

**IDENTIFICAÇÃO DE MARCADORES SNPSTRs NO GENE *F8* E A SUA  
UTILIZAÇÃO PARA A DETERMINAÇÃO DA FASE GAMÉTICA DE  
CROMOSSOMOS X EM MULHERES PORTADORAS DE MUTAÇÕES  
CAUSADORAS DE HEMOFILIA A**

**GRAZIELA DE SÁ MACHADO ARAÚJO**

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY  
RIBEIRO  
CAMPOS DOS GOYTACAZES  
MARÇO 2013**



**IDENTIFICAÇÃO DE MARCADORES SNPSTRs NO GENE *F8* E A SUA  
UTILIZAÇÃO PARA A DETERMINAÇÃO DA FASE GAMÉTICA DE  
CROMOSSOMOS X EM MULHERES PORTADORAS DE MUTAÇÕES  
CAUSADORAS DE HEMOFILIA A**

**GRAZIELA DE SÁ MACHADO ARAÚJO**

Dissertação de Mestrado apresentado ao centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro como parte das exigências para obtenção do título de Mestre em Biociências e Biotecnologia, sob orientação do Prof. Dr. Enrique Medina-Acosta.

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY  
RIBEIRO  
CAMPOS DOS GOYTACAZES  
MARÇO 2013**

# UTILIZAÇÃO PARA A DETERMINAÇÃO DA FASE GAMÉTICA DE CROMOSSOMOS X EM MULHERES PORTADORAS DE MUTAÇÕES CAUSADORAS DE HEMOFILIA A

**GRAZIELA DE SÁ MACHADO ARAÚJO**

Dissertação de Mestrado apresentado ao centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro como parte das exigências para obtenção do título de Mestre em Biociências e Biotecnologia, sob orientação do Prof. Dr. Enrique Medina-Acosta.

**Aprovada em 25 de Março de 2013**

**Comissão Examinadora:**

---

Álvaro Fabrício Lopes Rios – (Doutor em Genética e Epigenética Animal) – UENF

---

Victor Martin Quintana Flores – (Doutor em Biociências e Biotecnologia) - UENF

---

Regina Célia de Souza Campos Fernandes (Doutora em Doenças Infecciosas) – Faculdade de Medicina de Campos

---

Enrique Medina-Acosta (Doutor em Parasitologia Médica e Molecular) – UENF. (Orientador).

## AGRADECIMENTOS

Agradeço a Deus por ter colocado pessoas tão especiais na minha vida.

Aos meus pais Gustavo e Josely por não terem medido esforços para que eu completasse com sucesso mais uma etapa da minha vida.

Aos meus irmãos Marcelle, Gustavo Filho, Gustavo Neto pelo convívio harmonioso e pela união.

Aos meus avós Antônio e Marly por estarem sempre presentes na minha vida.

Ao meu Tio Marquinho e minha madrinha Giselma minha gratidão.

Ao Mestre Prof. Enrique Medina-Acosta meu sincero agradecimento pela lição de sabedoria, pela orientação no momento exato, pela dedicação diária, pelo incentivo nas horas difíceis com palavras de força e em especial por compartilhar suas experiências de vida e me auxiliar a trilhar meu caminho.

A minha colega e amiga Thais pelo exemplo de companheirismo no laboratório, pela paciência em compartilhar sua experiência não medindo esforços para que fosse possível alcançar minhas metas nos experimentos.

A mãezona do NUDIM, Maria e a todos amigos : Thais, Cláudia, Viviane, Laís, Fabrício, Thiago, Viviane, Antônio, Priscila , Katrine e Milena pelo compartilhamento de experiências e conhecimento.

Aos professores da UENF pelas aulas e ensinamentos.

À Universidade Estadual Norte Fluminense Darcy Ribeiro pela bolsa fornecida e pelo ensino transmitido.

Para finalizar gostaria de agradecer a todos que contribuíram também indiretamente para a conclusão deste trabalho.

“Um dos paradoxos dolorosos do nosso tempo reside no fato de serem os estúpidos os que têm a certeza, enquanto os que possuem imaginação e inteligência se debatem em dúvidas e indecisões.”

(Bertrand Russell)

"A mente que se abre a uma nova idéia jamais voltará ao seu tamanho original."

(Albert Einstein 1879-1955)

## SUMÁRIO

<b>LISTA DE FIGURAS E GRÁFICOS.....</b>	<b>ix</b>
<b>LISTAS DE TABELAS.....</b>	<b>xi</b>
<b>RESUMO.....</b>	<b>xii</b>
<b>1. INTRODUÇÃO.....</b>	<b>1</b>
1.1. Hemofilia A.....	1
1.2. Marcadores moleculares.....	2
1.2.1. STR.....	2
1.2.2. SNP.....	2
1.2.3. SNPSTR.....	3
1.3. Haplótipo.....	5
1.4. Desequilíbrio Gamético.....	7
1.5. Diagnóstico molecular indireto em Hemofilia A.....	8
1.5.1. Aplicação dos marcadores para diagnóstico indireto da Hemofilia A.....	9
<b>2. OBJETIVOS.....</b>	<b>10</b>
2.1. Objetivo geral.....	10
2.1. Objetivos específicos.....	10
<b>3. MATERIAL E MÉTODOS.....</b>	<b>11</b>
3.1. Material biológico e extração de DNA.....	11
3.2. Aspectos éticos da pesquisa.....	11
3.3. Critérios de seleção para locos potencialmente polimórficos.....	11
3.4. Iniciadores marcador-específicos.....	12
3.5. Tipagem dos marcadores.....	12
3.6. Caracterização dos alelos amplificados.....	12
3.7. Digestão enzimática.....	13
3.8. Frequências alélicas e dados populacionais.....	13
<b>4. RESULTADOS.....</b>	<b>14</b>
4.1. Varredura e desenho de marcadores.....	14
4.2. Padronização do ensaio da PCR-RFLP para os marcadores STR.....	18
4.3. Padronização de marcadores SNP.....	27
4.4 Análise dos SNPSTRs.....	32

4.5. Dados populacionais relativos ao ensaio pentaplex de Hemofilia A .....	36
<b>5. DISCUSSÃO .....</b>	<b>39</b>
<b>6. CONCLUSÕES .....</b>	<b>41</b>
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>42</b>
<b>ANEXO .....</b>	<b>45</b>



## LISTA DE FIGURAS E GRÁFICOS

<b>Figura 1.</b> Esquema representativo da determinação da fase gamética de uma família utilizando cinco marcadores SNPSTR. ....	4
<b>Figura 2.</b> Eletroferograma representativo da genotipagem do microsatélite dinucleotídeo (CA) <sub>n</sub> e do SNP rs60186903 .....	5
<b>Figura 3.</b> Esquema representativo da segregação de haplótipos SNPSTR em duas gerações de Hemofilia A .....	6
<b>Figura 4.</b> Esquema representativo da marcação dos iniciadores para a genotipagem do SNPSTR por PCR-RFLP .....	12
<b>Figura 5.</b> Sequência do produto de amplificação <i>in silico</i> para os marcadores SNPSTR utilizados no diagnóstico indireto de Hemofilia A .....	18
<b>Figura 6.</b> Eletroferograma representativo da padronização das condições de amplificação do marcador DXS154082597. ....	20
<b>Figura 7.</b> Eletroferograma representativo da padronização das condições de amplificação do marcador DXS154087367. ....	21
<b>Figura 8.</b> Eletroferograma representativo da padronização das condições de amplificação do marcador DXS154088940. ....	22
<b>Figura 9.</b> Eletroferograma representativo da padronização das condições de amplificação do marcador DXS154103837. ....	23
<b>Figura 10.</b> Eletroferograma representativo da padronização das condições de amplificação do marcador DXS154124632. ....	24
<b>Figura 11.</b> Eletroferograma representativo do perfil alélico dos cinco marcadores STR em reação multiplex .....	25
<b>Figura 12.</b> Eletroferograma representativo do perfil alélico do pentaplex padronizado .	26
<b>Figura 13.</b> Eletroferogramas representativos da genotipagem do SNP rs28890032.....	28
<b>Figura 14.</b> Eletroferogramas representativos da genotipagem do SNP rs60186903.....	29
<b>Figura 15.</b> Eletroferogramas representativos da genotipagem do SNP rs5987052.....	30
<b>Figura 16.</b> Eletroferogramas representativos da genotipagem do SNP rs79260793.....	31
<b>Figura 17.</b> Eletroferogramas representativos da genotipagem do SNP rs112922881 ....	31
<b>Figura 18.</b> Eletroferograma representativo do perfil alélico de uma mulher genotipada com marcadores STR em reação pentaplex .....	33
<b>Figura 19.</b> Eletroferograma representativo do perfil alélico gerado por PCR-RFLP .....	34

