticos ligados ao gene da Hemofilia A (F8) na população Brasileira**. 2009. 156p. Tese (Doutorado) – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

A Hemofilia A é uma doença sanguínea condicionada por gene localizado no cromossomo X. É causada pela deficiência parcial ou total da atividade do Fator VIII (FVIII), uma glicoproteína plasmática cuja função é necessária para a coagulação normal do sangue. Devido às dificuldades encontradas para o reconhecimento direto da mutação no gene F8, o diagnóstico das portadoras é feito de forma indireta, isto é, por análise de ligação com marcadores polimórficos localizados dentro ou próximos ao gene que permite determinar a co-segregação do haplótipo e da mutação na família sob estudo e, desta maneira, detectar o estado de portadora e, eventualmente, auxiliar no diagnóstico pré-natal. O presente trabalho teve por objetivo avaliar o poder de alguns desses marcadores na diferenciação das populações e determinar o grau de sua informatividade para o diagnóstico e aconselhamento genético de famílias afetadas, bem como verificar o eventual uso forense de tais marcadores. Foram então determinadas as frequências alélicas e haplotípicas, diversidade genética, diferenciação populacional, desequilíbrio de ligação e composição ancestral de quatro microssatélites intragênicos (IVS 1, IVS 13, IVS 22, IVS 25.3), localizados em introns do F8, e um extragênico (IKBKG) em amostras de populações brasileiras urbanas (indivíduos normais de São Paulo, Rio Grande do Sul e Pernambuco), de quilombos (Mimbó, Sítio Velho e Gaucinha localizados no Estado do Piauí) e Ameríndios (Tikúna, Baníwa e Kashináwa). As análises, quando cabível, incluíram um grupo de pacientes hemofilicos. O DNA dos sítios polimórficos foi amplificado por PCR, os produtos separados em PAGE e corado por nitrato de prata. Para as análises estatísticas foram empregados programas já considerados de uso rotineiro. Os parâmetros de diversidade mostraram diferenças entre as amostras populacionais analisadas. Tais diferenças regionais nas frequências alélicas devem ser levadas em conta quando o diagnóstico indireto da Hemofilia A estiver sendo realizado. Com exceção do IKBKG, todos os demais microssatélites apresentaram altas taxas de heterozigose. Usando tais marcadores, o diagnóstico foi possível em 10 das 11 famílias analisadas. Os microssatélites IVS 22, IVS 1, IVS 13, IVS 25.3 e IKBKG foram informativos em 63,6% (7/11), 54.5% (6/11), 54.5% (6/11), 45.5% (5/11) e 18.2% (2/11) dos casos, respectivamente, demonstrando a eficácia do uso desses microssatélites no diagnóstico pré-natal e na identificação de portadoras na população brasileira.

ABSTRACT

Massaro, J. D. Dynamics of genetic polymorphisms linked to the gene for Hemophilia A (F8) in the Brazilian population. 2009. 156p. Thesis (Doctoral) – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

Hemophilia A is a bleeding disorder conditioned by a gene located on the X chromosome and caused by partial or total deficiency of the Factor VIII (FVIII) activity, a plasma glycoprotein whose function is necessary for normal blood clotting. Due to difficulties faced on direct recognition of the F8 gene mutation, carrier diagnosis is done indirectly by linkage analysis with polymorphic markers located within or near the gene. These markers may determine the haplotype and the mutation co-segregation within the studied family, and thus detect the carrier status and possibly assist in prenatal diagnosis. This study aimed to evaluate the power of some of these markers in population differentiation and to determine their degree of informativeness for diagnosis and genetic counseling of affected families, as well as to verify the possible forensic use of such markers. We then determined the allele and haplotype frequencies, genetic diversity, population differentiation, linkage disequilibrium and ancestral composition in Brazilian urban (healthy individuals from São Paulo, Rio Grande do Sul and Pernambuco), quilombo remnant (Mimbó, Sítio Velho and Gaucinha in the State of Piauí) and Amerindian (Tikúna, Baníwa and Kashináwa) population samples by the analysis of four intragenic microsatellites (IVS 1, IVS 13, IVS 22, IVS 25.3), located on F8 introns, and one extragenic (IKBKG). When appropriated, the analysis included a group of hemophilic patients. DNA polymorphisms were detected by PCR, PAGE and silver nitrate staining. Statistical analysis was implemented by programs already considered in routine use. Diversity parameters showed differences among the populational samples analyzed. Such regional differences in allele frequencies must be taken into account when conduct the indirect diagnosis of Hemophilia A. Except for IKBKG, all other microsatellites showed high rates of heterozygosity. Using these markers, the diagnosis was possible in 10 of the 11 families analyzed. The microsatellites IVS 22, IVS 1, IVS 13, IVS 25.3 and IKBKG were informative in 63.6% (7 / 11), 54.5% (6 / 11), 54.5% (6 / 11), 45.5% (5 / 11) and 18.2% (2 / 11) of the cases, respectively, demonstrating the effectiveness of the use of these microsatellites in prenatal diagnosis and on carrier identification in the Brazilian population.

SUMÁRIO

INTRODUÇÃO	
Hemofilia A. Quadro Clínico	16
Diagnóstico	17
Epidemiologia	17
Herança	
Estrutura do gene F8	
Mutações no gene F8	20
O diagnóstico molecular direto	25
Diagnóstico indireto	27
Polimorfismos ligados ao gene F8 na população brasileira	29
HIPÓTESE	31
OBJETIVOS	32
MATERIAL E MÉTODOS	33
Esquema do Trabalho	33
As amostras	33
Amostras Urbanas	33
Amostras de Afro-Derivados	35
Amostras de Tribos Indígenas da Amazônia Brasileira	
Amostras de pacientes com Hemofilia A	
Aspectos Éticos	
Coleta e Conservação do Material	40
Amostras de Populações Urbanas (SP, RS e PE)	40
Amostras de Afro-derivados (MI, SV, GA)	41
Amostras das tribos indígenas (TI, KA, BA)	41
Amostras de Hemofilicos A e familiares	42
Marcadores genéticos analisados	42
Análise Laboratorial	43
Extração do DNA Genômico	43
Reagentes e Soluções	43
Procedimento	43
Reação em Cadeia da Polimerase (PCR)	45
Reagentes e Soluções	45

Procedimento	45
Programas utilizados	47
Análise do Produto Amplificado	47
Reagentes e Soluções	47
Procedimento	48
Coloração com Nitrato de Prata e Secagem do Gel:	50
Reagentes e soluções	50
Procedimento	50
Determinação Fenotípica (Nomenclatura)	51
Tamanho Amostral	52
Análise Estatística	53
Frequências Gênicas em Populações Ancestrais	53
Estimativas das frequências alélicas e genotípicas	53
Aderência ao Equilíbrio de Hardy-Weinberg	54
Desequilíbrio de ligação entre <i>loci</i>	54
Diferenciação genética das populações	55
Diversidade haplotípica	55
Análise de Variância Molecular (AMOVA)	56
Diversidade genética	56
Análises de Componente Principal	59
Estimativas da composição ancestral	59
Dendrogramas	60
Parâmetros Forenses	61
Frequencias alelicas, equilibrio de Hardy-Weinberg e heterozigose	
Desequilibrio de ligação entre <i>loci</i>	
Diferenciação genetica das populações	
Diversidade haplotipica	
Analise de Variancia Molecular (AMOVA)	
Diversidade genetica	
Análises de Componente Principal	85
Estimativas de composição ancestral	86
Dendrograma	
Parāmetros Forenses	89
Aplicabilidade no aconselhamento genético	93

DISCUSSÃO	
Novos alelos	95
Freqüências alélicas e equilíbrio de Hardy-Weinberg	96
Comparação com populações de outros continentes	99
Comparação das freqüências alélicas entre as amostras	
Diversidade Genética	
Desequilíbrio de ligação	
Diversidade Haplotípica	103
Diferenciação Genética	105
Dendrograma	106
Análise do componente principal	107
Composição ancestral	
Parâmetros Forenses	109
Aplicabilidade no aconselhamento genético	110
CONCLUSÕES	112
CONCLUSÕES	112
CONCLUSÕES Alelos Novos Desequilíbrio	
CONCLUSÕES Alelos Novos Desequilíbrio Distribuição das freqüências por continente	 112 112 112 112 112
CONCLUSÕES Alelos Novos Desequilíbrio Distribuição das freqüências por continente Diversidade Genética	112 112112112112112113
CONCLUSÕES Alelos Novos Desequilíbrio Distribuição das freqüências por continente Diversidade Genética Desequilíbrio de ligação	112 112 112 112 112 112 113 113
CONCLUSÕES Alelos Novos Desequilíbrio Distribuição das freqüências por continente Diversidade Genética Desequilíbrio de ligação Diversidade haplotípica	112 112 112 112 112 112 113 113 113 113
CONCLUSÕES Alelos Novos Desequilíbrio Distribuição das freqüências por continente Diversidade Genética Desequilíbrio de ligação Diversidade haplotípica Parâmetros forenses	112 112 112 112 112 112 113 113 113 113 113
CONCLUSÕES Alelos Novos Desequilíbrio Distribuição das freqüências por continente Diversidade Genética Desequilíbrio de ligação Diversidade haplotípica Diversidade haplotípica Utilidade para o aconselhamento	112 112 112 112 112 112 113 113 113 113 113 113 113 113 113 113 113 113 113 113 113 113
CONCLUSÕES Alelos Novos Desequilíbrio Distribuição das freqüências por continente Diversidade Genética Desequilíbrio de ligação Diversidade haplotípica Parâmetros forenses Utilidade para o aconselhamento REFERÊNCIAS BIBLIOGRÁFICAS	112 112 112 112 112 112 113 1145
CONCLUSÕES Alelos Novos Desequilíbrio Distribuição das freqüências por continente Diversidade Genética Desequilíbrio de ligação Diversidade haplotípica Parâmetros forenses Utilidade para o aconselhamento REFERÊNCIAS BIBLIOGRÁFICAS MANUSCRITO	112 112 112 112 112 112 113 1145 115

LISTA DE TABELAS

Tabela 1. Lista dos loci analisados, seus respectivos primers, número de alelos e tamanho em pb43
Tabela 2. Condições de PCR para os <i>loci</i> analisados no presente trabalho (quantidade em μl para 1 reação)
Tabela 3. Condições específicas para eletroforese de cada locus analisado no presente trabalho49
Tabela 4. Distribuição das frequências alélicas (%) e parâmetros de diversidade genética dos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, de acordo com o sexo ((M: mulheres, H: homens, T: amostra total) em três amostras de populações urbanas brasileiras (SP: São Paulo, RS: Rio Grande do Sul, PE: Pernambuco) e uma amostra de Hemofilicos A(HA). Os valores significativos antes da correção de Bonferroni estão destacados em negrito e os valores significativos após a correção (α _{Bonf.} < 0,0025) estão indicados com *
Tabela 5. Distribuição das frequências alélicas (%) e parâmetros de diversidade genética dos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, de acordo com o sexo ((M: mulheres, H: homens, T: amostra total) em três amostras de quilombos (MI: Mimbó, SV: Sítio Velho, GA: Gaucinha). Os valores significativos antes da correção de Bonferroni estão destacados em negrito e os valores significativos após a correção (α _{Bonf.} < 0,003) estão indicados com *
Tabela 6. Distribuição das frequências alélicas (%) dos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, de acordo com o sexo (M: mulheres, H: homens, T: amostra total) em três amostras de tribos indígenas da Amazônia brasileira, Tikúna (UM: Umariaçu, FE: Feijoal, VE: Vendaval), Baníwa (BA) e Kashinawa (KA). Os valores significativos antes da correção de Bonferroni estão destacados em negrito e os valores significativos após a correção ($\alpha_{Bonf.} < 0,002$) estão indicados com *
Tabela 7. Diferenciação gênica e genotípica entre os pares possíveis das doze amostras populacionais analisadas, baseada nos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, de acordo com o sexo (H: homens, M: mulheres)72
Tabela 8 . Frequência absoluta dos haplótipos do cromossomo X observados em homens nas amostras de São Paulo (SP), Pernambuco (PE), Rio Grande do Sul (RS), Hemofilicos A (HA), Mimbó (MI), Sítio Velho (SV), Gaucinha (GA), Umariaçu (UM), Feijoal (FE), Vendaval (VE), Baníwa (BA) e Kashináwa (KA)
Tabela 9. Estimadores de variabilidade segundo a metodologia de Nei (1973,1975) para os STRs em homens, mulheres e total (homens e mulheres), nas doze amostras populacionais São Paulo (SP), Pernambuco (PE), Rio Grande do Sul (RS), Hemofilicos A (HA), Mimbó (MI), Sítio Velho (SV), Gaucinha (GA), Umariaçu (UM), Feijoal (FE), Vendaval (VE), Baníwa (BA) e Kashináwa (KA)
Tabela 10. Matriz de distância da sub população de homens baseada nos valores de F_{ST} (abaixo da diagonal) e valores de P (significante a 0,05) com seu respectivo erro padrão (acima da diagonal), entre os pares possíveis das amostras de São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE), pacientes Hemofilicos A (HA), Mimbó (MI), Sítio Velho (SV), Gaucinha (GA), Umariaçu (UM), Feijoal (FE), Vendaval (VE), Baníwa (BA) e Kashináwa (KA), baseado nos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG
Tabela 11. Matriz de distância da sub população de mulheres baseada nos valores de <i>F</i> _{ST} (abaixo da diagonal) e valores de P (significante a 0,05) com seu respectivo erro padrão (acima da diagonal), entre os pares possíveis das amostras de São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE), pacientes Hemofilicos A (HA), Mimbó (MI), Sítio Velho (SV), Gaucinha (GA), Umariaçu (UM), Feijoal (FE), Vendaval (VE), Baníwa (BA) e Kashináwa (KA), baseado nos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG
Tabela 12. Características dos componentes principais baseados nas frequências alélicas dequatro STRs (IVS 1, IVS 13, IVS 22, IVS 25.3)
Tabela 13. Contribuição dos componentes africano, europeu e ameríndio obtidos a partir de doisSTRs (IVS 13 e IVS 22), para as amostras São Paulo (SP), Pernambuco (PE), Rio Grande do Sul(RS), Hemofilicos A (HA)
Tabela 14. Valores de Parâmetros de utilidade forense para o IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, nas doze amostras de populações brasileiras (SP, PE, RS, HA, MI, SV, GA, UM, FE, VE, BA, KA) em homens, mulheres e total (homens e mulheres)90

LISTA DE FIGURAS

Figura 7. Localização geográfica das amostras urbanas das cidades de Recife (Estado de Pernambuco), Ribeirão Preto (Estado de São Paulo) e Porto Alegre (Estado do Rio Grande do Sul).34

Figura 8. Localização geográfica de comunidades afro-derivadas do Piauí. Neste trabalho foram analisadas as comunidades rurais: Mimbó (MI), Sítio Velho (SV) e Gaucinha (GA)......35

Figura 9. Localização	geográfica das	aldeias p	pertencentes	às tribos	Tikúna,	Baníwa e	Kashináwa e	
o respectivo o número	de indivíduos a	analisado	os					37

Figura 10. Posição física dos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG no cromossomo X. Os números correspondem à distância em pares de bases do Xq-Tel usando a sequência referência do NC_000023......42

INTRODUÇÃO

Hemofilia A. Quadro Clínico

A Hemofilia A (HA) ou Hemofilia Clássica (OMIM: 306700) é uma doença sanguínea causada pela deficiência parcial ou total da atividade do Fator VIII (FVIII), uma glicoproteína plasmática cuja função é necessária para a coagulação normal do sangue (Eyster et al., 1978). Quase a metade dos pacientes com HA possui o quadro grave e apresentam hemorragias espontâneas frequentes dentro das articulações, músculos e órgãos internos, além de hemartroses graves que podem evoluir para artropatias crônicas e incapacitantes (Rodgers e Greenberg, 1999). A HA moderada ocorre em quase 10% dos pacientes. Nestes casos, os hematomas e hemartroses estão associados a pequenos traumas, mas podem ocorrer eventos isolados de hemorragias espontâneas. Estas hemorragias não são tão intensas quando comparadas à HA grave, porém se não tratadas adequadamente o quadro clínico poderá ser mais acentuado do que poderia indicar o nível plasmático do Fator VIII. A HA leve, que ocorre em 30 a 40% dos pacientes, nunca provoca hemorragias espontâneas (é preciso um traumatismo intenso ou cirurgia), porém podem apresentar hemartroses, especialmente em articulações onde previamente ocorreu hemorragia pós-traumática. Muitas vezes, é reconhecida no adulto somente após cirurgia ou extração dentária. (Marques & Fristma, 2007).

A expressão clínica da HA varia de família para família, porém numa mesma família, a gravidade das manifestações e das alterações laboratoriais é relativamente constante. O aparecimento de manifestações clínicas mais graves, dentro da mesma família, pode ser decorrente do desenvolvimento de inibidor (anticorpo que neutraliza a função pró-coagulante do FVIII infundido) ou de lesão anatômica predisponente a sangramentos (Rodgers e Greenberg, 1999, Darby *et al.*, 2004).

Diagnóstico

O diagnóstico das doenças da coagulação, tais como a HA, requer avaliação da história pregressa pessoal e familiar e dos quadros clínicos e laboratoriais. Os testes de laboratório iniciais realizados em indivíduos com suspeita de deficiência congênita de um dos fatores de coagulação são: a análise do Tempo de Trombina (TT), Tempo de Tromboplastina Parcial Ativada (TTPA), Tempo de Protrombina (TP) e Quantificação do Fibrinogênio e subseqentemente, baseado nestes resultados, segue-se à análise dos fatores da coagulação. Indivíduos com TT e TP normais e TTPA prolongado são associados à deficiência de um desses três Fatores: VIII, IX e XI. Portanto, após a obtenção deste resultado é necessário a quantificação desses fatores para se chegar ao diagnóstico (Marques & Fristma, 2007). A HA é uma doença onde a frequência e a gravidade do quadro hemorrágico são inversamente proporcionais ao nível de atividade coagulante do FVIII (FVIIIc). Portanto sua classificação é feita da seguinte forma: 1) grave, com FVIIIc menor que 1% ; 2) moderada, com FVIIIc 1-5% ; ou 3) leve, com FVIIIc 5%-40% (Eyster *et al.*, 1978).

Epidemiologia

A HA afeta um em cada 10.000 homens no Brasil (Caio *et al.*, 2001) e segundo dados do Ministério da Saúde a sua prevalência por diagnóstico em julho de 2007 era de 6881 casos (2º Publicação do Perfil das Coagulopatias, Ministério da Saúde 2007).

A HA pode ocorrer em mulher, em decorrência da união de homem com hemofilia e mulher portadora. Essa condição é muito rara, por isso, são observadas apenas mulheres portadoras. Estas podem apresentar baixos níveis de atividade coagulante do FVIII, evento relacionado à inativação desproporcional do cromossomo X (Marques & Fristma, 2007).

Herança

A herança tem padrão recessivo ligado ao cromossomo X. Filhas de homem com hemofilia serão portadoras obrigatórias (Pandey & Mittal, 2001). Em 1/3 dos casos a HA surge por mutação *de novo*, não havendo relato de ocorrência da doença em outros membros da família. Nos 2/3 restantes dos casos, o gene é transmitido por mães portadoras a seus filhos do sexo masculino (Larner, 1987).

Estrutura do gene F8

O gene *F8* foi identificado em 1984 (Gitschier *et al.*) e a estrutura molecular da proteína determinada, possibilitando a compreensão da fisiopatologia da HA e, consequentemente, do seu diagnóstico, tratamento e aconselhamento (White II e Shoemaker, 1989).

Este gene, localizado na extremidade distal do braço longo do cromossomo X (Xq28), compreende 186 Kb e contém 26 *exon*s (Figura 1A e B). A extensão dos *exon*s varia de 69 a 262 nucleotídeos, exceto os *exons* 14 e 26 (Figura 1C) que apresentam 3106 e 1958 nucleotídeos, respectivamente. Existem seis grandes *introns*, dentre eles o Intron 22 (IVS 22) com 32 Kb (Figura 1D) (Gitschier *et al*, 1984; Wood *et al*, 1984; Toole *et al*, 1984).



Figura 1. A) Representação esquemática da localização cromossômica do gene *F8*. B) Escala de 100 Kb. O gene está local**i**zado cerca de 1000 Kb da região Xqter (região terminal do braço longo do cromossomo X); entre a extremidade 5' do gene e o telômero existem regiões (int1h-2 e int22h-2/-3) que são duplicações das regiões intragênicas int1h-1 e int22h-1. O gene G6PD encontra-se centromérico em relação ao gene *F8*. C) Escala de 10 Kb. Posicionamento relativo dos 26 *exons; exons* 14 e 26 são indicados por caixa preta por serem significantemente maiores que os demais; IVS 1 e IVS 22 indicam os *Intron* 1 e *Intron* 22, respectivamente. *Int1h-1*, região de 1041 pb no IVS 1 que possui sequência homóloga extragênica. D) IVS 22. *F8a* e ilha de CpG (9,5 Kb) contidos na sequência In22h-1; *F8b* em posição centromérica à int22h-1. As setas indicam a orientação das sequências (Adaptado de Antonarakis, 1995).

Dentro do IVS 22, existe uma ilha CpG (sequência em que C e G se repetem, considerada um ponto preferencial para a ocorrência de mutação), que atua como um promotor bidirecional de dois genes, *F8a* e *F8b* (Figura 1D) (Bowen, 2002). O *F8a* possui 1,8 Kb, está totalmente contido no IVS 22 e não possui *Introns*. Este gene é expresso abundantemente em uma ampla variedade de células e tem orientação de transcrição oposta ao gene *F8* (Figura 1D). O *F8b* possui 2,5 Kb e é transcrito na mesma direção do gene *F8*. Este gene compreende um *exon* 5' no IVS 22 e a região dos *exons* 23-26 do gene *F8* (Figura 1D) (Bowen, 2002, Bagnall *et al.,* 2006). A sequência que inclui a ilha CpG e o gene *F8a* compreende 9,5 Kb e é

outras regiões extragênicas, localizadas entre 400 e 500 Kb, anteriores à extremidade 5' do gene F8 (Figura 1B). A cópia proximal em relação ao gene F8 é denominada de *int22h-2*, e a cópia distal, denominada *int22h-3* (Figura 1B) (Bowen, 2001, De Brasi & Bowen, 2008). No *Intron* 1 (IVS 1), existe uma região de 1041 pb, denominada *int1h-1*, que possui uma sequência duplicada a ~140 kb anteriores a extremidade 5' do gene F8 (Bagnall *et al.*, 2002). Estas regiões de alta similaridade funcionam como sítios de recombinação homóloga e determinam rearranjos intragênicos (Antonarakis *et al.*, 1995).

Mutações no gene F8

Um banco de dados sobre a Hemofilia A é periodicamente publicado, e estas informações atualizadas podem ser obtidas no site http://europium.csc.mrc.ac.uk/ (HAMSTeRS, 2008). Este banco disponibiliza dados fenotípicos, modelos moleculares, métodos para o rastreamento de mutações, lista de polimorfismos, revisão sobre a patologia molecular da HA e a listagem completa das mutações publicadas. Esta lista possui atualmente por volta de 1000 mutações. A maioria delas são SNPs (do inglês "single nucleotide polymorphisms"), seguidas de inversões, deleções ou inserções na região codificadora do gene (Pandey & Mittal, 2001). Admite-se que este elevado número de mutações no gene F8 e o fato de 1/3 dos casos de HA decorrerem de mutação nova (não presente no cromossomo X da mãe do paciente) estariam ligados ao seu grande tamanho, à presença de regiões onde as mutações são mais comuns (ilhas CpG representam hot spots) e à sua organização peculiar, ou seja, a existência de cópias homólogas extragênicas de regiões do IVS 1 e IVS 22 (Figura 1), o que facilita a ocorrência de inversões (Rodgers & Greenberg, 1999).

Quase 50% dos pacientes severamente afetados possuem uma inversão parcial do gene F8 devido a uma recombinação homóloga intracromossômica, principalmente

durante a gametogênese masculina (Mantilla-Capacho *et al.*, 2007). Essa recombinação pode resultar em duplicação, deleção e em dois tipos de inversão: a Tipo I, que inclui a cópia extragênica distal (*int22h-3*) e a região *int22h-1* (no gene *F8*) e a Tipo II, que inclui a cópia extragênica proximal (*int22h-2*) e a região *int22h-1* (Figura 2) (Bagnall *et al.*, 2006).



B Inversão Tipo II



Figura 2. Eventos de recombinação homóloga entre as regiões *int22h*. O tipo de rearranjo é indicado no painel esquerdo. (A) e (B) Recombinações intracromossômicas entre a região *int22h-1* e a região *int22h-3* ou *int22h-2*, respectivamente, resultando em inversões. O painel direito mostra a estrutura dos alelos normais e com os rearranjos, bem como a localização e orientação das sequências de *primers* (indicada pelas setas) usadas para amplificação dos rearranjos (Adaptado de Bagnall *et al.*, 2006).

Inicialmente, acreditava-se que as regiões *int22h-2* e *int22h-3* estariam similarmente orientadas no cromossomo, mais especificamente, em direção oposta

ao int22h-1, devido ao fato de ambas participarem da inversão (Antonarakis et al., 1995). O sequenciamento do cromossomo X humano mostrou que a região int22h-2 tem a mesma orientação que a região int22h-1, mas orientação oposta a região int22h-3. Int22h-2 e int22h-3 formam a parte interna dos braços de um palíndromo imperfeito com um loop central único de 67,3 kb e braços de 50,5 kb, onde int22h-2, fica no braço centromérico (Figura 3). No entanto, se todos os cromossomos X tivessem a estrutura acima relatada, a recombinação ocorrida entre int22h-1 e int22h-2 deveria resultar exclusivamente em deleção ou duplicação da região de 400 kb que separa estas sequências (Figura 4). Para conciliar os dados das sequências acima descritas com as evidências da inversão causada pela recombinação homóloga int22h-1 int22h-2, foi de com proposta uma recombinação intracromossomal entre os braços do palíndromo, para que int22h-2 invertesse de posição com int22h-3 tornando-se inversamente orientado em relação a int22h-1, possibilitando a inversão entre essas duas regiões (Figura 3). Isso explicaria o fato da inversão Tipo II ser cinco vezes mais rara que a inversão Tipo I (Bagnall et al., 2006). Os rearranjos resultantes da recombinação entre as regiões int22h estão ilustrados nas Figuras 2 e 4.



Figura 3. Mecanismo de inversão proposto como causa de polimorfismo no palíndromo formado pelas regiões int22h-2 e int22h-3, onde int22h-2 é o braço centromérico. (A) Configuração mais comum do palíndromo, com a sequência do cromossomo X acima (NC_000023.8) mostrando o número de nucleotídeos nas posições relevantes. (B) Recombinação homóloga intracromossômica produzindo a configuração de palíndromo invertido. (C) Sequência resultante da inversão no palíndromo. Gene *F8*: caixa branca; braços do palíndromo: caixas cinza; sequências *Int22h*: caixas-pretas com setas para indicar as orientações e com os tamanhos dos fragmentos de restrição produzidos pela *BclI* indicado abaixo; setas brancas: orientação da sequência única que separa os braços do palíndromo; barras pretas verticais: *exons* a-d, *exon* e, e *exon* 2 do gene *CLIC2*; círculo preenchido: telômero (Bagnall *et al.*, 2006).

A Duplicação e deleção



a) Duplicação



B Deleção



Figura 4. Eventos de recombinação homóloga entre as regiões *int22h*. O tipo de rearranjo é indicado no painel esquerdo. (A) Recombinação homóloga inter-cromossômica entre as regiões *int22h* desalinhadas, que causam deleção e duplicação recíprocas. (B) Recombinação homóloga intra/cromátide causando deleção. O painel direito mostra a estrutura dos alelos normais e com os rearranjos, bem como a localização e orientação das sequências de *primers* (indicada pelas setas) usadas para amplificação dos rearranjos (Adaptado de Bagnall *et al.*, 2006).

A inversão no IVS 1, apresenta uma frequência de 5% em pacientes com HA grave (Mantilla-Capacho *et al.*, 2007). A recombinação intracromossômica ocorre entre a região intragênica, *int1h-1*, e sua cópia homóloga extragênica (*int1h-2*) e produz dois mRNA quiméricos. Um desses mRNA, provavelmente sob controle do promotor do gene *F8*, contem o primeiro *exon* do gene *F8* seguido por *exons* facultativos e pelos *exons* 2-6 do gene VBP1. O outro mRNA , transcrito sobre o controle do promotor C6.1A, possui o gene C6.1A, um número de *exons* facultativos, parte do IVS 1 e *exons* 2 a 26 do gene *F8* (Figura 5) (Brinke *et al.*, 1996, Bagnall *et al.*, 2002).



Figura 5. Esquema da Inversão do IVS 1 no gene F8. (1) Mapa genômico das regiões flanqueadoras das repetições *int1h*. Foram representados os *exons* 1 e 2 do gene F8, os *exons* 1-3 do gene VBP1, os genes C6.1A e C6.1B e *exons* facultativos (caixas cinza). (A) A barra horizontal representa o IVS 1 do gene F8, flanqueado pelos *exons* 1 e 2, que indicam a direção da transcrição. A caixa cinza da barra indica a região que participa da inversão, denominada *int1h-1*. A linha horizontal indica a região onde é encontrada a sequência homóloga ao *int1h-1* (localizada anteriormente a extremidade 5' do gene F8), denominada *int1h-2*; também representada por uma caixa cinza. As setas indicam a orientação das sequências repetidas. (B, C) Recombinação homóloga entre *int1h-1* e *int1h-2* proposta para explicar a origem da inversão. (D) Inversão resultante da recombinação. (A) e (D) Localização e orientação das sequências de *primers* (indicada pelas setas) usadas para amplificação dos rearranjos. CEN indica o centrômero e TEL o telômero (Adaptado de Bagnall *et al.*, 2002).

Ambos, o rearranjos do IVS 1 e IVS 22 impedem a formação completa do mRNA do gene *F8* e consequentemente resulta na ausência da proteína, o que conduz à HA grave (Mantilla-Capacho *et al.*, 2007).

O diagnóstico molecular direto

Diante do conhecimento de que as inversões do IVS 1 e IVS 22 apresentam alta

frequência (Pandey e Mittal, 2001), o diagnóstico deve ser inicialmente tentado pela

análise direta desses rearranjos.

O Southern blot, primeiro método de detecção proposto para análise e diagnóstico das inversões, é capaz de detectar todos os tipos de rearranjos resultantes da recombinação homóloga entre as regiões *int22h* (Lakich *et al.*, 1993). Liu *et al.* (1998) propuseram um método mais rápido e acessível que combina *Long-range*-PCR (LD-PCR) e *Overlapping*-PCR, porém não discrimina as inversões entre si e nem as inversões das deleções. Os métodos atualmente propostos utilizam as técnicas LD-PCR e *Inverse Shifting*-PCR (IS-PCR). Ambas as técnicas podem detectar e discriminar cada possível rearranjo no IVS 1 (Bagnall *et al.*, 2002, Rossetti *et al.*, 2008) e IVS 22 (Bagnall *et al.*, 2006; Rossetti *et al.*, 2008). No segundo método, proposto por Rossetti *et al.*, é possível a análise simultânea dos rearranjos no IVS 1 e IVS 22. Entretanto, estas técnicas não estão adaptadas ao diagnóstico de rotina por serem técnicas elaboradas e dispendiosas. Nas Figuras 2 e 4 são indicados os pares de *primers* usados na amplificação por LD-PCR dos rearranjos no IVS 22, e na Figura 5, do IVS 1.

As outras mutações responsáveis pela HA são, na sua maioria, mutações de ponto no gene F8, o que torna difícil o seu reconhecimento, devido ao grande número de análises a serem verificadas em um só indivíduo para se chegar ao diagnóstico. Nesses casos, a detecção da mutação exige técnicas, tais como SSCP (do inglês "single strand conformation polymorphism"), CSGE (do inglês "conformation sensitive gel electrophoresis"), AMD (do inglês "amplification and mismatch detection"), DGGE (do inglês "denaturing gradient gel electrophoresis") e sequenciamento do DNA. Entretanto, cada uma dessas técnicas possui aplicabilidade e eficiência variáveis (Pandey & Mittal, 2001). Atualmente, a técnica DHPLC (do inglês "denaturing high performance liquid chromatography") seguida por sequenciamento tem sido utilizada com frequência por permitir a análise de toda a região codificadora, como também dos sítios flanqueadores de *splicing* (Oldenburg *et al.*, 2001), embora também seja técnica de difícil aplicação na rotina. (Lin *et al.*, 2008).

Diagnóstico indireto

Na impossibilidade do uso rotineiro de métodos de diagnóstico direto, a maioria dos serviços de investigação se baseia na análise de ligação do gene *F8* e marcadores polimórficos (El-Maarri *et al.*, 1999). Ou seja, é determinada a fase de ligação entre polimorfismos de DNA situados próximos ao gene *F8* ou mesmo internamente a este (geralmente dentro dos seus *Introns*). Em seguida, o heredograma é analisado para distinguir qual é o haplótipo que acompanha a mutação e, desta forma, traçar a segregação do alelo mutante na família estudada (Kim *et al.*, 2005).

Neste processo, dois pontos devem ser cuidadosamente considerados: a taxa de recombinação e a heterozigose de cada marcador.

Como ocorre cerca de um evento de recombinação para cada milhão de pares de bases, genes próximos num cromossomo geralmente não são separados, e tendem a ser herdado juntos. Portanto, quanto mais próximo fisicamente o marcador estiver da região de interesse, menor a taxa de recombinação entre eles e maior a tendência ao desequilíbrio de ligação (LD). O LD é a ocorrência na população, de uma frequência de um determinado haplótipo, maior do que o esperado pelo produto das frequências alélicas individuais. Apesar de que possam ser usados marcadores muito próximos ao gene F8 e, portanto, com probabilidade de recombinação quase nula, fica claro que os marcadores intragênicos são mais seguros e precisos para este tipo de análise.

A detecção dos eventos de recombinação que permitem uma avaliação da ligação entre dois *loci*, só é possível quando esses dois *loci* são polimórficos e apresentam alta heterozigose. Quanto maior a heterozigose, maior a informatividade dos haplótipos, pois mais fácil se torna distinguir a ocorrência de recombinação. A heterozigose é estimada a partir das frequências alélicas cuja determinação na população de interesse se torna obrigatória (Bowen, 2002).

Para a análise de ligação de *loci* ligados ao cromossomo X o genótipo do pai dá informação direta sobre a fase de ligação dos genes em sua filha. Qualquer marcador ligado ao X presente no genótipo desta, mas ausente em seu pai, deve ter sido herdado de sua mãe. A Figura 6 exemplifica uma análise de ligação de uma família com doença recessiva ligada ao cromossomo X.



Figura 6. Informação de ligação com marcador polimórfico em heredograma, de uma doença recessiva ligada ao X (Adaptado de Kim *et al.*, 2005).

Neste heredograma, nota-se que o haplótipo formado pelos alelos 14, 19, 20 e 18 acompanha a mutação, ou seja, está em acoplamento com o gene da doença e é transmitido pela avó materna (I-2), portadora do alelo anormal, aos seus descendentes. Como se vê, conhecer este haplótipo, pela verificação da fase, permite saber se o indivíduo é ou não portador deste gene sem que seja necessário fazer um exame direto, ou seja, sem que se tenha feito a caracterização da alteração molecular na família. A fase de ligação varia entre as famílias e, portanto, deve ser estabelecida para cada caso em particular.

O uso da análise de ligação é o passo que se segue após a tentativa de detecção das inversões no IVS 1 e IVS 22. No entanto, este tipo de análise não é só usada nos casos em que a análise direta seja difícil ou impossível, mas às vezes, simplesmente para confirmar o diagnóstico. Sendo assim, o potencial de uso do desequilíbrio de ligação no diagnóstico indireto e na confirmação de eventuais resultados de exames diretos das mutações causadoras da deficiência do FVIII cria a necessidade de maior conhecimento de parâmetros de distribuição populacional de marcadores ligados ao gene, bem como do estabelecimento de dados eventualmente úteis a tais processos diagnósticos.

Polimorfismos ligados ao gene F8 na população brasileira

A maioria dos trabalhos realizados em populações brasileiras visando à determinação de frequências alélicas de polimorfismos próximos ao gene *F8* ou dentro dos seus *Introns*, bem como a detecção de portadoras e diagnóstico pré-natal baseou-se na técnica de RFLPs (Figueiredo *et al.*, 1993; Arruda *et al.*, 1993a; Arruda *et al.*, 1993b; da Silva Jr *et al.*, 1994; Soares *et al.*, 2001).

Os polimorfismos intragênicos com repetições dinucleotídicas (CA)n (Lalloz *et al.,* 1991 e 1994) nos IVS 13 e IVS 22 do gene *F8* foram analisados em uma amostra da cidade de São Paulo (Soares *et al.,* 2001), em uma amostra do Espírito Santo (De Carvalho *et al.,* 2007) e em três outras amostras urbanas de São Paulo, Pernambuco e Rio Grande do Sul (Massaro, 2005) e se mostraram bons marcadores multialélicos a serem usados em análises de ligação para detecção de portadoras e diagnóstico pré-natal.

No Brasil, devido às peculiaridades na formação de sua população, não basta fazer um simples levantamento de frequências alélicas. Há que se considerar a ancestralidade da população, pois nos primeiros 350 anos da colonização do Brasil, entre 1500 e 1850, chegaram aproximadamente 500.000 portugueses e quatro milhões de africanos que dividiram espaço na nova terra com 2,4 milhões de ameríndios então existentes (Curtin, 1969; IBGE 2000). No período seguinte, até o presente, chegaram oficialmente no Brasil mais de cinco milhões de imigrantes, a maioria portugueses (espanhóis, alemães, sírios, libaneses, japoneses e outros não ultrapassaram 30% deste total) (IBGE, 2000). Assim, a formação da população brasileira é resultado de cinco séculos de mistura entre populações de três diferentes continentes: colonizadores europeus, escravos africanos e ameríndios. As migrações inter-regionais, em massa durante o período colonial do Brasil promoveram a dispersão e a miscigenação desses três componentes ancestrais fundadores em diferentes proporções ao longo do país. Em decorrência disso, a população brasileira atual é considerada tri-híbrida, embora as proporções dos três grandes componentes ancestrais variem consideravelmente conforme a região geográfica.

Até o momento, não existem trabalhos no Brasil com marcadores ligados ao gene F8 em que tenham sido comparados os três componentes ancestrais (ameríndios, portugueses e africanos) formadores de nossa população. Também, não foi identificado nenhum haplótipo associado às mutações da Hemofilia A. Portanto, a determinação de um maior número de marcadores conhecidos e ligados ao gene F8 em amostras dos diferentes segmentos da população brasileira (população urbana, quilombos e ameríndios) fornecerá um painel de frequências alélicas que poderá ser utilizado como banco de dados para trabalhos nesta linha.

HIPÓTESE

Os cinco marcadores selecionados para este estudo (IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG) distam até 500 Kb, o que limita significativamente a probabilidade de recombinação entre eles. Nenhum deles tem distribuição alélica e haplotípica estabelecidas para todos os segmentos da população brasileira, mas quatro deles têm frequências descritas em outras populações (Tabelas A1. 1-4) o que torna os dados aqui obtidos comparáveis a outros estudos.

Haplótipos como os acima descritos podem ser utilizados não só no diagnóstico indireto da HA, auxiliando o aconselhamento genético, como também na determinação do tempo de miscigenação das populações analisadas. No primeiro caso, busca-se determinar a existência de haplótipos associados à doença ou, pelo menos, co-segregantes em famílias de afetados. No segundo caso, o desequilíbrio de ligação que eventualmente esteja presente pode ser usado para medir o tempo de formação (miscigenação) da população em estudo (Collins, 2009; Van Hout *et al.*, 2009).

OBJETIVOS

Diante destas considerações, propõem-se aqui os objetivos seguintes.

Caracterizar a distribuição das frequências alélicas e haplotípicas de cinco marcadores ligados ao gene *F8* em amostras de vários segmentos da população brasileira, a saber: quilombos localizados no Estado do Piauí, aldeias de ameríndios do Estado do Amazonas, amostras de populações urbanas dos Estados de São Paulo, do Rio Grande do Sul e de Pernambuco e em uma amostra de pacientes hemofílicos do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo (HCFMRP-USP);

Determinar parâmetros de diversidade genética dentro e entre as populações analisadas e comparar com os parâmetros obtidos anteriormente com marcadores autossômicos e outros marcadores ligados ao cromossomo X;

Estimar a ancestralidade genômica com base nestes marcadores;

Verificar o poder de descrição dos haplótipos, isto é, se existe combinações típicas dessas populações, como também verificar a eventual existência de um haplótipo associado à doença.

Verificar a possibilidade de uso desses haplótipos no aconselhamento genético a famílias dos afetados.

Estimar o eventual desequilíbrio de ligação e, em caso positivo, verificar a possibilidade de seu uso na determinação da composição relativa das ancestralidades (européia, africana e ameríndia) nas amostras populacionais estudadas.

MATERIAL E MÉTODOS

Esquema do Trabalho

Fenótipos de cinco loci polimórficos, ligados ao gene F8, foram determinados em 493 indivíduos nascidos em três estados brasileiros, São Paulo (SP), Rio Grande do Sul (RS) e Pernambuco (PE), em 126 habitantes de três quilombos (Mimbó, Gaucinha e Sítio Velho) do Estado do Piauí, 174 ameríndios de três tribos (Tikúna, Baníwa e Kaxináwa) do estado do Amazonas e em 72 pacientes hemofilicos do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo (HCFMRP-USP) que são atendidos no HEMOCENTRO de Ribeirão Preto e 10 mães (os fenótipos dos familiares de 11 desses pacientes foram determinados para o estudo familiar). Este levantamento, de frequências alélicas e haplotípicas, permitiu determinar parâmetros de diversidade genética; estimar a ancestralidade genômica; verificar o poder de descrição dos haplótipos, isto é, se existem combinações típicas dessas populações, como também, verificar a existência de desequilíbrio de ligação e em caso positivo se é possível seu uso na determinação do tempo de mistura populacional na formação das amostras populacionais estudadas. E com intuito de auxiliar no aconselhamento genético, esses dados permitiram a investigação da existência de um haplótipo associado à doença e a comparação dos resultados obtidos dos pacientes com aqueles das populações normais, para assim, verificar a possibilidade do uso desses haplótipos no diagnóstico da HA.

As amostras

Amostras Urbanas

1. **SP.** A primeira amostra, oriunda do Estado de São Paulo, é constituída por indivíduos que procuraram o serviço de Investigação de Paternidade do

Departamento de Genética do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo (HCFMRP-USP). Desta amostra foram selecionados 159 indivíduos independentes (79 homens e 80 mulheres) provenientes de cidades próximas ao serviço.

2. **RS.** A segunda amostra, oriunda do Estado do Rio Grande do Sul, totaliza 156 indivíduos não aparentados (107 homens e 49 mulheres) selecionados entre os pacientes que procuraram o Instituto de Cardiologia de Porto Alegre, para atendimento por diversos problemas.

3. **PE.** A terceira amostra, composta por 178 indivíduos não aparentados (150 homens e 28 mulheres) nascidos no Estado de Pernambuco, foi selecionada entre os doadores de sangue do Hemocentro de Pernambuco, HEMOPE, na cidade de Recife.

A Figura 7 mostra a localização geográfica das cidades do Brasil onde foram realizadas as coletas das amostras urbanas.



Figura 7. Localização geográfica das amostras urbanas das cidades de Recife (Estado de Pernambuco), Ribeirão Preto (Estado de São Paulo) e Porto Alegre (Estado do Rio Grande do Sul).

Amostras de Afro-Derivados

Esta amostra, composta por 126 indivíduos independentes de três comunidades afro-brasileiras situadas no estado do Piauí: Mimbó (30 homens e 26 mulheres), Sítio Velho (22 homens e 26 mulheres) e Gaucinha (sete homens e 15 mulheres) (Figura 8); foi coletada em 1993 por professores da Universidade do Piauí, liderados pela professora Dra. Zélia Arpini Sampaio, em projeto que incluía atendimento médico e alguns exames laboratoriais de assistência às comunidades estudadas. Fez-se uma visita a todos os domicílios das três comunidades. Para cada indivíduo com idade superior a três meses, preencheu-se uma ficha para a identificação e registros de dados pessoais e genealógicos Tais amostras foram caracterizadas com base em marcadores clássicos, para vários parâmetros, incluindo levantamento demográfico (Arpini-Sampaio 1998; Arpini-Sampaio *et al.*, 1999).



Figura 8. Localização geográfica de comunidades afro-derivadas do Piauí. Neste trabalho foram analisadas as comunidades rurais: Mimbó (MI), Sítio Velho (SV) e Gaucinha (GA).

 MI. Mimbó localiza-se a 6º 14' latitude Sul e 42º 50' longitude oeste a 22 Km da sede do município de Amarante e a 170 Km de Teresina, capital do Estado (Figura 8). Segundo moradores mais antigos, o povoado originou-se na época da abolição da escravatura. Grupos de negros abandonaram as fazendas e migraram em busca de
um local para se fixarem e viverem sua liberdade. De um desses locais, próximo ao município de Oeiras, no início do século 20, desmembraram-se do grupo três casais e se fixaram às margens do riacho Mimbó, onde gradativamente, foram incorporados ao grupo outros migrantes do povoado original ou negros oriundos de outras fazendas do Estado.

A liderança é sempre mantida pelos membros mais velhos e embora não existam leis escritas, vigora o bom senso na decisão sobre o que é permitido ou não. Eles utilizam a agricultura de subsistência, em especial a de mandioca. Os produtos obtidos na agricultura são divididos entre todos os moradores, que se consideram uma grande família.

2. **SV.** Sítio Velho localiza-se a 5º 20' latitude Sul e 41º 15' longitude Oeste a 280 Km de Teresina (Figura 8). Segundo moradores mais antigos, a comunidade surgiu no início do século 20 por aglomeração (nas margens de um riacho onde a caça era abundante) de africanos oriundos de diferentes pontos de uma região relativamente ampla da fronteira entre os Estados do Piauí e Ceará. Sua existência foi registrada por volta de 1984, quando cearenses, em busca de terras, chegaram ao povoado desencadeando uma guerrilha com seus moradores. Praticam a agricultura de subsistência.

3. **GA.** Gaucinha está localizada a 4º 49' latitude Sul e 42º 10' longitude Oeste, distante 80 Km da sede do município de Campo Maior e a 180 Km de Teresina. (Figura 8). Na época em que foi realizada a visita à comunidade, esta contava com apenas 11 choupanas com localização esparsa, em decorrência da emigração dos moradores para a sede de Campo Maior.

Amostras de Tribos Indígenas da Amazônia Brasileira

Esta amostra, composta por 174 indivíduos das tribos Tikúna (50 homens e 60 mulheres de três aldeias: Umariaçu, Feijoal e Vendaval), Baníwa (22 homens e 24 mulheres da aldeia Jandu Cachoeira) e Kaxináwa (seis homens e 12 mulheres da aldeia Canabrava), foi coletada no ano de 1976 durante a expedição "Alpha-Hélix" (Neel *et al.*, 1980; Mestriner *et al.*, 1980) destinada à Amazônia oriental brasileira e aprovada segundo processo FUNAI/BSB/4854/75 e autorizações 26/76, 27/76, 74/76 e 75/76. Descrição detalhada encontra-se em Simões (1980). As localizações geográficas das aldeias selecionadas para o presente estudo estão apresentadas na Figura 9.



Aldeias	Latitude	Longitude	n
Umariaçu	4°17' S	69°55' O	48
Feijoal	4°15' S	69°35' O	43
Vendaval	3°30' S	69°27' O	19
Jandu Cachoeira	1º33'S	68°44' O	46
Canabrava	8° 07' S	69°32' O	18
Total de indivi	íduos		174
	Aldeias Umariaçu Feijoal Vendaval Jandu Cachoeira Canabrava Total de indivi	Aldeias Latitude Umariaçu 4°17'S Feijoal 4°15'S Vendaval 3°30'S Jandu Cachoeira 1°33'S Canabrava 8° 07'S Total de indivision	Aldeias Latitude Longitude Umariaçu 4°17'S 69°55'O Feijoal 4°15'S 69°35'O Vendaval 3°30'S 69°27'O Jandu Cachoeira 1°33'S 68°44'O Canabrava 8°07'S 69°32'O

Figura 9. Localização geográfica das aldeias pertencentes às tribos Tikúna, Baníwa e Kashináwa e o respectivo o número de indivíduos analisados.

O critério para a seleção dos indivíduos da amostra analisada foi o menor grau de parentesco possível dentro da amostra total. A partir das fichas de coleta, que incluem o número de identificação, o nome do indivíduo, sexo, idade, notas sobre parentesco, mistura com brancos e o nome dos pais e/ou esposos, foi possível excluir primeiramente os filhos e netos da amostra. Em uma segunda triagem foi escolhido, de maneira aleatória, apenas um indivíduo dentre os que apresentavam irmãos ou o mesmo sobrenome, o que resultou em uma amostra de 323 indígenas (Mendes Jr., 2001). Dentre esses, selecionamos os 174 aqui analisados.

 TI. Os Tikúna habitam o Estado do Amazonas, mas representantes desta tribo também podem ser encontrados em território peruano e colombiano. Na época da coleta havia cerca de 11.000 membros desta tribo em território brasileiro. No entanto, o censo de 1994 indicou 23.000 membros, também em território brasileiro. Embora apresentem contatos com não-indígenas, esta tribo ainda mantém a sua identidade étnica e casamentos com pessoas de fora são raros (Simões, 1980).

2. KA. Os Kashináwa habitam o extremo sudoeste do Estado do Amazonas e uma porção do Estado do Acre. Apresentavam cerca de 2.000 indivíduos habitando as margens do Rio Embira e seus tributários na época da coleta (Mestriner *et al.,* 1980). Estimativas de 1994 apontaram 3.387 membros desta tribo em território brasileiro. O grau de aculturação desta tribo é bastante variado (Simões, 1980).

3. **BA.** Os Baníwa habitam a região pertencente aos municípios de São Gabriel da Cachoeira e Japurá (AM), noroeste do Estado do Amazonas. Estimativas na época da coleta (1976) indicavam a existência de aproximadamente 1.500 indígenas distribuídos em 16 aldeias, consideradas semi-aculturadas. Em censo realizado em 1995 foram apontados 3.189 indivíduos (Oliveira, 1999).

Amostras de pacientes com Hemofilia A

HA. Esta amostra, recentemente coletada, é composta por 72 pacientes hemofilicos do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo (HCFMRP-USP) que frequentam o Ambulatório de Hemostasia do HEMOCENTRO de Ribeirão Preto, USP, SP e 29 familiares. Destes 72 pacientes, 42 (58,33 %) são hemofilicos grave (oito produzem inibidor), 14 (19,45

%) são hemofílicos moderado (um produz inibidor) e 16 (22,22 %) são hemofílicos leve (um produz inibidor).

Aspectos Éticos

Neste trabalho foram analisadas seis amostras populacionais brasileiras, dentre essas, cinco já haviam sido coletadas (exceto HA), fazem parte de bancos de amostras, encontram-se armazenadas em nosso laboratório e foram utilizadas em estudos anteriores aprovados pelo Comitê de Ética do HCFMRP-USP.

A coleta de informações demográficas e de material biológico dessas populações ocorreu anteriormente, durante o desenvolvimento de outros projetos, alguns mais antigos que a própria resolução CNS 196/96 do Ministério da Saúde. Em todos eles foram seguidos rigorosamente os preceitos de conscientização e autorização dos participantes e das lideranças das comunidades, aquiescência voluntária e garantia de anonimato. A visita às tribos indígenas que deu origem às amostras que aqui utilizamos, foi também aprovada pela FUNAI (processos FUNAI/BSB/4854/75 e autorizações 26/76, 27/76, 74/76 e 75/76). Em todos os processos acima relacionados à aprovação incluiu isenção do Termo de Consentimento Livre e Esclarecido, diante da total impossibilidade de reencontro dos indivíduos amostrados.

Nos estudos já realizados, foi analisado certo número de marcadores genéticos. O presente trabalho apenas ampliou este número, não alterando, portanto sua natureza, nem acrescentando nenhum outro tipo de análise que implique alterações dos aspectos éticos. Reafirmamos que está garantido o anonimato dos participantes, não sendo previsível qualquer possibilidade de prejuízo, lesão física ou de direitos de membros destas amostras populacionais e comunidades, individual ou coletivamente considerados. Portanto, como se tratava de amostras já

coletadas e não estava programado nenhum contato direto ou indireto com estes indivíduos, foi concedida à isenção da aplicação do Termo de Consentimento Livre e Esclarecido para este trabalho também.

Aos indivíduos da amostra HA, recentemente coletada, composta por pacientes hemofílicos, foi aplicado o Termo de Consentimento Livre e Esclarecido no momento da coleta.

O presente projeto foi analisado pelo Comitê de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo (HCFMRP-USP) e enquadrado na categoria aprovado de acordo com o Processo HCRP nº 6115/2007.

Coleta e Conservação do Material

Amostras de Populações Urbanas (SP, RS e PE)

O sangue das amostras SP, PE e RS foi coletado em tubo *VacuTainer* com anticoagulante (EDTA), sendo os tubos devidamente identificados. As amostras de sangue foram submetidas à centrifugação, sendo o sobrenadante (plasma) descartado. A fração celular foi lavada três vezes em solução salina a 0.9% e ao sedimento foi adicionado igual volume de Tampão de Estocagem de Eritrócitos (32,5g de citrato de tripotássio + 6,0g de fosfato dibásico de potássio + 4,7g de fosfato monobásico de potássio + água destilada qsp. 1 litro) em glicerol a 40%. Após este procedimento, essa fração celular foi congelada a -20° C e, posteriormente, foi realizada a extração de DNA.

Para estocagem do sangue utilizou-se:

Solução salina: NaCl 9% em água
Tampão de Estocagem de Eritrócitos: Citrato tripotássio 0,1M, KH₂PO₄ 0,0345M, K₂HPO₄ 0,0344M.

Amostras de Afro-derivados (MI, SV, GA)

Coletou-se 5 mL de sangue em tubo *VacuTainer*, com EDTA, os quais foram mantidos e transportados em caixas de isopor e transferidos para a geladeira, até 12 horas após a coleta.

Foi coletado um total de 407 amostras sendo: 183 na comunidade de Mimbó, 182 na comunidade de Sítio Velho e 42 na comunidade de Gaucinha.

Todas as amostras foram processadas no laboratório de Biologia Geral da Universidade Federal do Piauí, em Teresina, em até 72 horas após a coleta. O plasma foi estocado diretamente a –20°C, o *buffy coat* (cerca de 500 µL) foi recolhido em tubo de polipropileno (*eppendorf*) de 1,5 mL e mantido em geladeira para a extração de DNA. Os eritrócitos restantes foram lavados em solução salina e estocados a - 20°C em glicerol 40% v/v em tampão contendo Citrato tripotássico KH2PO4, K2HPO4 e H2O. As amostras devidamente processadas e congeladas foram transportadas por via aérea para o laboratório de Genética Bioquímica do Departamento de Genética da FMRP.

Amostras das tribos indígenas (TI, KA, BA)

Durante a expedição "Alpha-Helix" em 1976, as amostras de sangue foram coletadas em tubos *VacuTainer* contendo 2,0 mL de anticoagulante ACD e foram refrigeradas cerca de 12 horas após a coleta. Cerca de três dias após a coleta este material foi processado para estocagem. As amostras foram então centrifugadas (1.500 rpm, durante 10 minutos) e o plasma retirado. As hemácias foram lavadas três vezes em solução salina isotônica, e o sedimento resultante após a última centrifugação foi ressuspenso em igual volume de solução de glicerol a 40% em tampão de estocagem (32,5g de citrato de tripotássio + 6,0g de fosfato dibásico de potássio + 4,7g de fosfato monobásico de potássio + água destilada qsp. 1 litro). As amostras foram então mantidas a -20° C até o presente.

Amostras de Hemofílicos A e familiares

A amostra HA, recentemente coletada, inclui sangue total e *swab* bucal. Para o sangue, seguimos o mesmo processo de coleta e conservação descrito nas amostras urbanas (SP, PE e RS). As amostras de swab bucal foram coletadas com escova cervical estéril a óxido de etileno (Kolplast ci Ltda). O procedimento consistiu em um esfregaço da mucosa bucal do indivíduo. O material coletado foi transferido para um tubo *VacuTainer*, com EDTA e 5 mL de soro fisiológico. Esse processo foi repetido até a visualização do precipitado de células. Após este procedimento, o material foi congelado a –20° C e, posteriormente, foi realizada a extração de DNA.

Marcadores genéticos analisados

Para este trabalho foi analisado um conjunto de cinco marcadores: quatro microssatélites intragênicos (IVS 1, IVS 13, IVS 22 e IVS 25.3) e um microssatélite extragênico (IKBKG). Todos esses cinco marcadores distam até 500 Kb (Figura 10). A lista de *primers*, o número de alelos e o tamanho em pb são descritos na Tabela 1.



Figura 10. Posição física dos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG no cromossomo X. Os números correspondem à distância em pares de bases do Xq-Tel usando a sequência referência do NC_000023.

Loci	Sequências dos <i>Primers</i> (5' - 3')*	Tamanho (pb)	Alelos ^a
	(F) TTT ACC ATT GGC ATA TAT TT	97 102	0 (12 20)
IVSI (CA)n	(R) TTT ATA CAT CAA TTG TAT TA	87-103	9 (12-20)
	(F) TGC ATC ACT GTA CAT ATG TAT CTT		14 (14 40 00)
1VS 13 (CA) _n	(R) CCA AAT TAC ATA TGA ATA AGC C	129, 133-157	14 (14,16-28)
	(F) TTC TAA GAA TGT AGT GTG TG	71.00	10 (20, 20)
1V5 22 (G1)n(AG)n	(R) TAA TGC CCA CAT TAT AGA	71-69	10 (20-29)
	(F) AGT CCA AGA TCA AGG GGT AGG	110 101 100 146	10 (0 10 14 00)
1V3 25.3 (CA)n	(R) CAT CAC ATT CCA GCC TGG ACT	110, 124, 120-140	12 (9, 12,14-23)
	(F) TGG GGG ATA CAT TTT AAC AGG ATT TC	160,100	c
	(R) CAA AAC CCA GAT CTA CCC AAG GC	100-100	0

Tabela 1. Lista dos *loci* analisados, seus respectivos *primers*, número de alelos e tamanho em pb.

* - Sequências dos *primers*: IVS 1 e IVS 25.3 (Kim *et al.*, 2005); IVS 13 (Lalloz *et al.*, 1991); IVS 22 (Lalloz *et al.*, 1994); e IKBKG (Fang *et al.*, 2006).

^a – Tamanho do alelo: o número inteiro refere-se ao número de alelos já descritos e os números entre parênteses referem-se ao número de repetições do menor alelo e do maior alelo.

Análise Laboratorial

Extração do DNA Genômico

Reagentes e Soluções

Detergente Nonidet (polyethoxyethanol)

Triton X

Tween 20

Tampão de lise de eritrócitos: Tris/HCl 0,01 M pH 7,6; sacarose 0,32 M; MgCl 5,0 mM; e Triton X-100 1%

Tampão de lise de leucócitos: Tris/HCl 0,01M pH 8,5; KCl 50 mM; MgCl2 2,5 mM; NP-40 0,45%; Tween 20 - 0,45%

Proteinase K (10 mg/mL)

Procedimento

A extração de DNA para as amostras de sangue foi realizada a partir de uma adaptação ao método de HIGUCHI (1989). Aproximadamente 50µL de sangue total foi colocada em um microtubo de polipropileno de 1,5 mL (tipo *eppendorf*), utilizando-se uma micropipeta e ponteiras estéreis. Em seguida, adicionou-se 1,0 mL de tampão de lise de eritrócitos a cada microtubo. Esta mistura foi homogeneizada e a centrifugação do conteúdo dos microtubos foi realizada a 6.000 giros durante 2 minutos. Retirou-se o sobrenadante com o auxílio de um sugador, e acrescentou-se, novamente, 1,0 mL de tampão de lise de eritrócitos ao precipitado avermelhado. Homogeneizou-se e centrifugou-se cada microtubo novamente. Estes procedimentos foram repetidos até que o precipitado apresentasse cor branca (de 3-4 repetições), indicando a presença de glóbulos brancos (leucócitos) e a ausência de hemácias.

Quando o precipitado adquiriu esta aparência branca, retirou-se o sobrenadante, suspendeu-se o precipitado em 300 μ L de tampão de lise de leucócitos, e adicionou-se 5 μ L de Proteinase K em cada microtubo.

Terminadas estas etapas da extração, as amostras foram colocadas em uma estufa a 65°C por um período de 1 hora e, em seguida, em uma estufa a 37°C por um período de aproximadamente 12 horas. Por fim, as amostras foram submetidas a um passo de desnaturação da Proteinase K (94°C por 10 min) e então armazenadas em um refrigerador com temperatura aproximada de -20°C.

Para a extração de *swab* bucal, após homogeneização do tubo *VacuTainer*, foi colocado aproximadamente 1 mL da amostra em um microtubo de polipropileno de 1,5 mL (tipo *eppendorf*), utilizando-se uma micropipeta e ponteiras estéreis. O conteúdo dos microtubos foi centrifugado a 6.000 giros durante 2 minutos. Retirouse o sobrenadante com o auxílio de um sugador, suspendeu-se o precipitado em 300 µL de tampão de lise de leucócitos, e adicionou-se 5 µL de Proteinase K em cada microtubo. Os passos seguintes seguiram-se como na extração de sangue.

Reação em Cadeia da Polimerase (PCR)

Reagentes e Soluções

- **DNA polimerase (Taq):** 1 U/µL de tampão de estocagem (BIOTOOLS B & M Labs, AS)
- Tampão de estocagem da DNA polimerase: 10 mM de Tris/HCl (pH 8,0), 50 mM KCl, 1 mM EDTA, 01% Triton x-100, 50% de glicerol (BIOTOOLS B & M Labs, AS)
- **dNTP solução estoque:** quatro soluções separadas de 100mM de cada base (dATP, dCTP, dGTP, dTTP), pH 8,3 (BIOTOOLS B & M Labs, AS)
- dNTP solução trabalho (20 mM): obtida diluindo-se com água a solução estoque (100mM) de cada dNTP para uma solução única de concentração 20mM (160 μL de água de MiliQ autoclavada mais 10 μL da solução estoque de cada dNTP).
- dNTP solução trabalho (5 mM): obtida diluindo-se com água a solução estoque (100mM) de cada dNTP para uma solução única de concentração 5 mM (160 μL de água de MiliQ autoclavada mais 10 μL da solução estoque de cada dNTP)
- Iniciadores (*Primers*) específicos solução estoque (50 μM): os primers liofilizados (Bio-Synthesis) A e B específicos para cada *locus* foram diluídos em água autoclavada, e estocados separadamente.
- **Iniciadores (***Primers***) solução trabalho (2,5 μM):** 10 μL do *primer* A, 10 μL do *primer* B (solução estoque, 50 μM) e 180 μL de água autoclavada.
- Tampão PCR livre de MgCl₂: Tris/HCl 75 mM pH 9.0; KCl 50 mM; (NH₄)2SO₄ 20 mM. (BIOTOOLS – B & M Labs, AS).

MgCl₂: Concentração de 50 mM (BIOTOOLS – B & M Labs, AS).

Dimethyl Sulfoxide (DMSO) 10%: SIGMA

Procedimento

Os ensaios da PCR foram realizados em um volume total de 25 μL. Todos os reagentes (Tampão de PCR livre de cloreto, dNTP, *primers* – solução trabalho, MgCl₂ e a Taq – DNA polimerase), com exceção do DNA, foram misturados em quantidade

específica para cada *locus* (Tabela 2) em um único tubo (mistura de reação) para garantir a homogeneidade das reações. As sequências de todos os *primers* estão na Tabela 1.

Loci	Água	Tampão livre de cloreto	dNTP (5 mM)	<i>Primers</i> (solução trabalho)	MgCl2 (50 mM)	Taq	DMSO
IVS 1	10,5	2,5	0,25	6,0	1,25	0,5	-
IVS 13	13,6	2,5	0,50	3,0	1,0	0,4	-
IVS 22	13,0	2,5	0,25	2,0	1,0	0,5	1,75
IVS 25.3	14,75	2,5	0,25	2,0	1,0	0,5	-
IKBKG	14,75	2,5	0,25	2,0	1,0	0,5	-

Tabela 2. Condições de PCR para os *loci* analisados no presente trabalho (quantidade em μ l para 1 reação).

Em cada microtubo de 0,5 mL ou em cada tubo de 0,2 mL da placa de 96 tubos para PCR, foram pipetados 4 µL do DNA genômico, previamente extraído. Para cada análise foi usado um controle negativo, contendo água no lugar do DNA genômico. Em cada microtubo, sob a amostra, foram pipetados 21 µL da mistura de reação. Quando foram utilizados os termocicladores MJ Research (Inc PTC - 100TM) e, consequentemente, microtubos de 0,5 mL, uma gota de óleo mineral foi adicionada para evitar a evaporação da mistura de reação. Este procedimento não foi necessário quando da utilização dos termocicladores Techne Genius ou GeneAmp® PCR system 9700 (Applied Biosystems), pois estes possuem tampas hermeticamente fechadas. Foram nestes termocicladores que as placas de PCR foram utilizadas ao invés dos tubos. A este passo seguiu-se o programa do termociclador correspondente ao locus (listado abaixo). Após o término da reação o produto da PCR foi guardado em geladeira (4C°) até sua utilização.

Programas utilizados

<u>IVS 1:</u>

35 ciclos: 94°C por 30 segundos, 45°C por 30 segundos, 72°C por 30 segundos; 1 ciclo: 72°C por 10 minutos, 20°C indefinidamente.

<u>IVS 13:</u>

1 ciclo: 94°C por 5 minutos; 33 ciclos: 94°C por 35 segundos, 59°C por 1 minuto, 72°C por 1 minuto e 30 segundos; 1 ciclo: 72°C por 10 minutos, 20°C indefinidamente.

<u>IVS 22:</u>

1 ciclo: 94°C por 5 minutos; 30 ciclos: 94°C por 1 minuto, 46°C por 1 minuto, 72°C por 1 minuto; 1 ciclo: 72°C por 10 minutos, 20°C indefinidamente.

<u>IVS 25.3:</u>

1 ciclo: 94°C por 5 minutos; 33 ciclos: 94°C por 35 segundos, 59°C por 1 minuto, 72°C por 1 minutos e 30 segundos; 1 ciclo: 72°C por 10 minutos, 20°C indefinidamente.

IKBKG:

1 ciclo: 95°C por 5 minutos; 30 ciclos: 95°C por 30 segundos, 60°C por 30 segundos, 72°C por 45 segundos; 1 ciclo: 72°C por 5 minutos, 20°C indefinidamente.

Análise do Produto Amplificado

O produto amplificado foi separado por eletroforese em géis de poliacrilamida

desnaturantes e não-desnaturantes.

Reagentes e Soluções

TEMED: tetrametiletilenodiamina (Pharmacia Biotech).

Formamida

- **Solução de acrilamida/bis-acrilamida (29:1):** 29 g de acrilamida; 1 g de bisacrilamida diluídas em 100 mL H₂O.
- **Solução de EDTA pH 8,0:** 186 g de EDTA; 11 de H₂O. Acertar o pH com pastilhas de hidróxido de sódio (NaOH).

- **Solução saturada de Persulfato de potássio:** 650 mg de persulfato de potássio; 6,5 mL de H₂O.
- **Tampão TBE 10X (0,9M) pH 8,0:** 108 g de Tris (PM=121,1); 53 g ácido bórico; 40 mL de solução de EDTA; H₂O qsp 1L.
- **Tampão TBE cubas (1X):** 100 mL do tampão TBE (10X); 900 mL H₂O.
- Tampão de amostra (Loadding buffer): 900 μL de brofenol; 900 μL xilenocianol; 900 μL TBE; 4,5 mL Ficol 30% diluído em H₂O; 1,8 mL EDTA 0,5 M pH 8,0; 3,6 g sacarose. Misturar tudo até a dissolução completa da sacarose.
- **Tampão de amostra desnaturante:** 300 μl de tampão de amostra; 900 μL formamida.
- **Gel desnaturante 12%:** 9,6 g uréia; 4 mL de H2O; 8 mL solução de bisacrilamida; 2 mL TBE (10x); 15 μL TEMED; 300 μL solução de persulfato de potássio.
- **Gel desnaturante 10%:** 9,6 g uréia; 9,743 mL de H₂O; 6,667 mL solução de bisacrilamida; 2 mL TBE (10x); 15 μL TEMED; 300 μL solução de persulfato de potássio.
- Gel não desnaturante 10%: 21,9 mL de H₂O; 13,333 mL solução de bisacrilamida; 3,75 mL TBE (10x); 31,25 μL TEMED; 1 mL μL solução de persulfato de potássio.

Procedimento

O produto amplificado dos cinco microssatélites ligados ao gene *F8* foi separado por eletroforese em géis de poliacrilamida a 12% (PAGE) desnaturante (IVS 13 e IVS 25.3) e 10% (PAGE) desnaturante (IKBKG) e a 10% não-desnaturante (IVS 1 e IVS 22). Os géis desnaturantes foram feitos dissolvendo a uréia com a solução de acrilamida/bisacrilamida e água em "banho Maria" a 50°C, antes da adição do TBE, adicionado apenas após o resfriamento da solução.

Os catalizadores da reação de polimerização do gel, TEMED e persulfato de potássio foram adicionados à mistura do gel imediatamente antes de vertê-la em um cassete previamente montado, composto de duas placas de vidro separadas por espaçadores de *teflon* e presas com grampos. O tamanho das placas é determinado de acordo com o tipo de marcador (Tabela 3). Logo após, um pente de *teflon* foi colocado na borda superior, formando poços no gel, onde posteriormente foram aplicadas as amostras de DNA amplificado por PCR. Aguardou-se a polimerização por no mínimo 30 minutos.

Tabela 3. Condições específicas para eletroforese de cada *locus* analisado no presente trabalho.

Loci	Gel	Tamanho da placa (cm)	Voltagem	Tempo (h)
IVS 1 e IVS 22	10% não-desnaturante	22,0 x 17,0	240V	5:30
IVS 13	12% desnaturante	22,0 x 17,0	24mA/gel	5:00
IVS 25.3	12% desnaturante	14,5 x 16,5	24mA/gel	3:00
IKBKG	10% desnaturante	22,0 x 17,0	24mA/gel	4:30

Após a polimerização do gel o pente foi retirado e os poços foram lavados com água. O gel polimerizado foi montado em cuba de eletroforese vertical contendo tampão TBE (1x), em ambos os pólos (porção superior e inferior). Esta cuba foi conectada a uma fonte de voltagem, Amershan Pharmacia Biotech (EPS 1001), ajustada à voltagem ou corrente constante necessária para uma boa separação dos fragmentos amplificados (Tabela 3). Foi feita uma pré-corrida de pelo menos 15 minutos, onde as cubas com os géis foram ligados às fontes antes da aplicação das amostras e submetidas às condições de eletroforese. A fonte foi desligada e as amostras foram aplicadas nos poços.

Nos *loci* em que os fragmentos amplificados precisam ser submetidos a condições desnaturantes (IVS 13, IVS 25.3 e IKBKG), antes da aplicação das amostras no gel foi necessário um tratamento prévio com formamida, que auxilia no processo de desnaturação das amostras. Para isso foram colocados 7 µL de Tampão de Amostra desnaturante em um tubo *eppendorf*, junto com 7 µL do produto amplificado. Estes

tubos foram aquecidos a 94°C por 15 minutos e colocados imediatamente em banho de gelo (tratamento desnaturante), seguindo-se a aplicação no gel (Tabela 3).

Para os IVS 1 e IVS 22 não foi necessário tratamento desnaturante, sendo aplicado 7 μ L de amostra de DNA amplificado oriundo da PCR, juntamente com 7 μ L de Tampão amostra, em géis não desnaturantes 10% (Tabela 3).

Após a aplicação das amostras as fontes foram novamente ligadas e a eletroforese prosseguiu da maneira descrita na Tabela 3. Com o término da corrida eletroforética, o gel foi retirado cuidadosamente das placas de vidro e submetido aos procedimentos de coloração e secagem.

Coloração com Nitrato de Prata e Secagem do Gel:

Reagentes e soluções

- **Solução de nitrato de prata:** 10 g nitrato de prata; 100 mL de H₂O. Dissolver a prata em uma parte da água e depois completar com o restante, manter a solução ao abrigo da luz (volume final 100 mL).
- **Solução fixadora:** 160 mL etanol (PA) e 7 mL de ácido acético glacial (PA); 833 mL de H2O (volume final 1 L).
- **Solução reveladora:** 22,5 g de NaOH; 1 L de H₂O. Dissolver em um agitador o hidróxido de sódio em uma parte da água e depois completar com o restante (volume final 1 L). Na hora da coloração adicionar 1 mL de formaldeído para cada 100 mL da solução.

Procedimento

- **Fixação:** Após a retirada das placas de vidro e dos espaçadores o gel foi colocado em um recipiente de vidro contendo 100 mL de solução fixadora.
- **Impregnação com Nitrato de prata:** adicionou-se 2,0 mL de solução de nitrato de prata, e agitou-se por 5 minutos. A solução foi então descartada e o gel lavado em água quente por cerca de 10 segundos, agitando levemente e, ao final, descartando a água.

- **Revelação:** A solução reveladora foi despejada cuidadosamente no recipiente contendo o gel, que foi submetido à agitação por alguns minutos até que as bandas aparecessem nitidamente. A solução foi pré-aquecida em estufa a 65°C, para facilitar a reação de coloração.
- **Bloqueio da reação:** Após ter sido revelado, a solução reveladora foi descartada e a reação bloqueada com a lavagem direta do gel em 100 mL de solução fixadora.
- **Secagem do gel:** Após a leitura, todos os géis passaram por um processo de secagem simples para que pudessem ser armazenados para análises e confirmações posteriores. Duas folhas de papel celofane foram molhadas; uma placa de vidro, com a área maior que a do gel, foi coberta com uma das folhas; o gel foi colocado sobre a placa com o celofane sem deixar bolhas; o gel foi então bem molhado e coberto com a outra folha de celofane, também com cuidado de não deixar bolhas; este gel foi deixado secando à temperatura ambiente por dois ou três dias até a secagem completa, sendo então devidamente identificado e arquivado.