**Figura 20.** Heredograma da segregação de haplótipos em núcleo familiar com um acometido pela Hemofilia A.....35

**Gráfico 1.** Relação entre a pontuação e a heterozigose observada.....38

**LISTAS DE TABELAS**

<b>Tabela 1.</b> Marcadores SNPSTR selecionados por rastreamento in silico. ....	15
<b>Tabela 2.</b> Informações dos cinco marcadores SNPSTR utilizados para o gene <i>F8</i> . ....	16
<b>Tabela 3.</b> Efeito das variáveis temperatura de anelamento e concentração de MgCl <sub>2</sub> na genotipagem do marcador DXS154082597. ....	20
<b>Tabela 4.</b> Efeito das variáveis temperatura de anelamento e concentração de MgCl <sub>2</sub> na genotipagem do marcador DXS154087367. ....	21
<b>Tabela 5.</b> Efeito das variáveis temperatura de anelamento e concentração de MgCl <sub>2</sub> na genotipagem do marcador DXS154088940. ....	22
<b>Tabela 6.</b> Efeito das variáveis temperatura de anelamento e concentração de MgCl <sub>2</sub> na genotipagem do marcador DXS154103837. ....	23
<b>Tabela 7.</b> Efeito das variáveis temperatura de anelamento e concentração de MgCl <sub>2</sub> na genotipagem do marcador DXS154124632. ....	24
<b>Tabela 8.</b> Informações sobre os cinco marcadores SNPSTRs utilizados neste estudo.	32
<b>Tabela 9.</b> Frequências alélicas utilizando o ensaio pentaplex desenvolvido (n=73). ....	37
<b>Tabela 10.</b> Taxas de heterozigose e número de alelos de 5 microssatélites utilizados na análise de ligação para mutações no gene <i>F8</i> . ....	37
<b>Tabela 11.</b> Frequência dos marcadores SNPs. ....	38

## RESUMO

A Hemofilia A é uma coagulopatia recessiva, ligada ao cromossomo X causada por mutações no gene *F8*, que codifica o fator VIII de coagulação. Os homens são afetados e as mulheres raramente. Inversões envolvendo introns 22 e 1 são responsáveis por cerca de 50% dos casos de deficiência de fator VIII. A identificação de portadores do sexo feminino é importante para o aconselhamento genético. As mutações podem ser identificadas por sequenciamento. A portadora pode também ser identificada por métodos moleculares indiretos como a análise de ligação de marcadores polimórficos em núcleos familiares com pelo menos um homem afetado. Esta análise baseia-se fortemente em duas premissas: em primeiro lugar, que a mutação presente no homem afetado está presente em um dos cromossomos maternos e, segundo, que não há nenhuma recombinação entre os marcadores polimórficos do gene *F8*. É importante notar que em mulheres, a análise de ligação de marcadores polimórficos define diplotipos e não haplótipos, porque este método não pode determinar a fase gamética do alelo ligado ao gene *F8* defeituoso. Portanto, a utilização dessas premissas implica em risco de erro de diagnóstico com relação à portabilidade do gene defeituoso. Além disso, a análise de ligação se torna mais difícil nos casos que faltam membros familiares. A genotipagem com marcadores combinados contendo um polimorfismo de um único nucleótido (SNP), ligado a um polimorfismo de sequências curtas repetidas em série (STR), chamados SNPSTRs, permitem diferenciar a fase gamética de cromossomos. Este projeto visa à identificação, validação e desenvolvimento de SNPSTRs marcadores localizados no gene *F8* para a determinação da fase gamética em portadoras de Hemofilia A. Computacionalmente foram identificados locos SNPSTR candidatos, utilizando sequências do genoma de referência. STRs foram genotipados por PCR quantitativa por fluorescência de SNPs por PCR-RFLP e analisado por eletroforese em capilar de alta performance. Cinco novos locos SNPSTR foram identificados e desenvolveu-se um ensaio de tipagem de um único tubo de Pentaplex. Ao utilizar o ensaio de Pentaplex a fase gamética de todos os cromossomos X estudados foi determinada. As frequências dos alelos e haplótipos também foram estimadas em mulheres não aparentadas. Estes resultados mostraram que a genotipagem com locos SNPSTR é um método rápido e preciso de determinação da fase gamética do cromossomo X. Este método é um ativo em diagnóstico molecular em que confere maior resolução para o aconselhamento genético de portadoras heterozigotas de Hemofilia A.

**Palavras-chave:** Fase Gamética, Hemofilia A, SNP, STR, variação.

## ABSTRACT

Hemophilia A is a recessive, X-linked coagulopathy caused by mutations in the *F8* gene, which codes for coagulation factor VIII. It affects men and rarely females. Inversions involving introns 22 and 1 together are responsible for about 50% of cases of factor VIII deficiency. The identification of female carriers is important for genetic counseling. Mutations can be identified by sequencing. Carrier status can also be indirectly determined by linkage analysis of polymorphic markers in nuclear families with at least one affected male. This analysis relies heavily on two assumptions: firstly, that the mutation present in the affected male is also present in one of the maternal chromosomes and secondly, that there is no recombination between the polymorphic markers and the *F8* gene. Importantly, in women, linkage analysis of polymorphic markers defines diplotypes and not haplotypes, because this method cannot determine the gametic phase of allele linked to the defective *F8* gene. Therefore, those assumptions imposed substantial risk of misdiagnosis with regards to the portability of the defective gene. Moreover, linkage analysis becomes more difficult in cases with missing family members. Genotyping with combined markers containing a single nucleotide polymorphism (SNP) linked to a polymorphic short tandem repeat (STR), called SNPSTRs, allow differentiating the gametic phase of chromosomes. This project was aimed at the identification, validation and development of SNPSTRs markers located in the *F8* gene for the determination of the gametic phase in female carriers of hemophilia A. Candidate SNPSTR loci were identified computationally using reference genome sequences. STRs were typed by fluorescence quantitative PCR and SNPs by PCR-RFLP and analyzed by high performance capillary electrophoresis. Five novel SNPSTR loci were identified and developed into a single-tube pentaplex typing assay. Using the pentaplex assay the gametic phases of all studied X chromosomes were determined. Allele and haplotype frequencies were also estimated in a small number of unrelated females. These results showed that genotyping with SNPSTR loci is a fast and precise method of determination of gametic phase of X chromosomes. This method is an asset in molecular diagnosis as it grants higher resolution for genetic counseling of heterozygous female carriers of hemophilia A.

Keywords: Gametic phase, Hemophilia A, SNP, STR, variation.

## 1. INTRODUÇÃO

### 1.1. Hemofilia A

A hemofilia A (OMIM 306700) é a desordem hemorrágica mais comum em humanos. Ela é causada por diversas mutações no gene *F8* que codifica uma glicoproteína fundamental (fator VIII) de uma rede funcional de proteínas estruturais e enzimas que integram a cascata de coagulação sanguínea (Goodeve *et al.*, 2003; Oldenburg *et al.*, 2004; Kemball-Cook, 2005; Machado e Medina-Acosta, 2009). Hemofilia A afeta aproximadamente um em cada 10.000 homens no Brasil (Saiki, 2010). O padrão de herança da hemofilia A é recessivo ligado ao cromossomo X.

O diagnóstico da doença é realizado pela classificação do nível de atividade coagulante do fator VIII (grave, com fator VIIIc menor que 1%; moderada, com fator VIIIc de 1-5%; leve, com fator VIIIc de 5-40%) (Barbosa De Souza *et al.*, 2011). Tendo em vista a alta incidência da doença e morbidade associada a ela, o aconselhamento genético é requerido (Saiki, 2010).

O diagnóstico molecular direto nos acometidos inclui a investigação de inversões, deleções, duplicações envolvendo os íntrons 22 e 1 do gene *F8* ou o sequenciamento do gene *F8* (Machado e Medina-Acosta, 2009). A identificação de portadoras é importante para aconselhamento genético. As portadoras podem ser identificadas por métodos moleculares indiretos como a análise de ligação de marcadores polimórficos em núcleos familiares com pelo menos um acometido. Essa análise se baseia fortemente em duas premissas: a primeira, que a mutação no acometido está presente em um dos cromossomos maternos; a segunda, que não há recombinação entre nenhum dos marcadores polimórficos e o gene *F8* (Machado, Alves *et al.*, 2009). Essas análises em mulheres resultam em diplótipos e não haplótipos, pois não se consegue determinar a fase gamética dos alelos ligados ao gene *F8* defeituoso. Portanto, a utilização dessas premissas implica em importante risco de erro diagnóstico. Todavia, essa análise se torna mais difícil nos casos em que faltam membros familiares.

## 1.2. Marcadores moleculares

### 1.2.1. STR

Os polimorfismos de sequências curtas de DNA repetidas em série (STR, microssatélites) são formados por unidades de repetição constituídas por mono-, di-, tri-, tetra-, penta- ou hexanucleotídeos (Butler, 2005).

Os locos STR variam em extensão devido a inserções ou deleções (indels) de uma ou mais unidades de repetição, que são causadas por erros de replicação pelo DNA polimerase num evento molecular conhecido como polimerase “*slippage*”. Assim sendo, STRs são sistemas multialélicos. Este evento é responsável por elevadas taxas de mutação (Leclercq *et al.*, 2007).

Os locos STR apresentam elevada taxa de heterozigose e presença de alelos múltiplos. Esta característica faz desses locos, sequências altamente informativas, e por isso, esses são priorizados em mapeamento gênico, investigação de parentesco, desequilíbrio gamético, e diagnóstico direto e indireto de anomalias genéticas (Ellegren, 2004).

A tipagem simultânea de microssatélites localizados no gene *F8* pode propiciar uma melhor forma no estudo da segregação do alelo afetado nas famílias. O nosso grupo já desenvolveu dois ensaios da PCR quantitativa fluorescente (QF-PCR) para microssatélites com informatividade combinada de 0,91 e 0,97 (Machado, Alves *et al.*, 2009; Machado, Duarte *et al.*, 2009a)

### 1.2.2. SNP

Os polimorfismos de um único nucleotídeo (SNP) são marcadores com baixa taxa de mutação ( $2,0-2,5 \times 10^{-8}$ ) por geração, isso significa que esses podem ser específicos de uma região geográfica e grupos populacionais. Portanto, são posições do genoma que estão bem conservadas na evolução (Odriozola *et al.*, 2009).

SNPs são em sua maioria bi-alélicos, onde a frequência do alelo mais recente é superior a 0,01. Esse tipo de marcador pode ser utilizado para distinguir mudanças em alelos nos descendentes (Odriozola *et al.*, 2009). Existem bancos de dados de SNPs contendo frequências dos alelos, onde as variabilidades em cada população são observadas (Agrafioti *et al.*, 2007).

### 1.2.3. SNPSTR

Os marcadores SNPSTRs são uma nova classe de polimorfismos sintênicos (isto é, na mesma fita de DNA) que consistem na combinação de um SNP e um STR fisicamente próximos (< 300 bp) (Mountain *et al.*, 2002). Por estarem tão próximos há pouca probabilidade de recombinação entre eles e, portanto, eles constituem sistemas haplotípicos, ou seja, co-segregam (Odriozola *et al.*, 2009). SNPSTRs são utilizados para determinação empírica rápida da fase gamética em indivíduos duplo heterozigotas (Mountain *et al.*, 2002). A determinação empírica da fase gamética é de extrema importância para aconselhamento genético de possíveis portadores de genes defeituosos causadores de doenças recessivas.

A taxa de mutação para cada STR por geração é maior do que o SNP, sendo aproximadamente  $1,5 \times 10^{-3}$ . Estas diferentes taxas demonstram que os STRs são marcadores mais informativos (Mountain *et al.*, 2002).

A determinação da fase gamética pela PCR-RFLP para marcadores SNPSTRs só é possível quando os indivíduos são heterozigotos para cada marcador. Devido às diferenças nas frequências alélicas entre SNPs e STRs, a resolução de fase gamética exige a tipagem de múltiplos SNPSTRs (Figura 1 e figura 2).



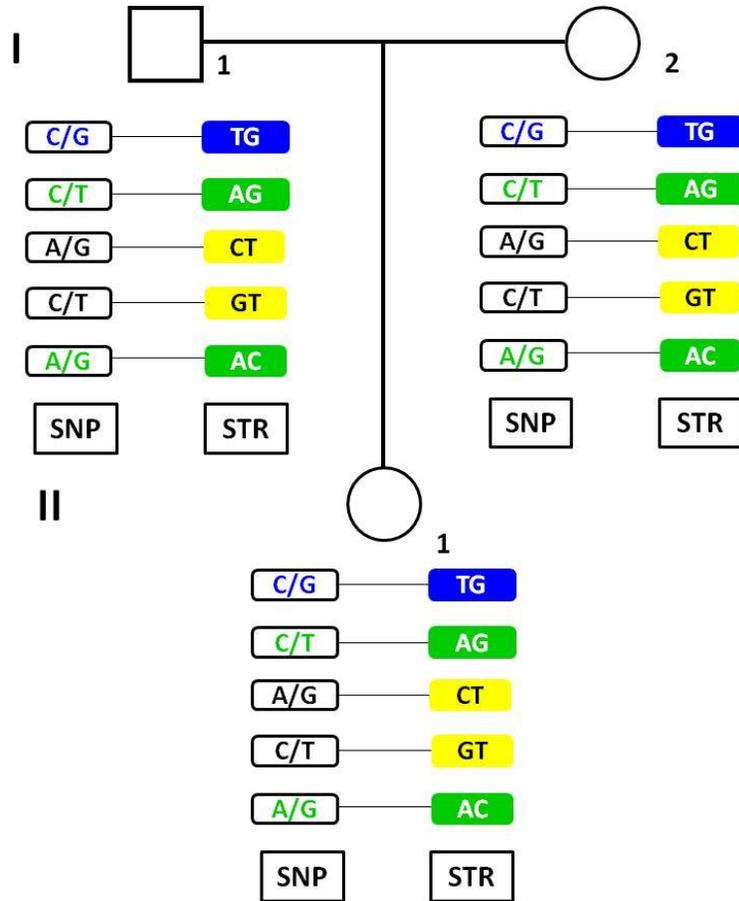
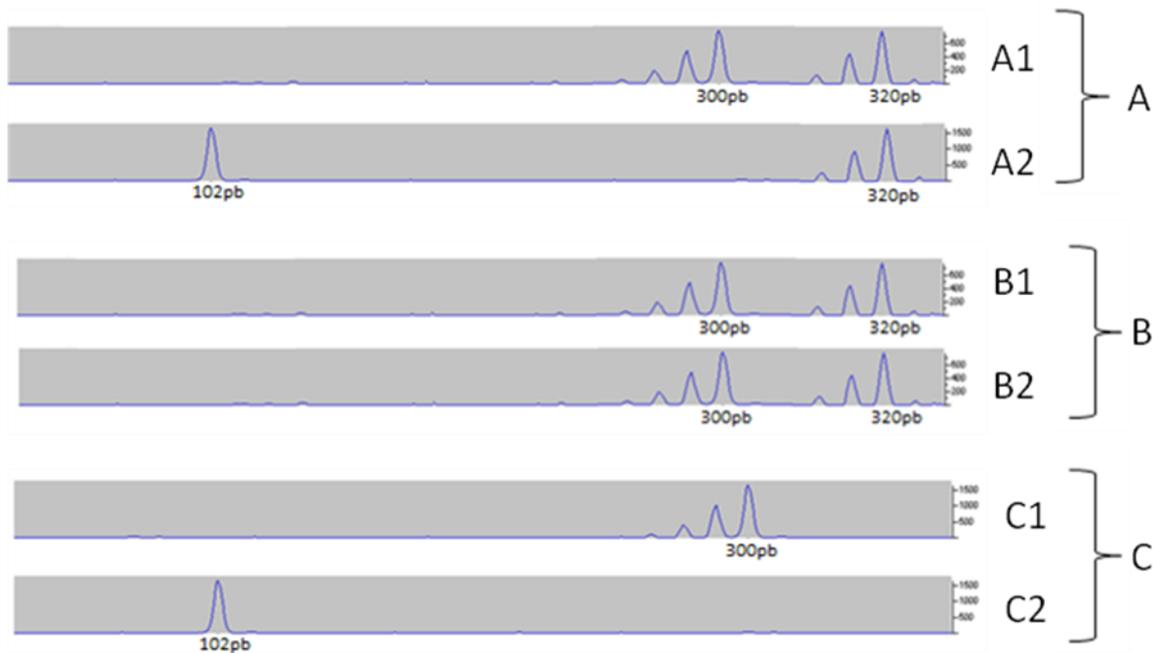


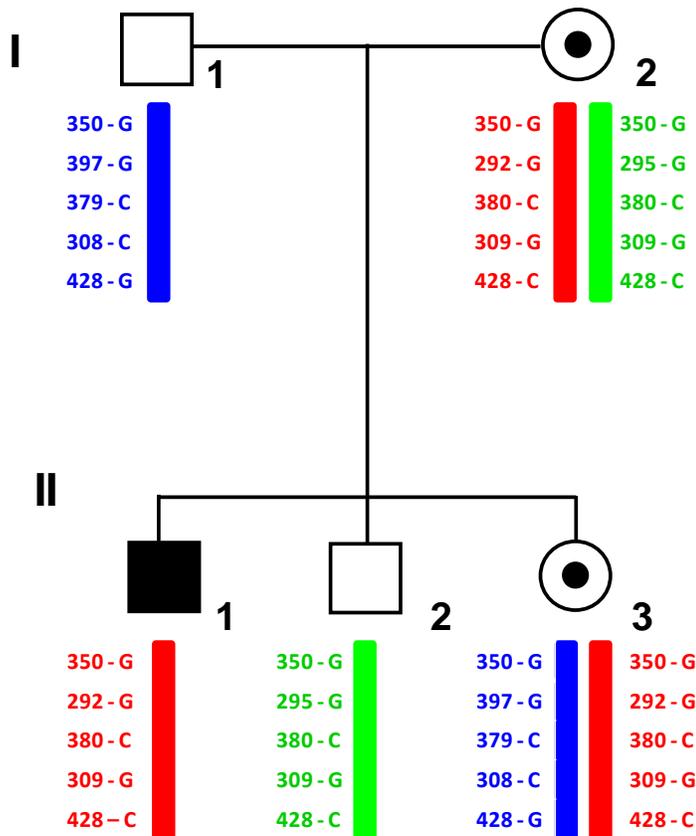
Figura 2. Esquema representativo da determinação da fase gamética de uma família utilizando cinco marcadores SNP/STR.