Determinação Fenotípica (Nomenclatura)

Os alelos dos microssatélites ligados ao gene *F8* foram designados segundo os autores citados na Tabela 1. Onde foi proposto que a nomenclatura dos alelos deva ser baseada no número de repetições: o número que denomina o alelo representa o número de repetições presente no mesmo. Por exemplo, um alelo com 10 repetições é denominado alelo *10. Os alelos foram identificados por ordem crescente de tamanho, definida pela migração eletroforética. Padrões diferentes de leitura foram intercalados a cada 3-5 amostras de maneira que o gel sempre começa e termina com um padrão de leitura. Os alelos destes padrões foram definidos por sequenciamento ou comparados lado a lado com amostras padrões previamente sequenciadas.

Após a eletroforese e coloração do gel as amostras foram comparadas com os padrões, de acordo com seu tamanho, e a leitura de seus alelos foi anotada. No caso

de alguma dúvida as amostras que apresentaram mesma leitura foram comparadas lado a lado em outro gel para confirmação de seus alelos. A Figura 11 mostra o padrão eletroforético de cada microssatélite analisado.



Figura 11. Fotos ilustrativas dos padrões de corrida eletroforética visualizados em cada microssatélite. O nome do marcador e as leituras dos alelos encontram-se, respectivamente, acima e abaixo no gel.

Tamanho Amostral

Os microssatélites deste trabalho foram analisados em 159 indivíduos de SP, 156 indivíduos do RS, 178 indivíduos de PE, 56 indivíduos de MI, 48 indivíduos de SV, 22 indivíduos de GA, 174 indivíduos de três tribos indígenas (Tikúna=110, Kashináwa=18 e Baníwa=46), 72 pacientes e 10 mães (os fenótipos dos familiares de 11 desses pacientes foram determinados para o estudo familiar). O tamanho amostral para cada *locus* variou, pois alguns indivíduos não amplificaram para todos os *loci.* O tamanho exato de cada amostra para cada *locus* é indicado nas Tabelas na medida em que os resultados são apresentados. Para análise dos haplótipos do cromossomo X só foram considerados aqueles indivíduos que amplificaram para todos os *loci.*

Análise Estatística

Frequências Gênicas em Populações Ancestrais

A partir de frequências alélicas de populações parentais disponíveis na literatura foi possível realizar as análises de componente principal e composição ancestral. Quando havia dados de mais de uma população foram feitas às médias das frequências, ponderadas pelos tamanhos amostrais. Essas populações e suas respectivas médias ponderadas podem ser encontradas nas Tabelas A1. 1-4.

Estimativas das frequências alélicas e genotípicas

Os microssatélites analisados apresentam alelos codominantes, o que permite inferir os genótipos a partir dos respectivos fenótipos. As frequências alélicas (x_i) e genotípicas (X_{ii}) de cada *locus* em cada amostra foram estimadas por contagem direta, utilizando-se o programa GENEPOP® (Raymond e Rousset, 1995a) versão 3.4 (disponível em http://wbiomed.curtin.edu.au/genepop), segundo as equações:

$$x_i = \frac{2n_{ii} + \sum n_{ij}}{2n} \quad e \qquad X_{ii} = \frac{n_{ii}}{n}$$

Em que:

x_i é a frequência do alelo i;
X_{ii} é a frequência do genótipo ii;
n_{ii} e n_{ij} correspondem ao número de homozigotos e heterozigotos observados para o alelo i, respectivamente;
n corresponde ao número de indivíduos analisados.

A comparação das distribuições das frequências alélicas foi realizada pelo teste exato de Fisher. Todos os alelos foram comparados entre as doze amostras populacionais e entre as populações disponíveis na literatura (Tabelas A1. 1-4).

Aderência ao Equilíbrio de Hardy-Weinberg

Segundo o teorema de Hardy-Weinberg as frequências genotípicas esperadas no equilíbrio podem ser estimadas a partir da expansão do seguinte binômio:

$$(x_i + x_j)^2 = x_i^2 + 2x_ix_j + x_j^2$$

Em que:

x_i² é a frequência esperada dos homozigotos do alelo i;
2 x_ix_j é a frequência esperada dos heterozigoto ij;
2x_j² é a frequência esperada dos homozigotos para o alelo j.

A aderência das frequências genotípicas observadas às proporções teóricas de Hardy-Weinberg foi verificada com o emprego do programa GENEPOP (Raymond e Rousset, 1995a) versão 3.4. Foram realizados três testes baseados na hipótese nula de união aleatória dos gametas: teste exato de probabilidade, teste para detecção da deficiência e para detecção do excesso de heterozigotos.

No teste exato de probabilidade (teste global) o valor de p corresponde à soma de probabilidades de todas as tabelas com probabilidade menor ou igual ao observado.

O segundo e o terceiro são testes mais sensíveis do que o de probabilidade e utilizam uma hipótese alternativa (H1) de excesso ou de deficiência de heterozigotos, respectivamente.

A aderência ao equilíbrio de Hardy-Weinberg foi verificada somente na amostra de mulheres.

Desequilíbrio de ligação entre loci

A análise de associações par-a-par entre *loci* foi realizada utilizando-se o programa GENEPOP® 3.4 (Raymond & Rousset, 1995a). A hipótese nula é a de que a distribuição genotípica em um *locus* é independente da distribuição em outro *locus*. Esta análise foi aplicada para verificar desvios do esperado pela regra de multiplicação entre pares de *loci* localizados em diferentes cromossomos. A palavra "ligação" neste caso não está relacionada com associação física entre alelos de *loci* de um mesmo cromossomo.

Diferenciação genética das populações

Os testes exatos para diferenciação populacional foram realizados com o uso do programa GENEPOP® 3.4 (Raymond & Rousset, 1995a). Este utiliza Tabelas de contingência RxC geradas automaticamente para cada *locus*, em que R é o número de populações e C é o número de alelos no *locus*.

Este procedimento compara cada *locus* em pares de populações, para determinar se existem diferenças nas frequências alélicas e genotípicas observadas, onde a hipótese nula testada é a de que a distribuição alélica é idêntica entre as populações (Raymond & Rousset, 1995b).

Diversidade haplotípica

A diversidade haplotípica (h), equivalente à heterozigose em dados diplóides, foi estimada considerando a frequência dos haplótipos como definida por Nei (1987). No caso dos genomas haplóides, h representa a probabilidade de que dois haplótipos escolhidos aleatoriamente sejam diferentes na amostra. As fórmulas da diversidade haplotípica (h) e seu desvio padrão (*DP*) são:

$$h = \frac{n}{n-1} \left(1 - \sum_{i=1}^{k} p_i^2 \right)$$

$$DP = \left\{ \frac{2}{n(n-1)} \left\{ 2(n-2) \left[\sum p_i^3 - \left(\sum p_i^2 \right)^2 \right] + \sum p_i^2 - \left(\sum p_i^2 \right)^2 \right\} \right\}^{\frac{1}{2}}$$

Em que:

p_i é a frequência do haplótipo *i*; *n* é o número de cromossomos da amostra; *k* é o número de haplótipos.

Os cálculos foram realizados com o programa ARLEQUIN (Schneider et al., 2000).

Análise de Variância Molecular (AMOVA)

O Programa ARLEQUIN® (Schneider *et al.*, 2000) foi usado para estimar a diferenciação genética das populações pela análise de variância molecular (AMOVA), o que permite análises hierárquicas de três componentes da variância genética, ou seja, aquela devido a diferenças genéticas entre indivíduos dentro das populações (Φ_{ST}), entre populações dentro dos grupos (Φ_{SC}) e entre grupos (Φ_{CT}). O teste de significância dos valores de variância genética foi estimado com o uso de 10.000 permutações.

Diversidade genética

A análise de diversidade genética foi estimada usando os parâmetros H_{K} , $H_{T'}$ e $G_{ST'}$ (Nei, 1987) usando o programa FSTAT 2.8 (Goudet, 1999).

Em que:

 $H_{\rm K}$ = heterozigose média dentro das populações, ou diversidade gênica; $H_{\rm T}$ ' = heterozigose média total, ou diversidade gênica total e independe do número de populações; $G_{\rm ST}$ ' = coeficiente de diferenciação genética entre as populações e independe do número de populações.

Estes parâmetros foram calculados para os cinco *loci* nas populações estudadas. A diversidade gênica também foi calculada para cada *locus* em cada população.

A significância do parâmetro G_{ST} foi estimada pelo teste do chi-quadrado, segundo Workman e Niswander (1970):

$$\chi^2 = 2NF_{ST}(k-1)_{com}(k-1)(s-1)_{graus de liberdade.}$$

Em que:

N é o número total de indivíduos amostrados,
k é o número de alelos, e
s é o número de subpopulações analisadas para o locus gênico.

A estimativa não-viciada da diversidade genética (H_K) por *locus* e por população foi obtida de acordo com a equação 7.39 proposta por Nei (1987), utilizando-se o programa FSTAT 2.8 (Goudet, 1999):

$$H_{sk} = \frac{n_k}{n_k - 1} (1 - \sum p_{ik}^2 - H_{ok} / 2n_k)$$

Em que:

n_k é o tamanho da k-ésima amostra,
p_{ik} é a frequência do alelo i na amostra k, e
H_{ok} é a proporção de heterozigotos observada na amostra da população k.

A diversidade gênica média (\overline{H}_S) com o respectivo erro padrão foi calculada para cada amostra populacional utilizando o programa DISPAN® (Ota, 1993), conforme a equação 8.6 apresentada em NEI (1987):

$$H = \sum_{j=i}^{r} h_j / r$$

Em que:

r é o número de loci utilizados;

 h_{j} , de acordo com a equação 8.1 de Nei (1987), é a heterozigose esperada para cada locus na j^{ésima} população, estimada por:

$$h = 1 - \sum_{i=1}^{m} x_i^2$$

Em que:

m é o número de alelos.

Esta medida de heterozigose média é equivalente à proporção média de heterozigotos por *locus* em uma população com padrão de acasalamento aleatório e, também, é igual à proporção de *loci* heterozigotos em um indivíduo escolhido aleatoriamente. O desvio padrão desta estimativa é descrito pela seguinte equação adaptada (Nei, 1987):

$$H = \left[\sum_{j=1}^{r} h_j - H\right]^2 / (r-1)r \left[\frac{1}{2}\right]^{1/2}$$

O parâmetro *Fst* (θ_W) foi uma das ferramentas utilizadas para se estimar a diversidade genética entre duas ou mais populações. Os cálculos, feitos por meio do programa Arlequin versão 2.000 (Schneider *et al.*, 2000), utilizando-se método baseado na correlação entre alelos (Weir e Cockerham, 1984). Embora os resultados não sejam exatamente os mesmos, o F_{ST} também pode ser definido em termos de probabilidade de identidade por descendência entre pares de alelos (Nei 1987):

$$F_{ST} = \frac{\left(J_{S} - J_{T}\right)}{\left(1 - J_{T}\right)}$$

onde, considerando-se uma população total dividida em k sub populações S, sendo p_{ki} a frequência do alelo i na sub população k e p_i a frequência do alelo i na

população total, temos: $J_s = \frac{\sum\limits_k J_k}{S} = \frac{\sum\limits_k \left(\sum\limits_i p_{ki}^2\right)}{S}$ correspondente à probabilidade média de que dois alelos amostrados ao acaso em uma mesma sub população sejam idênticos por descendência e $J_T = \sum\limits_i p_i^2$ à probabilidade de que dois alelos amostrados ao acaso na população total sejam idênticos por descendência (Nei 1987).

Análises de Componente Principal

O programa MVSP® (Multivariate Statistical Package for Windows, version 3.1; http://www.kovcomp.com/mvsp/) foi utilizado para a obtenção de Análises de Componente Principal relacionando as doze amostras do presente estudo com populações asiáticas (Lin *et al.*, 1995; Sawada *et al.*, 2002; Liang *et al.*, 2009) européias (Goodeve *et al.*, 1996; Sánchez-García *et al.*, 2005) americanas (Windsor *et al.*, 1994; Soares *et al.*, 2001; Gallegos *et al.*, 2004) e africana (Dangerfield *et al.*, 1997) disponíveis na literatura (Tabelas A1.1-4). As análises foram realizadas utilizando dados dos STRs IVS 1, IVS 13, IVS 22 e IVS 25.3.

Estimativas da composição ancestral

As estimativas de composição ancestral foram obtidas segundo o método de identidade gênica (Chakraborty, 1985) e foram realizadas com o uso do programa ADMIX® 3.

A estimativa foi realizada utilizando as frequências das três populações parentais, admitindo-se um modelo triíbrido de miscigenação.

A análise de composição ancestral foi realizada com os dados dos STRs IVS 13 e IVS 22, pois somente esses possuíam frequências, na literatura, das populações parentais européia e africana. Os europeus foram representados por uma amostra de Caucasianos, uma de Ingleses, uma de Italianos e uma de Espanhóis (Lalloz *et al.*, 1994; Goodeve *et al.*, 1996; Sánchez-García *et al.*, 2005). Como parental africana foi utilizada uma amostra Sul-Africana (Dangerfield *et al.*, 1997) e como parentais ameríndios foram utilizados os dados produzidos neste trabalho.

Nos casos em que houve disponibilidade de mais de uma estimativa de frequências para uma mesma população, foram utilizadas frequências médias, ponderadas pelos tamanhos amostrais.

Dendrogramas

Considerando-se o tipo de marcador utilizado (STRs) e as populações híbridas aqui analisadas, o método UPGMA (Unweighted Pair Group Method with Arithmetic Mean - Sokal e Michener 1958) foi considerado mais adequado para a inferência do dendrograma aqui apresentado, pois reflete de maneira adequada a informação de similaridade existente, uma vez que pares de populações são agrupados sucessivamente, de acordo com a similaridade entre elas. Visto que as populações aqui estudadas não se originaram sob a perspectiva de ramificação filogenética (não apresentando uma população ancestral recente), mas sim pela miscigenação de indivíduos de origens étnicas distintas em proporções bastante variáveis, o dendrograma é apresentado apenas com o objetivo de se analisar a similaridade entre as populações estudadas, não visando a obtenção de relações de parentesco e ancestralidade entre elas.

O cálculo de distâncias genéticas (Reynolds *et al.*, 1983) e a construção de dendrogramas UPGMA foram feitos utilizando-se 4 programas que fazem parte do pacote PHYLIP versão 3.5c (Felsenstein, 1989): SEQBOOT, GENDIST, NEIGHBOR E CONSENSE.

O programa SEQBOOT foi utilizado para construir um banco de dados contendo mil matrizes distintas obtidas a partir do banco de dados original (contendo frequências alélicas), pelo procedimento de *bootstrap* (Felsenstein, 1985).

Com a utilização do programa GENDIST, foram calculadas as distâncias genéticas entre as populações, para cada um das 1.000 matrizes distintas (devido a grande quantidade de matrizes, estes dados não são apresentados). Para cada uma destas matrizes, foram construídos dendrogramas distintos pelo método UPGMA, com o uso do programa NEIGHBOR. Em seguida, os mil dendrogramas obtidos foram analisados pelo programa Consense, obtendo-se um dendrograma "ótimo", com maior suporte estatístico. Por fim este dendrograma consenso foi visualizado utilizando-se o programa TreeView (Page, 1996), versão Win32.

Parâmetros Forenses

Poder de Discriminação (PD)

O poder de discriminação fornece a probabilidade de que dois indivíduos escolhidos ao acaso, dentro de uma população, tenham fenótipos diferentes (Desmarais *et al.*, 1998), e é calculado da seguinte forma:

PDf: Poder de discriminação para uma amostra de origem feminina.

$$PD_M = 1 - 2(\sum_i p_i^2)^2 + \sum_i p_i^4$$

PDm: Poder de discriminação para uma amostra de origem masculina.

$$PD_{H} = 1 - \sum_{i} p_{i}^{2}$$

Probabilidade Média de Exclusão (MEC)

A Probabilidade Média de Exclusão fornece a probabilidade de excluir um indivíduo erroneamente acusado em uma investigação de paternidade (Desmarais *et al.,* 1998), e é calculada da seguinte forma:

MEC Trio – PE₍₁₎: Probabilidade média de exclusão para testes de parentesco de uma menina, quando o trio inteiro é testado (mãe, filho e suposto pai).

$$PE_{(1)} = 1 - \sum_{i} p_{i}^{2} + \sum_{i} p_{i}^{4} - (\sum_{i} p_{i}^{2})^{2}$$

MEC Duo – PE₍₂₎: Probabilidade média de exclusão para testes de parentesco de uma menina, quando apenas a criança e o suposto pai são testados.

$$PE_{(2)} = 1 - 2\sum_{i} p_{i}^{2} + \sum_{i} p_{i}^{3}$$

RESULTADOS

Frequências alélicas, equilíbrio de Hardy-Weinberg e heterozigose

São aqui apresentadas as frequências alélicas, heterozigose e probabilidade de aderência ao Equilíbrio de Hardy-Weinberg para cada um dos cinco microssatélites sob estudo. (Tabelas 4, 5 e 6). Os dados relativos a dois desses marcadores (IVS 13 e IVS 22) já foram anteriormente estimados (Massaro, 2005) para três (São Paulo, Pernambuco e Rio Grande do Sul) das 12 amostras estudadas neste trabalho, mas são reapresentados para facilitar a análise do haplótipo.

Tabela 4. Distribuição das frequências alélicas (%) e parâmetros de diversidade genética dos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, de acordo com o sexo ((M: mulheres, H: homens, T: amostra total) em três amostras de populações urbanas brasileiras (SP: São Paulo, RS: Rio Grande do Sul, PE: Pernambuco) e uma amostra de Hemofilicos A(HA). Os valores significativos antes da correção de Bonferroni estão destacados em negrito e os valores significativos após a correção ($\alpha_{Bonf.} < 0,0025$) estão indicados com *.

стр ^а		SP			RS			PE			HA	
JIK	М	н	т	М	н	т	М	Н	т	М	н	т
IVS 1												
10								0,7	0,5			
12		1,3	0,4	1,0	4,7	2,9		0,7	0,5			
13					1,9	1,0						
14	0,6		0,4					0,7	0,5			
15	0,6		0,4		0,9	0,5	1,8	0,7	1,0			
16	1,9	1,3	1,7	8,2	4,7	6,3	8,9	9,3	9,2	10,0	9,7	9,8
17	51,4	53,1	52,1	62,3	59,8	61,1	48,3	50,6	50,0	65,0	63,8	64,2
18	28,7	17,7	25,0	24,5	23,4	23,8	19,6	22,7	21,8	10,0	15,3	14,1
19	8,7	19,0	12,1	2,0	3,7	2,9	10,7	5,3	6,8	5,0	4,2	4,3
20	5,6	6,3	5,8	2,0	0,9	1,5	10,7	8,0	8,7	10,0	5,6	6,5
21	2,5	1,3	2,1					1,3	1,0		1,4	1,1
n ^b	160	79	239	98	107	205	56	150	206	20	72	92
p °	0,219			0,472			0,003			0,511		
H_k^{d}	0,647	0,650	0,650	0,550	0,591	0,566	0,714	0,678	0,685	0,572	0,563	0,560
IVS 13												
16				1,0	0,9	1,0						
18	1,3		0,8					2,7	1,9	5,0	1,4	2,2

		SP			RS			PF			НΔ	
STR ^a	м	H	т	м	н	т	м	н	т	м	Н	т
19	3.1	2.5	2.9	4.1	7.5	5.8	5.4	7.3	6.8	5.0	8.3	7.6
20	43,7	38,0	42,1	56,1	53,3	54,4	39,3	48,5	46,2	65.0	45,9	50,0
21	25,0	22,8	24,2	28,6	26,2	27,7	30,4	14,7	18,9	10,0	22,2	19,6
22	12,5	12,7	12,5	5,1	6,5	5,8	7,1	10,7	9,7	5,0	5,6	5,4
23	6,9	13,9	9,2	5,1	5,6	5,3	8,9	6,7	7,3	5,0	9,7	8,7
24	5,6	7,6	6,2				8,9	8,7	8,7	5,0	6,9	6,5
25	1,3	2,5	1,7					0,7	0,5			
26	0,6		0,4									
n ^b	160	79	239	98	107	205	56	150	206	20	72	92
p°	0,093			0,075			0,003			0,274		
H_k^{d}	0,726	0,770	0,739	0,603	0,644	0,621	0,749	0,717	0,728	0,589	0,727	0,699
IVS 22												
18								0,7	0,5			
22	2,5	2,5	2,5					1,3	1,0			
23	6,9	6,3	6,7		1,0	0,5		3,3	2,4	10,0	4,2	5,4
24	8,7	7,6	8,3	5,1	3,8	4,5	10,7	10,0	10,2	10,0	2,8	4,3
25	36,9	40,5	37,9	37,8	34,3	35,8	50,1	31,3	36,4	20,0	23,6	22,8
26	39,4	41,8	40,4	51,0	57,1	54,7	32,1	50,0	45,1	50,0	63,8	61,0
27	5,0	1,3	3,8	3,1	3,8	3,0	7,1	2,7	3,9	5,0	4,2	4,3
28				2,0		1,0		0,7	0,5	5,0	1,4	2,2
29	0,6		0,4	1,0		0,5						
n ^b	160	79	239	98	105	203	56	150	206	20	72	92
р°	0,556			0,631			0,079			0,922		
H_k^{d}	0,698	0,658	0,682	0,599	0,555	0,573	0,643	0,644	0,654	0,717	0,540	0,577
IVS 25.3												
12				1,0		0,5						
15	0,6		0,4									
16	6,3	10,3	7,6	3,1	1,0	2,0	12,5	13,3	13,1		8,3	6,5
17	5,1	3,8	4,7	1,0	3,8	2,5	1,8	2,0	1,9	10,0	4,2	5,4
18	65,8	57,7	63,1	68,4	63,5	67,0	60,7	58,7	59,2	50,0	68,0	64,2
19	8,2	14,1	10,2	11,2	12,5	11,0	14,3	10,7	11,7	25,0	2,8	7,6
20	12,7	11,5	12,3	14,3	15,4	15,0	7,1	12,0	10,7	15,0	13,9	14,1
21	1,3	1,3	1,3	1,0	3,8	2,0	3,6	2,0	2,4		1,4	1,1
22		1,3	0,4					1,3	1,0		1,4	1,1
n ^b	158	78	236	98	104	202	56	150	206	20	72	92
p°	0,358			0,403			0,019			0,239		
H_k^{d}	0,541	0,631	0,570	0,503	0,544	0,518	0,603	0,615	0,609	0,700	0,515	0,563

етр а		SP			RS			PE			HA	
SIK	М	н	т	М	Н	т	М	н	т	М	н	т
IKBKG												
25								1,3	1,0			
27	1,3		0,8	1,0		0,5		1,3	0,5		1,4	1,1
28	5,6	7,6	6,2	10,2	3,7	6,8		8,7	6,3	10,0	6,9	7,6
29	88,1	82,3	86,3	81,7	93,5	87,8	98,2	88,0	91,2	90,0	88,9	89,1
30	5,0	10,1	6,7	6,1	2,8	4,4	1,8	0,7	1,0		2,8	2,2
33				1,0		0,5						
n ^b	160	79	239	98	107	205	56	150	206	20	72	92
р ^с	0,162			0,084			0,000*			1,000		
H_k^{d}	0,219	0,309	0,249	0,323	0,127	0,223	0,036	0,208	0,164	0,189	0,207	0,201
H _S [€]	0,566	0,605	0,577	0,515	0,494	0,500	0,546	0,575	0,568	0,552	0,510	0,519
Erro Padrã o	0,092	0,078	0,087	0,051	0,093	0,071	0,130	0,091	0,103	0,095	0,084	0,083

^a = Loci/Alelos/Parâmetros.

^b n = Número de cromossomos X

^c *p* = Probabilidade de aderência às proporções do Equilíbrio de Hardy-Weinberg

^d H_k = Heterozigose esperada

^e H_S = Heterozigose média esperada

Tabela 5. Distribuição das frequências alélicas (%) e parâmetros de diversidade genética dos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, de acordo com o sexo ((M: mulheres, H: homens, T: amostra total) em três amostras de quilombos (MI: Mimbó, SV: Sítio Velho, GA: Gaucinha). Os valores significativos antes da correção de Bonferroni estão destacados em negrito e os valores significativos após a correção ($\alpha_{Bonf.} < 0,003$) estão indicados com *.

етр а		МІ			SV			GA	
SIK	М	н	т	м	н	т	м	н	т
IVS 1									
9	1,9	3,4	2.5				10,0	14,3	10.5
10							6,7		5.3
12	3,8		2.5				6,7		5.3
14	1,9		1.3						
15				2,0		1.4			
16	5,8	3,4	5.0	6,0	4,5	5.6	6,7		5.3
17	32,7	45,0	37.4	34,0	36,4	34.7	13,3	57,1	23.7
18	21,2	24,1	22.5	22,0	18,2	20.8	40,0	14,3	34.1
19	11,5	17,2	13,8	24,0	18,2	22.2	13,3	14,3	13.2
20	13,5	6,9	10.0	12,0	18,2	13.9	3,3		2.6
21	7,7		5.0		4,5	1.4			
n ^b	52	29	81	50	22	72	30	7	37
р ^с	0,005			0,623			0,022		
H_k^{d}	0,825	0,717	0,786	0,776	0,809	0,775	0,817	0,625	0,817
IVS 13									
19	26,9	16,7	23.2	17,3	13,6	16.2	16,7	14,3	15.8
20	11,5	16,7	13.4	13,5	13,6	13.5	20,0	14,3	18.4
21	30,8	20,0	26.7	17,3	22,8	18.9	26,7	14,3	23.7
22	7,7	6,7	7.3	21,2	18,2	20.1	20,0	28,5	23.6
23	5,8	16,7	9.8	13,5	9,1	12.2	13,3	14,3	13.2
24	13,5	19,9	15.9	11,5	13,6	12.2	3,3	14,3	5.3
25	3,8	3,3	3.7	1,9	9,1	4,1			
26				1,9		1.4			
27				1,9		1.4			
n ^b	52	30	82	52	22	74	30	7	37
р ^с	0,096			0,309			0,056		
H_k^d	0,809	0,867	0,827	0,865	0,882	0,860	0,833	0,875	0,832
IVS 22									
19								14,3	2,6
22	5,8		3,7						
23	28,8	20,0	25,6	13,5		9,5	13,3		10,5

етр а		МІ			sv			GA	
SIK	М	н	т	м	н	т	м	н	т
24	15,4	13,3	14,6	25,0	27,3	25,7	10,0	28,6	13,2
25	36,6	56,7	43.9	50,0	54,5	51,2	56,7	28,6	50,0
26	9,6	10,0	9,8	9,6	18,2	12,2	10,0	28,5	15,8
27				1,9		1,4	6,7		5,3
28									
29	3,8		2,4				3,3		2,6
n ^b	52	30	82	52	22	74	30	7	37
р ^с	0,064			0,429			0,349		
H_k^{d}	0,763	0,640	0,720	0,674	0,623	0,656	0,660	0,792	0,712
IVS 25.3									
14				2,0		1,4			
16	17,3	24,1	18,8	24,0	19,0	22,9	3,6		2,9
17		6,9	2,5	6,0		4,3	7,1	16,7	5,9
18	25,0	34,5	28,7	34,0	57,2	40,0	53,6	33,3	53,0
19	9,6	6,9	8,7	22,0	19,0	21,4	14,3	33,3	17,6
20	48,1	27,6	41,3	12,0	4,8	10,0	14,3	16,7	14,7
21							7,1		5,9
n ^b	52	29	81	50	21	71	28	6	34
р°	0,047			0,251			0,794		
H_k^{d}	0,683	0,764	0,714	0,778	0,639	0,741	0,684	0,750	0,678
IKBKG									
25				7,7		5,4			
28	7,7	10,0	7,3	3,8	4,5	4,1	3,3		2,6
29	90,4	90,0	91,5	86,6	95,5	89,1	90,0	57,1	81,6
30	1,9		1,2	1,9		1,4	6,7	28,6	13,2
31								14,3	2,6
n ^b	52	30	82	52	22	74	30	7	37
р ^с	1,000			0,003			0,036		
H_k^d	0,180	0,129	0,160	0,251	0,091	0,204	0,195	0,625	0,326
H _s ^e	0,649	0,635	0,640	0,666	0,610	0,647	0,634	0,812	0,672
Erro Padrão	0,120	0,117	0,122	0,109	0,138	0,116	0,116	0,053	0,092

^a = *Loci*/Alelos/Parâmetros.

^{*b*} n =Número de cromossomos X

^c *p* = Probabilidade de aderência às proporções do Equilíbrio de Hardy-Weinberg

 $^{d}H_{k}$ = Heterozigose esperada

^e H_S = Heterozigose média esperada

етр а		UM			FE			VE			BA			KA	
31K	М	н	т	М	н	т	М	н	т	М	н	т	М	н	т
IVS 1															
10										6,2	9,5	7.4		16,7	3.3
13		4,8	1.4												
16				5,0	15,8	8.6	9,1	37,5	16.7	4,2		2.9			
17	38,0	28,6	34.3	45,0	21,1	36.2	50,0	25,0	43.3	62,5	47,6	57.4	79,2	33,3	70.0
18	34,0	42,8	37.1	32,5	31,6	32.8	27,3	37,5	30.0	27,1	28,6	27.9	20,8	50,0	26.7
19	18,0	19,0	18.6	17,5	31,5	22.4	13,6		10.0		14,3	4.4			
20	10,0	4,8	8.6												
n ^b	50	21	71	40	19	59	22	8	30	48	21	69	24	6	30
p°	0,179			0,199			0,001*			0,077			0,032		
H_k^d	0,713	0,728	0,713	0,679	0,771	0,719	0,705	0,750	0,721	0,543	0,711	0,594	0,356	0,750	0,462
IVS 13															
18										2,1		1.4			
19	10,0	9,1	9.7	10,0		6.9	13,6	37,5	10.0	14,6	4,5	11.4	33,3	50,0	36.7
20	44,0	50,0	45.8	45,0	47,4	46.6	45,5	62,5	43.3	52,1	22,7	43,0	54,2	33,3	50.0
21	40,0	40,9	40.3	40,0	47,3	41.4	40,9		46.7	14,6	54,7	27.1	12,5	16,7	13.3
22				2,5		1.7				8,3	13,6	10.0			
23	4,0		2.8	2,5	5,3	3.4				8,3	4,5	7.1			
26	2,0		1.4												
n ^b	50	22	72	40	19	59	22	8	30	48	22	70	24	6	30

Tabela 6. Distribuição das frequências alélicas (%) dos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, de acordo com o sexo (M: mulheres, H: homens, T: amostra total) em três amostras de tribos indígenas da Amazônia brasileira, Tikúna (UM: Umariaçu, FE: Feijoal, VE: Vendaval), Baníwa (BA) e Kashinawa (KA). Os valores significativos antes da correção de Bonferroni estão destacados em negrito e os valores significativos após a correção ($\alpha_{Bonf.} < 0,002$) estão indicados com *.

0--- ³		UM			FE			VE		-	BA			KA	
STR "	М	н	т	М	н	т	м	н	т	М	н	т	м	н	т
p°	0,247			0,355			0,015			0,744			0,2171		
<i>H</i> _k ^d	0,649	0,595	0,626	0,642	0,604	0,618	0,650	0,583	0,614	0,688	0,664	0,726	0,606	0,750	0,619
IVS 22															
24	3,8		2,7								4,5	1,4			
25	59,7	31,8	51,4	32,6		25,8	54,5	37,5	50,0	22,9	27,3	24,3	45,5	50,0	41,6
26	34,6	63,7	43,2	52,2	10,0	56,0	36,4	50,0	40,0	52,1	45,5	50,0	36,4		33,3
27	1,9	4,5	2,7	10,9	65,0	15,2	9,1	12,5	10,0	18,8	18,2	18,6	9,1	50,0	16,7
28				4,3	25,0	3,0				6,2	4,5	5,7	4,5		4,2
29													4,5		4,2
n ^b	52	22	74	46	20	66	22	8	30	48	22	70	22	4	24
р°	0,019			0,001*			0,000*			0,435			0,128		
H_k^d	0,538	0,509	0,557	0,628	0,528	0,608	0,618	0,667	0,614	0,653	0,714	0,663	0,691	-	0,727
IVS 25.3															
17	5,8		4,2							6,2	4,5	5,7	18,2		14,3
18	40,4	57,1	44,4	50,0	50,0	50,0	68,8	25,0	54,1	48,0	36,4	44,3	9,1	33,3	14,3
19	5,8	4,8	5,6	9,1	10,0	9,4		12,5	4,2	6,2	18,2	10,0	13,6	16,7	14,3
20	46,1	38,1	44,4	34,1	40,0	35,9	31,2	62,5	41,7	39,6	40,9	40,0	59,1	50,0	57,1
21	1,9		1,4	6,8		4,7									
22															
n ^b	52	21	73	44	20	64	16	8	24	48	22	70	22	6	28
р°	0,031			0,005			1,000			0,042			0,079		
H_k^d	0,632	0,556	0,610	0,635	0,600	0,618	0,464	0,583	0,557	0,620	0,695	0,640	0,623	0,833	0,640
IKBKG															

ето а		UM			FE			VE			BA			KA	
JIK	М	н	т	м	н	т	М	н	т	м	н	т	М	Н	т
28										4,2		2,9			
29	96,0	100,0	97,2	100,0	100,0	100,0	100,0	62,5	90,0	95,8	95,5	95,7	91,7	100,0	93,3
30	4,0		2,8					37,5	10,0		4,5	1,4	8,3		6,7
n ^b	50	22	72	46	20	66	22	8	30	48	22	70	24	6	30
p°	0,019									1,000			0,044		
H_k^d	0,080	0,000	0,056	0,000	0,000	0,000	0,000	0,500	0,186	0,082	0,091	0,084	0,167	0,000	0,133
Hs ^e	0,520	0,480	0,512	0,515	0,499	0,511	0,473	0,621	0,531	0,516	0,570	0,541	0,481	0,0573	0,510
Erro Padrão	0,114	0,125	0,117	0,129	0,131	0,130	0,124	0,042	0,090	0,111	0,120	0,116	0,100	0,144	0,104

^a = *Loci*/Alelos/Parâmetros.

^{*b*} n =Número de cromossomos X

^c p = Probabilidade de aderência às proporções do Equilíbrio de Hardy-Weinberg

 d *H*_k = Heterozigose esperada

 e H_S = Heterozigose média esperada

Dado o grande número de testes independentes, torna-se necessário realizar a correção de Bonferroni (Hartl & Clark, 2007). Por este procedimento, o limite α de 5% foi recalculado pelo número de testes: Urbanas (20), de onde se derivou o valor de $\alpha_{Bonf.} = 0,0025$; quilombos (15), de onde se derivou o valor de $\alpha_{Bonf.} = 0,002$ (Tabelas 4, 5 e 6).

É o primeiro relato dos alelos *IVS 1*9* (MI, GA), *IVS 1*10* (PE, GA, BA, KA), *IVS 1*21* (SP, PE, HA, MI, SV) e *IVS 25.3*12* (RS), como também das frequências alélicas do microssatélite IKBKG.

Todas as 12 amostras populacionais compartilharam pelo menos um dos dois alelos mais frequentes para os cinco *loci* analisados: IVS 1 (*17, *18), IVS 13 (*20, *21), IVS 22 (*25, *26), IVS 25.3 (*18, *20) e IKBKG (*29, *28 ou *30). O *locus* IKBKG mostrou-se monomórfico na subpopulação de mulheres das aldeias Feijoal e Vendaval e na de homens das aldeias Umariaçu, Feijoal e Kashináwa (Tabela 6).

Alguns alelos ocorreram de forma restrita, isto é, em apenas uma ou duas amostras: *IVS 1*9* em MI e GA, *IVS 1*13* em RS e UM, *IVS 13*16* em RS, *IVS 13*27* em SV, *IVS 22*18* PE e BA, *IVS 22*19* em GA, *IVS 25.3*12* em RS, *IVS 25.3*14* em SV e *IKBKG *25* em PE e SV, *IKBKG*31* em GA e *IKBKG*33* em RS.

As frequências genotípicas foram comparadas aos valores esperados pelo equilíbrio de Hardy-Weinberg pelo Teste Exato (Tabelas 4, 5, 6).

O STR IVS 1 mostrou desvio significativo (*P*<0,05) em relação às proporções esperadas pelo equilíbrio de Hardy-Weinberg nas amostras de PE, MI, GA, VE e KA; o IVS 13, em PE e VE; o IVS 22, na tribo Tikúna (considerando as três aldeias conjuntamente); o IVS 25.3 em PE, MI, UM, FE e BA e o IKBKG em PE, SV, GA, na tribo Tikúna e KA. O teste sensível à deficiência de heterozigotos indica

desequilíbrio determinado por déficit de heterozigotos (*P*<0,05) destes microssatélites, nestas amostras.

A heterozigose média (H_s nas Tabelas 4, 5, 6) foi similar para homens e mulheres, exceto em GA (homens: 0,812 e mulheres: 0,634) e Vendaval (homens: 0,621 e mulheres: 0,473).

Desequilíbrio de ligação entre loci

A análise de associação entre todos os pares de microssatélites intrônicos (excluindo, portanto, o IKBKG), considerando homens e mulheres, separada ou conjuntamente, demonstrou associação altamente significativa (P<0,001) em todas as comparações, exceto IVS 22 x IVS 25.3 (P=0,02).