**Figura 3. Eletroferograma representativo da genotipagem do microssatélite dinucleotídeo (CA) $n$  e do SNP rs60186903.** O SNP apresenta sistema bi-alélico A/G. O alelo ancestral G consiste em sítio de restrição para enzima Cac8I. O indivíduo A é duplo heterozigoto para o marcador SNPSTR, apresentando os alelos **rs60186903** G/A e STR (CA) $n$  300/320 respectivamente. Em (A1) está ilustrado os alelos 300 e 320 antes da digestão com a enzima Cac8I, no entanto após a digestão com a enzima de restrição Cac8I (A2) o alelo 300pb é reduzido para 102 pb podendo inferir a fase gamética desses alelos, onde o alelo G está ligado ao alelo 300 e o alelo A está ligado ao 320. O indivíduo B é heterozigoto apenas para o marcador STR em (B1) visto que não houve clivagem para os alelos 300 e 320 pela enzima de restrição Cac8I pode inferir que é homozigoto para o alelo A do SNP (B2). O indivíduo C é homozigoto para o marcador SNPSTR. Em (C1) está ilustrado os alelos 300 antes da digestão, no entanto após a digestão os alelos são reduzidos a 102 podendo inferir que o alelo é G(C2).

### 1.3. Haplótipo

Haplótipo pode ser definido como um conjunto de alelos que são herdados como um grupo de um dos pais, ou seja, os genótipos de *loco* que segregam em conjunto formam haplótipos (Machado, 2008). O haplótipo constitui um padrão identificável de alelos que pode ser utilizado para rastrear a herança de todos os genes nele contidos. Esse conjunto de alelos é, em até certo ponto, independente ao que diz respeito a *desequilíbrio gamético* ou de recombinação (Nothnagel *et al.*, 2005). Normalmente os alelos são herdados de forma independente, mas em diferentes situações, pois estes não obedecem aos padrões de herança mendeliana. Marcadores em *desequilíbrio gamético* segregam em blocos estáveis, formando haplótipos (figura 3) (Szibor, 2007; Medina-Acosta, 2010).



**Figura 4. Esquema representativo da segregação de haplótipos SNPSTR em duas gerações de Hemofilia A.** Homens (quadrados); mulheres (círculos); indivíduos afetados (totalmente sombreadas); portador (sombreamento médio); Determinação da fase gamética de cromossomos X em mulheres suspeitas portadoras de Hemofilia A utilizando 5 marcadores SNPSTR hipotéticos.

Em princípio existem poucos haplótipos nos fundadores, mas eventos de recombinação entre pares de cromossomos homólogos tendem a desassociar haplótipos ao longo das gerações, com a inevitável consequência que alelos localizados muito perto uns dos outros permanecem juntos por maior tempo. A abordagem básica para estudar taxas de recombinação no genoma é construir um mapa genético por genotipagem, com uma alta densidade de marcadores, um grande número de indivíduos em famílias e então agrupar os dados no mapa físico correspondente (Kong *et al.*, 2002).

A análise de haplótipos é um método de mapeio de genes associados com doenças, podendo ser útil na identificação de regiões no genoma de indivíduos que se associa à ocorrência de alguma doença genética herdada, relacionada a mutações em um só gene (Ren *et al.*, 2007). O haplótipo que é constantemente herdado junto com a

doença pode então ser analisado com maior resolução molecular para encurtar a região em que o gene defeituoso deve se localizar. Isso significa que podemos procurar por regiões ou segmentos de DNA que são compartilhados entre os indivíduos acometidos para ajudar identificar a localização do gene defeituoso. Similarmente, a comparação de haplótipos entre acometidos pode servir para identificar o haplótipo fundador, isto é, aquele do ancestral comum, por que vários dos acometidos compartilham alelos comuns lado e lado do gene defeituoso (Medina-Acosta, 2010).

Considerando que o diagnóstico indireto de portadoras de mutações por meio da análise de ligação de marcadores polimórficos se baseia na identificação de haplótipos que co-segregam com o fenótipo no acometido (inferindo uma mutação herdável e excluindo mutação de novo), a análise de ligação exige a tipagem do maior número possível de marcadores informativos em cada família sob estudo (Machado e Medina-Acosta, 2009).

#### **1.4. Desequilíbrio Gamético**

O desequilíbrio gamético (DG) é a associação não aleatória entre diferentes locus, ou seja, há uma frequência de uma determinada combinação entre alelos, diferente à esperada pelo produto de suas frequências individuais. Portanto, marcadores em DG segregam em bloco formando haplótipos. Se os marcadores se mostrarem estar em forte desequilíbrio gamético, a frequência haplotípica não pode ser calculada com base nas frequências alélicas de cada locus individual, mas deve ser estimada diretamente (Hering, 2006). Os níveis de DG dependem da frequência do alelo. A frequência de um alelo tende a ser maior quando o alelo é mais ancestral (Payseur. *et al.*, 2008).

O mapa de ligação indica a distância genética em cM dos marcadores, que em porcentagem indica a frequência de recombinação. Quanto maior a taxa de recombinação menor a taxa de DG. O DG entre marcadores SNP e STR ao longo do Genoma Humano revela processos evolutivos e fornece informações cruciais para a associação de mapeamento de genes à doenças (Machado e Medina-Acosta, 2009) .

O DG está relacionado a polimorfismos ao longo do DNA na população. É importante ressaltar que o DG é reduzido por mutações recorrentes e depende do tipo de repetições do STR. Os marcadores SNPSTR são altamente informativos, pois têm

alta taxa de heterozigose, através desses marcadores haplótipos são observados para investigação do desequilíbrio gamético. As análises envolvem indivíduos não relacionados. A fase gamética dos haplótipos é determinada para cada par de alelos SNPSTR (Payseur. *et al.*, 2008).

### **1.5. Diagnóstico molecular indireto em Hemofilia A**

A análise indireta consiste na verificação de co-segregação de haplótipo, dentro de uma família, pela utilização de locos polimórficos para um gene marcador, para estabelecer se o marcador e a característica estão fisicamente ligados. Este método tem exercido um papel importante nas tentativas de mapear doenças complexas, que supostamente possuem um mecanismo mendeliano, com efeito significativo sobre a característica. A análise indireta requer a elaboração de heredograma, contendo informações quanto às linhagens paternas e maternas dos acometidos de maneira ascendente e descendente (Davies *et al.*, 1983; Schwartz *et al.*, 1992).

A identificação indireta de mutações por polimorfismos intragênicos é essencial para famílias onde nenhuma deleção é encontrada, e permite um aumento na confiabilidade no diagnóstico de portadoras de mutações em famílias com acometidos. No entanto, a possibilidade de recombinação intragênica entre os polimorfismos pode levar a erros no diagnóstico (Oudet *et al.*, 1991), problema esse que pode ser resolvido aumentando o número de marcadores polimórficos intragênicos para o gene analisado. E por isso uma análise de haplótipo é muito mais informativa quando comparada a uma investigação utilizando apenas um único marcador (Hong *et al.*, 1995).

Entretanto, a análise de ligação de marcadores polimórficos apresenta importantes limitações, pois ela pode ser utilizada apenas em famílias com histórico prévio da doença, não é utilizada em casos esporádicos (mutações *de novo*) (Kim *et al.*, 2002). O diagnóstico indireto da hemofilia A envolve a análise de ligação para marcadores polimórficos localizados dentro ou adjacentes do gene *F8* em indivíduos de famílias com pelo menos um acometido pela doença (Machado e Medina-Acosta, 2009).

### **1.5.1. Aplicação dos marcadores para diagnóstico indireto da Hemofilia A**

Em mulheres, a análise de ligação de marcadores polimórficos define diplótipos e não haplótipos, pois essa metodologia não consegue determinar a fase gamética dos alelos ligados ao gene *F8* defeituoso. Partindo das premissas de que a mutação no acometido está presente em um dos cromossomos maternos e de não haver recombinação entre nenhum dos marcadores polimórficos, implica em importante risco de erro diagnóstico na portabilidade do gene defeituoso. Todavia, essa análise se torna mais difícil nos casos em que faltam membros familiares. A tipagem de marcadores combinados SNPSTRs permite diferenciar a fase gamética dos cromossomos.

## **2. OBJETIVOS**

### **2.1. Objetivo geral**

Este projeto visou à identificação, validação e desenvolvimento de marcadores combinados SNPSTRs localizados no gene *F8* para a determinação da fase gamética de cromossomo X humanos em mulheres duplo heterozigotas e portadoras de mutações no gene *F8* causadoras de hemofilia A.

### **2.1. Objetivos específicos**

- Realizar uma varredura *in silico* do gene *F8* para identificação de marcadores combinados SNPSTR para determinação da fase gamética.
- Desenhar iniciadores específicos para os locos de SNPSTR encontrados no gene *F8*.
- Validar ensaio multiplex da QF-PCR para os novos marcadores desenhados para SNPSTR.
- Identificar perfis alélicos em mulheres da Região Norte Fluminense.
- Determinar parâmetros populacionais relevantes em amostras da Região Norte Fluminense

### **3. MATERIAL E MÉTODOS**

#### **3.1. Material biológico e extração de DNA**

Foram utilizadas 73 amostras de DNA genômico do banco de DNA do Núcleo de Diagnóstico e Investigação Molecular – NUDIM, extraídas de *swabs* (esfregaço bucal) utilizando o kit de extração e purificação de DNA genômico *Illustra blood genomicPrep*, GE Healthcare, UK, conforme a especificação do fabricante. Todas as amostras extraídas serão quantificadas pelo Thermo Scientific NanoDrop 2000c (Thermo Fisher Scientific) e estocadas a -20°C.

#### **3.2. Aspectos éticos da pesquisa**

O presente projeto foi submetido para autorização ao Comitê de Ética em pesquisa envolvendo seres humanos da Faculdade de Medicina de Campos, RJ, sob FR-278769. A inclusão neste projeto foi voluntária por meio de assinatura de Termo de Consentimento Livre e Esclarecido. Considerando que o propósito primário da genotipagem dos SNPSTRs é o aconselhamento genético das mulheres portadoras do gene *F8* com mutação para a Hemofilia A.

#### **3.3. Critérios de seleção para locos potencialmente polimórficos**

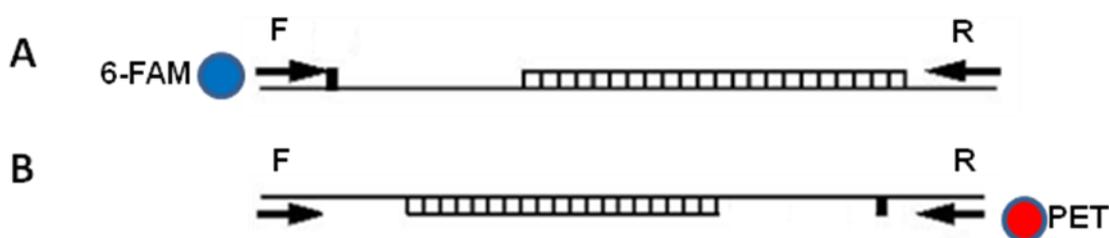
Para selecionar novos marcadores potencialmente polimórficos foram avaliados os elementos STR *F8*-intragênicos constantes no site HEMApSTR ((Medina-Acosta *et al.*, 2008), que lista STR com pontuação mínima de 80 e identidade de 90%; estes critérios correspondem aos valores mínimos apresentados por 97% e 79% dos tetranucleotídeos validados na Genética Forense, respectivamente. Outros critérios foram utilizados para incluir os dinucleotídeos no mapa de alta resolução de marcadores intragênicos e adjacentes ao Fator VIII (*F8*) de coagulação sanguínea (Machado e Medina-Acosta, 2009). As regiões pseudo-autossômicas não foram avaliadas.

Para selecionar os SNPs ligados ao STR, foram selecionados SNPs que se encontram numa distância de aproximadamente 400 pb do STR. A sequência franqueadora dos SNPs foi analisada quanto à disponibilidade de pelo menos uma enzima de restrição alelo específicas para a determinação do alelo.



### 3.4. Iniciadores marcador-específicos

Para o desenho de iniciadores específicos dos genes foram utilizados os programas livres e online *OligoPerfect™ Designer* da *Invitrogen™* (<http://www.invitrogen.com/>) e o *OligoCalc Oligonucleotide Properties Calculator* (<http://www.basic.northwestern.edu/biotools/oligoalc.html>). Os iniciadores *forward* foram modificados com fluóroforos (6-FAM, VIC, NED, PET) na extremidade mais próxima ao SNP (5' ou 3') para facilitar a sua detecção (Figura 4). Para a identificação de enzimas de restrição com sítios alelo específicos nos SNPs alvos foi utilizado o programa online NEBcutter (<http://tools.neb.com/NEBcutter2/>)



**Figura 5. Esquema representativo da marcação dos iniciadores para a genotipagem do SNPSTR por PCR-RFLP. A.** A localização do SNP é anterior ao STR, portanto o primer forward será marcado com fluóroforo. **B.** A localização do SNP é depois do STR, portanto o primer reverse será marcado com fluóroforo.

### 3.5. Tipagem dos marcadores

Os microssatélites foram tipados por PCR quantitativa por fluorescência (Machado, Duarte *et al.*, 2009a, b). E os SNPs foram tipados pela metodologia de PCR-RFLP (Tasleem Raza *et al.*, 2009) onde após a amplificação da região contendo o SNP, uma enzima de restrição alelo específica é utilizada para a determinação do alelo. Para padronização dos ensaios foram realizados gradientes com as seguintes variáveis: temperatura de anelamento, concentração de iniciadores, concentração de  $MgCl_2$  e concentração de DNA alvo.

### 3.6. Caracterização dos alelos amplificados

Aos produtos da amplificação da PCR quantitativa de fluorescência e RFLP foram adicionados formamida (Hi-Di Formamida, Applied Biosystems) e o padrão de massa molecular GeneScan LIZ 500 marcado com o fluóroforo LIZ™ (fluorescência laranja) (Applied Biosystems). Os produtos de amplificação foram separados por eletroforese capilar na plataforma ABI Prism™ 310 Genetic Analyzer, utilizando o polímero POP4 (contendo 4-6% de acrilamida). Os perfis eletroforéticos foram

analisados utilizando os programas GeneScan™ e Genotyper™ (Applied Biosystems).

### **3.7. Digestão enzimática**

Para digestão enzimática foi utilizado o DNA amplificado da reação de QF-PCR feita para amplificação dos marcadores STR. Cada amostra foi digerida com as enzimas BanII, FatI, Cac8I, Sau3AI e Faul.

A digestão pelas enzimas BanII, Cac8I e Sau3AI é feita a 37°C por 1h e 80°C por 20min. E as enzimas Faul e FatI são digeridas a 55°C por 1h e 80°C por 20min.

Para genotipagem por SNP foi realizada a técnica de RFLP-PCR utilizando as mesmas amostras genotipadas por QF-PCR para os marcadores STR.

### **3.8. Frequências alélicas e dados populacionais**

Os genótipos das mães foram utilizados para a determinação da taxa de heterozigose destes marcadores utilizando o software online livre PowerStatsV12 (<http://www.promega.com/geneticidtools/powerstats/>), a análise do equilíbrio de Hardy-Weinberg foi feita utilizando o *software* DNView V28.1. A determinação da segregação dos haplótipos foi realizada através do confronto entre os genótipos dos genitores.

## 4. RESULTADOS

### 4.1. Varredura e desenho de marcadores

Computacionalmente foi realizada a varredura do gene *F8* com objetivo de encontrar SNPs próximos a STRs. Foram rastreados SNPSTR *in silico* com os seguintes critérios de seleção:

A) A distância entre o SNP e STR não poderia ser maior que 300 pares de base.

B) O *amplicon* não poderia ser maior que 500 pares de base, os marcadores não poderiam ter complementaridade.

C) o alelo ancestral deveria ser parte de um sítio de restrição para enzima.