O *locus* IKBKG, extragênico, está em equilíbrio com todos os outros marcadores na amostra de mulheres. Na amostra de homens está em desequilíbrio com o IVS 22, mas em equilíbrio com os outros três, e na amostra total está em desequilíbrio com o IVS 1 e IVS 22 e em equilíbrio com o IVS 13 e IVS 25.3.

Diferenciação genética das populações

A comparação par a par das frequências alélicas entre as doze amostras populacionais estimada pelo teste exato de Fisher (Tabela 7), considerando todos os *loci*, revelou diferença significativa (P<0,05) em 48 das 66 comparações possíveis na amostra de homens (27 das 48, altamente significativas, P<0,001) e em 59 das 66 comparações na amostra de mulheres (46 delas com P<0,001).

O subgrupo de mulheres pode ser comparado também quanto às suas frequências genotípicas: 50 das 66 comparações apresentaram diferenças significativas (P<0,05), 38 das quais com P<0,001.
	Gêr	nica	Genotípica		Gêı	nica	Genotípica
	н	М	М		н	М	М
SP x PE	*			HA x UM	*	***	**
SP x RS	***	***	***	HA x FE	***	*	*
SP x HA	*			HA x VE	*	*	
SP x MI	*	***	***	HA x BA	*	*	
SP x SV		***	***	HA x KA	*	*	*
SP x GA		***	***	MI x SV		***	*
SP x UM	**	***	***	MI x GA	*	**	*
SP x FE	***	***	***	MI x UM	***	***	***
SP x VE	*	*		MI x FE	***	***	***
SP x BA	*	***	***	MI x VE	***	***	***
SP x KA	**	***	***	MI x BA	***	***	***
PE x RS	***	***	*	MI x KA	*	***	***
PE x HA		*		SV x GA		*	
PE x MI	***	***	***	SV x UM	***	***	***
PE x SV		***	**	SV x FE	***	***	***
PE x GA	*	***	*	SV x VE	***	***	***
PE x UM	*	***	***	SV x BA	**	***	***
PE x FE	***	***	*	SV x KA	*	***	***
PE x VE	**			GA x UM	***	***	***
PE x BA	***	***	***	GA x FE	***	***	***
PE x KA	*	***	***	GA x VE		*	
RS x HA	*			GA x BA		***	***
RS x MI	***	***	***	GA x KA		***	***
RS x SV	***	***	***	UM x FE		*	
RS x GA	***	***	***	UM x VE			
RS x UM		***	***	UM x BA		***	**
RS x FE	***	**	*	UM x KA	*	***	*
RS x VE	*			FE x VE			
RS x BA	*	***	***	FE x BA		*	*
RS x KA	*	***	***	FE x KA	*	***	**
HA x MI	***	***	***	VE x BA		*	
HA x SV	***	***	*	VE x KA		**	*
HA x GA	*	*	*	BA x KA		*	

Tabela 7. Diferenciação gênica e genotípica entre os pares possíveis das doze amostras populacionais analisadas, baseada nos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, de acordo com o sexo (H: homens, M: mulheres).

(***) valores significativos após a correção de Bonferroni ($\alpha_{Bonf.}$ < 0,0007), (**) p \leq 0,001 e (*) 0,001 0,05.

Diversidade haplotípica

A análise haplotípica, realizada na amostra de homens, considerando apenas os indivíduos cujo DNA amplificou em todos os *loci*, definiu 210 haplótipos diferentes (Tabela 8), sendo que 49 foram compartilhados por pelo menos duas populações, enquanto 22 ocorreram exclusivamente em São Paulo, 52 em Pernambuco, 22 no Rio Grande do Sul, 21 em Hemofílicos A, 11 em Mimbó, 10 em SV, 4 em Gaucinha, 2 em Umariaçu, 4 em Feijoal, 4 em Vendaval, 7 em Baníwa e 2 em Kashináwa.

Dos 46 haplótipos encontrados na população de SP, 12 acorreram 44 vezes e 34 foram de ocorrência única, como também, de 83 em PE, 16 ocorreram 83 vezes e 67 foram de ocorrência única. No RS dos 50 haplótipos encontrados, 12 acorreram 65 vezes e 38 foram de ocorrência única. HA, de 45, oito ocorreram 34 vezes e 38 foram de ocorrência única. MI, de 22, três ocorreram nove vezes e 19 foram de ocorrência única. Em SV, de 19, dois ocorreram quatro vezes e 17 foram de ocorrência única. GA apresentou seis haplótipos diferentes em seis indivíduos. Em UM, de 10, três ocorreram 13 vezes e sete foram de ocorrência única. Feijoal apresentou 14 haplótipos, dois ocorreram seis vezes e 12 foram de ocorrência única. Vendaval apresentou oito haplótipos de ocorrência única. Baníwa, dos 17 haplótipos, três ocorreram sete vezes e 14 foram de ocorrência única. Kashináwa apresentou quatro haplótipos, um ocorreu duas vezes e dois foram de ocorrência única.

			STR	s/Alelos							N	lúmero de c	romossomo	s X				
Haplótipo	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	PE	RS	НА	МІ	sv	GA	UM	FE	VE	ВА	KA
2	17	20	26	18	29	96	15	32	28	15				5	1			
13	17	21	26	18	29	24	3	4	6	6		2				1	2	
3	18	21	25	20	29	17	1	1	6	1	3			1	1	1	2	
39	17	19	26	18	29	15	1	6	4	3	1							
1	17	21	25	20	29	14	3	3	4	1	1			1			1	
36	18	20	26	18	29	14	1	3	2	1				5	2			
42	20	24	25	16	29	12	1	6		2	1	2						
58	16	20	26	18	29	12		6	3	2					1			
30	17	20	26	19	29	11	2	5	1	1					1	1		
20	18	23	25	18	29	9	3		3	1	1	1						
35	18	21	26	20	29	8	1	1		1				1		1	3	
27	17	21	25	19	29	7	3	2	1		1							
33	18	20	25	20	29	7	2	1	1	1								2
17	19	24	25	16	29	6	1	1			4							
101	17	20	25	18	29	6		1	2	2		1						
12	17	22	25	19	29	5	1		3	1								
18	19	22	24	18	29	5	3		1			1						
135	17	20	27	18	29	5			1	2					1		1	
140	19	21	25	20	29	5			1					3	1			
193	19	21	26	20	29	5								1	4			
14	19	23	25	18	29	4	4											
26	17	20	26	18	28	4	2		1	1								
34	17	20	26	18	30	4	2		1				1					
45	18	21	26	18	29	4	1	1	1						1			
50	18	20	25	18	29	4		3				1						
66	18	22	25	20	29	4		3									1	
113	18	21	25	21	29	4		1	2	1								

Tabela 8. Frequência absoluta dos haplótipos do cromossomo X observados em homens nas amostras de São Paulo (SP), Pernambuco (PE), Rio Grande do Sul (RS), Hemofilicos A (HA), Mimbó (MI), Sítio Velho (SV), Gaucinha (GA), Umariaçu (UM), Feijoal (FE), Vendaval (VE), Baníwa (BA) e Kashináwa (KA).

			STR	s/Alelos							N	úmero de c	romossomo	os X				
Haplótipo	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	PE	RS	НА	МІ	SV	GA	UM	FE	VE	ВА	KA
23	17	22	25	16	29	3	2	1										
51	18	20	26	20	29	3		1		2								
57	18	22	24	18	29	3		2					1					
77	17	21	26	19	29	3		2									1	
91	17	18	26	18	29	3		2		1								
93	17	23	25	16	29	3		1		1	1							
112	19	21	24	18	29	3		1		1		1						
151	17	21	26	20	29	3				1					1		1	
6	19	21	25	21	29	2	1		1									
16	17	20	23	22	29	2	1			1								
19	21	24	25	16	29	2	1	1										
25	18	21	25	18	29	2	1		1									
37	18	21	25	19	29	2	1							1				
60	17	20	26	20	29	2		2										
63	18	21	24	18	29	2		1				1						
79	20	23	25	16	29	2		2										
86	10	23	25	18	29	2		1									1	
90	18	19	26	18	29	2		1			1							
109	17	20	28	18	29	2		1		1								
117	18	23	26	18	29	2			1	1								
120	12	20	26	18	29	2			2									
126	17	19	26	19	29	2			1		1							
133	17	22	26	19	29	2			1			1						
137	17	20	27	19	29	2			1								1	
141	17	20	25	20	29	2			1		1							
170	17	20	23	17	29	2					2							
82	16	20	26	19	29	2		1	1									
4	12	21	26	19	29	1	1											
5	19	22	24	17	30	1	1											
7	19	23	26	18	30	1	1											

			SIR	s/Alelos							N	lúmero de cr	omossomo	s X				
Haplótipo	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	PE	RS	НА	м	sv	GA	UM	FE	VE	BA	KA
8	17	21	25	20	30	1	1											
9	18	24	25	18	30	1	1											
10	19	23	24	19	28	1	1											
11	17	25	25	16	29	1	1											
15	17	20	26	16	29	1	1											
21	17	23	26	16	28	1	1											
22	20	22	25	19	29	1	1											
24	20	20	23	17	29	1	1											
28	19	22	22	18	28	1	1											
29	19	24	22	18	29	1	1											
31	17	21	27	18	28	1	1											
32	19	20	25	18	30	1	1											
38	17	25	23	19	29	1	1											
40	20	23	25	18	29	1	1											
41	18	20	23	17	29	1	1											
43	16	19	24	18	29	1	1											
44	17	22	23	20	29	1	1											
46	18	20	26	18	30	1	1											
47	18	19	24	19	29	1		1										
48	19	25	23	18	28	1		1										
49	18	23	25	19	29	1		1										
52	17	22	25	17	29	1		1										
53	17	18	25	18	29	1		1										
54	19	20	25	21	29	1		1										
55	20	20	25	18	29	1		1										
56	17	24	22	16	29	1		1										
59	16	20	26	20	29	1		1										
61	18	22	24	16	29	1		1										
62	18	22	22	18	28	1		1										
64	18	21	25	20	28	1		1										

			STR	s/Alelos							N	lúmero de cro	omossomos	s X				
Haplótipo	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	PE	RS	НА	мі	SV	GA	UM	FE	VE	ВА	KA
65	17	19	24	18	29	1		1										
67	17	20	27	17	30	1		1										
68	16	23	25	16	29	1		1										
69	16	23	24	22	28	1		1										
70	17	21	24	20	28	1		1										
71	14	23	24	18	28	1		1										
72	17	20	18	18	29	1		1										
73	16	22	25	16	29	1		1										
74	17	24	26	18	29	1		1										
75	19	20	25	18	29	1		1										
76	12	24	25	16	29	1		1										
78	20	20	26	18	29	1		1										
80	17	20	25	21	29	1		1										
81	15	23	23	18	28	1		1										
83	20	24	25	16	28	1		1										
84	16	20	24	18	28	1		1										
85	17	20	27	16	29	1		1										
87	17	23	25	19	29	1		1										
88	17	22	25	18	29	1		1										
89	19	22	25	18	28	1		1										
92	16	21	27	18	27	1		1										
94	19	22	25	19	29	1		1										
95	18	20	24	18	25	1		1										
96	18	21	26	20	27	1		1										
97	18	22	25	19	29	1		1										
98	18	18	24	18	29	1		1										
99	16	22	28	22	29	1		1										
100	20	21	24	18	29	1		1										
102	18	20	23	20	28	1		1										
103	18	19	24	18	25	1		1										

			STR	s/Alelos							N	úmero de c	romossomo	os X				
Haplótipo	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	PE	RS	HA	МІ	SV	GA	UM	FE	VE	ВА	KA
104	17	20	25	17	29	1		1										
105	21	20	25	16	29	1		1										
106	18	22	26	18	29	1		1										
107	19	20	26	18	29	1		1										
108	18	24	23	19	29	1		1										
110	17	19	26	18	28	1		1										
111	18	20	25	20	28	1		1										
114	18	20	27	18	29	1		1										
115	12	20	27	18	29	1			1									
116	19	20	25	20	29	1			1									
118	18	20	26	17	29	1			1									
119	18	21	26	19	29	1			1									
121	12	20	25	18	29	1			1									
122	15	20	27	18	30	1			1									
123	12	20	25	21	29	1			1									
124	17	21	24	19	29	1			1									
125	18	21	25	17	29	1			1									
127	13	21	23	19	29	1			1									
128	18	23	25	20	29	1			1									
129	18	22	26	20	29	1			1									
130	17	21	24	18	29	1			1									
131	20	20	25	17	29	1			1									
132	18	20	26	18	28	1			1									
134	17	16	26	16	29	1			1									
136	18	23	25	18	30	1			1									
138	17	20	26	19	28	1			1									
139	17	20	24	17	29	1			1									
142	16	22	25	18	29	1			1									
143	17	19	25	18	29	1			1									
144	17	21	25	18	30	1				1								

			STR	s/Alelos							N	úmero de cro	mossomos	х				
Haplótipo	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	PE	RS	НА	мі	sv	GA	UM	FE	VE	ВА	KA
145	17	19	26	16	29	1				1								
146	16	19	26	18	29	1				1								
147	16	20	25	18	28	1				1								
148	16	22	26	18	29	1				1								
149	19	24	25	18	29	1				1								
150	18	24	26	16	29	1				1								
152	20	24	26	16	29	1				1								
153	17	23	26	18	29	1				1								
154	19	22	25	17	29	1				1								
155	17	22	26	18	29	1				1								
156	17	21	25	17	29	1				1								
157	17	20	27	18	27	1				1								
158	16	21	26	18	28	1				1								
159	16	20	26	18	28	1				1								
160	17	23	25	18	29	1				1								
161	17	20	23	20	28	1				1								
162	17	23	23	20	29	1				1								
163	21	19	24	20	29	1				1								
164	18	21	26	18	30	1				1								
165	20	23	26	17	29	1				1								
166	17	22	25	20	29	1					1							
167	17	20	24	18	28	1					1							
168	17	21	24	20	29	1					1							
169	16	24	25	16	26	1					1							
171	18	25	25	18	29	1					1							
172	17	19	23	18	29	1					1							
173	9	23	24	18	29	1					1							
174	20	19	23	20	29	1					1							
175	18	20	23	18	29	1					1							
176	17	23	24	18	29	1					1							

			STR	s/Alelos							N	úmero de cr	omossomo	s X				
Haplótipo	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	PE	RS	HA	МІ	SV	GA	UM	FE	VE	BA	KA
177	18	22	26	16	29	1						1						
178	21	23	25	18	29	1						1						
179	20	19	24	18	29	1						1						
180	17	21	25	18	29	1						1						
181	19	19	25	19	28	1						1						
182	19	22	25	18	29	1						1						
183	17	25	24	19	29	1						1						
184	17	19	25	16	29	1						1						
185	16	25	24	20	29	1						1						
186	17	20	25	19	29	1						1						
187	17	19	19	19	31	1							1					
188	19	22	25	20	29	1							1					
189	17	21	26	17	29	1							1					
190	9	24	25	19	29	1							1					
191	20	21	25	20	29	1								1				
192	13	20	26	18	29	1								1				
194	18	23	27	18	29	1									1			
195	19	21	27	18	29	1									1			
196	16	20	27	19	29	1									1			
197	16	20	27	20	29	1									1			
198	18	21	26	20	30	1										1		
199	16	21	25	20	30	1										1		
200	16	20	27	18	30	1										1		
201	10	20	25	20	29	1										1	4	
202	19	19	25	18	29	1											1	
203	19	21 20	24	20	29 20	1											1	
204	19	20 21	20 27	10	29 20	1											1	
203	17	21	21	10	29	1											1	
200	10	20	20	19	3U 20	1											1	
207	19	22	20	10	29	1											I	

			STR	s/Alelos							N	úmero de cr	omossomo	s X				
Haplótipo	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	PE	RS	HA	МІ	SV	GA	UM	FE	VE	BA	KA
208	17	22	27	19	29	1											1	
209	17	19	27	18	29	1												1
210	10	19	27	18	29	1												1
	N	lúmero de	indivíduc	s		529	78	150	103	72	28	21	6	20	18	8	21	4
	N	úmero de	haplótipo	os		210	46	83	50	46	22	19	6	10	14	8	17	3
	Γ	Diversida	de Haplo	otípica (<i>h</i>)			0.957± 0.016	0.949± 0.014	0.918± 0.022	0.950± 0.019	0.974± 0.020	0.991± 0.018	1.000± 0.096	0.880± 0.048	0.954± 0.039	1.000± 0.063	0.976± 0.023	0.833± 0.22

A diversidade haplotípica (*h*) estimada com base nos cinco STRs analisados teve valor máximo nas amostras de Gaucinha e Vendaval e mais baixo (0,833) na amostra de Kashináwa.

Análise de Variância Molecular (AMOVA)

A análise de variância molecular (AMOVA) estimada com base em três grupos (1° grupo composto pelas populações urbanas SP, PE, RS e HA; 2° grupo pelos quilombos MI, SV e GA; e o 3° grupo pelos Ameríndios TI, BA e KA) mostrou que a maior parte (93,53%) da diversidade haplotípica foi observada dentro das populações.

Diversidade genética

As estimativas de diversidade (Tabela 9) baseadas nas frequências alélicas foram calculadas segundo a metodologia de Nei (1987).

O índice de subdivisão populacional (G_{ST}) variou de 0,014 (IKBKG) na amostra total a 0,168 (IKBKG) na amostra de homens.

O IKBKG apresentou os menores valores de diversidade genética total, sendo que o menor ($H_T = 0,145$) foi encontrado em mulheres; os outros quatro microssatélites apresentaram valores entre 0,663 (IVS 25.3) e 0,764 (IVS 13), ambos em homens. A H_T não apresentou diferença significativa entre homens e mulheres.

Tabela 9. Estim	nadores de variabilidade segu	undo a metodologia de l	Nei (1973,1975) para
os STRs em h	omens, mulheres e total	(homens e mulheres),	nas doze amostras
populacionais Sa	ão Paulo (SP), Pernambuco	(PE), Rio Grande do Su	1 (RS), Hemofilicos A
(HA), Mimbó (MI	I), Sítio Velho (SV), Gaucinha	a (GA), Umariaçu (UM), I	Feijoal (FE), Vendaval
(VE), Baníwa (BA	Á) e Kashináwa (KA).		

	Diversida	de Gênica
Amostra/Loci	$H_{T'}$	G _{ST'}
Homens		
IVS 1	0,720	0,034
IVS 13	0,764	0,055
IVS 22	0,714	0,160
IVS 25.3	0,663	0,040
IKBKG	0,231	0.168
Total	0,618	0,080
Mulheres		
IVS 1	0,686	0,042
IVS 13	0,740	0,053
IVS 22	0.697	0.059
IVS 25.3	0.681	0,087
IKBKG	0,145	0,087
Total	0,590	0,057
Homens/Mulheres		
IVS 1	0,694	0,035
IVS 13	0,749	0,053
IVS 22	0,690	0.066
IVS 25.3	0,673	0,076
IKBKG	0,168	0,014
Total	0,595	0.055

Os resultados de distância genética, entre os pares das doze amostras, baseados nos valores de F_{ST} são apresentados nas Tabelas 10 (homens) e 11 (mulheres).

Das 66 comparações possíveis, entre os pares de amostras da sub população de homens, 34 apresentaram diferenças significativas (13 altamente significativas, P<0,001) e na sub população de mulheres 52 (33 altamente significativas).

Tabela 10. Matriz de distância da sub população de homens baseada nos valores de F_{ST} (abaixo da diagonal) e valores de P (significante a 0,05) com seu respectivo erro padrão (acima da diagonal), entre os pares possíveis das amostras de São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE), pacientes Hemofilicos A (HA), Mimbó (MI), Sítio Velho (SV), Gaucinha (GA), Umariaçu (UM), Feijoal (FE), Vendaval (VE), Baníwa (BA) e Kashináwa (KA), baseado nos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG.

Homens	SP	PE	RS	HA	МІ	sv	GA	UM	FE	VE	ВА	KA
SP			0,024	0,016	0,005			0,017	<u>0,001</u>	0,010	0,010	
PE					<u>0,000*</u>	0,012		0,014	<u>0,001</u>	0,006	<u>0,001</u>	
RS	0,014				<u>0,000*</u>	<u>0,001</u>			0,004	0,009	0,004	0,044
HA	0,017				<u>0,000*</u>	<u>0,000*</u>		0,010	<u>0,001</u>	0,002	<u>0,000*</u>	0,021
МІ	0,041	<u>0,064*</u>	<u>0,105*</u>	<u>0,107*</u>				0,002	<u>0,000*</u>	0,011	0,009	
SV		0,039	<u>0,080</u>	<u>0,084*</u>				0,002	<u>0,000*</u>	<u>0,001</u>	0,010	
GA												
UM	0,043	0,040		0,051	0,105	0,102						
FE	<u>0,071</u>	<u>0,072</u>	0,074	<u>0,082</u>	<u>0,122*</u>	<u>0,119*</u>						
VE	0,090	0,107	0,119	0,133	0,101	<u>0,131</u>						
BA	0,043	<u>0,065</u>	0,062	<u>0,078*</u>	0,055	0,059						
KA			0,112	0,135						0,079	0,057	

(*) valores significativos após a correção de Bonferroni ($\alpha_{Bonf.} < 0,0007$), (itálico e sublinhado) p< 0,001 e (negrito) 0,001 < p > 0,05.

Tabela 11. Matriz de distância da sub população de mulheres baseada nos valores de F_{ST} (abaixo da diagonal) e valores de P (significante a 0,05) com seu respectivo erro padrão (acima da diagonal), entre os pares possíveis das amostras de São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE), pacientes Hemofilicos A (HA), Mimbó (MI), Sítio Velho (SV), Gaucinha (GA), Umariaçu (UM), Feijoal (FE), Vendaval (VE), Baníwa (BA) e Kashináwa (KA), baseado nos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG

Mulheres	SP	PE	RS	HA	МІ	SV	GA	UM	FE	VE	BA	KA
SP			0,047		<u>0,000*</u>	<u>0,000*</u>	<u>0,000*</u>	<u>0,000*</u>	0,026		<u>0,001</u>	<u>0,000*</u>
PE			0,009	0,037	<u>0,000*</u>	<u>0,000*</u>	<u>0,000*</u>	0,003	0,038		<u>0,000*</u>	<u>0,000*</u>
RS	0,007	0,022			<u>0,000*</u>	<u>0,000*</u>	<u>0,000*</u>	<u>0,000*</u>	0,011		0,004	<u>0,000*</u>
HA		0,031			<u>0,000*</u>	<u>0,000*</u>	<u>0,000*</u>	0,002	0,040			0,006
МІ	<u>0,093*</u>	<u>0,082*</u>	<u>0,140*</u>	<u>0,111*</u>		0,003	<u>0,001</u>	<u>0,000*</u>	<u>0,000*</u>	0,002	<u>0,000*</u>	<u>0,000*</u>
SV	<u>0,067*</u>	<u>0,042*</u>	<u>0,119*</u>	<u>0,094*</u>	0,026			<u>0,000*</u>	<u>0,000*</u>	<u>0,001</u>	<u>0,000*</u>	<u>0,000*</u>
GA	<u>0,052*</u>	<u>0,037*</u>	<u>0,104*</u>	<u>0,112*</u>	<u>0,041</u>			0,005	<u>0,000*</u>	0,036	<u>0,000*</u>	<u>0,000*</u>
UM	<u>0,043*</u>	0,038	<u>0,064*</u>	0,085	<u>0,058*</u>	<u>0,074*</u>	0,048				<u>0,000*</u>	0,003
FE	0,015	0,020	0,027	0,038	<u>0,077*</u>	<u>0,088*</u>	<u>0,065*</u>					<u>0,000*</u>
VE					0,060	<u>0,061</u>	0,038					
BA	<u>0,031</u>	<u>0,054*</u>	0,029		<u>0,106*</u>	<u>0,119*</u>	<u>0,112*</u>	<u>0,056*</u>				
KA	<u>0,104*</u>	<u>0,117*</u>	<u>0,110*</u>	0,072	<u>0,092*</u>	<u>0,128*</u>	<u>0,156*</u>	0,073	<u>0,091*</u>			

(*) valores significativos após a correção de Bonferroni ($\alpha_{Bonf.} < 0,0007$), (itálico e sublinhado) p< 0,001 e (negrito) 0,001 < p > 0,05.

Análises de Componente Principal

A análise de componente principal (ACP) baseada nas frequências alélicas dos três grupos aqui estudados (Ameríndios, quilombos e populações urbanas) e de um conjunto de amostras africana, européias e asiáticas (Ver Material e Métodos) está representada na Figura 12, cujos dados estão listados na Tabela 12. O *locus* IKBKG não foi incluído por não existir dados na literatura sobre frequências nas populações aqui consideradas como parentais.

As populações urbanas e os ameríndios estão agrupados entre si e com uma amostra da Espanha e próximas de uma amostra da China. As amostras de quilombos estão agrupadas entre si e no mesmo quadrante que México, Canadá e África.

Componente	Eigenvalue	Proporção explicada (%)	Proporção cumulativa explicada (%)
Eixo 1	1957,261	40,732	40,732
Eixo 2	921,652	19,18	59,912
Eixo 3	691,588	14,392	74,305
Eixo 4	351,861	7,322	81,627
Eixo 5	236,749	4,927	86,554
Eixo 6	163,286	3,398	89,952
Eixo 7	147,97	3,079	93,032
Eixo 8	102,753	2,138	95,15

Tabela 12. Características dos componentes principais baseados nas frequências alélicas de quatro STRs (IVS 1, IVS 13, IVS 22, IVS 25.3).



Figura 12. Componentes principais obtidos a partir do IVS 1, IVS 13, IVS 22 e IVS 25.3 e que relacionam as frequências alélicas das populações analisadas: Urbanas (SP: São Paulo, PE: Pernambuco, RS: Rio Grande do Sul e HA: Hemofilicos A), Ameríndios (Tikúna: Umariaçu-UM, Feijoal-FE e Vendaval-VE; Baníwa-BA e Kashináwa-KA) e Quilombos (Mimbó-MI, Sítio Velho-SV e Gaucinha-GA) com populações mundiais (Referências Material e Métodos).

Estimativas de composição ancestral

A Tabela 13 apresenta a contribuição dos componentes ameríndio, africano e

europeu nas amostras estudadas.

		Não po	onderada	
Componentes	SP	PE	RS	HA
Africano	0.2897+/- 1.#INF	0.2291 +/- 1.#INF	-0.0032 +/- 1.#INF	- 0.0197 +/- 1.#INF
Ameríndio	0.2172 +/- 1.#INF	0.2435 +/- 1.#INF	0.1233 +/- 1.#INF	1.1349 +/- 1.#INF
Europeu	0.4931 +/- 0.0000	0.5274 +/- 0.0000	0.8798 +/- 0.0000	- 0.1152 +/- 0.0000
R2	0,997424	0,99245	0,978477	0,693651

Tabela 13. Contribuição dos componentes africano, europeu e ameríndio obtidos a partir de dois STRs (IVS 13 e IVS 22), para as amostras São Paulo (SP), Pernambuco (PE), Rio Grande do Sul (RS), Hemofilicos A (HA).

As estimativas de mistura realizadas com apenas dois dos STRs (IVS 13 e IVS 22) foram consistentes com o modelo triíbrido e indicam uma contribuição européia preponderante nas amostras de SP, PE e RS, sendo que em RS atingiu o maior valor, no entanto apresentou valor negativo para o componente africano. As variações com relação aos valores de contribuição ameríndia esperados estão condizentes com o tipo de marcador (matrilinear).

Dendrograma

O dendrograma (Figura 13), baseados nas frequências alélicas, relaciona as 12 amostras analisadas entre si. Ele demonstra a formação de três ramos principais. O primeiro ramo agrupa as amostras de quilombos; o segundo, as tribos Tikúna e Baníwa; e o terceiro, as amostras urbanas. A tribo Kashináwa encontra-se separada de todas as outras amostras.



Figura 13. Dendrograma UPGMA, obtidos pelos cálculos de distância de Reynolds, relacionando as amostras de São Paulo (SP), Pernambuco (PE), Rio Grande do Sul (RS), Hemofilicos A (HA), Mimbó (MI), Sítio Velho (SV), Gaucinha (GA). Umariaçu (UM), Feijoal (FE), Vendaval (VE), Baníwa (BA) e Kashináwa (KA). Os números correspondem os valores de Bootstrap.

Parâmetros Forenses

Para verificar o efeito das diferenças nas frequências gênicas dos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG sobre os parâmetros para aplicações forenses, foram estimados, para a amostra total e separadamente para homens e mulheres: o Poder de Discriminação (PD) e a Probabilidade Média de Exclusão (MEC) (Desmarais *et al.*, 1998) (Tabela 14).

O Cálculo do Poder de Discriminação (PD) indica a probabilidade de dois indivíduos não aparentados apresentarem fenótipos diferentes. PDf (Poder de discriminação de uma amostra de origem feminina) mostrou os maiores valores nas amostras de quilombos, variando de 95,30% em SV, no IVS 13 (Amostra Total) a 74,70% em GA, no IKBKG (Amostra de Homens). Os menores valores foram encontrados em ameríndios, variando de 77,00% em HA (Amostra de homens) a 0% em FE (Amostra de homens, mulheres e total). O PDm (Poder de discriminação de uma amostra de origem do sexo masculino) também mostrou os maiores valores na amostra de quilombos, variando de 84,80% em SV, no IVS 13 (Amostra total) a 57,20% em GA, no IKBKG (Amostra de homens). Os menores foram encontrados em ameríndios, variando de 84,80% em SV, no IVS 13 (Amostra total) a 57,20% em GA, no IKBKG (Amostra de homens). Os menores foram encontrados em ameríndios, variando de 49,10% em UM, no IVS 22 (Amostra de homens) a 0% em FE, no IKBKG (Amostra de homens, mulheres e total). Todos os quatro marcadores intrônicos apresentaram valores altos de PD, para a maioria das doze amostras. O microssatélite IKBKG apresentou os menores valores.

A Probabilidade Média de Exclusão (MEC) indica a capacidade de um determinado marcador genético excluir um indivíduo erroneamente indicado como pai em uma investigação de paternidade. Este parâmetro pode ser calculado considerando a MEC para testes de parentesco de uma menina, quando o trio inteiro é testado (mãe, filho e suposto pai) (MEC trio) e a MEC quando apenas a criança e suposto pai são testados (MEC duo). Como encontrado nos cálculos de PD, os maiores valores foram obtidos entre as amostras de quilombos e os menores nas amostras de ameríndios. Todos os microssatélites intrônicos apresentaram valores superiores a 0,5 na maioria das amostras. O IKBKG mais uma vez apresentou os menores valores.

		Hon	nens			Mulł	Mulheres				Total			
	PDf	PDm	MEC trio	MEC duo	PDf	PDm	MEC trio	MEC duo	PDf	PDm	MEC trio	MEC duo		
IVS 1														
SP	0,831	0,646	0,603	0,455	0,820	0,642	0,590	0,444	0,829	0,647	0,601	0,454		
PE	0,856	0,674	0,636	0,491	0,873	0,697	0,662	0,518	0,862	0,682	0,645	0,500		
RS	0,780	0,581	0,537	0,390	0,739	0,544	0,491	0,346	0,763	0,564	0,517	0,370		
HA	0,770	0,555	0,523	0,375	0,765	0,545	0,517	0,368	0,769	0,552	0,522	0,373		
MI	0,869	0,703	0,660	0,516	0,938	0,805	0,781	0,660	0,921	0,774	0,746	0,616		
SV	0,910	0,764	0,729	0,595	0,904	0,760	0,722	0,586	0,908	0,764	0,728	0,593		
GA	0,807	0,613	0,570	0,420	0,930	0,780	0,758	0,631	0,929	0,790	0,763	0,637		
UM	0,855	0,694	0,642	0,498	0,853	0,698	0,642	0,496	0,857	0,703	0,648	0,504		
FE	0,878	0,731	0,682	0,539	0,820	0,659	0,595	0,448	0,856	0,704	0,647	0,502		
VE	0,807	0,656	0,582	0,434	0,822	0,649	0,594	0,446	0,845	0,685	0,629	0,483		
BA	0,830	0,662	0,607	0,459	0,717	0,530	0,468	0,325	0,769	0,584	0,526	0,380		
KA	0,773	0,611	0,536	0,389	0,496	0,330	0,275	0,165	0,613	0,438	0,365	0,237		
IVS 13														
SP	0,910	0,761	0,728	0,594	0,886	0,722	0,685	0,545	0,895	0,735	0,700	0,562		
PE	0,892	0,714	0,688	0,547	0,886	0,729	0,689	0,549	0,894	0,724	0,694	0,555		
RS	0,818	0,634	0,586	0,439	0,780	0,595	0,539	0,393	0,801	0,618	0,565	0,419		
HA	0,885	0,715	0,682	0,541	0,783	0,555	0,535	0,386	0,873	0,691	0,659	0,516		
MI	0,949	0,831	0,808	0,693	0,927	0,791	0,762	0,635	0,941	0,815	0,790	0,670		
SV	0,955	0,843	0,823	0,713	0,956	0,844	0,825	0,715	0,958	0,848	0,829	0,720		
GA	0,941	0,815	0,792	0,671	0,931	0,802	0,773	0,646	0,935	0,809	0,781	0,657		
UM	0,728	0,574	0,484	0,344	0,796	0,634	0,564	0,419	0,778	0,617	0,542	0,397		

Tabela 14. Valores de Parâmetros de utilidade forense para o IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG, nas doze amostras de populações brasileiras (SP, PE, RS, HA, MI, SV, GA, UM, FE, VE, BA, KA) em homens, mulheres e total (homens e mulheres).

		Hon	nens			Mulł	neres			Total			
	PDf	PDm	MEC trio	MEC duo	PDf	PDm	MEC trio	MEC duo	PDf	PDm	MEC trio	MEC duo	
FE	0,693	0,549	0,446	0,311	0,787	0,626	0,553	0,409	0,765	0,605	0,526	0,383	
VE	0,608	0,469	0,359	0,235	0,763	0,607	0,524	0,380	0,738	0,584	0,495	0,353	
BA	0,814	0,627	0,580	0,432	0,859	0,672	0,639	0,493	0,876	0,713	0,671	0,529	
KA	0,773	0,611	0,536	0,390	0,745	0,580	0,502	0,358	0,758	0,598	0,517	0,372	
IVS 22													
SP	0,814	0,651	0,586	0,442	0,854	0,693	0,642	0,499	0,842	0,680	0,624	0,48 ⁻	
PE	0,813	0,640	0,583	0,437	0,799	0,630	0,566	0,419	0,816	0,652	0,589	0,444	
RS	0,721	0,553	0,474	0,333	0,757	0,593	0,515	0,373	0,735	0,570	0,490	0,349	
HA	0,732	0,533	0,483	0,339	0,866	0,685	0,650	0,505	0,769	0,569	0,524	0,377	
MI	0,802	0,611	0,565	0,415	0,896	0,745	0,706	0,568	0,873	0,709	0,666	0,523	
SV	0,767	0,595	0,526	0,380	0,835	0,660	0,611	0,463	0,825	0,648	0,597	0,449	
GA	0,880	0,735	0,685	0,542	0,838	0,635	0,606	0,457	0,874	0,692	0,661	0,517	
UM	0,657	0,491	0,407	0,273	0,685	0,522	0,435	0,298	0,696	0,548	0,448	0,312	
FE	0,693	0,505	0,443	0,301	0,778	0,608	0,539	0,393	0,777	0,596	0,536	0,388	
VE	0,752	0,594	0,511	0,367	0,723	0,562	0,476	0,335	0,735	0,580	0,492	0,350	
BA	0,846	0,681	0,629	0,483	0,814	0,637	0,583	0,434	0,826	0,653	0,600	0,452	
KA	0,625	0,500	0,375	0,250	0,813	0,648	0,585	0,440	0,844	0,685	0,628	0,483	
IVS 25.3													
SP	0,825	0,622	0,590	0,441	0,760	0,537	0,511	0,363	0,786	0,568	0,541	0,39 [,]	
PE	0,817	0,611	0,579	0,429	0,798	0,589	0,556	0,407	0,813	0,606	0,575	0,428	
RS	0,766	0,554	0,519	0,371	0,714	0,497	0,464	0,319	0,732	0,515	0,482	0,336	
HA	0,731	0,509	0,481	0,335	0,829	0,655	0,603	0,455	0,774	0,555	0,527	0,378	
MI	0,885	0,737	0,692	0,551	0,837	0,667	0,615	0,467	0,861	0,704	0,653	0,509	
SV	0,787	0,598	0,547	0,398	0,904	0,760	0,722	0,586	0,884	0,730	0,687	0,546	

		Hon	nens		Mulheres				Total				
	PDf	PDm	MEC trio	MEC duo	PDf	PDm	MEC trio	MEC duo	PDf	PDm	MEC trio	MEC duo	
GA	0,872	0,722	0,672	0,528	0,853	0,660	0,629	0,482	0,847	0,659	0,623	0,475	
UM	0,679	0,527	0,430	0,295	0,779	0,617	0,542	0,399	0,759	0,601	0,519	0,377	
FE	0,735	0,580	0,492	0,350	0,789	0,621	0,553	0,407	0,775	0,610	0,537	0,392	
VE	0,717	0,531	0,468	0,324	0,582	0,430	0,337	0,215	0,677	0,532	0,428	0,294	
BA	0,822	0,665	0,600	0,453	0,766	0,605	0,527	0,383	0,791	0,631	0,558	0,413	
KA	0,773	0,611	0,536	0,389	0,789	0,591	0,547	0,397	0,807	0,613	0,570	0,420	
IKBKG													
SP	0,498	0,307	0,285	0,172	0,380	0,218	0,209	0,120	0,420	0,247	0,234	0,137	
PE	0,376	0,218	0,205	0,117	0,069	0,035	0,035	0,018	0,294	0,165	0,157	0,087	
RS	0,228	0,124	0,120	0,065	0,516	0,318	0,299	0,183	0,385	0,223	0,212	0,122	
HA	0,357	0,204	0,195	0,111	0,311	0,180	0,164	0,090	0,349	0,200	0,190	0,107	
MI	0,227	0,125	0,117	0,063	0,312	0,177	0,166	0,092	0,281	0,157	0,148	0,081	
SV	0,161	0,086	0,082	0,043	0,414	0,242	0,231	0,135	0,355	0,201	0,194	0,110	
GA	0,747	0,572	0,502	0,356	0,326	0,184	0,175	0,098	0,506	0,315	0,290	0,176	
UM					0,145	0,077	0,074	0,038	0,104	0,054	0,053	0,027	
FE													
VE	0,608	0,469	0,359	0,234					0,311	0,180	0,164	0,090	
BA	0,161	0,086	0,082	0,043	0,151	0,081	0,077	0,040	0,157	0,083	0,081	0,043	
KA	0,000	0,000	0,000	0,000	0,270	0,152	0,141	0,076	0,227	0,125	0,117	0,063	

Aplicabilidade no aconselhamento genético

Onze dos 72 pacientes HA apresentaram informações familiares que permitissem a elaboração de um heredograma (Figura 14) correspondente. Foi possível determinar o haplótipo que acompanha a mutação em 10 dessas famílias. No único caso em que isto não foi possível, a mãe do paciente era homozigota para todos os marcadores do bloco de haplótipos considerado (Figura 12.C). Em outro caso (Figura 12.A), uma vez que temos um heredograma pequeno com um único paciente afetado, não é possível descartar uma nova mutação no IVS 1 como causa da hemofilia neste paciente em particular. Em um terceiro caso (Figura 12.D), a paternidade alegada foi excluída, mas este fato não interferiu no diagnóstico.