Essa triagem rendeu 14 alvos (Tabela 1). Após a análise *in silico*, o critério de escolha de alvos para começar a padronização foi a menor fração de recombinação entre marcadores (isto é, a menor distância genética interpolada) para reduzir o risco de recombinação entre os marcadores. Esse critério permitiu identificar cinco alvos (Tabela 2). Os iniciadores 5' (*Forward*) para os alvos selecionados foram marcados com os fluóforos (VIC, 6-FAM, NED). Os SNPSTRs foram tipados por PCR-RFLP quantitativa por fluorescência. Os alelos foram analisados por eletroforese capilar de alta resolução.

Todos os marcadores foram também validados por iPCR. A figura 5 mostra as sequências do produto de amplificação *in silico* dos marcadores SNPs. O programa utilizado foi o *iPCR* do *browser* da UCSC. Este programa mostra as regiões de amplificação com o par de iniciador submetido para análise. Para que o par de iniciadores seja considerado específico, deve ser mostrado pelo programa, apenas um produto de amplificação.

Tabela 1. Marcadores SNPSTR selecionados por rastreamento in silico.

SNP	Enzima	STR	Sentido	Anti-sentido	SNP	Produto
rs6643711	Ddel	TG	TCTCTCTTTGCTTGCTTTCATT	GCCTGGACTACAGAGGGAGA	[C/T]	439
rs28890032	Ban II	TG	TCCAATCAGTTGAAGGCCTAA	TCATGGTTCCCAAAGAAACA	[C/G]	343
rs112922881	FauI	AG	GTCTGCCTTTGGACTGGAAC	AAAATTAGCCAAGCGTGGTG	[C/T]	382
rs5987052	FatI	TCCT/CT	TCCCAAATCCCTCTTAACCA	AGGCAACAGTGCCTGACTCT	[A/G]	435
rs79260793	Sau3AI	GT	TGGGCACCATGTCAGTTCTA	GGTGTCAAATCCCACGTTTT	[C/T]	309
rs60186903	Cac8I	AC	GCAAGTGAAGCAGTCACCAA	TTCAGGAGAGGGGAGGAAAT	[A/G]	299
rs5945255	HaeIII	AC	GTAAATGTGGCCTTGGGCTA	TGGATGGCTTCAAGCTCTTT	[C/G]	454
rs12014480	HpyCH4IV	AT	GCCAACACACACTGAAGGAA	CTCCCAATGCAAAGACCTTC	[A/G]	264
rs113261176	NÃO TEM	GA	GAGGCAGCCCTTTTGTAGTTT	TGCAGCTTCCATGTTATTGC	[G/T]	287
rs113244055	HpyCH4III	AG	CTTCAGCCCAGGAAGTTGAG	TCCCATTCACAATTGCTTCA	[C/G]	430
rs147382988	HpyCH4IV	CA	ACCATGCCTGGCTAATTCAA	CATCTTCTTTCATTGTACTTCTTCAA	[C/T]	471
rs73567715	NÃO TEM	TG	TTCTCCTGCTTCAGCCTCTC	AAGGCATTTGACAAAATTCAA	[T/A]	474
rs137852394	Ndel	TTG	TCATCTCAATCCTACGCTTTCA	GGGCAACAGAGTGAGACTCC	[T/C]	407
rs34173607	BsrI	AC	GTGGGCATTATCCATTCCAC	TGGAGAAATTGGAACCCTTG	[-/G]	393

Tabela 2. Informações dos cinco marcadores SNPSTR utilizados para o gene *F8*.

SNPSTR	STR	SNP	Enzima	STR	Sentido	Anti-sentido	SNP	Produto
<b>DXS154082597</b>	<i>F8</i> Int25.2	rs28890032	Ban II	TG	TCCAATCAGTTGAAGGCCTAA	TCATGGTTCCCAAAGAAACA	[C/G]	343
<b>DXS154087367</b>	<i>F8</i> Int25.1	rs112922881	FauI	AG	GTCTGCCTTTGGACTGGAAC	AAAATTAGCCAAGCGTGGTG	[C/T]	382
<b>DXS154088940</b>	<i>F8</i> Int24.1	rs5987052	FatI	TCCT/CT	TCCCAAATCCCTCTTAACCA	AGGCAACAGTGCCTGACTCT	[A/G]	435
<b>DXS154103837</b>	<i>F8</i> Int22	rs79260793	Sau3AI	GT	TGGGCACCATGTCAGTTCTA	GGTGTCAAATCCCACGTTTT	[C/T]	309
<b>DXS154124632</b>	<i>F8</i> Int21	rs60186903	Cac8I	AC	GCAAGTGAAGCAGTCACCAA	TTCAGGAGAGGGGAGGAAAT	[A/G]	299





O ensaio pentaplex foi inicialmente validado utilizando 0,8  $\mu\text{M}$  de cada iniciador em amostra de DNA de uma mulher (figura 11). A mulher foi heterozigota para três marcadores e o homem, como esperado, hemizigótico para todos os marcadores. Esse teste piloto também indica a necessidade de um pequeno ajuste nas concentrações dos iniciadores para o marcador DXS154124632 visando obter um melhor balanço dos alelos. Validado em DNA de um homem, um melhor balanço foi obtido para DXS154124632 a 0,64  $\mu\text{M}$  (figura 12). Como esperado, homens são hemizigoto para marcadores do cromossomo X.



**Tabela 3. Efeito das variáveis temperatura de anelamento e concentração de MgCl<sub>2</sub> na genotipagem do marcador DXS154082597.**

MgCl <sub>2</sub> / °C	58°C	59°C	60°C	61°C	62°C	63°C
1 mM	++	++	++	++	++	++
1,5 mM	++	++	++	++	++	++
2 mM	++	++	++	++	++	+++

(++) O marcador amplificou, mas não nas condições ideais.

(+++) O marcador amplificou nas condições ideais de anelamento.



**Figura 7. Eletroferograma representativo da padronização das condições de amplificação do marcador DXS154082597.** Foi utilizada a temperatura de 63°C e 2mM de MgCl<sub>2</sub>. Amostra referente a tipagem de um controle masculino.

**Tabela 4. Efeito das variáveis temperatura de anelamento e concentração de MgCl<sub>2</sub> na genotipagem do marcador DXS154087367.**

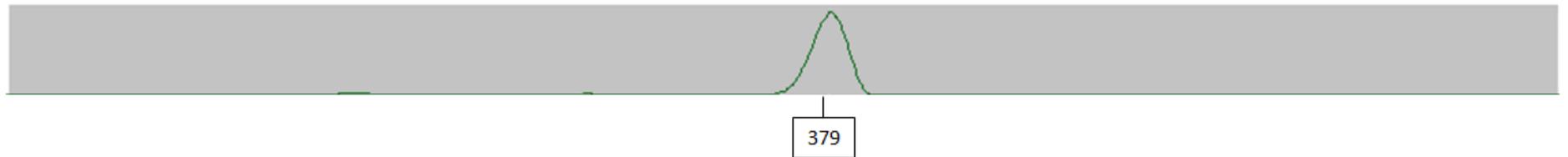
Codificação: (+) amplificação boa; (++) muito boa; (+++) ótima.

MgCl <sub>2</sub>	58°C	59°C	60°C	61°C	62°C	63°C
1mM	+	+	++	++	++	++
1,5mM	+	+	+	++	++	++
2mM	+	+	+	+	++	+++

(+) O marcador amplificou, mas em uma escala de fluorescência baixa.

(++) O marcador amplificou, mas não nas condições ideais.

(+++) O marcador amplificou nas condições ideais de anelamento.



**Figura 8. Eletroferograma representativo da padronização das condições de amplificação do marcador DXS154087367.** Foi utilizada a temperatura de 63°C e 2mM de MgCl<sub>2</sub>. Amostra referente a tipagem de um controle masculino.

**Tabela 5. Efeito das variáveis temperatura de anelamento e concentração de MgCl<sub>2</sub> na genotipagem do marcador DXS154088940.**

Codificação: (+) amplificação boa; (++) muito boa; (+++) ótima.

	58°C	59°C	60°C	61°C	62°C	63°C
1mM	++	++	++	++	++	++
1,5mM	++	++	++	++	++	++
2mM	++	++	++	++	++	+++

(++) O marcador amplificou, mas não nas condições ideais.

(+++) O marcador amplificou nas condições ideais de anelamento.



**Figura 9. Eletroferograma representativo da padronização das condições de amplificação do marcador DXS154088940.** Foi utilizada a temperatura de 63°C e 2mM de MgCl<sub>2</sub>. Amostra referente a tipagem de um controle masculino.