Figura 14. Heredograma e análise do haplótipo formado pelos microssatélites IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG em 11 famílias de Hemofilicos A (A a K). Os Box em negrito representam o haplótipo com a mutação.

DISCUSSÃO

Novos alelos

A ocorrência dos alelos *IVS 1*9*, *IVS 1*10*, *IVS 1*21* e *IVS 25.3*12*, descrita aqui pela primeira vez, foi confirmada pelos procedimentos de clonagem e seqüenciamento. Esta descoberta não surpreende, visto que não existe um grande número de estudos populacionais realizados até o momento (Tabelas A1.1-4).

O alelo *IVS 25.3 *12*, de ocorrência única, teria como causa mais provável a mutação *de novo*. Porém, esta causa, não pode ser atribuída à ocorrência de quatro cópias do alelo *IVS 1 *21*. Este alelo, presente em amostras urbanas e quilombolas, apresenta similaridades entre e dentro dos grupos de indivíduos, com relação à composição alélica do haplótipo formado pelos cinco marcadores analisados aqui. Em Mimbó, por exemplo, este alelo ocorreu em quatro mulheres, sendo que três delas apresentavam praticamente a mesma composição alélica (Tabela A2.1), o que aumenta a chance de apresentarem o mesmo haplótipo. Nas populações urbanas, das oito ocorrências duas apresentaram os mesmos alelos. Isto sugere que uma mutação africana, embora não provável, possa ser considerada, já que não foi encontrado no Rio Grande do Sul (região na qual a composição ancestral africana é pequena), como também em ameríndios. A comparação dos haplótipos possíveis revela que pode-se aceitar um número mínimo de nove haplótipos. Porém não há dados disponíveis na literatuta para comparação e possível confirmação desta suposição.

O alelo *IVS 1 *9*, apesar de ocorrer apenas em quilombos (MI e GA), não pode ser considerado um alelo africano, pois neste caso seria encontrado nas outras amostras populacionais. A hipótese mais aceitável é que o surgimento deste alelo se deve a uma mutação regional "fundadora", comum a esses dois quilombos. A

comparação entre os haplótipos possíveis em que tal alelo ocorre revela que, dos cinco indivíduos portadores do alelo *9, dois são homens e apresentam haplótipos diferentes (Tabela A2.1). Nas três mulheres, o máximo que pode ter ocorrido é que um haplótipo possível em Mimbó seja igual a um dos haplótipos possíveis de Gaucinha. Portanto, a hipótese mais provável é que haja, não mínimo, quatro haplótipos diferentes que seja portador do alelo *IVS 1 *9*.

Foram contadas nove ocorrências do alelo *IVS 1 *10* (Tabela A2.1) em oito indivíduos (um deles, de Gaucinha, homozigoto para todos os cinco marcadores). Considerando que um dos haplótipos é único em KA, outro repetiu-se quatro vezes (uma em PE, uma em BA e duas na mulher homozigota de GA) e que encontram-se dois pares de haplótipos possíveis, conclui-se que cinco seria o número mínimo possível desses haplótipos. Neste caso, a hipótese de que tal haplótipo tenha origem ameríndia não pode ser descartada.

Freqüências alélicas e equilíbrio de Hardy-Weinberg

Há, na literatura, um único relato, parcial, de freqüências alélicas do STR IKBKG (Fang *et al.*, 2006), em que são descritos seis alelos (168-188 pb) e H_k = 56,06. Este é o primeiro relato completo de freqüências deste *locus* (Tabelas 4, 5 e 6). Dentre as amostras do presente trabalho, foram observados sete alelos (*IKBKG *25, *27, *28, *29, *30, *31* e **33*). Os alelos **28, *29, *30* foram compartilhados por todos os grupos de amostras. O alelo **25* (180 pb) ocorreu em PE e SV, o alelo **27* ocorreu nas amostras urbanas, enquanto que os alelos **31* e **33* (196 pb) ocorreram exclusivamente em GA e RS, respectivamente. O maior valor de H_k foi encontrado em Gaucinha (62,5) e o menor em ameríndios (zero).

O que chama a atenção quanto às freqüências gênicas encontradas é a magnitude (34%) do número de testes que revelaram desequilíbrio de H-W (20 comparações significativas em um total de 58 possíveis), todos eles por déficit de heterozigotos.

Com exceção de PE, todas as outras amostras urbanas estão em equilíbrio (Tabela 4) e, com exeção do *locus* IVS 22, todos os outros *loci* estão em desequilíbrio na amostra PE. No total de 20 testes possíveis, quatro apresentaram-se em desequilíbrio, mas após a correção de Bonferroni ($\alpha_{Bonf.}=0,0025$), apenas o IKBKG se manteve em desequilíbrio.

Nas comunidades quilombolas, foram detectados cinco casos de desvio em relação ao equilíbrio (Tabela 5) em um total de 15 testes possíveis: IVS 1 em MI e GA, IVS 25.3 em MI e IKGKB em SV e GA. Não há, portanto, *locus* ou comunidade com excesso de desequilíbrio, o que sugere que tais desvios sejam casuais. Após a correção de Bonferroni ($\alpha_{Bonf.}=0,003$), nenhum desses desequilíbrios se mantiveram.

Entre as amostras indígenas (Tabela 6) também foram encontrados casos de desequilíbrio: 11 em um total de 23 testes possíveis. Aqui também, sua distribuição não sugere concentração por *locus* ou amostra. Após a correção de Bonferroni ($\alpha_{Bonf.}$ =0,002) o IVS 1 na amostra VE e IVS 22 nas amostras FE e VE se mantiveram em desequilíbrio.

No conjunto, após a correção, permanecem quatro amostras (~7%) com pelo menos um locus em desequilíbrio, num total de 58 comparações, indicando um ligeiro excesso. Todavia, o método de Bonferroni tem sido considerado muito conservativo por alguns autores que sugerem que a correção seja sempre apresentada, porém acompanhada de alguma análise ou comentários dos dados sem ela.

A amostra PE, única amostra urbana a apresentar desequilíbrio, foi anteriormente estudada por Leboute (2000) e outras amostras de mesma extração, por (Maurício-

da-Silva (2000) e Dellalibera (2004) que não relataram quaisquer características especiais que pudessem explicar desequilíbrios. O desvio aqui observado pode ser atribuível ao pequeno tamanho efetivo (n = 28) da subpopulação de mulheres nas quais os cálculos presentes foram baseados.

Estudo anterior (Wiezel, 2003) de marcadores de X nas mesmas três comunidades quilombolas evidenciou desvio em MI, nos *loci DXS548* e FRAXAC1, em SV, no *locus DXS548* e em GA, no *locus* FRAXAC1, sendo que o *locus* DXS548 em SV foi o único que apresentou desvio originado por deficit de heterozigotos.

Estudos de outros marcadores anteriormente realizados com as mesmas amostras ameríndias (Mendes Junior, 2001; Luizon *et al.*, 2007) não evidenciaram características que pudesem explicar o desequilíbrio aqui relatado em ameríndios.

O elevado número (34%) de exemplos de desequilíbrio diminui muito, mas permanece significativo (7%) após a correção de Bonferroni.

Vícios estatísticos (amostras e/ou populações pequenas, erros de amostragem, etc.) não podem ser afastados, entretanto, além da alta freqüência de casos de desequilíbrio, chama a atenção o fato de que em todos eles há déficit de heterozigotos.

Isto sugere que, concomitante com os fatores mencionados, haja também um efeito próprio desses marcadores decorrente do fato de serem ligados ao cromossomo X. É sabido, pela própria descrição do teorema do equilíbrio H-W, que cada *locus* autossômico, isoladamente, entra em equilíbrio na primeira geração de uma população em panmixia. Isto, entretanto, não acontece para os *loci* ligados ao cromossomo X. Consequentemente, esses desequilíbrios sugerem que este tipo de marcador pode dar uma contribuição diferente da dos marcadores autossômicos ao estudo da história recente destas populações.

Comparação com populações de outros continentes

Fica evidente a dificuldade em se estabelecer uma tendência continental para distribuição das freqüências alélicas diante dos ainda poucos relatos de freqüências alélicas dos marcadores aqui analisados, conforme Tabelas A1.1-4 em que estão realçados os alelos mais frequentes e os segundos mais freqüentes em amostras populacionais de vários continentes.

No IVS 1 todas as populações: Espanhóis (Tizzano *et al.*, 2005), Coreanos (Kim *et al.*, 2005), Indianos (Saha *et al.*, 2007) e Chineses (Liang *et al.*, 2009) estudadas até agora são diferentes umas das outras em relação ao alelo mais freqüente (*16, *14, *15 e *17, respectivamente, Tabela A1.1), enquanto o alelo *17 é o mais freqüente nas amostras aqui descritas (exceto GA, na subamostra de mulheres e UM, FE, VE e KA, na subamostra de homens, que apresentaram o alelo *18). É dificil abordar uma explicação para essa variabilidade, uma vez que mesmo a hipótese mais provável, diferenças metodológicas na genotipagem, aparentemente não é suficiente para explicá-lo.

A média ponderada (Tabela A1.2) das freqüências de Europeus e de Asiáticos do IVS 13 encontra-se centrada nos alelos *20 e *21, mesmo incluindo amostras discrepantes como das populações da Nova Zelândia [Harraway *et al.*, 2006], China [Yip *et al.*, 1994] e Índia [Saha *et al.*, 2007], onde os alelos mais freqüentes foram: *17 e *18, *25 e *26, e *22-*24, respectivamente. Na única amostra de Africanos (Dangerfield *et al.*, 1997) ocorre uma distribuição homogênea dos alelos *19-*23, sendo que o alelo *22 apresenta freqüência um pouco maior. As amostras de quilombos analisadas aqui apresentam a mesma característica de distribuição ampla, mas alternam entre os alelos *21 e *22 como mais freqüentes. As distribuições alélicas nos Americanos divergem nas três amostras descritas anteriormente, incluindo uma de Brasileiros paulistanos [Soares *et al.*, 2001]. Isto dificulta a comparação entre os dados produzidos aqui, onde os alelos *20 e *21 são os mais freqüentes, com aqueles relatados na literatura.

No IVS 22 (Tabela A1.3), Europeus e Asiáticos apresentam distribuições centradas nos alelos *25 e *26, como também as amostras urbanas e de ameríndios analisadas aqui. Embora a amostra de Nova Zelândia (Harraway *et al.*, 2006) seja claramente diferente das outras, nós não sabemos precisamente a origem dos pacientes examinados. A mesma amostra do subcontinente Indiano que divergiu da média Asiática no IVS 13 também divergiu no IVS 22. As altas freqüências dos alelos *25 e *23 ou *24 nos quilombos coincidem com a de Africanos usada como referência, onde os alelos *24 e *25 são mais freqüentes. Estes alelos também são mais freqüentes em Canadenses (Windsor *et al.*, 1994) e Mexicanos (Gallegos *et al.*, 2004), enquanto na amostra de Brasileiros paulistanos, apesar da freqüência expressiva dos alelos *25 e *26, o alelo mais freqüente é o *27.

No IVS 25.3, o alelo mais freqüente nas amostras urbanas é o *18 e em quilombos e ameríndios os alelos se revesam entre o *18 e o *20. Em algumas amostras analisadas anteriormente (Tabela A1.4), o alelo *18 foi o mais freqüente.

Há como visto, marcantes diferenças de algumas amostras em relação a outras do mesmo continente. É o caso de Coreanos, Indianos e Espanhois no *locus* IVS 1. Também no *locus* IVS 13 observa-se a mesma heterogeneidade em Neozelandeses, Chineses, Indianos e Brasileiros paulistanos, bem como, no *locus* IVS 22, Neozelandeses, Indianos e Brasileiros paulistanos.

Esta heterogeneidade sugere um 'deslocamento' alélico, isto é, os autores teriam usado uma nomenclatura tal que todos os alelos seriam classificados algumas unidades a mais ou a menos. Em apenas alguns desses trabalhos há referência de terem confirmado a genotipagem por seqüenciamento, o que foi feito no presente estudo, em que pelo menos a metade dos alelos foram seqüenciados para servirem de referência.

No IVS 25.3, as amostras MI e KA se diferenciam das demais e esta diferença não pode ter a mesma explicação acima porque foram examinadas segundo a mesma técnica e com os mesmos referenciais. Trata-se, portanto, de uma diferença genuína.

Comparação das freqüências alélicas entre as amostras

A maioria dos alelos encontrados foi compartilhada pelas doze amostras Tabelas 4, 5 e 6. Entre os que ocorreram de forma restrita, destacam-se os alelos *IVS 1*9, IVS 13*27, IVS 22*19, IVS 25.3*14* e *IKBKG*31* que ocorreram apenas em quilombos; e os alelos *IVS 13*16* e *IKBKG*33*, de ocorrência exclusiva no RS. Porém, com exceção do alelo *IVS 1*9*, que apresentou freqüência entre 10% a 14% em GA, os demais alelos apresentaram baixas freqüências o que explica sua ocorrência exclusiva em algumas populações.

No entando, se por um lado há alelos exclusivos, por outro há alelos que, pela sua alta frequência média, teriam grande probabilidade de estarem representados em todas as amostras, mas não são observados em algumas delas. É o caso dos alelos *14 e *21 do *locus* IVS 1, alelos *18, *24 e *25 do *locus* IVS 13 e do alelo *22 do *locus* IVS 22 que exibem alta freqüência em quase todas as amostras, mas estão ausentes da amostra RS. Os alelos *IVS* 1 *14, *IVS* 13 *25 e *IVS* 22 *22 também não aparecem na amostra HA.

Diversidade Genética

Aparentemente, RS e HA possuem semelhanças com relação as frequências alélicas e, dentre as urbanas, ambas apresentam valores de H_K menores quando comparados com as demais. Esta suposição de que elas sejam mais assemelhadas

encontra suporte na análise pelo dendrograma IPGMA (Figura 13) e pelo componente principal (Figura 12), nas quais ambas as amostras estão sempre muito próximas no mesmo grupo.

Nota-se, também, que na amostra RS sempre há menor número de alelos, isto é, em todos os *loci* há alelos que aparecem nas amostras urbanas PE, SP e HA, mas não em RS e HA. Além disso, e também inexplicável, é a H_K em RS sempre significativamente menor do que em SP, PE e praticamente igual, mas quase sempre menor, do que no grupo masculino HA.

Excetuando-se o microssatélite IKBKG e o IVS 22, os maiores valores de H_K foram obtidos nas amostras quilombolas e os ameríndios apresentaram os menores valores. Não há uma diferença bem definida da heterozigose entre os grupos, embora se reconheça uma concentração de altos valores também em quilombos.

O microssatélite IKBKG apresentou os menores valores de H_K , comparado com os demais marcadores, em todos os grupos de amostras (urbanas, quilombos e ameríndios). Esta pouca diversidade não permite conclusões precisas, embora haja uma clara tendência a baixos valores entre os indígenas.

É característico de ameríndios ter uma reduzida diversidade intrapopulacional e haplotípica e um alto nível de diferenciação entre as populações, possivelmente como uma conseqüência do seu isolamento por longo período de tempo (Tarazona-Santos *et al.*, 2001).

Distintas características genéticas são, portanto, esperadas para o cromossomo X em relação aos autossomos principalmente uma menor diversidade genética devido à menor taxa de mutação em mulheres do que em homens e um menor tamanho populacional efetivo (Schaffner, 2004), uma estrutura genética mais pronunciada devido ao menor efeito do tamanho populacional e, conseqüentemente, uma deriva genética mais forte e um

maior desequilíbrio de ligação, uma vez que o cromossomo X recombina apenas na mulher (Pereira *et al.*, 2007).

Desequilíbrio de ligação

A análise de ligação entre todos os pares de microssatélites intragênicos demonstrou associação significativa entre IVS 22 x IVS 25.3 (P=0,02) e altamente significativa (P<0,001) em todas as outras comparações. A alta associação era esperada, uma vez que esses marcadores estão muito próximos entre si.

Por outro lado, apesar da proximidade do *locus* extragênico IKBKG, há equilíbrio dele com todos os outros marcadores na amostra de mulheres. Entretanto, há desequilíbrio com o IVS 22 na amostra de homens e com IVS 22 e IVS 1 na amostra total de homens e mulheres.

De acordo com o teorema de H-W, genes ligados ao cromossomo X demorariam para entrar em equilíbrio e, enquanto isso não ocorre, as freqüências femininas estão sempre mais próximas do equilíbrio do que as masculinas. Desta maneira, os dados sugerem que as amostras aqui estudadas são oriundas de populações ainda subestruturadas.

Ressalta-se, entretanto, que a ausência do DL não invalida o uso desses marcadores no aconselhamento, uma vez que o DL mede a ocorrência de *recombinação* ao longo das gerações. Para o aconselhamento o que importa é a probabilidade de recombinação em uma única meiose.

Diversidade Haplotípica

Ter encontrado haplótipos compartilhados por amostras do mesmo grupo (urbanas, quilombolas e ameríndias), tal como se observa na Tabela 8, não é um evento que cause surpresa. As freqüências de três deles apresentam valores intermediários e muito semelhantes nos três grupos de amostras. É razoável supor-se que se trate de haplótipos antigos, pois têm ocorrência generalizada e as pequenas diferenças de freqüência seriam decorrentes de deriva.

O haplótipo mais freqüente (17/20/26/18/29) ocorreu apenas em amostras urbanas e ameríndias (96 ocorrências em 529 possibilidades = 18%), não tendo sido observado nas amostras quilombolas. Os dados históricos não permitem supor que este haplótipo tenha se expandido dos ameríndios para as populações urbanas neobrasileiras. O sentido inverso também não tem apoio histórico, além do que os ameríndios aqui analisados são os mesmos de estudos anteriores (Mendes-Junior, 2005, Luizon *et al.*, 2007) em que a ancestralidade européia é considerada mínima, no máximo 2%. Portanto, levanta-se a hipótese de que o haplótipo presente nas amostras urbanas teria origem européia, porém sua forma presente nos Amerindios, apesar de ter frequência proporcionalmente elevada (5/20 = 25%), não seria idêntica por descendência, mas decorrente de mutações novas seguidas de fixação recente.

Todas as amostras analisadas apresentaram haplótipos de ocorrência única e exclusiva em número correspondente ao seu tamanho amostral. A hipótese mais provável é que sejam haplótipos fundadores ou originados por mutação *de novo*.

As amostras urbanas e ameríndias compartilharam 15 haplótipos. Com exceção do haplótipo mais freqüente (17/20/26/18/29), que ocorreu em ~22% dos indivíduos urbanos e em ~8% dos ameríndios, todos os demais haplótipos são mais freqüentes em ameríndios.

Também os haplótipos compartilhados entre amostras urbanas e quilombolas, com exceção do mais frequente, apresentam frequência muito menor nas amostras urbanas. A comparação haplotípica dos dados apresentados com os da literatura é possível apenas para o bloco de haplótipos formados pelos microssatélites IVS 13 e IVS 22. O haplótipo mais freqüente deste bloco nas populações urbanas foi 20/26 (35.24%). Em Turcos [El-Maari *et al.*, 1999], os haplótipos mais freqüentes, apresentados em 58 cromossomos normais de diferentes regiões geográficas e 26 pacientes com inversão no IVS 22, formados por quatro marcadores (dois SNPs: IVS 19 *Hind*III e IVS 22 *Xba*I e dois microssatélites: IVS 13 e IVS 22) foram -/+/20/26 e +/-/21/25, respectivamente. No entanto, um terceiro haplótipo (-/-/20/26) foi mais freqüente entre os pacientes (P = 0,034).

A distribuição do haplótipo 20/26 foi semelhante em indivíduos saudáveis e pacientes negros Sul-Africanos (com ou sem a inversão IVS 22) (Dangerfield *et al.*, 1997). A freqüência desse haplótipo em pacientes Sul-Africanos caucasóides, sem a inversão no IVS 22 foi maior que em indivíduos saudáveis (*P*<0,01). A baixa freqüência da inversão no IVS 22 neste grupo, juntamente com a representação desproporcional de um haplótipo em cromossomos sem a inversão, sugere a presença de uma mutação fundadora na Hemofilia A (Dangerfield *et al.*, 1997).

Na análise do bloco de haplótipos formado pelos marcadores do presente trabalho não foi possível observar nenhum haplótipo com maior freqüência em HA que pudesse caracterizar uma associação.

Diferenciação Genética

A diferenciação entre as amostras urbanas, de quilombos e de ameríndios medida pelos valores do parâmetro *Fst* (Tabelas 10 e 11) são concordantes, na maioria das comparações, com as estimativas de diferenciação gênica e genotípica encontradas pelas comparações par a par entre todas estas amostras (Tabela 7).

Independente do marcador utilizado espera-se que a maior parte da diversidade seja intrapopulacional o que está de acordo com os dados obtidos no presente trabalho. Porém, apesar da estimativa interpopulacional ter apresentado valores baixos, estes foram significativos, o que indica uma alta estruturação populacional (Tabelas 9, 10 e 11).

A população brasileira é o resultado de inter-cruzamentos entre europeus (principalmente Portugueses), escravos africanos por eles trazidos e populações indígenas autóctones (Ribeiro, 1995), mas a proporção destes componentes variam entre as diferentes regiões [Ferreira *et al.*, 2006; Muniz *et al.*, 2008]. A região Sul, por exemplo, possui contribuição africana e ameríndia insignificantes. Portanto, a diferenciação alélica e genotípica (Tabela 7) observadas nas populações aqui analisadas pode ser devido à composição ancestral diferente e as diferenças nas freqüências alélicas dessas populações fundadoras. A amostra de pacientes é composta por indivíduos de vários estados brasileiros e não apenas da região de Ribeirão Preto. Sendo assim, algum grau de diferenciação entre HA e as outras amostras urbanas é esperada.

Dendrograma

Apesar de ser pequeno o número de *loci* analisados, como também o tamanho de algumas amostras, o dendrograma UPGMA, obtido pelos cálculos de distância de Reynolds (Figura 12), mostra agrupamento homogêneo, observando-se apenas que KA se colocou próximo ao grupo Ameríndio, mas não exatamente dentro dele.

Obserrva-se também que as tribos Tikúna e Baníwa agruparam-se em um único ramo, sendo que Vendaval ficou mais próximo de Umariaçu, seguido por Feijoal e subseqüentemente por Baníwa. O suporte estatístico não foi muito alto (bootstrap 0,639, 0,484 e 0,481 respectivamente). Os baixos valores de bootstrap exibidos no dendrograma e observados na tribo Tikúna pode ser interpretado como indicativo de homogeneidade da tribo, originada por efeito do fundador e mantido pelo contínuo fluxo gênico entre as aldeias. Kashínawa se encontra em um ramo separado de todas as outras populações. Esta separação das demais tribos pode ser resultado do baixo número amostral.

As populações urbanas agruparam-se em um único ramo, sendo que SP ficou mais próximo de PE, tanto quanto HA de RS. Este resultado não condiz com a localização geográfica dessas amostras e, apesar da baixa diversidade interpopulacional essa separação pode ser resultado da contribuição diferenciada dos componentes ancestrais.

As amostras de quilombos encontram-se no mesmo ramo, porém GA e SV estão mais próximas entre si que MI.

Análise do componente principal

A análise de componentes principais é um procedimento utilizado na simplificação de dados multivariados com mínima perda de informação (Cavalli-Sforza *et al.,* 1994). As populações do presente trabalho foram submetidas a esta análise em conjunto com populações mundiais com o intuito de verificar eventual agrupamento por similaridade das freqüências dos alelos examinados.

Esta análise teve como pontos positivos os agrupamentos entre as amostras de quilombos, urbanas e de ameríndios entre si. Porém, as populações urbanas e de ameríndios dividem a mesma área no gráfico. As amostras de Asiáticos se encontram dispersas, as de americanos no mesmo quadrante das amostras de quilombos analisados aqui e de africanos; e os europeus, excluindo a amostra de Espanhóis, que se agrupou com as amostras urbanas e de ameríndios, as demais se agruparam com asiáticos.
Entretanto, os resultados da análise não permitiram uma conclusão realista e/ou eficiente, uma vez que oito eixos foram necessários para explicar 95% da variância dos *loci* analisados.

Composição ancestral

Dentre os problemas encontrados na análise de composição ancestral, estão a determinação das populações que devem ser consideradas como ancestrais, as diferenças dos períodos de chegada das populações à América e o momento em que as populações são analisadas. Também o número de marcadores e a omissão de estimadores de erros-padrão das frequências devem ser considerados (Sans, 2000). Estas dificuldades foram particularmente presentes neste trabalho, principalmente em decorrência do pouco número de relatos publicados sobre a distribuição populacional destes marcadores.

No presente trabalho, as estimativas de mistura foram realizadas com dados de apenas dois *loci:* IVS 13 e IVS 22 (Tabela 13), o que limita em muito a possibilidade de obtenção de estimativas fidedignas. Além disso, as populações disponíveis na literatura não são as mais adequadas, já que não temos dados de freqüência alélica em Portugueses para estes marcadores, a principal população européia que contribuiu na formação da população brasileira. Além disso, os africanos são representados por uma única amostra da África do Sul. As freqüências parentais em Ameríndios foram representadas pelos dados produzidos aqui em tribos indígenas da Amazônia Brasileira.

Nenhuma estimativa de mistura conta a historia verdadeira da composição ancestral da população e, as presentes estimativas não permitem que façamos inferências fidedignas sobre a formação das populações analisadas. Entretanto, apesar de não apresentarem valores de desvio-padrão confiáveis, as presentes estimativas estão condizentes com a história da formação e contribuição de cada componente ancestral, quando considerado o tipo de marcador, matrilinear.

Parâmetros Forenses

Os marcadores do tipo STR se caracterizam por ter uma maior taxa de mutação que os SNPs e, portanto, um menor LD (Hering *et al.*, 2004). No entanto, devido à peculiar herança do cromossomo X, na qual apenas 2/3 destes se recombinam a cada geração, o LD pode ocorrer entre STRs relativamente distantes, sendo necessário que o grau de ligação entre os marcadores seja avaliado em cada população para sua posterior utilização na prática forense (Robino, 2006).

Os STRs do cromossomo X são capazes de complementar a análise dos marcadores autossômicos e do cromossomo Y de forma muito eficiente, pois se caracterizam por apresentarem Probabilidade média de exclusão (MEC) maior do que os autossômicos (Szibor *et al.*, 2003). A maior capacidade de exclusão dos STR do cromossomo X se dá pela diferença de alelos existente entre os tradicionais marcadores autossômicos e os do cromossomo X para os indivíduos do sexo masculino. Estes apresentam sempre dois alelos para os marcadores autossômicos e, apenas um, para os do cromossomo X. O pai transfere seu perfil haplotípico (100%) do cromossomo X para a filha, o que não ocorre com os autosômicos, nos quais apenas 50% da informação genética paterna é herdada.

Em virtude dessa propriedade do cromossomo X, a maior vantagem da sua aplicação reside nos casos de paternidade, nos quais o material biológico do suposto pai não está disponível e o DNA de seus parentes é analisado para a reconstrução do seu perfil genético (Gomes *et al.*, 2007a). Com a análise dos tradicionais STRs autossômicos, muitas vezes estes casos permanecem inconclusivos, pois para se ter uma boa reconstrução do perfil do suposto pai, vários parentes deste precisam ser tipados.

Gusmão *et al.* (2009), que também estudou marcadores do cromossomo X, observou valores de PD (mulheres entre 0.9999999992 e 0.99999999999998, e homens entre 0.9999997 and 0.9999996) maiores que os observados aqui. No entanto, mesmo que menores, os resultados obtidos aqui mostram que o uso combinado desses cinco STRs é uma útil adição de informação aos STRs autossômicos e do cromossomo Y, melhorando potencialmente a discriminação na identificação forense e teste de parentesco.

Aplicabilidade no aconselhamento genético

Muitas mulheres, parentes de pacientes hemofilicos desejam saber suas reais chances de transmitir a doença, embora algumas possam ser identificadas como portadoras baseando-se apenas na história familiar (filhas de pais hemofilicos).

A utilidade de um polimorfismo no diagnóstico de portadora da doença e / ou prénatal é diretamente proporcional aos seus níveis de heterozigose esperada (H_K).

Para o microssatélite IVS 1, os valores de H_K (Tabelas A1.1-4) variaram de 0,56 (Espanhóis) a 0,66 (Asiáticos). O maior valor de H_K foi encontrado no IVS 13, em Americanos (0,81), onde só os asiáticos (0,58) apresentaram um valor mais baixo. O IVS 22 e o IVS 25.3 apresentaram H_K variando de 0,63 (Africanos) para 0,71 (Americanos) e de 0,38 (Asiáticos) para 0,57 (presente estudo, consideradas as populações urbanas), respectivamente. Estas diferenças entre amostras de populações de todo o mundo demonstram a importância de levantamentos regionais das freqüências alélicas.

O marcador IKBKG apresentou uma baixa taxa de heterozigose, contribuiu para a análise familiar em apenas dois (18,2%) casos e não se apresenta em desequilíbrio de ligação com os marcadores intrônicos.

A caracterização desses cinco marcadores ligados ao gene *F8* mostra, embora o diagnóstico por análise de ligação tenha limitações, a eficácia do uso desses microssatélites no diagnóstico pré-natal e na identificação de portadores na população brasileira.

CONCLUSÕES

Alelos Novos

Quatro alelos novos foram aqui descritos. Um deles (*IVS 25.3 *21*) ocorreu em uma única cópia. Os outros três (*IVS 1 *9*, *IVS 1 *10* e *IVS 1 *21*) foram detectados em seis, nove e treze cromossomos, respectivamente. Pela análise dos haplótipos de que participam, é provável que não tenham sido originados de mutações locais novas e, portanto, representem alelos fundadores.

Desequilíbrio

O exceso de exemplos de desequilíbrio de H-W sugere a ação, provavelmente não isolada, mas concomitante a outros fatores, de efeito próprio desses marcadores, isto é, o fato de serem ligados ao cromossomo X, permitindo supor-se que este tipo de marcador pode dar uma contribuição diferente da dos marcadores autossômicos ao estudo da história recente destas populações.

Distribuição das freqüências por continente

O pequeno número de relatos na bibliografia sobre os *loci* aqui estudados, torna dificil, as vezes impossível, a comparação com outras populações. Além disso, algumas populações têm distribuição alélica claramente discrepante da tendência para o continente a que pertencem. Tendo em vista que os alelos de referência podem ser seqüenciados com relativa facilidade, não se pode admitir a explicação mais simples, que seria a ocorrência um erro técnico de tipagem decorrente do uso de amostras de referência com peso molecular acima ou abaixo dos valores habituais de referência.

Os dados ainda são insuficientes para fundamentar uma hipótese que explique tais achados. Entretanto, deve-se considerar a possibilidade de que seja um efeito provocado pelo tipo de marcadores, isto é, marcadores ligados ao cromossomo X que, pela sua dinâmica populacional única, poderiam mostrar aspectos da migração populacional humana ainda não revelados por marcadores autossômicos, ligados ao cromosso Y ou mitocondriais, cujos relatos são abundantes.

Diversidade Genética

Há um menor número de alelos nas amostras de RS e HA, bem como menor heterozigose, especialmente notável em RS. As três comunidades quilombolas do Piauí (Mimbó, Sítio Velho e Gaucinha) apresentaram, de um modo geral, valores de H_k nitidamente maiores do que as outras amostras.

Desequilíbrio de ligação

A análise de associação revelou equilíbrio de ligação generalizado em mulheres, mas alguns desequilíbrios na amostra total de homens e mulheres, o que sugere que as amostras aqui estudadas são oriundas de populações ainda subestruturadas.

Diversidade haplotípica

Na análise do bloco de haplótipos formado pelos marcadores do presente trabalho não foi possível observar nenhum com maior freqüência em HA que pudesse caracterizar uma associação.

Parâmetros forenses

Os dados obtidos aqui mostram que o uso combinado desses cinco STRs é uma útil adição de informação aos STRs autossômicos e do cromossomo Y, melhorando potencialmente a discriminação na identificação forense e nos testes de parentesco.

Utilidade para o aconselhamento

O marcador IKBKG apresentou uma baixa heterozigose, mas mesmo assim contribui para análises familiares. Portanto, a caracterização destes cinco marcadores ligados ao gene *F8* mostra que, embora o diagnóstico por análise de ligação tenha limitações, estes microssatélites são suficientemente informativos para serem empregados no diagnóstico indireto pré-natal ou na identificação de portadoras na população brasileira.

REFERÊNCIAS BIBLIOGRÁFICAS

- ARPINI-SAMPAIO Z. Estrutura Genética de Três Isolados Afro-Brasileiros (156 p). Tese de Doutorado. Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil, 1998.
- 2. ARPINI-SAMPAIO Z, COSTA M C, MELO A A, CARVALHO M F, DEUS M S, SIMÕES AL. Genetic polymorphisms and ethnic admixture in African-derived black communities of northeastern Brazil. *Hum Biol.* 71: 69-85, 1999.
- 3. ANTONARAKIS SE. Molecular genetics of coagulation factor VIII gene and hemophilia A. *Thromb Haemost.* 74 (1): 322-8, 1995.
- ARRUDA VR, ANNICHINO-BIZZACCHI JM COSTA FF. Detection of hemophilia A gene carriers in a Brazilian population. *Rev Assoc Med Bras.* 39 (3): 126-30, 1993. Portuguese. Erratum in: *Rev Assoc Med Bras.* 39 (4): 259, 1993a.
- 5. ARRUDA VR, ANNICHINO-BIZZACHI JM, SONATI MDE F, COSTA FF. Factor VIII and IX genes polymorphisms in a Brazilian black population. *Thromb Haemost.* 2; 70 (2): 371, 1993b.
- ASEEV M, SURIN V, BABOEV K, GORNOSTAEVA N, KUZNETZOVA T, KASCHEEVA T, IVASCHENKO T, SOLOVYEV G, MIKHAILOV A, LEBEDEV V, *ET AL.* Allele frequencies and molecular diagnosis in haemophilia A and B patients from Russia and from some Asian republics of the former U.S.S.R. *Prenat Diagn.* 14 (7): 513-22, 1994.
- 7. BAGNALL RD, WASEEM N, GREEN PM, GIANNELLI F. Recurrent inversion breaking intron 1of the factor VIII gene is a frequent cause of severe hemophilia A. *Blood.* 99: 168-74, 2002.
- 8. BAGNALL RD, GIANNELLI F, GREEN PM. Int22h-related inversions causing hemophilia A: a novel insight into their origin and a new more discriminant PCR test for their detection. *J Thromb Haemost.* 4 (3): 591-8, 2006.
- 9. BOWEN DJ. Haemophilia A and haemophilia B: molecular insights. *Mol Pathol.* 2002 Apr; 55(2):127-44. Review. Erratum in: *Mol Pathol.* 55 (3): 208, 2002.
- 10. BRINKE A, TAGLIAVACCA L, NAYLOR J, GREEN P, GIANGRANDE P, GIANNELLI F. Two chimaeric transcription units result from an inversion breaking intron 1 of the factor VIII gene and a region reportedly affected by

reciprocal translocations in T-cell leukaemia. Hum Mol Genet. 5 (12): 1945-51, 1996.