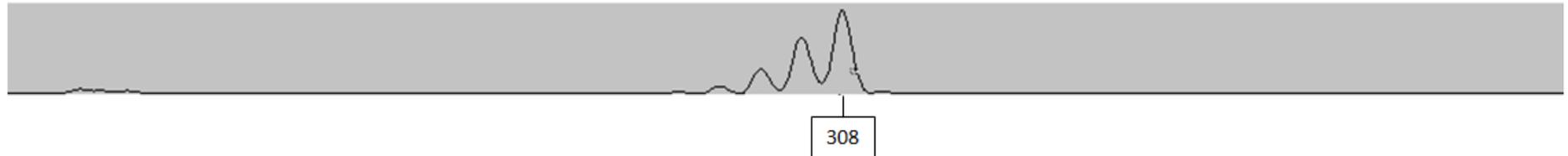
**Tabela 6. Efeito das variáveis temperatura de anelamento e concentração de MgCl<sub>2</sub> na genotipagem do marcador DXS154103837.**

Codificação: (+) amplificação boa; (++) muito boa; (+++) ótima.

MgCl <sub>2</sub>	58°C	59°C	60°C	61°C	62°C	63°C
1mM	++	++	++	++	++	++
1,5mM	++	++	++	++	++	++
2mM	++	++	++	++	+++	++

<sup>(+)</sup> O marcador amplificou, mas não nas condições ideais.

<sup>(++)</sup> O marcador amplificou nas condições ideais de anelamento.



**Figura 10. Eletroferograma representativo da padronização das condições de amplificação do marcador DXS154103837.** Foi utilizada a temperatura de 63°C e 2mM de MgCl<sub>2</sub>. Amostra referente a tipagem de um controle masculino.

**Tabela 7. Efeito das variáveis temperatura de anelamento e concentração de MgCl<sub>2</sub> na genotipagem do marcador DXS154124632.**

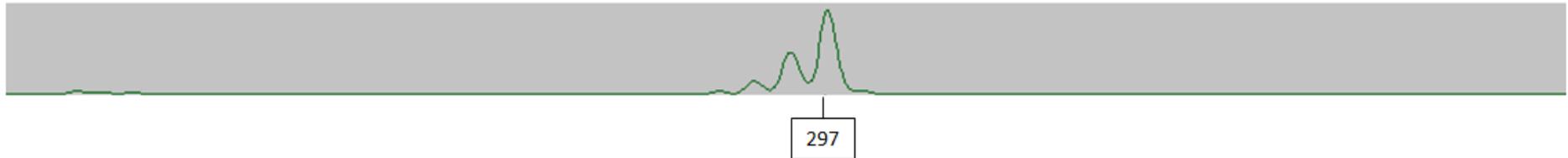
Codificação: (+) amplificação boa; (++) muito boa; (+++) ótima.

MgCl <sub>2</sub>	58°C	59°C	60°C	61°C	62°C	63°C
1mM	++ <sup>(b)</sup>	++	++	++	++	++
1,5mM	++	++	++	++	++	++
2mM	++	++	++	+	++	+++ <sup>(c)</sup>

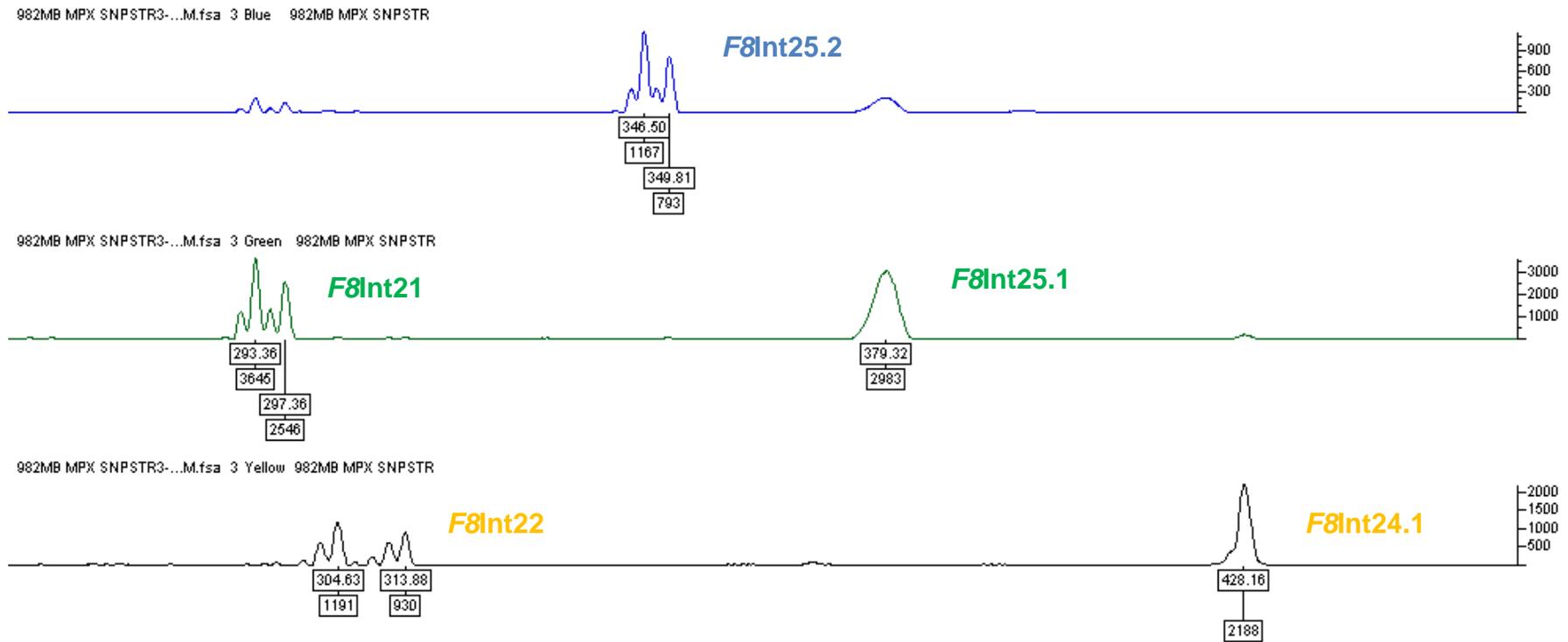
<sup>(+)</sup> O marcador amplificou, mas em uma escala de fluorescência baixa.

<sup>(++)</sup> O marcador amplificou, mas não nas condições ideais.

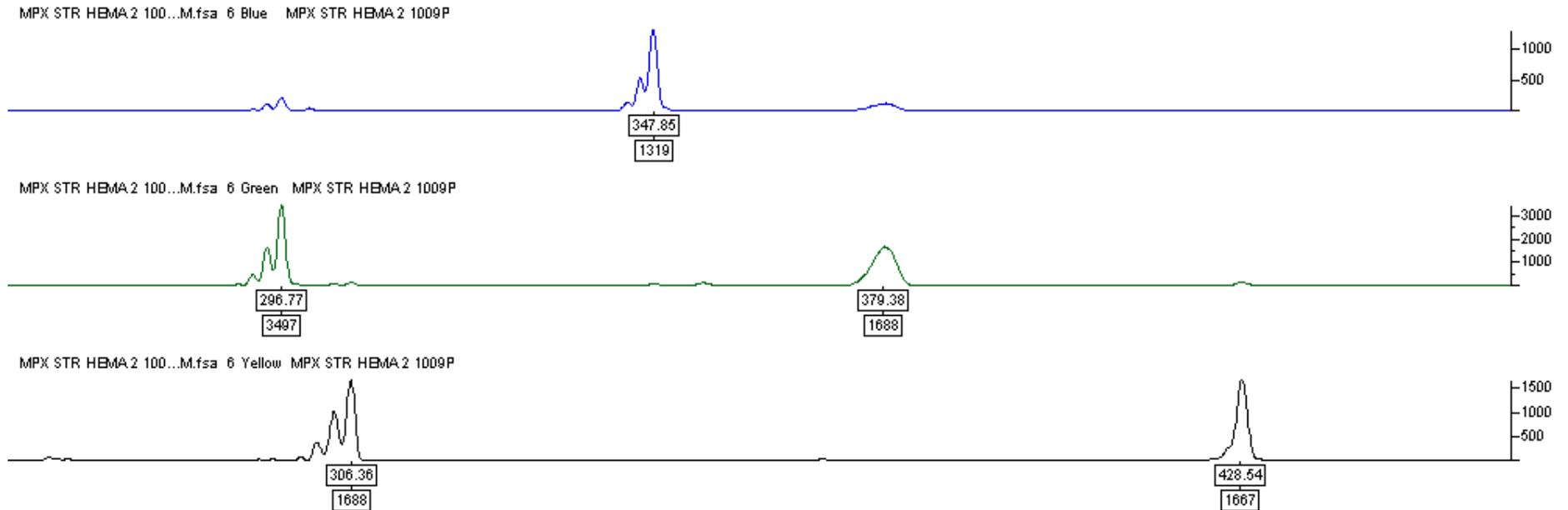
<sup>(+++)</sup> O marcador amplificou nas condições ideais de anelamento.