- 11. CAIO VM, PAIVA E SILVA RB. MAGNA LA, RAMALHO AS. Community genetics and hemophilia in a Brazilian population. *Caderno Saúde Pública*. Rio de Janeiro. 17: 595-605, 2001.
- 12. CAVALLI-SFORZA L L, MENOZZI P, PIAZZA A. History and geography of human genes. Princeton: *Princeton University Press*, 1994.
- 13. CHAKRABORTY R. Gene identity in racial hybrids and estimation of admixture rates. In Ahuja YR, Neel JV, editors. Genetic diferentiation in human and others animal populations. *Indian Anthropological Association*. 171-180, 1985.
- 14. CHAKRAVARTI A. Population genetics making sense out of sequence. Nat Genet. 21: 56-60, 1999.
- 15. CHOWDHURY MR, HERRMANN FH, SCHRODER W, LAMBERT CT, LALLOZ MR, LAYTON M, *ET AL.* Factor VIII gene polymorphisms in the Asian Indian population. *Haemophilia.* 6 (6): 625-30, 2000.
- 16. COLLINS A. Allelic association: linkage disequilibrium structure and gene mapping. *Mol Biotechnol.* 41 (1): 83-9, 2009.
- 17. CURTIN PD. The Atlantic Slave Trade: A Census. University of Winsconsin Press, Madison, 1969.
- DANGERFIELD BT, MANGA P, FIELD SP, HARTMAN E, JEMKINS T, KRAUSE A. Feasibility of prenatal diagnosis and carrier detection in South African haemophilia A patients. *Br J Haematol.* 97 (3): 558-60, 1997.
- 19. DA SILVA JUNIOR WA, FIGUEIREDO MS. Analysis of factor VIII gene polymorphisms in Brazilian blacks reveals further differences in the black population. *Hum Hered.* 44 (5): 252-60, 1994.
- 20. DARBY SC, KEELING DM, SPOONER RJ *ET AL*. The incidence of factor VIII and factor IX inhibitors in the hemophilia population of the UK and their effect on subsequent mortality, 1977-99. *J Thromb Haemost*. 2: 1047–1054, 2004.
- 21. DE BRASI CD, BOWEN DJ. Molecular characteristics of the intron 22 homologs of the coagulation factor VIII gene: an update. *J Thromb Haemost.* 6 (10): 1822-4, 2008.

- 22. DE CARVALHO FM, DE VARGAS WOLFGRAMM E, PANETO GG, DE PAULA CARETA F, SPAGNOL PERRONE AM, DE PAULA F, LOURO ID. Analysis of Factor VIII polymorphic markers as a means for carrier detection in Brazilian families with haemophilia A. *Haemophilia*. 13 (4): 409-12, 2007.
- 23. DELLALIBERA E, HAVRO ML, SOUZA M, KAJIHARA K, MAURICIO-DA-SILVA L, DOS SANTOS SILVA R. Genetic analysis of 13 STR loci in the population from the State of Pernambuco, northeast Brazil. *Forensic Sci Int.* 10; 146 (1): 57-9, 2004.
- 24 DESMARAIS D, ZHONG Y, CHAKRABORTY Y, PERREAULT C, BUSQUE L. Development of a Highly Polimorphic STR Marker for Identity Testing Purposes at the Human Androgen Receptor Gene (HUMARA). *J Forensic Sci.* 43 (5): 1046-1049, 1998.
- 25. EL-MAARRI O, KAVAKLI K, ÇAGLAYAN H. Intron 22 inversions in the Turkish haemophilia A patients: prevalence and haplotype analysis. *Haemophilia*. 5: 169-73, 1999.
- 26. EYSTER ME, LEWIS JK, SHAPIRO SS, GILL F, KAJANI M, PRAGER D. The Pennsylvania hemophilia program. *Am J Hematol.* 9: 277-86, 1978.
- 27. FANG Y, WANG XF, DAI J, WANG HL. A rapid multifluorescent polymerase chain reaction for genetic counseling in Chinese haemophilia A families. *Haemophilia.* 12: 62-67, 2006.
- 28. FELSENSTEIN J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 39: 783-791, 1985.
- 29. FELSENSTEIN J. PHYLIP-phylogeny inference package (version 3.2). *Cladistics*. 5: 164-166, 1989.
- FERREIRA LB, MENDES-JUNIOR CT, WIEZEL CEV, LUIZON MR, SIMÕES AL. Genomic ancestry of a sample population from the state of São Paulo, Brazil. *Am J Hum Biol.* 18 (5): 702-5, 2006.
- FIGUEIREDO MS, FRANCO RF, ZAGO MA. Detection of Hemophilia A carriers by DNA analysis. Comparison with coagulation tests. *Rev Brasil Genet.* 16 (1), 187-202, 1993.
- 32. GALLEGOS RM, ARANDA HB, NAVARRETE CP, ESPINOZA R, GOMEZ FS, ARANDA DA. Polymorphism distribution of Int13, Int22, and St14 VNTRs in a Mexican population and their application in carrier diagnosis of hemophilia A. *Am J Hematol.* 77 (1): 1-6, 2004.

- GITSCHIER J, WOOD WI, GORALKA TM, WION KL, CHEN EY, EATON DE, VEHAR GA, CAPON DJ, LAWN RM. Caracterization of the human factor VIII gene. *Nature*. 312: 326-330, 1984.
- 34. GOMES I, PRINZ M, PEREIRA R, MEYERS C, MIKULASOVICH RS, AMORIM A, CARRACEDO A, GUSMÃO L. Genetic analysis of three US population groups using an X-chromosomal STR decaplex. *Forensic Sci Int.* 121 (3): 198-203, 2007.
- 35. GOODEVE AC, TAGARIELLO G, CHUANSUMRIT A, PRESTON FE, PEAKE IR. A rapid and cost effective method for analysis of dinucleotide repeat polymorphisms in the factor VIII gene. *Blood Coagul Fibrinolysis.* 7 (7): 672-7, 1996.
- 36. GOUDET J. A program to estimate and test gene diversities and fixation indices FSTAT (version 2.8). Updated from Godet (1995), 1999.
- 37. GUSMÃO L, SÁNCHEZ-DIZ P, ALVES C, GOMES I, ZARRABEITIA MT, ABOVICH M, ET AL. A GEP-ISFG collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations. Int J Legal Med. 123 (3): 227-34, 2009.
- 38. HARRAWAY JR, SMITH MP, GEORGE PM. A highly informative, multiplexed assay for theindirect detection of hemophilia A using five-linked microsatellites. *J Thromb Haemost.* 4 (3): 587-90, 2006.
- 39. HARTL DL, CLARK AG. Principles of population genetics. 4th ed. Sinauer Associates, Inc. Publishers. Sunderland, Massachusetts, 2007.
- 40. HERING S, BRUNDIRS N, KUHLISCH E, EDELMANN J, PLATE I, BENECKE M, VAN PH, MICHAEL M, SZIBOR R. DXS10011: studies on structure, allele distribution in three populations and genetic linkage to further q-telomeric chromosome X markers. *Int J Legal Med.* 118: 313-319, 2004.
- HIGUCHI R. Simple and rapid preparation of samples for PCR. In: Erlich HA, (Ed) PCRtechnology – principles and applications for DNA amplification. New York: Stockton, Press. 31-38, 1989.
- 42. IBGE CENSO 2000. Site do Instituto Brasileiro de Geografia e Estatística. Disponível na Internet via WWW. URL: www.ibge.gov.br. Acesso em setembro de 2009.
- 43. JARJANAZI H, TIMUR AA, EL-MAARRI O, CAGLAYAN SH. Analysis of the two microsatellite repeat polymorphisms of the factor VIII gene in the Turkish population. *Br J Haematol.* 100 (3): 589-93, 1998.

- 44. JAYANDHARAN G, SHAJI V, GEORGE B, CHANDY M, SRIVASTAVA A. Informativeness of linkage analysis for genetic diagnosis of haemophilia A in India. *Haemophilia*. 10 (5): 553-9, 2004.
- 45. KEMBALL-COOK G. The Hemophilia A Mutation, Structures, Test and Resource Site: HAMSTERS, 2008. Disponível na Internet via: http://europium.csc.mrc.ac.uk/WebPages/Main/main.htm. Acesso em setembro de 2009.
- 46. KIM JW, PARK SY, KIM YM, KIM DJ, RYU HM. Identification of new dinucleotide-repeat polymorphisms in factor VIIII gene using fluorescent PCR. *Haemophilia*. 11: 38-42, 2005.
- 47. LAKICH D, KAZAZIAN HH JR, ANTONARAKIS SE, GITSCHIER J. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nat Genet.* 5 (3): 236-41, 1993.
- 48. LALLOZ MRA, MCVEY JH, PATTINSON JK, TUDDENHAM EGD. Haemophilia A diagnosis by analysis of a hypervariable dinucleotide repeat within the factor VIII gene. *Lancet.* 338: 207-11, 1991.
- 49. LALLOZ MRA, SCHWAAB R, MCVEY JH, MICHAELIDES K, TUDDENHAM EGD. Haemophilia A diagnosis by simultaneous analysis of two variable dinucleotide tandem repeats withim the factor VIII gene. *Br J Haematol.* 86: 804-9, 1994.
- 50. LARNER AJ. The molecular pathology of haemophilia. *Q J Med.* 63 (242): 473-91; 1987.
- 51. LEBOUTE APM. Polimorfismos Genéticos em três amostras da população urbana brasileira. 93 p. Tese de Doutorado. Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil, 2000.
- 52. LIANG Y, ZHAO Y, YAN M, FAN XP, XIAO B, LIU JZ. Prenatal diagnosis of haemophilia A in China. *Prenat Diagn.* 27, 2009.
- LIU Q, NOZARI G, SOMMER SS. Single-tube polymerase chain reaction for rapid diagnosis of the inversion hotspot of mutation in hemophilia A. *Blood*. 92: 1458–9, 1998.
- 54. LIN SR, CHANG SC, LEE CC, SHEN MC, LIN SW. Genetic diagnosis of haemophilia A of Chinese origin. *Br J Haematol.* 91 (3): 722-7, 1995.

- 55. LIN SY, SU YN, HUNG CC, TSAY W, CHIOU SS, CHANG CT, HO HN, LEE CN. Mutation spectrum of 122 hemophilia A families from Taiwanese population by LD-PCR, DHPLC, multiplex PCR and evaluating the clinical application of HRM. *BMC Med Genet.* 20; 9:53, 2008.
- 56. LUIZON MR. Polimorfismos de DNA população-específicos em indígenas da Amazônia brasileira. Dissertação de Mestrado. Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil, 2003.
- 57. LUIZON MR, MENDES-JUNIOR CT, DE OLIVEIRA SF, SIMÕES AL. Ancestry informative markers in Amerindians from Brazilian Amazon. *Am J Hum Biol.* 20 (1): 86-90, 2007.
- 58. MANTILLA-CAPACHO JM, BELTRÁN-MIRANDA CP, LUNA-ZÁIZAR H, AGUILAR-LÓPEZ L, ESPARZA-FLORES MA, LÓPEZ-GUIDO B, TROYO-SANROMÁN R, JALOMA-CRUZ AR. Frequency of intron 1 and 22 inversions of Factor VIII gene in Mexican patients with severe hemophilia A. *Am J Hematol.* 82 (4): 283-7, 2007.
- MÁRQUEZ MB, FRISTMA GA. "Hematology: Clinical principles and applications". Third edition. Bernadette F. Rodak; George A. Fritsma; Kathryn Doig. Capítulo: 41 chapter - "Hemorrhagic coagulation disorders". 588-604, 2007.
- 60. MASSARO JD. Microssatélites do gene do FVIII em populações brasileiras. 102 p. Dissertação de Mestrado. Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil, 2005.
- 61. MAURICIO-DA-SILVA L, SILVA RS, DELLALIBERA E, DONADI EA. Population genetics of HPRTB, F13B, and LPL in Pernambuco, Northeast Brazil. *J Forensic Sci.* 45 (3): 684-6, 2000.
- 62. MENDES-JUNIOR CT. Frequências alélicas de polimorfismos de DNA do tipo STR em indígenas da Amazônia brasileira (163 p). Dissertação de Mestrado. Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil, 2001.
- 63. MESTRINER MA, SIMÕES AL, SALZANO FM. New studies on the esterase D polymorphism in South American Indians. *Am J Phys Anthropol.* 52 (1): 95-101, 1980.
- 64. MUNIZ YC, FERREIRA LB, MENDES-JUNIOR CT, WIEZEL CE, SIMÕES AL. Genomic ancestry in urban Afro-Brazilians. Ann Hum Biol. 35 (1): 104-11, 2008.

- 65. NEEL JV, GERSHOWITZ H, MOHRENWEISER HW, AMOS B, KOSTYU DD, SALZANO F M, MESTRINER MA, LAWRENCE D, SIMOES AL, SMOUSE PE, OLIVER WJ, SPIELMAN RS, NEEL Jr JV. Genetic studies on the Ticuna, an enigmatic tribe of Central Amazonas. *Ann Hum Genet.* 44 (1): 37-54, 1980.
- 66. NEI M. Molecular evolutionary genetics. New York: Columbia University Press. 512, 1987.
- 67. OLDENBURG J, IVASKEVICIUS V, ROST S, FREGIN A, WHITE K, HOLINSKI-FEDER E, MLLER CR, WEBER BH. Evaluation of DHPLC in the analysis of hemophilia A. *J Biochem Biophys Methods*. 30; 47 (1-2): 39-51, 2001.
- OLIVEIRA SF. Inserções Alu em populações indígenas da Amazônia Brasileira.
 135 p. Tese de Doutorado Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil, 1999.
- 69. OTA T. DISPAN: genetic distance and phylogenetic analysis. State College, PA: Institute of Molecular Evolutionary Genetics, Pennsylvania State University, 1993.
- 70. PANDEY GS, MITTAL B. Molecular diagnosis in haemophilia A. J Postgrad Med. 47: 274-80, 2001.
- 71. PEREIRA R, GOMES I, AMORIM A, GUSMÃO L. Genetic diversity of 10 X chromosome STRs in northern Portugal. *Int. J. Legal Med.* 121: 192-197, 2007.
- 72. PERFIL DAS COAGULOPATIAS HEREDITÁRIAS NO BRASIL: 2007 / Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Atenção Especializada. – Brasília. Brasil. 96 p: il. – (Série G. Estatística e Informação em Saúde). Editora do Ministério da Saúde, 2008.
- 73. RABBANI B, REZAEIAN A, KHANAHMAD H, BAGHERI R, KAMALI E, ZEINALI S. Analysing two dinucleotide repeats of FVIII gene in Iranian population. *Haemophilia.* 13 (6): 740-4, 2007.
- 74. RAYMOND M, ROUSSET F. GENEPOP (version 1.2): population genetics software forexact test and ecumenicism. *J Hered.* 86: 248–249, 1995.
- 75. REYNOLDS J, WEIR BS, COCKERHAM CC. Estimation of the coancestry coefficient: basis for a short term genetic distance. *Genetics*. 105: 767-779, 1983.

- 76. RIBEIRO D. O povo brasileiro: a formação e o sentido do Brasil. Companhia das Letras, São Paulo, 1995.
- 77. ROBINO C, GIOLITTI A, GINO S, TORRE C. Development of two multiplex PCR systems for the analysis of 12 X-chromosomal STR loci in a northwestern Italian population sample. *Int J Legal Med.* 120: 315-318, 2006.
- RODGERS GM, GREENBERG CS. Inherited coagulation disorders. In: Lee G.R., Foerster J., Lukens J., Paraskevas F., Greer J.P., Rodgers G.G., editors. *Wintrobe's Clinical Hematology*. 10Th. Baltimore: Williams & Wilkins, 1882-732, 1999.
- 79. ROSSETTI LC, RADIC P, LARRIPA IB, DE BRASI CD. Developing a new generation of tests for genotyping hemophilia-causative rearrangements involving int22h and int1h hotspots in the factor VIII gene. J Thromb Haemost. 6: 830–6, 2008.
- 80. SAHA A, MUKHERJEE S, MAULIK M, CHANDAK GR. Indian Genome Variation Consortium,Ray K. Evaluation of genetic markers linked to hemophilia A locus: an Indian experience. *Haematologica*. 92 (12): 1725-6, 2007.
- 81. SÁNCHEZ-GARCÍA JF, GALLARDO D, RAMÍREZ L, VIDAL F. Multiplex fluorescent analysis of four short tandem repeats for rapid haemophilia A molecular diagnosis. *Thromb Haemost.* 94 (5): 1099-103, 2005.
- 82. SAWADA A, SUMITA C, HIGASA S, UEDA M, SUEHIRO A, KAKISHITA E. Suitability of four polymorphic DNA markers for indirect genetic diagnosis of haemophilia A in Japanese subject. *Thromb Res.* 1; 105 (3): 271-6, 2002.
- 83. SCHAFFNER SF. The X chromosome in population genetics. *Nat Rev Genet.*. 5 (1): 43-51, 2004.
- 84. SCHENEIDER S, ROESSLI D, EXCOFFIER L. Arlequin ver 2000: a software for population genetics data analysis. Switzerland, Genetics and Biometry Laboratory, University of Geneva, 2000.
- 85. SHETTY S, GHOSH K, PATHARE A, COLAH R, BADAKARE S, MOHANTY D. Factor VIII and IX gene polymorphisms and carrier analysis in Indian population. *Am J Hematol.* 54 (4): 271-5, 1997.
- 86. SIMÕES AL. O polimorfismo da esterase D em indígenas sul-americanos. 115 p. Tese de Doutorado – Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil, 1980.

- 87. SOARES RP, CHAMONE DA, BYDLOWSKI SP. Factor VIII gene inversions and polymorphisms in Brazilian patients with haemophilia A: carrier detection and prenatal diagnosis. *Haemophilia*. 7: 299-305, 2001.
- SZIBOR R, KRAWCZAK M, HERING S, EDELMANN J, KUHLISCH E, KRAUSE D. Use of X-linked markers for forensic purposes. Int J Legal Med. 117 (2): 67-74, 2003.
- 89. TARAZONA-SANTOS E, CARVALHO-SILVA DR, PETTENER D, LUISELLI D; DE STEFANO GF, LABARGA CM, RICKARDS O, TYLER-SMITH C, PENA SD, SANTOS FR. Genetic differentiation in South Amerindians is related to environmental and cultural diversity: evidence from the Y chromosome. Am J Hum Genet. 68 (6):1485-96, 2001.
- 90. TIZZANO E, VENCESLÁ A, CORNET M, BAENA M, BAIGET M. Utility of a (GT) dinucleotide repeat in intron 1 of the factor 8 gene for haemophilia A carrier diagnosis. *Haemophilia*. 11 (2): 142-4, 2005.
- 91. TOOLE JJ, KNOPF JL, WOZNEY JM, SULTZMAN LA, BUECKER JL, PITTMAN DD, *ET AL*. Molecular cloning of a cDNA encoding human antihaemophilic factor. *Nature*. 312: 342-347, 1984.
- 92. VAN HOUT CV, LEVIN AM, RAMPERSAUD E, SHEN H, O'CONNELL JR, MITCHELL BD, SHULDINER AR, DOUGLAS JA. Extent and distribution of linkage disequilibrium in the Old Order Amish. *Genet Epidemiol.* 20, 2009.
- 93. VENCESLÁ A, BAENA M, FARES TAIE L, CORNET M, BAIGET M, TIZZANO EF. Application of intron 9 and intron 25 dinucleotide repeats of the factor VIII gene for carrier diagnosis in haemophilia A. *Haemophilia*. 14 (3): 489-93, 2008.
- 94. YIP B, CHAN V, CHAN TK. Intragenic dinucleotide repeats in factor VIII gene for the diagnosis of haemophilia A. *Br J Haematol.* 88 (4): 889-91, 1994.
- 95. WEIR BS, COCKERHAM CC. Estimating F-statistics for the analysis of population structure. *Evolution.* 38 (6): 1358-1370, 1984.
- 96. WIEZEL CEV. Diversidade genética e composição étnica em remanescentes de quilombos do estado do Piauí. 110 p. Tese de Doutorado. Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil, 2003.
- WINDSOR S, TAYLOR SA, LILLICRAP D. Direct detection of a common inversion mutation in the genetic diagnosis of severe hemophilia A. *Blood.* 84 (7): 2202-5, 1994.

- 98. WHITE II GC AND SHOEMAKER CB. Factor VIII gene and hemophilia A. *Blood*. 73: 1-12, 1989.
- 99. WOOD WI, CAPON DJ, SIMONSEN CC, EATON DL, GITSCHIER J, KEYT B, SEEBURG PH, *ET AL*. Expression of active human factor VIII from recombinant DNA clones. *Nature*. 312: 330-337, 1984.
- 100. WORKMAN PL, NISWANDER JD. Populations studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. Am J Hum Genet. 22: 24-29, 1970.

MANUSCRITO

GENETIC POLYMORPHISMS FOR INDIRECT DIAGNOSIS OF HEMOPHILIA A

Juliana Doblas Massaro¹, Cláudia Emília Vieira Wiezel¹, Yara Costa Netto Muniz¹, Eduardo Magalhães Rego², Luciana Corrêa Oliveira de Oliveira², Celso Teixeira Mendes-Junior³, and Aguinaldo Luiz Simões¹.

¹ Department of Genetics, School of Medicine of Ribeirão Preto, University of São Paulo, 14049-900, Ribeirão Preto-SP, Brazil.

² Department of Medical Clinic, School of Medicine of Ribeirão Preto, University of São Paulo, 14049-900, Ribeirão Preto-SP, Brazil.

³ Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, 14040-901, Ribeirão Preto-SP, Brazil.

ABSTRACT

Hemophilia A is an X-linked inherited bleeding disorder. caused by partial or total inactivity of the coagulation Factor VIII (FVIII), a plasma glycoprotein whose function is necessary for normal blood clotting. Due to the difficulties found for the direct recognition of the disease mutation in the coagulation F8 gene, the indirect diagnosis using polymorphic markers, located inside or close to the gene, is used as an alternative for determining the segregation of the mutant gene within families, and thus to detect carrier individuals and/or assist in the prenatal diagnosis. The present study characterized the allelic and haplotype frequencies, genetic diversity, population differentiation and linkage disequilibrium of five microsatellites (IVS 1, IVS 13, IVS 22, IVS 25.3 and IKBKG) in samples of healthy individuals from São Paulo, Rio Grande do Sul, Pernambuco and patients from São Paulo with Hemophilia A, to determine the degree of informativeness of these microsatellites for diagnostic purposes. The interpopulational diversity parameters highlight the differences among the analyzed population samples. Such regional differences in allelic frequencies must be taken into account when conducting indirect diagnosis of Hemophilia A. With the exception of IKBKG, all of the microsatellites presented high heterozygosity levels. Using such markers, the diagnosis was possible in 10

out of 11 families. The IVS 22, IVS 1, IVS 13, IVS 25.3 and IKBKG *loci* were informative in seven, six, six, five and two of the cases, respectively, demonstrating the effectiveness of the use of these microsatellites in the prenatal diagnosis and for the identification of carriers in the Brazilian population.

Running title: Indirect diagnosis of Hemophilia A.

Keywords: Brazilian populations, Factor VIII, haplotypes, Hemophilia A, indirect diagnosis, microsatellites.

Correspondence: Aguinaldo L. Simões, Ph.D. Department of Genetics School of Medicine in Ribeirão Preto. University of São Paulo. Avenida: Bandeirantes, 3900. 14049-900 Ribeirão Preto-São Paulo Brazil Tel. and fax: +55 16 3602 3050 E-mail: alsimoes@fmrp.usp.br

INTRODUCTION

Hemophilia A (HEMA; OMIM 306700) is a recessive disease caused by partial or total inactivity of the coagulation Factor VIII (FVIII), a plasma glycoprotein coded by an 186 Kb gene with 26 *exons*, located at the Xq28 position [1].

One third of the cases are caused by new mutations, that is, the gene was not inherited from their mothers [2]. The high number of new mutations (about 1.000 have been described so far) identified as causing HEMA [3] is attributed to the size of the gene, to the existence of a CpG island inside it and to its peculiar organization (presence of copies of the gene F8a in its interior and in the neighborhood of the FVIII gene, which facilitates the occurrence of inversions) [4].

In cases of severe HEMA, the most frequent mutations are the inversions in IVS 22 (50%) and IVS 1 (5%) [5-6]. All other mutations have very low frequencies [7]. Therefore, it is very difficult to directly determine the mutation in each particular case. In these circumstances, the use of a haplotype formed by polymorphic markers located near or within the gene is a good alternative for determining the segregation of the mutant gene within the family being studied [8], despite the fact that this method is hardly applicable to the *de novo* mutation cases. This method is also useful even when the direct diagnosis is available. Hence, more knowledge of population specific parameters of useful markers linked to the FVIII gene is needed.

Four intronic (IVS 1, IVS 13, IVS 22 and IVS 25.3) and one extragenic (IKBKG) microsatellites were analyzed in three samples of urban Brazilian populations as well as in a group of Hemophilia A patients to test the usefulness of such a set of markers in the indirect diagnosis of carriers.

MATERIALS AND METHODS

Samples

Blood samples or buccal swabs were obtained from individuals of urban Brazilian populations: São Paulo (SP): 160 unrelated individuals (80 males and 80 females) from cases of paternity investigation at the Hospital of the School of Medicine of Ribeirão Preto, University of São Paulo (HCFMRP-USP); Pernambuco (PE): 178 blood donors (150 males and 28 females) from Hemocentro de Pernambuco (HEMOPE) from the city of the Recife); Rio Grande do Sul (RS): 154 individuals (105 males and 49 females) from the city of Porto Alegre; Hemophilic-A (HA): 82 individuals (72 Hemophilia A patients and 10 mothers of the patients) of the HEMOCENTRO of HCFMRP-USP (individuals from 11 families were collected for family analysis). This research was approved by the National Ethics Research Committee (Process HCRP-FMRP, USP n° 6115/2007).

Laboratory Analysis

Allelic and haplotype frequencies of an extragenic microsatellite (IKBKG) and four intronic microsatellites (IVS 1, IVS 13, IVS 22 and IVS 25.3) (Figure 1) were obtained for each sample. For that purpose, DNA was extracted from blood or buccal swab [9) (Higuchi, 1989). The *loci, primers* and genotyping have been previously described [8, 10, 11, 12-13).

The amplified products were separated by electrophoresis on 12% denaturing (IVS 13, IVS 25.3 and IKBKG) or 10% non-denaturing (IVS 1 and IVS 22) polyacrylamide gels, and staining was done with silver nitrate. The alleles were defined by sequencing and used as size standards on subsequent gels.

Statistic Analysis

The allelic and genotypic frequencies, the Hardy-Weinberg equilibrium, the exact tests of population differentiation and the linkage disequilibrium between pairs of microsatellites were estimated by GENEPOP software version 3.4 [14]. The haplotype diversity and the analysis of molecular variance (AMOVA) were calculated with the ARLEQUIN software [15]. The haplotypes were designated as follows: repeat numbers separated by bars always in the following order: IVS 1 / IVS 13 / IVS 22 / IVS 25.3 / IKBKG. The genetic diversity with the standard error was calculated for each sample using the DISPAN software [16]. Others parameters of diversity (*hk*, *HT* and *GST*) were estimated by the Fstat 2.8 software [17], while F-statistics were calculated according to Weir and Cockerham (1984), using the software GDA [18]. The bootstrap procedure with 1000 replicates was performed in order to test if the values of *FST* differed significantly from zero. If the confidence intervals of 95% and 99% did not include the zero, the estimate was considered significantly different from zero with α =5% or 1%, respectively.

RESULTS

The allelic frequencies, heterozygosity and probabilities of adherence to the Hardy-Weinberg Equilibrium of the population samples were analyzed and are listed in Table 1. Some alleles occurred only once: *IVS 1*10* in PE, *IVS 13*26* in SP, *IVS 13*16* in RS, *IVS 22*18* in PE, *IVS 25.3*12* in RS, *IKBKG*25* in PE and *IKBKG*33* in the RS. The allelic frequencies of IKBKG, as well as the alleles of *IVS 1*10, IVS 1*21* and *IVS 25.3*12*, are described here for the first time. All the microsatellites in SP, RS and HA female fractions exhibited genotypic proportions as expected by the Hardy-Weinberg Equilibrium. In the PE sample, all *loci*, except IVS 22, presented disequilibrium determined by a deficit of heterozygotes.

Considering the five microsatellites and each population sample, the average H_s were similar for male and female, grouped or not. The IKBKG microsatellite presented the smaller values of total genetic diversity (H_T = 0,209) compared with the other four microsatellites, whose values of H_T varied between 0.567 and 0.699. The H_T did not present a significant difference between males (54.9%) and females (55.5%).

The allelic differentiation (Table 2), tested by the Fisher's exact test for five markers in the four samples pairwise comparisons, showed that of the 30 possible comparisons among male fractions, 13 presented significant differences, whereas between female fractions, only 7 did. The population pairwise F_{ST} values (Table 3), based on the five analyzed microsatellites, showed significant differences between males in the pairs SP/RS, RS/HA, RS/PE, and SP/HA and between females in the pairs SP/HA and RS/PE.

The association analyses between all intronic microsatellites pairs, considering males and females separately, revealed highly significant p-values. The IKBKG loci, the only extragenic microsatellite, did not show linkage disequilibrium with the

others. A total of 163 different haplotypes (Table 4) were found in the male fraction of the samples described here: 149 in the sample of PE, 101 in RS, 78 in SP, and 73 in HA. The haplotype 17/20/26/18/29 was the most frequent in the four population samples. Moreover, 33 of the 163 different haplotypes were present in all of the samples, while 58 occurred exclusively in PE, 25 in RS, 24 in SP and 23 in HA. The analysis of molecular variance (AMOVA) showed that almost all diversity (99.25%) was found within samples (p=0.195).

Eleven of the 73 hemophilic patients presented relatives for pedigree analysis. It was possible to recognize the haplotype that follows the mutation in 10 of these families. In the unique case in which it was not possible to recognize the haplotype that followed the mutation, the patient's mother was homozygous for the most frequent haplotype (Figure 2A). In another case (Figure 2B) there was clearly a mutation at IVS 1; moreover, since it is a small pedigree with a single affected patient, it is not possible to discard a new mutation causing the hemophilia in this particular patient. In a third case, alleged paternity was excluded, but this fact did not interfere with the diagnosis. The microsatellites IVS 22, IVS 1, IVS 13, IVS 25.3 and IKBKG were informative in 7 (63.6%), 6 (54.5%), 6 (54.5%), 5 (45.4%) and 2 (18.2%) cases, respectively. The haplotype 17/20/26/18/29, the most frequent in the four population samples, occurred in four of these families.

DISCUSSION

The occurrence of alleles *IVS 1*10, IVS 1*21* and *IVS 25.3*12*, which were described for the first time herein, was confirmed by cloning and sequencing procedures. The presence of exclusive alleles in single population samples can be explained by their low frequency (Table 5).

IKBKG allelic frequencies, briefly related by Fang *et al.* (2006) [19], are detailed here for the first time (Table 1). The low heterozygosity ($H_T = 0.209$) suggests a reduced usefulness in the indirect diagnosis.

All of the microsatellites, except IVS 22, exhibited significant deviation from the Hardy-Weinberg due to the deficit of heterozygotes in the sample of PE (Table 1). Genotyping errors can be discarded because otherwise the same disequilibrium would be present in the other samples. Previous studies [20, 21, 22] in the same region (PE) did not show any special characteristics that could explain such a disequilibrium. This is probably due to the small size (n = 28) of the female subsample in which our computations were based.

Weighted allele frequencies obtained for each continental group (European, Africans, Asiatics and Americans) were used for the comparison with the present study results (Table 5). It is noteworthy that in IVS 1, all population (Spanish, Korean, Indian, Chinese) so far studied are different to each other in relation to the most frequent allele (16, 14, 15 and 17, respectively, Table 5) while the allele 17 is the most frequent in the populations here described. It is difficult to address an explanation for this variability, since even the more likely hypothesis, methodological differences in genotyping, apparently is not sufficient to explain it.

The European and Asian average frequencies of IVS 13 alleles centered at 20 and 21 repeat alleles, although discrepant samples from New Zealand [23], China [24] and Indian [25] populations have been reported, where the most frequent alleles were 17 and 18, 25 and 26, and 22-24, respectively. Although the allele 22 is the most frequent in the only sample of Africans (Table 5) reported so far [26], alleles 19-23 also have high frequency.

The allelic distributions in Americans diverge in the three previously described samples, including a Brazilian one [27]. This makes a comparison between present data, where alleles 20 and 21 are the most frequent, with those reported in the literature difficult.

In IVS 22, Europeans and Asians present distributions centered in alleles 25 and 26. Although the New Zealand sample (Table 5) is clearly different from the other ones, we do not precisely know the origin of the examined patients [23]. The same sample of subcontinent Indians [25] that diverged from the Asian average in IVS 13 also diverged sufficiently from the Asian frequencies of IVS 22. The only sample of Africans (Table 5) described so far presents the 24 and 25 alleles as the most frequent ones. These alleles are also more frequent in Canadians and Mexicans, while in the São Paulo sample, in spite of the expressive frequencies of 25 and 26, the most frequent allele is 27.

In the samples of the present study and in the few samples previously analyzed for the IVS 25.3 microsatellite, the highest frequency was achieved by allele 18.

The utility of one polymorphism in the prenatal disease diagnosis and/or of carrier diagnosis is directly proportional to its expected heterozygosity levels (h_k) in the population under study. For the IVS 1 microsatellite, the values of h_k ranged from 0.56 (Spanish) to 0.66 (Asian). The highest values of h_k were found at IVS 13 for Americans (0.81), where only the Asians (0.58) showed a lesser values of h_k . The IVS 22 and the IVS 25.3 presented h_k ranging from 0.63 (African) to 0.71 (American) and from 0.38 (Asian) to 0.57 (present study), respectively. These differences between worldwide populations demonstrate the importance of regional surveys of allelic frequencies.

The Brazilian population is the result of interethnic crosses of European (mainly Portuguese), African and Amerindian populations, but the proportions of these components vary among the different regions. The African and Amerindian contribution to the southern Brazilian population is small [28-29]. Allelic and genotypic differentiation (Table 2) observed here could be due to the different ethnic composition and allelic frequency differences of such founder populations. The sample of patients analyzed herein is composed by individuals from several Brazilian States and not only the region of Ribeirão Preto city. Therefore, some degree of differentiation between HA and the other samples could be expected.

The exact test of population differentiation based on allele frequencies (Table 2) and pairwise F_{ST} analysis (Table 3) indicate an increased number of significant differences when males are considered in the comparison. The differences in the interpopulational diversity parameters are in agreement with the analysis of molecular variance. These results highlight the differences among the analyzed population samples, which could be the consequence of differences in the contribution of founder populations to each one of the regional samples [28-29]. Thus, the regional differences on allelic frequencies must be taken into account when conducting an indirect diagnosis of Hemophilia A.

The population differences were less pronounced when considering the IKBKG microsatellite, which could be explained by the low heterozygosity of this microsatellite. Also, it is the only one that did not show linkage disequilibrium with the other microsatellites analyzed, which is probably due to it being most distant from the others, which would increase the recombination possibilities.

Haplotype comparison of the present data with the literature is possible only for the IVS 13 and IVS 22 microsatellites. The most frequent haplotype in the present data was 20/26 (35.24%), respectively. The most frequent haplotypes (H*ind*III, IVS 13,

IVS 22, Xbal) observed in Turkish patients and healthy individuals are formed by the alleles -/20/26/+ and +/21/25/-, respectively. On the other hand, a third haplotype (-/20/26/-) was more frequent among the sample of patients (p = 0.034) [30].

The haplotypic distribution was similar in South-African black patients and healthy individuals (with or without the IVS 22 inversion). The frequency of the haplotype 20/26 (IVS 13/IVS 22) in South-African Caucasoid patients without the IVS 22 inversion was higher than in healthy individuals (p<0.01). The finding of a low inversion frequency in this group, together with a disproportional representation of one haplotype on noninversion chromosomes, is suggestive of the presence of a founder Hemophilia A mutation (26).

Due to convergent evolution in microsatellites, the interpretation of haplotype frequencies across populations should be conducted with caution. Two or more populations can share haplotypes at a high frequency by convergent mutations, drift, gene flow or common origin [31]. For this reason, the research must be careful when surpassing conclusions based on samples of restricted areas. However, the existence of significant differences in allele frequencies between populations suggests that local patterns of variability of these microsatellites should be explored in order to evaluate their forensic usefulness. Regarding Hemophilia A, although diagnosis from linkage analysis has limitations, the results demonstrate the effectiveness of use these microsatellites at prenatal diagnosis and identification of carriers in the Brazilian population. The IKBKG marker presented a low rate of heterozygosity, contributed to family analysis in just 2 (18.2%) cases and did not show linkage disequilibrium with the intronic markers. Therefore, it must be considered with care in situations in which the diagnosis is based exclusively on data from this microsatellite.

Acknowledgments

We thank Maria do Carmo Tomitão Canas and Ana Lúcia Pimentel for the support

technician during the standardizations of these microsatellites.

REFERENCES

1. Gitschier J, Wood WI, Goralka TM, et al. Characterization of the human factor VIII gene. Nature 1984; 312: 326-30.

2. Larner AJ. The molecular pathology of haemophilia. Q J Med 1987; 63 (242): 473-91.

3. Kemball-Cook G. The Hemophilia A Mutation, Structures, Test and Resource Site: HAMSTeRS, 2008. Available at:http://europium.csc.mrc.ac.uk/WebPages/Main/main.htm.

4. Rodgers GM, Greenberg CS. Inherited Coagulation Disorders. In: Lee GR, Foerster J, Lukens J, Paraskevas F, Greer JP, Rodgers GM, Editors. *Wintrobe's Clinical Hematology*, 10th. Baltimore: Williams & Wilkins 1999; 1682-732.

5. Antonarakis SE, Rossiter JP, Young M, *et al.* Factor VIII gene inversions in severe hemophilia A: results of an international consortium study. *Blood* 1995; 86: 2206-12.

6. Bagnall RD, Waseem N, Green PM, Giannelli F. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe hemophilia A. *Blood* 2002; 99: 168-74.

7. Pandey GS, Mittal B. Molecular diagnosis in Haemophilia A. J Postgrad Med 2001; 47 (4): 274-80.

8. Kim JW, Park SY, Kim JM, Kim DJ, Ryu HM. Identification of new dinucleotide-repeat polymorphisms in factor VIII gene using fluorescent PCR. *Haemophilia* 2005; 11 (1): 38-42.

9. Higuchi R. Simple and rapid preparation of samples for PCR. In: Erlich HA, (Ed) PCR technology – principles and applications for DNA amplification. New York: Stockton, Press 1989; 31-38.

10. Lalloz MRA, Mcvey JH, Pattinson JK, Tuddenham EGD. Haemophilia A diagnosis by analysis of a hypervariable dinucleotide repeat within the factor VIII gene. *Lancet* 1991; 338: 207-11.

11. Lalloz MRA, Schwaab R, Mcvey JH, Michaelides K, Tuddenham EGD. Haemophilia A diagnosis by simultaneous analysis of two variable dinucleotide tandem repeats withim the factor VIII gene. *Br J Haematol* 1994; 86: 804-9.

12. Venceslá A, Baena M, Fares Taie L, Cornet M, Baiget M, Tizzano EF. Application of intron 9 and intron 25 dinucleotide repeats of the factor VIII gene for carrier diagnosis in haemophilia A. *Haemophilia* 2008; 14 (3): 489-93.

13. Machado FB, Medina-Acosta E. An IKBKG, and not a G6PD, short tandem repeat marker is used in indirect diagnosis of haemophilia A. *Haemophilia* 2008; 14 (4): 849-50.

14. Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact test and ecumenicism. *J Hered* 1995, 86: 248–249.

15. Scheneider S, Roessli D, Excoffier L. Arlequin ver 2000: a software for population genetics data analysis. Switzerland, Genetics and Biometry Laboratory, University of Geneva, 2000.

16. Ota T. 1993. DISPAN: genetic distance and phylogenetic analysis. State College, PA: Institute of Molecular Evolutionary Genetics, Pennsylvania State University.

17. Goudet J. A program to estimate and test gene diversities and fixation indices FSTAT (version 2.8). Updated from Godet (1995), 1999.

18. Lewis PO, Zaykin D. 2000. Genetic data analysis: computer program for the analysis of allelic data [distributed by the authors]. Storrs, CT: Department of Ecology and Evolution, University of Connecticut.

19. Fang Y, Wang XF, Dai J, Wang HL. A rapid multifluorescent polymerase chain reaction for genetic counselling in Chinese haemophilia A families. *Haemophilia* 2006; 12 (1): 62-7.

20. Leboute APM. 2000. Polimorfismos Genéticos em três amostras da população urbana brasileira. Tese de Doutorado. Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, USP, SP, Brazil.

21. Mauricio-da-Silva L, Silva RS, Dellalibera E, Donadi EA. Population genetics of HPRTB, F13B, and LPL in Pernambuco, Northeast Brazil. *J Forensic Sci* 2000; 45 (3): 684-6.

22. Dellalibera E, Havro ML, Souza M, Kajihara K, Mauricio-da-Silva L, dos Santos Silva R. Genetic analysis of 13 STR loci in the population from the State of Pernambuco, northeast Brazil. *Forensic Sci Int* 2004; 146 (1): 57-9.

23. Harraway JR, Smith MP, George PM. A highly informative, multiplexed assay for the indirect detection of hemophilia A using five-linked microsatellites. *J Thromb Haemost* 2006; 4 (3): 587-90.

24. Yip B, Chan V, Chan TK. Intragenic dinucleotide repeats in factor VIII gene for the diagnosis of haemophilia A. Br J Haematol 1994; 88 (4): 889-91.

25. Saha A, Mukherjee S, Maulik M, Chandak GR; Indian Genome Variation Consortium, Ray K. Evaluation of genetic markers linked to hemophilia A locus: an Indian experience. *Haematologica* 2007; 92 (12): 1725-6.

26. Dangerfield BT, Manga P, Field SP, Hartman E, Jemkins T, Krause A. Feasibility of prenatal diagnosis and carrier detection in South African haemophilia A patients. *Br J Haematol* 1997; 97 (3): 558-60.

27. Soares RP, Chamone DA, Bydlowski SP. Factor VIII gene inversions and polymorphisms in Brazilian patients with haemophilia A: carrier detection and prenatal diagnosis. *Haemophilia* 2001; 7 (3): 299-305.

28. Ferreira LB, Mendes-Junior CT, Wiezel CEV, Luizon R, Simões AL. Genomic ancestry of a sample population from the state of São Paulo, Brazil. *Am J Hum Biol* 2006; 18 (5): 702-5.

29. Muniz YC, Ferreira LB, Mendes-Junior CT, Wiezel CE, Simões AL. Genomic ancestry in urban Afro-Brazilians. *Ann Hum Biol* 2008; 35 (1): 104-11.

30. El-Maarri O, Kavakli K, Çaglayan H. Intron 22 inversions in the Turkish haemophilia A patients: prevalence and haplotype analysis. *Haemophilia* 1999 ; 5(3): 169-73.

31. Thomas MG, Parfitt T, Weiss DA, Skorecki K, Wilson JF, le Roux M, Bradman N, Goldstein DB. Y chromosomes traveling south: the cohen modal haplotype and the origins of the Lemba--the "Black Jews of Southern Africa". *Am J Hum Genet* 2000; 66 (2): 674-86.

32. Tizzano E, Venceslá A, Cornet M, Baena M, Baiget M. Utility of a (GT) dinucleotide repeat in intron 1 of the factor 8 gene for haemophilia A carrier diagnosis. *Haemophilia* 2005; 11 (2): 142-4.

33. Liang Y, Zhao Y, Yan M, Fan XP, Xiao B, Liu JZ. Prenatal diagnosis of haemophilia A in China. *Prenat Diagn* 2009; 27.

34. Aseev M, Surin V, Baboev K, Gornostaeva N, Kuznetzova T, Kascheeva T, Ivaschenko T, Solovyev G, Mikhailov A, Lebedev V, *et al.* Allele frequencies and molecular diagnosis in haemophilia A and B patients from Russia and from some Asian republics of the former U.S.S.R. *Prenat Diagn* 1994 ; 14 (7): 513-22.

35. Goodeve AC, Tagariello G, Chuansumrit A, Preston FE, Peake IR. A rapid and cost effective method for analysis of dinucleotide repeat polymorphisms in the factor VIII gene. *Blood Coagul Fibrinolysis* 1996; 7 (7): 672-7.

36. Sánchez-García JF, Gallardo D, Ramírez L, Vidal F. Multiplex fluorescent analysis of four short tandem repeats for rapid haemophilia A molecular diagnosis. *Thromb Haemost* 2005 ; 94 (5): 1099-103.

37. Lin SR, Chang SC, Lee CC, Shen MC, Lin SW. Genetic diagnosis of haemophilia A of Chinese origin. *Br J Haematol* 1995; 91 (3): 722-7.

38. Jarjanazi H, Timur AA, El-Maarri O, Caglayan SH. Analysis of the two microsatellite repeat polymorphisms of the factor VIII gene in the Turkish population. *Br J Haematol* 1998; 100 (3): 589-93.

39. Sawada A, Sumita C, Higasa S; Ueda M, Suehiro A, Kakishita E. Suitability of four polymorphic DNA markers for indirect genetic diagnosis of haemophilia A in Japanese subject. *Thromb Res* 2002; 105 (3): 271-6.

40. Jayandharan G, Shaji V, George B, Chandy M, Srivastava A. Informativeness of linkage analysis for genetic diagnosis of haemophilia A in India. *Haemophilia* 2004; 10 (5): 553-9.

41. Rabbani B, Rezaeian A, Khanahmad H, Bagheri R, Kamali E, Zeinali S. Analysing two dinucleotide repeats of FVIII gene in Iranian population. *Haemophilia* 2007; 13 (6): 740-4.

42. Windsor S, Taylor SA, Lillicrap D. Direct detection of a common inversion mutation in the genetic diagnosis of severe hemophilia A. *Blood.* 1994; 84 (7): 2202-5.

43. Gallegos RM, Aranda HB, Navarrete CP, Espinoza R, Gomez FS, Aranda DA. Polymorphism distribution of Int13, Int22, and St14 VNTRs in a Mexican population and their application in carrier diagnosis of hemophilia A. *Am J Hematol* 2004; 77 (1): 1-6.

44. Chowdhury MR, Herrmann FH, Schroder W, Lambert CT, Lalloz MR, Layton M, Kumbnani HK, Kabra M, Menon PS, Verma IC. Factor VIII gene polymorphisms in the Asian Indian population. *Haemophilia* 2000 ; 6 (6): 625-30.

TABLES:

STRs		SP			RS			PE			НА	
IVS 1	F	М	т	F	М	т	F	м	т	F	м	т
10								0.7	0.5			
12		1.3	0.4	1.0	4.9	3.0		0.7	0.5			
14	0.6		0.4					0.7	0.5			
15	0.6		0.4		1.0	0.5	1.8	0.7	1.0			
16	1.9	1.3	1.7	8.2	4.9	6.4	8.9	9.3	9.2	10,0	9,7	9,8
17	51.4	51.8	51.5	62.3	60.0	61.5	48.3	50.6	50.0	65,0	63,8	64,2
18	28.7	19.0	25.5	24.5	24.3	24.1	19.6	22.7	21.8	10,0	15,3	14,1
19	8.7	17.7	11.7	2.0	3.9	3.0	10.7	5.3	6.8	5,0	4,2	4,3
20	5.6	7.6	6.3	2.0	1.0	1.5	10.7	8.0	8.7	10,0	5,6	6,5
21	2.5	1.3	2.1					1.3	1.0		1,4	1,1
nª	160	79	239	98	103	201	56	150	206	20	72	92
p	0.201			0.454			0.002*			0,511		
H_k^c	0.659	0.658	0.651	0.550	0.574	0.558	0.714	0.674	0.682	0,572	0,563	0,560
IVS 13												
16				1.0		1.0						
18	1.3		0.8					2.7	1.9	5,0	1,4	2,2
19	3.1	2.5	2.9	4.1	6.7	5.9	5.4	7.3	6.8	5,0	8,3	7,6
20	43.7	38.7	42.1	56.1	54.2	54.5	39.3	48.5	46.2	65,0	45,9	50,0
21	25.0	22.5	24.2	28.6	26.7	27.3	30.4	14.7	18.9	10,0	22,2	19,6
22	12.5	12.5	12.5	5.1	6.7	5.9	7.1	10.7	9.7	5,0	5,6	5,4
23	6.9	13.8	9.2	5.1	5.7	5.4	8.9	6.7	7.3	5,0	9,7	8,7
24	5.6	7.5	6.2				8.9	8.7	8.7	5,0	6,9	6,5
25	1.3	2.5	1.7					0.7	0.5			
26	0.6		0.4									
nª	160	80	240	98	105	203	56	150	206	20	72	92
p ^b	0.070			0.085			0.003*			0,274		
H _k °	0.723	0.758	0.735	0.603	0.623	0.618	0.749	0.714	0.724	0,589	0,727	0,699
IVS 22												
18								0.7	0.5			
22	2.5	2.5	2.5					1.3	1.0			
23	6.7	6.2	6.7		1.0	0.5		3.3	2.4	10,0	4,2	5,4
24	8.3	7.5	8.3	5.1	4.0	4.5	10.7	10.0	10.2	10,0	2,8	4,3
25	36.9	40.0	37.9	37.8	33.6	35.8	50.1	31.3	36.4	20,0	23,6	22,8
26	39.4	42.5	40.4	51.0	58.4	54.7	32.1	50.0	45.1	50,0	63,8	61,0

Table 1: Allelic frequencies (%) of the IVS 1, IVS 13, IVS 22, IVS 25.3 and IKBKG microsatellites, according the sex (F=female, M=male, T=Total sample) in the four samples (SP=São Paulo, RS=Rio Grande do Sul, PE=Pernambuco and HA=Hemophilia A patients).

070-		07						05				
SIRS		SP			RS			PE			HA	
27	5.0	1.3	3.8	3.1	3.0	3.0	7.1	2.7	3.9	5,0	4,2	4,3
28				2.0		1.0		0.7	0.5	5,0	1,4	2,2
29	0.6		0.4	1.0		0.5						
n"	160	80	240	98	101	199	56	150	206	20	72	92
p	0.576			0.608			0.077			0,922		
H_k^c	0.701	0.649	0.680	0.599	0.543	0.570	0.643	0.640	0.651	0,717	0,540	0,577
IVS 25.3												
12				1.0		0.5						
15	0.6		0.4									
16	6.3	10.3	7.6	3.1		2.0	12.5	13.3	13.1		8,3	6,5
17	5.1	3.8	4.7	1.0	3.9	2.5	1.8	2.0	1.9	10,0	4,2	5,4
18	65.8	57.7	63.1	68.4	63.8	65.7	60.7	58.7	59.2	50,0	68,0	64,2
19	8.2	14.1	10.2	11.2	12.7	11.9	14.3	10.7	11.7	25,0	2,8	7,6
20	12.7	11.5	12.3	14.3	15.7	14.9	7.1	12.0	10.7	15,0	13,9	14,1
21	1.3	1.3	1.3	1.0	3.9	2.5	3.6	2.0	2.4		1,4	1,1
22		1.3	0.4					1.3	1.0		1,4	1,1
nª	158	78	236	98	102	200	56	150	206	20	72	92
р ^ь	0.394			0.409			0.015*			0,239		
H _k °	0.552	0.622	0.568	0.503	0.549	0.530	0.603	0.611	0.606	0,700	0,515	0,563
IKBKG												
25								1.3	1.0			
27	1.3		0.8	1.0		0.5		1.3	1.0		1,4	1,1
28	5.6	7.5	6.2	10.2	2.9	6.8		8.1	5.9	10,0	6,9	7,6
29	88.1	82.5	86.3	81.7	94.2	87.8	98.2	88.6	91.1	90,0	88,9	89,1
30	5.0	10.0	6.7	6.1	2.9	4.4	1.8	0.7	1.0		2,8	2,2
33				1.0		0.5						
nª	160	80	240	98	105	203	56	149	205	20	72	92
р ^ь	0.179			0.104			0,000*			1,000		
H _k °	0.232	0.304	0.247	0.323	0.111	0.223	0.036	0.208	0.166	0,189	0,207	0,201
<i>H</i> s ^d	0.566	0.602	0.577	0.516	0.487	0.500	0.545	0.571	0.567	0,552	0,510	0,519
Standard error	0.092	0.078	0.087	0.051	0.092	0.071	0.130	0.092	0.102	0,095	0,084	0,083

*significant at α =0.05. ^a n = number of X chromosome ^b p = Probability of adherence to the Hardy-Weinberg Equilibrium proportions ^c H_k = Expected heterozygosity ^d H_S = Average expected heterozygosity

		IVS 1	IVS 13	IVS 22	IVS 25.3	IKGKB
Male	SP x RS	*	**	**	**	**
	SP x PE	**	0.103	0.670	0.940	**
	SP x HA	**	0.325	**	0.252	0.213
	RS x PE	**	**	0.350	*	0.084
	RS x HA	0.097	0.092	0.375	**	0.334
	PE x HA	0.922	0.652	0.236	0.270	0.715
Female	SP x RS	*	*	*	0.411	0.389
(Genic)	SP x PE	0.121	0.805	0.184	0.255	0.199
	SP x HA	0.192	0.418	0.269	0.201	0.606
	RS x PE	*	*	0.080	0.121	*
	RS x HA	0.249	0.095	0.052	0.151	0.828
	PE x HA	0.829	0.224	*	0.146	0.067
Female	SP x RS	*	*	*	0.426	0.555
(Genotypic)	SP x PE	0.196	0.802	0.089	0.406	0.077
	SP x HA	0.325	0.673	0.386	0.287	0.594
	RS x PE	0.054	0.131	0.149	0.288	*
	RS x HA	0.480	0.222	0.099	0.213	0.683
	PE x HA	0.828	0.524	*	0.216	0.065

Table 2. Allelic and genotypic differentiation between the possible pairs of the samples of São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE) and Haemophilia A patients (HA) based in the microsatellites: IVS 1, IVS 13, IVS 22, IVS 25.3 and IKBKG.

 $(**) p \le 0.001 (*) 0.001$

Table 3. Pairwise F_{st} values between population samples of São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE) and Haemophilia A patients (HA) and their 99% and 95% Confidence Intervals (CI), based in IVS 1, IVS 13, IVS 22, IVS 25.3 and IKBKG microsatellites. Significant values (p < 0.05) are in boldface:

Populations	Male	99% CI	95% CI	Female	99% CI	95% CI
SP x RS	0.017	0.005/0.031	0.009/0.026	0.006	-0.002/0.011	-0.000/0.010
SP x PE	0.004	-0.005/0.011	-0.002/0.010	0.001	-0.006/0.018	-0.004/0.012
SP x HA	0.017	-0.002/0.043	0.001/0.038	0.013	-0.010/0.021	0.001/0.020
RS x PE	0.007	0.003/0.013	0.004/0.012	0.019	0.007/0.051	0.010/0.036
RS x HA	0.003	0.000/0.007	0.000/0.006	0.007	-0.008/0.024	-0.005/0.018
PE x HA	0.005	-0.003/0.013	-0.000/0.011	0.023	-0.007/0.066	-0.004/0.054

HaplotypeIVS 11IVS 22IVS 25.3IKBKGFreqSPRSPEHA1102325182911121220251829111312202518291114122002618291115122027182911161224251629111714232428281119152323182911110161924182911111116202618281111121620261828111113162026182811111416202618291111151620261829111116212718271111162225182911112016222518291112116222518 <th></th> <th>N</th> <th>lumber o</th> <th>f repeats</th> <th>s in the mic</th> <th colspan="5">Number of X chromosomes</th>		N	lumber o	f repeats	s in the mic	Number of X chromosomes					
1 10 23 25 18 29 1 1 1 2 12 20 25 18 29 1 1 1 3 12 20 25 18 29 2 22 2 5 12 20 26 18 29 1 1 1 6 12 24 25 16 29 1 1 1 7 14 23 24 28 28 1 1 1 9 15 23 23 18 28 1 1 1 11 16 19 24 18 28 1 1 1 12 16 20 26 18 29 1 1 1 13 16 20 26 18 28 1 1 1 14 16 20 26 18 29 1 1 1 14 16 20 26	Haplotype	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	RS	PE	HA
2 12 20 25 18 29 1 1 3 12 20 25 21 29 1 1 4 12 20 27 18 29 1 1 5 12 24 25 16 29 1 1 1 7 14 23 24 28 28 1 1 1 9 15 23 23 18 28 1 1 1 10 16 19 24 18 29 1 1 1 1 11 16 19 26 18 29 1 1 1 1 13 16 20 26 18 28 1 1 1 1 14 16 20 26 18 28 1 1 1 1 15 16 20 26 18 29 1 1 1 1 16 21	1	10	23	25	18	29	1			1	
3 12 20 25 21 29 1 1 4 12 20 26 18 29 2 2 5 12 20 27 18 29 1 1 1 6 12 24 25 16 29 1 1 1 7 14 23 24 28 28 1 1 1 9 15 23 23 18 29 1 1 1 9 15 23 23 18 29 1 1 1 11 16 19 24 18 28 1 1 1 12 16 20 26 18 28 1 1 1 14 16 20 26 18 28 1 1 1 16 16 20 26 18 29 1 1 1 17 16 22 25 16 29	2	12	20	25	18	29	1		1		
4 12 20 26 18 29 2 2 5 12 24 25 16 29 1 1 6 12 24 25 16 29 1 1 7 14 23 24 28 28 1 1 1 9 15 23 23 18 28 1 1 1 10 16 19 24 18 29 1 1 1 11 16 20 24 18 28 1 1 1 12 16 20 26 18 29 1 3 6 2 15 16 20 26 18 28 1 1 1 16 16 20 26 18 28 1 1 1 16 12 27 18 27 1 1 1 1 18 16 21 25 16 29	3	12	20	25	21	29	1		1		
5 12 20 27 18 29 1 1 6 12 24 25 16 29 1 1 7 14 23 24 28 28 1 1 1 9 15 23 23 18 28 1 1 1 9 16 19 24 18 29 1 1 1 11 16 19 26 18 28 1 1 1 13 16 20 26 18 28 1 1 1 14 16 20 26 18 28 1 1 1 14 16 20 26 18 29 1 1 1 15 16 20 26 18 29 1 1 1 16 122 23 22 18 29 1 1 1 16 22 25 16 29 1 <th>4</th> <th>12</th> <th>20</th> <th>26</th> <th>18</th> <th>29</th> <th>2</th> <th></th> <th>2</th> <th></th> <th></th>	4	12	20	26	18	29	2		2		
6 12 24 25 16 29 1 1 7 14 23 24 28 28 1 1 9 15 23 23 18 28 1 1 1 10 16 19 24 18 29 1 1 1 1 11 16 19 26 18 29 1 1 1 1 12 16 20 24 18 28 1 1 1 1 13 16 20 26 18 29 11 3 6 2 16 16 20 26 18 29 1 1 1 1 17 16 20 26 20 29 1 1 1 1 1 18 16 21 26 18 29 1 1 1 1 1 1 1 1 1 1 1 1 1 1	5	12	20	27	18	29	1		1		
7 14 23 24 28 28 1 1 8 15 20 27 18 30 1 1 9 15 23 23 18 28 1 1 1 10 16 19 24 18 29 1 1 1 11 16 19 26 18 28 1 1 1 13 16 20 26 18 28 1 1 1 14 16 20 26 18 29 1 3 6 2 16 16 20 26 19 29 2 1 1 1 17 16 21 27 18 27 1 1 1 20 16 21 27 18 27 1 1 1 22 16 22 25 18 29 1 1 1 23 16 22 25 <th>6</th> <th>12</th> <th>24</th> <th>25</th> <th>16</th> <th>29</th> <th>1</th> <th></th> <th></th> <th>1</th> <th></th>	6	12	24	25	16	29	1			1	
8 15 20 27 18 30 1 1 1 9 15 23 23 18 28 1 1 1 10 16 19 24 18 29 1 1 1 1 11 16 19 26 18 29 1 1 1 1 12 16 20 24 18 28 1 1 1 13 16 20 26 18 28 1 1 1 16 16 20 26 18 28 1 1 1 17 16 20 26 29 1 1 1 1 20 16 21 27 18 27 1 1 1 1 21 16 22 25 16 29 1 1 1 1 22 16 23 24 22 28 1 1 1 1	7	14	23	_== 24	28	28	1			1	
9 15 20 17 18 28 1 1 10 16 19 24 18 29 1 1 1 11 16 19 24 18 29 1 1 1 12 16 19 26 18 29 1 1 13 16 20 25 18 28 1 1 13 16 20 26 18 29 11 3 6 2 16 16 20 26 19 29 2 1 1 17 16 20 26 18 28 1 1 1 18 16 21 27 18 27 1 1 1 20 16 22 25 18 29 1 1 1 21 16 22 25 18 2	8	15	20	27	_0 18	-0 30	1		1	•	
10 16 19 24 18 29 1 1 1 11 16 19 26 18 29 1 1 1 13 16 20 24 18 28 1 1 1 13 16 20 26 18 28 1 1 1 14 16 20 26 18 28 1 1 1 16 16 20 26 18 29 1 1 1 1 17 16 20 26 18 28 1 1 1 1 18 16 21 26 18 29 1 1 1 1 20 16 22 25 16 29 1 1 1 1 21 16 22 25 16 29 1 1 1 1 22 16 23 25 16 29 1 1 1	9	15	23	23	18	28	1			1	
11 16 19 26 18 29 1 1 12 16 20 24 18 28 1 1 13 16 20 25 18 28 1 1 14 16 20 26 18 29 11 3 6 2 16 16 20 26 18 29 1 1 1 17 16 20 26 18 29 1 1 1 18 16 21 26 18 28 1 1 1 20 16 21 27 18 27 1 1 1 21 16 22 25 16 29 1 1 1 21 16 22 25 18 29 1 1 1 23 16 23 25 16 29 1 1 1 24 18 29 1 1 1<	10	16	10	20	18	20	1	1			
12 16 20 24 18 28 1 1 13 16 20 25 18 28 1 1 14 16 20 26 18 28 1 1 15 16 20 26 18 29 11 3 6 2 16 16 20 26 19 29 2 1 1 1 17 16 20 26 18 28 1 1 1 18 16 21 27 18 27 1 1 1 20 16 22 23 22 29 1 1 1 21 16 22 25 16 29 1 1 1 22 16 22 25 18 29 1 1 1 23 16 22 26 18 29 1 1 1 24 16 23 25 1	10	16	10	27	18	20	1				1
12 10 20 24 10 20 1 1 13 16 20 25 18 28 1 1 14 16 20 26 18 29 11 3 6 2 16 16 20 26 19 29 2 1 1 1 17 16 20 26 18 28 1 1 1 1 18 16 21 27 18 27 1 1 1 1 20 16 22 23 22 29 1 1 1 1 21 16 22 25 18 29 1 1 1 1 22 16 22 25 18 29 1 1 1 1 1 24 16 23 24 22 28 1 1 1 1 1 1 1 1 1 1 1 1 <t< th=""><th>12</th><th>16</th><th>20</th><th>20</th><th>10</th><th>23</th><th>1</th><th></th><th></th><th>1</th><th>I</th></t<>	12	16	20	20	10	23	1			1	I
13 10 20 26 18 29 1 3 6 2 16 16 20 26 18 29 11 3 6 2 16 16 20 26 19 29 2 1 1 17 16 20 26 18 29 1 1 1 18 16 21 26 18 28 1 1 1 20 16 22 23 22 29 1 1 1 21 16 22 25 18 29 1 1 1 22 16 22 25 18 29 1 1 1 23 16 22 26 18 29 1 1 1 24 16 23 25 16 29 1 1 1 26 17 18 26 18 29 1 1 1 27 17 </th <th>12</th> <th>16</th> <th>20</th> <th>24</th> <th>10</th> <th>20</th> <th>1</th> <th></th> <th></th> <th>I</th> <th>1</th>	12	16	20	24	10	20	1			I	1
14 16 20 26 18 29 1 3 6 2 16 16 20 26 19 29 2 1 1 17 16 20 26 20 29 1 1 1 18 16 21 27 18 28 1 1 1 20 16 22 23 22 29 1 1 1 20 16 22 25 16 29 1 1 1 21 16 22 25 16 29 1 1 1 22 16 22 26 18 29 1 1 1 24 16 23 24 22 28 1 1 1 25 16 23 24 17 29 1 1 1 26 17 18 26 18 29 1 1 1 30 17 19 </th <th>13</th> <th>10</th> <th>20</th> <th>20</th> <th>10</th> <th>20</th> <th>1</th> <th></th> <th></th> <th></th> <th>1</th>	13	10	20	20	10	20	1				1
15 10 20 26 18 29 11 3 0 2 16 16 20 26 19 2 1 1 1 17 16 20 26 20 29 1 1 1 18 16 21 26 18 28 1 1 1 20 16 22 23 22 29 1 1 1 20 16 22 25 16 29 1 1 1 21 16 22 25 18 29 1 1 1 22 16 22 25 16 29 1 1 1 23 16 23 25 16 29 1 1 1 26 17 18 26 18 29 1 1 1 29 17 19 26 18 29 1 1 1 30 17 19 </th <th>14</th> <th>10</th> <th>20</th> <th>20</th> <th>10</th> <th>20</th> <th>1</th> <th></th> <th>0</th> <th><u>^</u></th> <th></th>	14	10	20	20	10	20	1		0	<u>^</u>	
16 20 26 19 29 2 1 1 17 16 20 26 29 1 1 1 18 16 21 27 18 27 1 1 1 20 16 22 23 22 29 1 1 1 20 16 22 25 18 29 1 1 1 21 16 22 25 18 29 1 1 1 23 16 22 26 18 29 1 1 1 24 16 23 24 22 28 1 1 1 25 16 23 25 16 29 1 1 1 26 17 18 26 18 29 1 1 1 29 17 19 26 18 28 1 1 1 30 17 19 26 18 29<	15	10	20	20	18	29			3	0	Z
17 16 20 26 20 29 1 1 18 16 21 26 18 28 1 1 20 16 22 23 22 29 1 1 20 16 22 23 22 29 1 1 21 16 22 25 18 29 1 1 23 16 22 26 18 29 1 1 23 16 23 24 22 28 1 1 24 16 23 24 22 28 1 1 26 17 16 26 16 29 1 1 27 17 18 26 18 28 1 1 1 28 17 19 26 16 29 1 1 1 30 17 19 26 18 28 1 1 1 33 17	16	16	20	26	19	29	2		1	1	
18 16 21 26 18 28 1 1 19 16 21 27 18 27 1 1 1 20 16 22 23 22 29 1 1 1 21 16 22 25 18 29 1 1 1 22 16 22 25 18 29 1 1 1 23 16 22 26 18 29 1 1 1 24 16 23 25 16 29 1 1 1 26 17 16 26 16 29 1 1 1 28 17 18 26 18 29 1 1 1 30 17 19 26 18 28 1 1 1 31 17 19 26 18 29 1 1 1 33 17 20 23 20	17	16	20	26	20	29	1			1	
19 16 21 27 18 27 1 1 20 16 22 23 22 29 1 1 21 16 22 25 16 29 1 1 22 16 22 25 18 29 1 1 23 16 22 26 18 29 1 1 24 16 23 24 22 28 1 1 26 17 16 26 16 29 1 1 27 17 18 25 17 29 1 1 28 17 18 26 18 29 3 2 1 30 17 19 26 16 29 1 1 1 31 17 19 26 18 28 1 1 1 33 17 19 26 18 29 1 1 1 33 <	18	16	21	26	18	28	1				1
20 16 22 23 22 29 1 1 21 16 22 25 16 29 1 1 22 16 22 25 18 29 1 1 23 16 22 26 18 29 1 1 1 24 16 23 24 22 28 1 1 1 26 17 16 26 16 29 1 1 1 26 17 18 25 17 29 1 1 1 28 17 18 26 18 29 3 2 1 30 17 19 26 18 28 1 1 1 31 17 19 26 18 29 14 1 4 6 3 34 17 19 26 18 29 1 1 1 35 17 20 18 2	19	16	21	27	18	27	1			1	
21 16 22 25 16 29 1 1 22 16 22 25 18 29 1 1 1 23 16 22 26 18 29 1 1 1 24 16 23 24 22 28 1 1 1 25 16 23 25 16 29 1 1 1 26 17 16 26 16 29 1 1 1 28 17 18 26 18 29 3 2 1 30 17 19 26 16 29 1 1 1 31 17 19 26 18 28 1 1 1 33 17 19 26 18 29 14 1 4 6 3 34 17 19 26 18 29 1 1 1 35 17 20<	20	16	22	23	22	29	1			1	
22 16 22 25 18 29 1 1 23 16 22 26 18 29 1 1 24 16 23 24 22 28 1 1 25 16 23 25 16 29 1 1 26 17 16 26 16 29 1 1 27 17 18 25 17 29 1 1 28 17 19 24 18 29 3 2 1 30 17 19 25 18 28 1 1 1 31 17 19 26 16 29 1 1 1 32 17 19 26 18 28 1 1 1 33 17 20 18 18 29 1 1 1 35 17 20 23 22 29 1 1 1	21	16	22	25	16	29	1			1	
23 16 22 26 18 29 1 1 24 16 23 24 22 28 1 1 25 16 23 25 16 29 1 1 26 17 18 26 16 29 1 1 26 17 18 25 17 29 1 1 28 17 18 26 18 29 3 2 1 29 17 19 24 18 29 1 1 1 30 17 19 26 16 29 1 1 1 31 17 19 26 18 28 1 1 1 33 17 19 26 18 29 1 1 1 35 17 20 23 20 28 1 1 1 36 17 20 25 18 29 1 1	22	16	22	25	18	29	1		1		
24 16 23 24 22 28 1 1 25 16 23 25 16 29 1 1 26 17 16 26 16 29 1 1 26 17 18 25 17 29 1 1 28 17 18 26 18 29 3 2 1 29 17 19 24 18 29 1 1 1 30 17 19 26 16 29 1 1 1 31 17 19 26 18 28 1 1 1 32 17 19 26 18 29 1 1 1 33 17 19 26 18 29 1 1 1 35 17 20 18 18 29 1 1 1 36 17 20 23 20 28 1	23	16	22	26	18	29	1				1
25 16 23 25 16 29 1 1 26 17 16 26 16 29 1 1 27 17 18 25 17 29 1 1 28 17 18 26 18 29 3 2 1 29 17 19 24 18 29 1 1 1 30 17 19 25 18 28 1 1 1 31 17 19 26 18 28 1 1 4 6 3 33 17 19 26 18 29 1 1 1 1 35 17 20 18 18 29 1 1 1 36 17 20 23 20 28 1 1 1 36 17 20 25 18 29 1 1 1 39 17 20 25<	24	16	23	24	22	28	1			1	
26 17 16 26 16 29 1 1 27 17 18 25 17 29 1 1 28 17 18 26 18 29 3 2 1 29 17 19 24 18 29 1 1 1 30 17 19 25 18 28 1 1 1 31 17 19 26 16 29 1 1 4 6 3 32 17 19 26 18 28 1 1 4 6 3 33 17 19 26 18 29 1 1 4 6 3 34 17 19 26 18 29 1 1 1 1 3 35 17 20 18 18 29 1 1 1 1 1 1 1 1 1 1 1 1 <td< th=""><th>25</th><th>16</th><th>23</th><th>25</th><th>16</th><th>29</th><th>1</th><th></th><th></th><th>1</th><th></th></td<>	25	16	23	25	16	29	1			1	
27 17 18 25 17 29 1 1 28 17 18 26 18 29 3 2 1 29 17 19 24 18 29 1 1 1 30 17 19 25 18 28 1 1 1 31 17 19 26 16 29 1 1 1 32 17 19 26 18 28 1 1 4 6 3 33 17 19 26 18 29 14 1 4 6 3 34 17 19 26 18 29 1 1 1 1 35 17 20 18 8 29 1	26	17	16	26	16	29	1		1		
28 17 18 26 18 29 3 2 1 29 17 19 24 18 29 1 1 1 30 17 19 25 18 28 1 1 1 1 31 17 19 26 16 29 1 1 1 1 32 17 19 26 18 28 1 1 4 6 3 34 17 19 26 18 29 1 1 4 6 3 34 17 19 26 19 29 1 1 1 3 35 17 20 18 18 29 1 1 1 1 3 1 3 1 <t< th=""><th>27</th><th>17</th><th>18</th><th>25</th><th>17</th><th>29</th><th>1</th><th></th><th></th><th>1</th><th></th></t<>	27	17	18	25	17	29	1			1	
29 17 19 24 18 29 1 1 30 17 19 25 18 28 1 1 1 31 17 19 26 16 29 1 1 1 32 17 19 26 18 28 1 1 1 1 33 17 19 26 18 29 14 1 4 6 3 34 17 19 26 18 29 1 1 1 1 35 17 20 18 18 29 1 1 1 1 36 17 20 23 20 28 1 1 1 1 37 17 20 23 22 29 1 1 1 1 1 38 17 20 25 18 29 5 2 1 2 1 2 40 17 20 26	28	17	18	26	18	29	3			2	1
30 17 19 25 18 28 1 1 31 17 19 26 16 29 1 1 32 17 19 26 18 28 1 1 1 33 17 19 26 18 29 14 1 4 6 3 34 17 19 26 18 29 1 1 1 1 35 17 20 18 18 29 1 1 1 1 36 17 20 23 20 28 1 1 1 1 37 17 20 23 22 29 1 1 1 1 38 17 20 25 18 29 5 2 1 2 40 17 20 25 21 29 1 1 1 41 17 20 26 16 29 1 1 1	29	17	19	24	18	29	1			1	
31 17 19 26 16 29 1 1 32 17 19 26 18 28 1 1 4 6 3 33 17 19 26 18 29 14 1 4 6 3 34 17 19 26 19 29 1 1 1 3 35 17 20 18 18 29 1 1 1 3 36 17 20 23 20 28 1 1 1 1 37 17 20 23 22 29 1 1 1 1 38 17 20 23 22 29 1 <	30	17	19	25	18	28	1		1		
32 17 19 26 18 28 1 1 1 33 17 19 26 18 29 14 1 4 6 3 34 17 19 26 19 29 1 1 4 6 3 35 17 20 18 18 29 1 1 1 1 36 17 20 23 20 28 1 1 1 1 37 17 20 23 22 29 1 1 1 1 38 17 20 24 17 29 1 1 1 2 39 17 20 25 18 29 5 2 1 2 40 17 20 25 21 29 1 1 1 42 17 20 26 16 29 1 1 1 43 17 20 26 18	31	17	19	26	16	29	1				1
33 17 19 26 18 29 14 1 4 6 3 34 17 19 26 19 29 1 1 1 1 35 17 20 18 18 29 1 1 1 1 36 17 20 23 20 28 1 1 1 1 37 17 20 23 22 29 1 1 1 1 38 17 20 24 17 29 1 1 1 2 40 17 20 25 18 29 5 2 1 2 41 17 20 25 20 29 1 1 1 1 42 17 20 26 16 29 1 1 1 1 43 17 20 26 18 28 4 2 1 1 44 17 20	32	17	19	26	18	28	1			1	
34 17 19 26 19 29 1 1 35 17 20 18 18 29 1 1 36 17 20 23 20 28 1 1 37 17 20 23 22 29 1 1 38 17 20 23 22 29 1 1 38 17 20 24 17 29 1 1 39 17 20 25 18 29 5 2 1 2 40 17 20 25 20 29 1 1 1 1 41 17 20 25 21 29 1 1 1 1 43 17 20 26 16 29 1 1 1 1 43 17 20 26 18 28 4 2 1 1 44 17 20 26 18	33	17	19	26	18	29	14	1	4	6	3
35 17 20 18 18 29 1 1 36 17 20 23 20 28 1 1 37 17 20 23 22 29 1 1 38 17 20 24 17 29 1 1 1 39 17 20 25 18 29 5 2 1 2 40 17 20 25 20 29 1 1 1 41 17 20 25 21 29 1 1 1 42 17 20 25 21 29 1 1 1 43 17 20 26 16 29 1 1 1 44 17 20 26 18 28 4 2 1 1 45 17 20 26 18 29 90 15 28 32 15 46 17 <td< th=""><th>34</th><th>17</th><th>19</th><th>26</th><th>19</th><th>29</th><th>1</th><th></th><th>1</th><th></th><th></th></td<>	34	17	19	26	19	29	1		1		
36 17 20 23 20 28 1 1 37 17 20 23 22 29 1 1 38 17 20 24 17 29 1 1 1 39 17 20 25 18 29 5 2 1 2 40 17 20 25 20 29 1 1 1 41 17 20 25 21 29 1 1 1 41 17 20 25 21 29 1 1 1 42 17 20 26 16 29 1 1 1 43 17 20 26 18 28 4 2 1 1 44 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 29 1 1 1 47	35	17	20	18	18	29	1			1	
37 17 20 23 22 29 1 1 38 17 20 24 17 29 1 1 1 39 17 20 25 18 29 5 2 1 2 40 17 20 25 20 29 1 1 1 41 17 20 25 21 29 1 1 1 41 17 20 25 21 29 1 1 1 42 17 20 26 16 29 1 1 1 43 17 20 26 17 29 1 1 1 44 17 20 26 18 28 4 2 1 1 45 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 29 1 1 1 4	36	17	20	23	20	28	1				1
38 17 20 24 17 29 1 1 39 17 20 25 18 29 5 2 1 2 40 17 20 25 20 29 1 1 2 41 17 20 25 21 29 1 1 1 41 17 20 25 21 29 1 1 1 42 17 20 26 16 29 1 1 1 43 17 20 26 17 29 1 1 1 44 17 20 26 18 28 4 2 1 1 45 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 29 1 1 1 47 17 20 26 18 30 3 2 1 1 <	37	17	20	23	22	29	1				1
39 17 20 25 18 29 5 2 1 2 40 17 20 25 20 29 1 1 1 1 41 17 20 25 21 29 1 1 1 1 42 17 20 26 16 29 1 1 1 1 43 17 20 26 16 29 1 1 1 1 44 17 20 26 18 28 4 2 1 1 1 44 17 20 26 18 29 90 15 28 32 15 45 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 30 3 2 1 1 47 17 20 26 18 30 3 2 1 1 48 </th <th>38</th> <th>17</th> <th>20</th> <th>24</th> <th>17</th> <th>29</th> <th>1</th> <th></th> <th>1</th> <th></th> <th></th>	38	17	20	24	17	29	1		1		
40 17 20 25 20 29 1 1 41 17 20 25 21 29 1 1 1 42 17 20 26 16 29 1 1 1 43 17 20 26 16 29 1 1 1 44 17 20 26 17 29 1 1 1 44 17 20 26 18 28 4 2 1 1 45 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 29 1 1 1 47 17 20 26 18 30 3 2 1 1 48 17 20 26 18 30 3 2 1 1	39	17	20	25	18	29	5		2	1	2
41 17 20 25 21 29 1 1 1 42 17 20 26 16 29 1 1 1 43 17 20 26 17 29 1 1 1 44 17 20 26 18 28 4 2 1 1 45 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 29 1 1 1 47 17 20 26 18 29 1 1 1 48 17 20 26 18 30 3 2 1 1	40	17	20	25	20	29	1		1		
42 17 20 26 16 29 1 1 43 17 20 26 17 29 1 1 44 17 20 26 18 28 4 2 1 1 45 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 29 1 1 1 47 17 20 26 18 30 3 2 1 1 48 17 20 26 18 30 3 2 1 1	41	17	20	25	21	29	1			1	
43 17 20 26 17 29 1 1 44 17 20 26 18 28 4 2 1 1 45 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 29 1 1 1 47 17 20 26 18 30 3 2 1 1 48 17 20 26 19 28 1 1 1	42	17	20	26	16	29	1	1			
44 17 20 26 18 28 4 2 1 1 45 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 29 1 1 1 47 17 20 26 18 30 3 2 1 1 48 17 20 26 19 28 1 1 1	43	17	20	26	17	29	1			1	
45 17 20 26 18 29 90 15 28 32 15 46 17 20 26 18 29 1 1 47 17 20 26 18 30 3 2 1 48 17 20 26 19 28 1 1	44	17	20	26	18	28	4	2	1		1
46 17 20 26 18 29 1 1 47 17 20 26 18 30 3 2 1 48 17 20 26 19 28 1 1	45	17	20	26	18	29	90	15	28	32	15
47 17 20 26 18 30 3 2 1 48 17 20 26 19 28 1 1	46	17	20	26	18	29	1			1	
48 17 20 26 19 28 1 1	47	17	20	26	18	30	3	2	1	•	
	48	17	20	_0 26	19	28	1	-	1		

Table 4. Absolute frequency of the haplotypes of chromosome X observed in males in the samples of São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE) and Haemophilia A patients (HA):

	N	umber o	f repeats	in the mi	crosatelli	tes	Nu	mber of X o	hromosor	nes
Haplotype	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	RS	PE	HA
49	17	20	26	19	29	9	2	1	5	1
50	17	20	26	20	29	2			2	
51	17	20	27	16	29	1			1	
52	17	20	27	17	30	1			1	
53	17	20	27	18	27	1				1
54	17	20	27	18	29	2				2
55	17	20	27	19	29	1		1		
56	17	20	28	18	29	2			1	1
57	17	21	24	18	29	1		1		
58	17	21	24	19	29	1		1		
59	17	21	24	20	28	1			1	
60	17	21	25	17	29	1				1
61	17	21	25	18	30	1				1
62	17	21	25	19	29	6	3	1	2	
63	17	21	25	20	29	11	3	4	3	1
64	17	21	25	20	30	1	1			
65	17	21	26	18	29	20	3	6	4	7
66	17	21	26	19	29	2			2	
67	17	21	27	18	28	1	1			
68	17	22	23	20	29	1	1			
69	17	22	25	16	29	3	2		1	
70	17	22	25	17	29	1			1	
71	17	22	25	18	29	1			1	
72	17	22	25	19	29	5	1	3		1
73	17	22	26	18	29	1				1
74	17	22	26	19	29	1		1		
75	17	22	26	19	29	1		1		
76	17	23	23	20	29	1				1
77	17	23	25	16	29	2			1	1
78	17	23	25	18	29	1				1
79	17	23	25	19	29	1			1	
80	17	23	26	16	28	1	1			
81	17	23	26	18	29	1				1
82	17	24	22	16	29	1			1	
83	17	24	26	18	29	1			1	
84	17	25	23	19	29	1	1			
85	17	25	25	16	29	1	1			
86	18	18	24	18	29	1			1	
87	18	19	24	18	25	1			1	
88	18	19	24	19	29	1			1	
89	18	19	26	18	29	1			1	
90	18	20	23	17	29	1	1			
91	18	20	23	20	28	1	4		1	
92	18	20	23	22	29	1	1		4	
93	18	20	24	28	25	1			1	
94	18	20	25	18	29	3 F	0	4	చ 4	4
90	18	20	25	20	29	5 ∡	2	1	1	1
90 07	10 10	20 20	20	10	29	1	4	ן ר	2	4
31	10	∠∪ 20	20	10	29	1	1	2	3	Ţ
30	10	20 20	20 26	10	3U 20	1 2	I		1	n
99 100	18	20 20	20 27	20 18	29 20	5 1			1	2

	Number of repeats in the microsatellites							Number of X chromosomes				
Haplotype	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	RS	PE	HA		
101	18	21	24	18	29	1			1			
102	18	21	25	17	29	1		1				
103	18	21	25	18	29	2	1	1				
104	18	21	25	19	29	1	1					
105	18	21	25	20	28	1			1			
106	18	21	25	20	29	9	1	6	1	1		
107	18	21	25	_s 21	29	4		2	1	1		
108	18	21	_6	18	29	3	1	1	1			
109	18	21	26	18	30	1	·	•	•	1		
110	18	21	26	19	29	1		1		•		
110	18	21	26	20	27	1		•	1			
112	18	21	26	20	29	4	1		1	2		
113	18	21	20	18	28	1			1	2		
114	18	22	24	16	20	1			1			
115	18	22	24	18	20	י 2			י 2			
116	10	22	24	10	29	<u>-</u> 1			<u>د</u> 1			
117	10 19	~~ 22	25 25	20	29 20	י 2			י 2			
112	10	22	20	20 18	29	- J - 1			1			
110	10	22	20	20	29	1		1	I			
119	10	22	20	20	29	7	2	ו ס		1		
120	10	20	20	10	29	1	3	3		I		
121	10	23	20	10	30	1		I	4			
122	18	23	25	19	29	1		4	1			
123	18	23	25	20	29	1		1				
124	18	23	26	18	29	2		1	4	1		
125	18	24	23	19	29	1	4		1			
126	18	24	25	18	30	1	1					
127	18	24	26	16	29	1			4	1		
128	19	20	25	18	29	1			1			
129	19	20	25	18	30	1	1					
130	19	20	25	20	29	1		1				
131	19	20	25	21	29	1			1			
132	19	20	26	18	29	1			1			
133	19	21	24	18	29	2			1	1		
134	19	21	25	20	29	1		1				
135	19	21	25	21	29	2	1	1				
136	19	22	22	18	28	1	1					
137	19	22	24	17	30	1	1					
138	19	22	24	18	29	4	3	1				
139	19	22	25	17	29	1				1		
140	19	22	25	18	28	1			1			
141	19	22	25	19	29	1			1			
142	19	23	24	19	28	1	1					
143	19	23	25	18	29	3	3					
144	19	23	36	18	30	1	1					
145	19	24	22	18	29	1	1					
146	19	24	25	16	29	2	1		1			
147	19	24	25	18	29	1				1		
148	19	25	23	18	29	1			1			
149	20	20	23	17	29	1	1					
150	20	20	25	17	29	1		1				
151	20	20	25	18	29	1			1			
152	20	21	24	18	29	1			1			
	Ν	lumber o	f repeats	in the mic	es	Nur	nber of X o	chromosor	nes			
-----------	-------	-----------	-----------	------------------	-------	------	--------------------	--------------------	--------------------	--------------------		
Haplotype	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	Freq	SP	RS	PE	HA		
153	20	21	26	19	29	1	1					
154	20	22	25	19	29	1	1					
155	20	23	25	16	29	2			2			
156	20	23	25	18	29	2	2					
157	20	23	26	17	29	1				1		
158	20	24	25	16	28	1			1			
159	20	24	25	16	29	9	1		6	2		
160	20	24	26	16	29	1				1		
161	21	19	24	20	29	1				1		
162	21	20	25	16	29	1			1			
163	21	24	25	16	29	2	1		1			
	Nui	nber of i	ndividua	ls		403						
	Nur	nber of h	aplotype	es		163	78	101	149	73		
	Hapl	otype Di	versity (h ^a)			0.9570 ± 0.0158	0.9143 ± 0.0230	0.9479 ± 0.0137	0.9486 ± 0.0182		

^ьН_к ^an 12 15 16 17 18 19 21 22 24 25 26 27 28 29 9 10 13 14 20 23 IVS 1 Spanish (32) 1.4 62.7 22 7.0 5.7 1.4 282 56.0 33.6 Korean (8) 0.1 0.5 79.2 19.2 1.0 1000 Indian (25) 1.5 1.5 47.9 37.3 10.3 1.5 126 46.0 Chinese (33) 0.2 0.5 0.9 75.9 21.1 0.9 0.5 440 37.9 Asian (WM) 0.1 0.4 50.7 16.3 3.9 22.2 6.0 0.3 0.1 1566 66.2 *SP 0.4 0.4 0.4 1.7 51.5 25.5 11.7 6.3 2.1 239 65.1 *RS 3.0 0.5 6.4 61.5 24.1 3.0 1.5 201 55.8 *PE 0.5 0.5 0.5 1.0 9.2 50.0 21.8 6.8 8.7 1.0 68.2 206 *HA 9.8 64.2 14.1 4.3 6.5 1.1 92 56.0 WM 0.1 1.1 0.3 0.5 6.1 55.3 22.8 7.0 5.7 1.1 738 63.0 **IVS 13** Slavs (34) 3.0 57.0 19.0 11.0 10.0 96 63.0 English (35) 3.1 52.5 25.8 9.3 7.2 2.1 97 64.3 Italian (35) 5.6 61.0 22.2 2.8 5.6 2.8 36 57.1 New Zealand (23) 2.0 48.0 30.0 5.0 15.0 100 72.0 Spanish (36) 1.2 52.2 30.5 8.5 7.3 1.2 82 46.3 European (WM) 0.5 11.7 7.3 3.4 44.8 18.6 6.7 6.0 1.0 411 73.4 Chinese (24)** 1.0 1.0 9.0 53.0 32.0 3.0 1.0 100 61.0 Chinese (37) 3.3 69.1 24.1 2.5 1.0 245 68.0 Chinese (33) 0.2 0.2 1.1 5.7 66.5 17.0 2.0 5.0 0.5 1.8 440 52.4 Uzbekistan (34) 80.0 14.0 3.0 3.0 63 34.0 Thailand (35) 20.5 59.1 11.4 9.0 88 58.8 Turkish (38) 1.8 51.3 34.0 7.8 4.3 0.8 397 62.0 Japonese (39) 2.8 70.0 17.3 0.4 6.0 3.5 91 45.6 Korean (8) 0.2 0.3 2.3 67.6 19.3 2.1 5.7 0.1 2.4 1000 50.1

Table 5: Allelic frequencies and Weighted allele frequencies obtained for each ethnic group (European, Africans, Asiatic and Americans) used for the comparison with the four populations (SP= São Paulo, RS= Rio Grande Do Sul, PE=Pernambuco and HA= Haemophilia A patients) in IVS 1, IVS 13, IVS 22 and IVS 25.3 microsatellites:

146 Manus<u>crito</u>

	9	10	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	an	^ь Н к
Indian (40)									1.0		53.0	36.0		10.0							50	57.0
Indian (25)											2.0	4.8	34.5	22.3	29.5	2.7	1.4	1.4	1.4		126	46.0
Iranian (41)									0.9	5.2	46.9	36.2	8.2	1.7	0.9						116	64.0
Asian (WM)								0.1	0.4	3.4	60.7	21.5	5.0	5.3	1.8	1.5	0.1	0.1	0.1		2616	57.7
Africans Cau (26)							1.0	1.0		5.0	52.0	31.0	5.0	4.0	1.0						161	62.0
African Neg (26)									2.0	13.0	11.0	13.0	29.0	19.0	9.0		4.0				47	82.0
Africans (WM)							0.8	0.8	0.5	6.8	42.7	26.9	10.6	7.4	2.8		0.9				208	72.3
Canadian (42)								1.7	1.2	4.0	7.0	25.0	47.7	13.0	0.4						243	68.0
Mexican (43)										2.6	9.3	16.0	28.8	14.0	14.0	10.0	5.3				166	83.0
São Paulo (27)												7.8	3.1	4.7	21.9	14.1	25.0	17.2	6.2		64	83.0
American (WM)								0.9	0.6	3.0	6.9	19.5	35.1	12.2	8.1	5.4	5.2	2.3	0.8		473	80.6
*SP									0.8	2.9	42.1	24.2	12.5	9.2	6.2	1.7	0.4				240	73.5
*RS							1.0			5.9	54.5	27.3	5.9	5.4							203	61.8
*PE									1.9	6.8	46.2	18.9	9.7	7.3	8.7	0.5					206	72.4
*HA									2.2	7.6	50.0	19.6	5.4	8.7	6.5						92	69.9
WM							0.3		1.1	5.4	34.3	28.8	14.9	7.7	6.7	0.7	0.1				741	75.4
IVS 22																						
Caucasian (11)														1.3		30.7	66.7		1.3		75	45.0
English (35)																40.0	53.3	5.6	1.1		180	55.3
Italian (35)														2.5	5.0	22.5	65.0	5.0			80	52.1
New Zealand (23)											3.0	30.0	60.0	3.0	4.0						100	62.0
Spanish (36)															7.3	41.5	50.0	1.2			82	46.3
European (WM)											0.6	5.8	12.0	1.2	2.7	28.4	46.2	2.9	0.6		517	68.7
Chinese (24)											1.0					29.0	63.0	7.0			100	51.0
Chinese (37)															0.37	18.1	75.4	5.4	0.37	0.36	260	68.0
Chinese (33)														0.2	0.2	24.1	68.4	4.8	1.8	0.5	440	47.1
Thailand (35)																30.0	66.0	4.0			100	47.3
Turkish (38)														0.4	1.2	38.8	56.0	3.2	0.4		252	53.0
Indian (44)														1.0		30.0	65.0	3.0	1.0		50	48.0
N Indian (44)																33.0	67.0				37	44.0

147 Manuscrito

	9	10	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	an	^ь Н К
Indian (25)									38.3	51.7	2.9	7.1									126	42.8
Korean (8)														0.1	20.2	71.4	7.9	0.2	0.2		1000	44.3
Iranian (41)															5.1	21.8	48.7	20.5	3.9		116	67.0
Asian (WM)									1.9	2.6	0.2	0.4		0.1	8.6	43.4	38.7	3.3	0.7	0.1	2481	65.3
African Cau (26)															2.0	35.0	57.0	4.0	2.0		161	54.0
African Neg (26)									2.0				2.0	11.0	23.0	47.0	9.0	6.0			47	70.0
Africans (WM)									0.5				0.5	2.5	6.7	37.7	46.1	4.5	1.5		208	63.7
Canadian (42)														6.0	54.0	35.0	3.0	2.0			116	59.0
Mexican (43)														13.1	48.8	34.2	3.9				166	62.8
São Paulo (27)												2.5	5.1	6.3	7.6	22.8	21.5	31.7	2.5		79	78.8
American (WM)												0.5	1.1	9.3	41.5	32	7.5	7.6	0.5		361	70.6
*SP													2.5	6.7	8.3	37.9	40.4	3.8		0.4	240	68
*RS														0.5	4.5	35.8	54.7	3.0	1.0	0.5	199	57
*PE									0.5				1.0	2.4	10.2	36.4	45.1	3.9	0.5		206	65.1
*HA														5.4	4.3	22.8	61.0	4.3	2.2		92	57.7
WM									0.1				1.1	3.7	7.3	35.0	48.1	3.7	0.7	0.3	737	63.7
IVS 25.3																						
Spanish (12)							3.2		75.8	9.0	10.5	1.5									100	41.0
Korean (8)	0.1					2.1	2.6	1.1	79.1	6.7	7.1	0.7	0.4	0.1							1000	33.6
Chinese (33)					0.2	2.7	3.4	0.7	76.1	5.7	9.8	0.9	0.5								440	40.5
Asi (WM)	0.1				0.1	2.3	2.8	1.0	78.1	6.4	7.9	0.8	0.4	0.1							1440	37.7
*SP						0.4	7.6	4.7	63.1	10.2	12.3	1.3	0.4								236	56.8
*RS			0.5				2.0	2.5	65.7	11.9	14.9	2.5									200	53
*PE							13.1	1.9	59.2	11.7	10.7	2.4	1.0								206	60.6
*HA							6.5	5.4	64.2	7.6	14.1	1.1	1.1								92	56.3
WM			0.1			0.1	7.5	3.4	62.9	10.8	12.8	1.9	0.5								734	57.0

Eur=Europeans. Asi=Asians. Afr=African. Ame=American and Spa=Spanish. ^an = number of X chromosome. ^bHs = Expected Heterozygosity WM = weighed means * present study data ** Not included in the weighed means



Figure 1: Physical positions in chromosome X of the IKBKG, IVS 25.3, IVS 22, IVS 13 and IVS 1 microsatellites. The numbers correspond to pairs of bases from the Xp-Tel using the NC_000023 reference sequence.



Figure 2: Pedigree and haplotype analysis of the IVS 1, IVS 13, IVS 22, IVS 25.3 and IKBKG microsattelites in two Hemophilic A families (A and B). The Gray box represents the mutation haplotype.

ANEXOS

Anexo I: Tabelas 1 a 4. Frequências alélicas e médias ponderadas das frequências alélicas obtidas para cada grupo continental (europeu, africano, asiático e americano), utilizados para a comparação com as doze populações (SP: São Paulo, PE: Pernambuco, RS: Rio Grande do Sul, HA: Hemofilicos A, Mi: Mimbó, SV: Sítio Velho, GA: Gaucinha, UM: Umariaçu, FE: Feijoal, VE: Vendaval, BA: Baníwa e KA: Kashináwa) nos microssatélites IVS 1, IVS 13, IVS 22 e IVS 25.3. Onde $n = n^{\circ}$ de cromossomos X, $H_K =$ Heterozigose esperada e MP = Média ponderada. O número entre parênteses, presentes ao lado de cada amostra, indica a numeração da referência bibliográfica.

IVS 1	9	10	12	13	14	15	16	17	18	19	20	21	n	Hκ
Espanhóis [90]						1,4	62,7	22	7,0	5,7	1,4		282	56,0
Europeus						1,4	62,7	22	7,0	5,7	1,4		282	56,0
Coreanos [46]			0,1	0,5	79,2	19,2	1,0						1000	33,6
Indianos [80]				1,5	1,5	47,9	37,3	10,3	1,5				126	46,0
Chineses [52]					0,2	0,5	0,9	75,9	21,1	0,9	0,5		440	37,9
Asiáticos (MP)			0,1	0,4	50,7	16,3	3,9	22,2	6,0	0,3	0,1		1566	66,2
SP			0,4		0,4	0,4	1,7	52,1	25,0	12,1	5,8	2,1	239	65,0
RS			2,9	1,0		0,5	6,3	61,1	23,8	2,9	1,5		205	56,6
PE		0,5	0,5		0,5	1,0	9,2	50,0	21,8	6,8	8,7	1,0	206	68,5
НА							9,8	64,2	14,1	4,3	6,5	1,1	92	56,0
Amostras Urbanas (MP)		0,1	1,1	0,3	0,3	0,5	6,1	55,5	22,4	7,1	5,5	1,1	742	63,0
MI	2,5		2,5		1,3		5,0	37,4	22,5	13,8	10,0	5,0	81	78,6
SV						1,4	5,6	34,7	20,8	22,2	13,9	1,4	72	77,5
GA	10,5	5,3	5,3				5,3	23,7	34,1	13,2	2,6		37	81,7
UM				1,4				34,3	37,1	18,6	8,6		71	71,3
FE							8,6	36,2	32,8	22,4			59	71,9
VE							16,7	43,3	30,0	10,0			30	72,1
BA		7,4					2,9	57,4	27,9	4,4			69	59,4
KA		3,3						70,0	26,7				30	46,2

Tabela A1.1. IVS 1.

Tabela A1.2. IVS 13.

IVS 13	14	16	17	18	19	20	21	22	23	24	25	26	27	28	n	Нк
Eslavos [6]					3,0	57,0	19,0	11,0	10,0						96	63,0
Ingleses [35]					3,1	52,5	25,8	9,3	7,2	2,1					97	64,3
Italianos [35]					5,6	61,0	22,2	2,8	5,6	2,8					36	57,1
Neozelandeses [38]		2,0	48,0	30,0	5,0	15,0									100	72,0
Espanhóis [81]					1,2	52,2	30,5	8,5	7,3	1,2					82	46,3
Europeus (MP)		0,5	11,7	7,3	3,4	44,8	18,6	6,7	6,0	1,0					411	73,4
Chineses [94] **								1,0	1,0	9,0	53,0	32,0	3,0	1,0	100	61,0
Chineses [54]					3,3	69,1	24,1	2,5		1,0					245	68,0
Chineses [52]	0,2		0,2	1,1	5,7	66,5	17,0	2,0	5,0	0,5	1,8				440	52,4
Uzbequistaneses [6]						80,0	14,0	3,0	3,0						63	34,0
Tailandeses [35]					20,5	59,1	11,4	9,0							88	58,8
Turcos [43]					1,8	51,3	34,0	7,8	4,3	0,8					397	62,0
Japoneses [82]					2,8	70,0	17,3	0,4	6,0		3,5				91	45,6
Coreanos [46]			0,2	0,3	2,3	67,6	19,3	2,1	5,7	0,1	2,4				1000	50,1
Indianos [44]				1,0		53,0	36,0		10,0						50	57,0
Indianos [80]						2,0	4,8	34,5	22,3	29,5	2,7	1,4	1,4	1,4	126	46,0
Iranianos [73]				0,9	5,2	46,9	36,2	8,2	1,7	0,9					116	64,0
Asiáticos (MP)			0,1	0,4	3,4	60,7	21,5	5,0	5,3	1,8	1,5	0,1	0,1	0,1	2616	57,7
Africanos Cau [18]		1,0	1,0		5,0	52,0	31,0	5,0	4,0	1,0					161	62,0
Africanos Neg [18]				2,0	13,0	11,0	13,0	29,0	19,0	9,0		4,0			47	82,0
Africanos (MP)		0,8	0,8	0,5	6,8	42,7	26,9	10,6	7,4	2,8		0,9			208	72,3
Canadenses [97]			1,7	1,2	4,0	7,0	25,0	47,7	13,0	0,4					243	68,0
Mexicanos [43]					2,6	9,3	16,0	28,8	14,0	14,0	10,0	5,3			166	83,0
Paulistanos [87]							7,8	3,1	4,7	21,9	14,1	25,0	17,2	6,2	64	83,0
Americanos (MP)			0,9	0,6	3,0	6,9	19,5	35,1	12,2	8,1	5,4	5,2	2,3	0,8	473	80,6
SP				0,8	2,9	42,1	24,2	12,5	9,2	6,2	1,7	0,4			239	73,9
RS		1,0			5,8	54,4	27,7	5,8	5,3						205	62,1
PE				1,9	6,8	46,2	18,9	9,7	7,3	8,7	0,5				206	72,8
HA				2,2	7,6	50,0	19,6	5,4	8,7	6,5					92	69,9
Amostras Urbanas (MP)		0,3		1,1	5,4	34,1	28,9	15,0	7,7	6,7	0,7	0,1			742	75,4
MI					23,2	13,4	26,7	7,3	9,8	15,9	3,7				82	82,7
SV					16,2	13,5	18,9	20,1	12,2	12,2	4,1	1,4	1,4		74	86,0
GA					15,8	18,4	23,7	23,6	13,2	5,3					37	83,2
UM					9,7	45,8	40,3		2,8			1,4			72	62,6
FE					6,9	46,6	41,4	1,7	3,4						59	61,8
VE					10,0	43,3	46,7								30	61,4
BA				1,4	11,4	43,0	27,1	10,0	7,1						70	72,6
KA					36,7	50,0	13,3								30	61,9

** Não incluída na média ponderada

Tabela A1.3. IVS 22.

IVS 22	18	19	20	21	22	23	24	25	26	27	28	29	n	Нк
Caucasianos [49]						1,3		30,7	66,7		1,3		75	45,0
Ingleses [35]								40,0	53,3	5,6	1,1		180	55,3
Italianos [35]						2,5	5,0	22,5	65,0	5,0			80	52,1
Neozelandeses [38]			3,0	30,0	60,0	3,0	4,0						100	62,0
Espanhóis [81]							7,3	41,5	50,0	1,2			82	46,3
Europeus (MP)			0,6	5,8	12,0	1,2	2,7	28,4	46,2	2,9	0,6		517	68,7
Chineses [94]			1,0					29,0	63,0	7,0			100	51,0
Chineses [54]							0,37	18,1	75,4	5,4	0,37	0,36	260	68,0
Chineses [52]						0,2	0,2	24,1	68,4	4,8	1,8	0,5	440	47,1
Tailandeses [35]								30,0	66,0	4,0			100	47,3
Turcos [43]						0,4	1,2	38,8	56,0	3,2	0,4		252	53,0
Indianos [44]						1,0		30,0	65,0	3,0	1,0		50	48,0
Indianos [15]								33,0	67,0				37	44,0
Indianos [80]	38,3	51,7	2,9	7,1									126	42,8
Coreanos [46]						0,1	20,2	71,4	7,9	0,2	0,2		1000	44,3
Iranianos [73]							5,1	21,8	48,7	20,5	3,9		116	67,0
Asiáticos (MP)	1,9	2,6	0,2	0,4		0,1	8,6	43,4	38,7	3,3	0,7	0,1	2481	65,3
Africanos Cau [18]							2,0	35,0	57,0	4,0	2,0		161	54,0
Africanos Neg [18]	2,0				2,0	11,0	23,0	47,0	9,0	6,0			47	70,0
Africanos (MP)	0,5				0,5	2,5	6,7	37,7	46,1	4,5	1,5		208	63,7
Canadenses [97]						6,0	54,0	35,0	3,0	2,0			116	59,0
Mexicanos [32]						13,1	48,8	34,2	3,9				166	62,8
Paulistanos [87]				2,5	5,1	6,3	7,6	22,8	21,5	31,7	2,5		79	78,8
Americanos (MP)				0,5	1,1	9,3	41,5	32	7,5	7,6	0,5		361	70,6
SP					2,5	6,7	8,3	37,9	40,4	3,8		0,4	239	68,2
RS						0,5	4,5	35,8	54,7	3,0	1,0	0,5	199	57,3
PE	0,5				1,0	2,4	10,2	36,4	45,1	3,9	0,5		206	65,4
HA						5,4	4,3	22,8	61,0	4,3	2,2		92	57,7
Amostras Urbanas (MP)	0,1				1,1	3,7	7,3	35,0	48,1	3,7	0,7	0,3	736	63,7
MI					3,7	25,6	14,6	43,9	9,8			2,4	82	72,0
SV						9,5	25,7	51,2	12,2	1,4			74	65,6
GA		2,6				10,5	13,2	50,0	15,8	5,3		2,6	37	71,2
UM							2,7	51,4	43,2	2,7			74	55,7
FE								25,8	56,0	15,2	3,0		66	60,8
VE								50,0	40,0	10,0			30	61,4
BA							1,4	24,3	50,0	18,6	5,7		70	66,3
KA								41,6	33,3	16,7	4,2	4,2	24	72,7

Tabela A1.4. IVS 25.3.

IVS 25.3	9	12	14	15	16	17	18	19	20	21	22	23	n	Ηκ
Europeus - Espanhóis [93]					3,2		75,8	9,0	10,5	1,5			100	41,0
Coreanos [46]	0,1			2,1	2,6	1,1	79,1	6,7	7,1	0,7	0,4	0,1	1000	33,6
Chineses [52]			0,2	2,7	3,4	0,7	76,1	5,7	9,8	0,9	0,5		440	40,5
Asiáticos (MP)	0,1		0,1	2,3	2,8	1,0	78,1	6,4	7,9	0,8	0,4	0,1	1440	37,7
SP				0,4	7,6	4,7	63,1	10,2	12,3	1,3	0,4		236	57,0
RS		0,5			2,0	2,5	67,0	11,0	15,0	2,0			202	51,8
PE					13,1	1,9	59,2	11,7	10,7	2,4	1,0		206	60,9
HÁ					6,5	5,4	64,2	7,6	14,1	1,1	1,1		92	56,3
Amostras Urbanas (MP)		0,1		0,1	7,5	3,4	63,2	10,5	12,8	1,8	0,6		736	56,6
MI					18,8	2,5	28,7	8,7	41,3				81	71,4
SV			1,4		22,9	4,3	40,0	21,4	10,0				71	74,1
GA					2,9	5,9	53,0	17,6	14,7	5,9			34	67,8
UM						4,2	44,4	5,6	44,4	1,4			73	61,0
FE							50,0	9,4	35,9	4,7			64	61,8
VE							54,1	4,2	41,7				24	55,7
BA						5,7	44,3	10,0	40,0				70	64,0
KA						14,3	14,3	14,3	57,1				28	64,0

Anexo 2.

Alelo/População IVS 1 *9	IVS 1	IVS 13	IVS 22	IVS 25.3	IKBKG	n	Haplótipos possíveis
Mimbó	0921	1922	2325	1820	2929	1	192318
Mimbó	0921					1	192320
Mimbó	0921					1	192518
Gaucinha	0909	1921	2525	1818	2929	2	192518
Mimbó	0921					1	192520
Gaucinha	0918	2022	2325	1718	2929	3	202317
Gaucinha	0918					3	202318
Gaucinha	0918					3	202517
Gaucinha	0918					3	202518
Gaucinha	0909					2	212518
Gaucinha	0918					3	222317
Mimbó	0921					1	222318
Gaucinha	0918					3	222318
Mimbó	0921					1	222320
Gaucinha	0918					3	222517
Mimbó	0921					1	222518
Gaucinha	0918					3	222518
Mimbó	0921					1	222520
Mimbó	09	23	24	18	29	4	232418
Gaucinha	09	24	25	19	29	5	242519
IVS 1 *10							
Baníwa						2	182518
Baníwa	1017	1820	2526	1820	2929	2	182520
Baníwa						2	182618
Baníwa						2	182620
Kashináwa	10	19	27	18	29	8	192718
Baníwa						2	202518
Baníwa	1017	2022	2526	1818	2929	3	202518
Baníwa						2	202520
Baníwa						2	202618
Baníwa						3	202618
Baníwa	1017	2023	2627	1919	2929	4	202619
Baníwa						2	202620
Baníwa						4	202719
Baníwa						3	222518
Baníwa						3	222618
Baníwa	10	22	26	18	29	6	222618
Gaucinha	1010	2323	2525	1818	2929	1	232518
Pernambuco	10	23	25	18	29	5	232518
Baníwa	10	23	25	18	29	7	232518
Baníwa						4	232619

Tabela A2.1. Lista dos Haplótipos passiveis de ocorrência criada a partir da composição alélica dos indivíduos que apresentaram os alelos novos *IVS 1 *9, IVS 1 *10, IVS 1 *21* e *IVS 25.3 *12*.

Baníwa						4	232719
IVS 1 *21							
Mimbó	1821	1919	2222	2020	2829	8	192220
Mimbó						8	192220
São Paulo						3	192317
São Paulo	1721	1921	2325	1820	2929	2	192318
São Paulo						2	192318
Mimbó	0921	1922	2325	1820	2929	5	192318
São Paulo						2	192320
São Paulo	2021	1924	2325	1720	2929	3	192320
Mimbó						5	192320
Mimbó	1821	1921	2325	2020	2929	6	192320
Mimbó	1821	1921	2325	2020	2929	7	192320
Hemofílicos A	21	19	24	20	29	12	192420
São Paulo						3	192517
São Paulo						2	192518
Mimbó						5	192518
São Paulo						2	192520
São Paulo						3	192520
Mimbó						5	192520
Mimbó						6	192520
Mimbó						7	192520
Pernambuco	21	20	25	16	29	10	202516
São Paulo						2	212318
São Paulo						2	212320
Mimbó						6	212320
Mimbó						7	212320
São Paulo						2	212518
São Paulo						2	212520
Mimbó						6	212520
Mimbó						7	212520
Mimbó						5	222318
Mimbó						5	222320
São Paulo	1721	2224	2525	1616	2929	1	222516
São Paulo	1821	2224	2525	1720	2929	4	222517
Mimbó						5	222518
São Paulo						4	222520
Mimbó						5	222520
Sítio Velho	21	23	25	18	29	13	232518
São Paulo						3	242317
São Paulo						3	242320
São Paulo						1	242516
São Paulo	21	24	25	16	29	9	242516
Pernambuco	21	24	25	16	29	11	242516
São Paulo						3	242517
São Paulo						4	242517
São Paulo						3	242520
São Paulo						4	242520

IVS 25.3 *12							
Rio Grande do Sul	1718	2020	2525	1218	2933	1	17202529
Rio Grande do Sul						1	17202533
Rio Grande do Sul						1	18202529
Rio Grande do Sul						1	18202533