**Figura 11. Eletoferograma representativo da padronização das condições de amplificação do marcador DXS154124632.** Foi utilizada a temperatura de 63°C e 2mM de MgCl<sub>2</sub>. Amostra referente a tipagem de um controle masculino.



**Figura 12. Eletroferograma representativo do perfil alélico dos cinco marcadores STR em reação multiplex.** As condições de amplificação foram temperatura de anelamento 63°C, iniciadores a 0,8 µM e DNA a 8ng/µL.



**Figura 13. Eletroferograma representativo do perfil alélico do pentaplex padronizado.** As condições de amplificação foram temperatura de anelamento 63°C, iniciadores a 0,8  $\mu\text{M}$  exceção do marcador DXS154124632 que foi reduzido para 0,64  $\mu\text{M}$ .

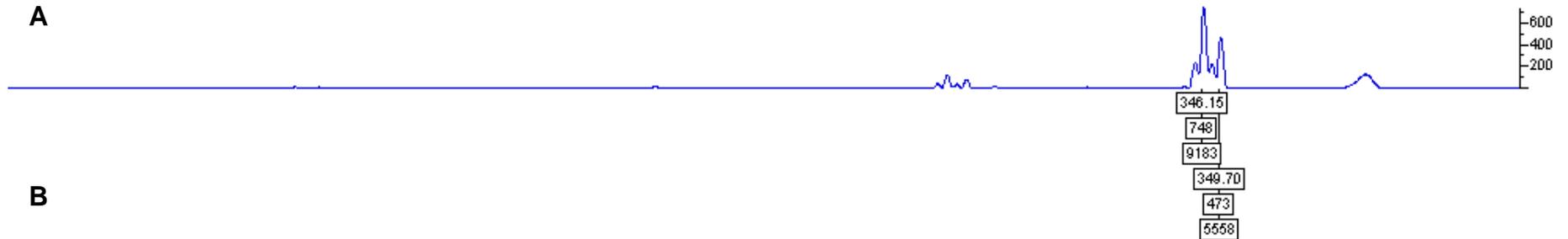
### **4.3. Padronização de marcadores SNP**

Em mulheres heterozigotas para o marcador STR em que um dos alelos contendo um determinado SNP alvo, a digestão dos produtos de amplificação com uma enzima de restrição para o SNP alvo resultará na redução do tamanho do alelo STR, assim determinando a fase gamética dos polimorfismos SNP e STR.

Cada enzima de restrição foi testada e suas condições de digestão padronizadas utilizando DNA plasmidial de referência (contendo pelo menos um sítio alvo). Eletroferogramas ilustrativos desses ensaios estão nas figuras 13-17.



MPX STR HEMA2 982...M.fsa 4 Blue MPX STR HEMA2 982M

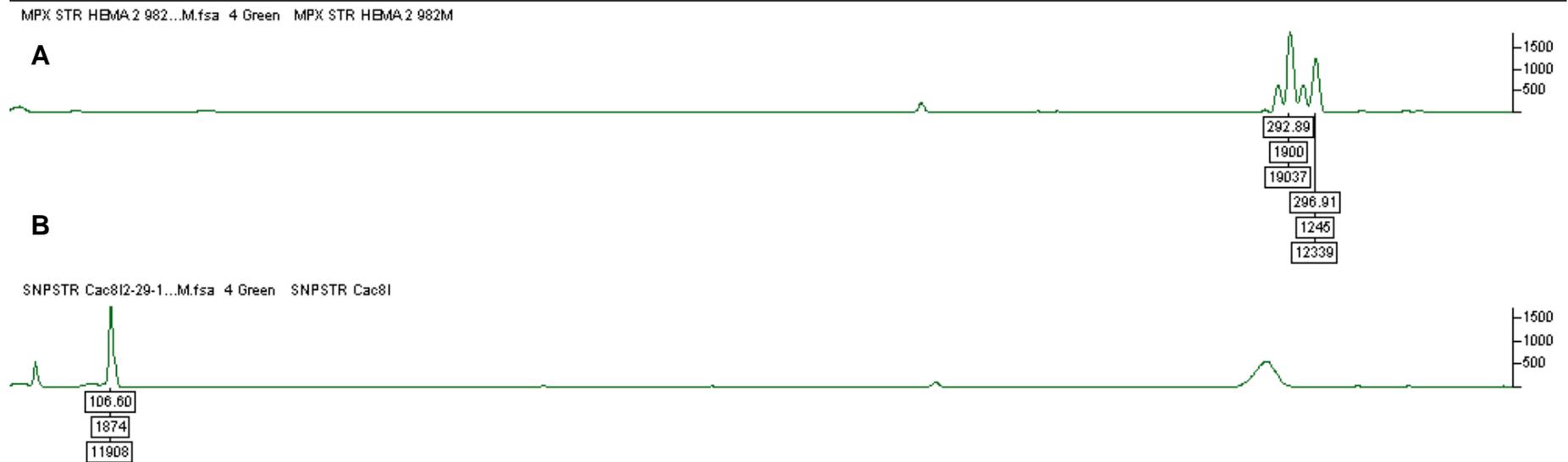
**A****B**

SNPSTR - BanII2-29...M.fsa 2 Blue SNPSTR - BanII



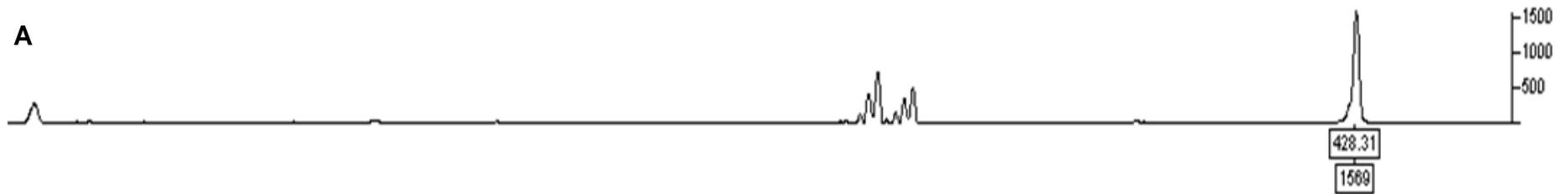
**Figura 14. Eletroferogramas representativos da genotipagem do SNP rs28890032. A.**

A mulher heterozigota para o marcador DXS154082597 (alelos 346 e 349pb). **B.** Após a digestão com a enzima de restrição BanII ambos os alelos do microsatélite são reduzidos para 126 pb, o que indica que a mulher é homozigota para o alelo G do SNP. Assim sendo, os haplótipos dos cromossomos X nessa mulher são SNPSTR G/346 e G/349.



**Figura 15. Eletroferogramas representativos da genotipagem do SNP rs60186903. A.** A mulher 982M é heterozigota para o marcador DXS154124632 (alelos 292 e 296pb). **B.** Após a digestão com a enzima de restrição Cac8I ambos os alelos do microssatélite são reduzidos para 106 pb, o que indica que a mulher é homozigota para o alelo G do SNP. Assim sendo, os haplótipos dos cromossomos X nessa mulher são SNPSTR G/292 e G/296.

MPX STR HBMA.2 982...M.fsa 4 Yellow MPX STR HBMA.2 982M

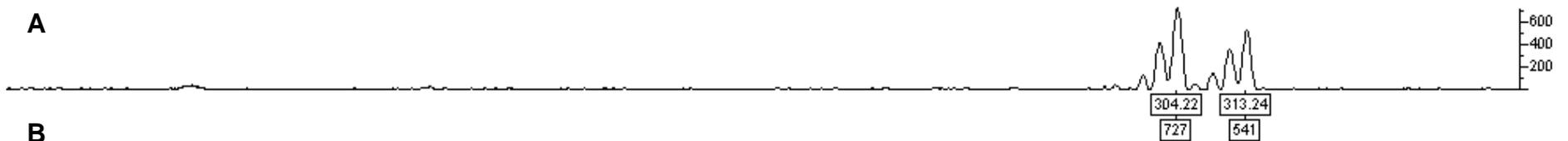
**A**

FatI3-1-12-1-39 PM.fsa 5 Yellow FatI

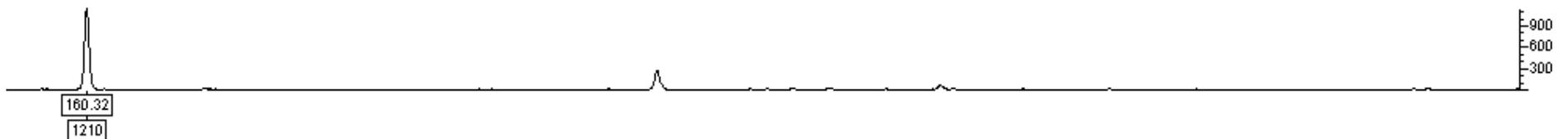
**B**

**Figura 16. Eletroferogramas representativos da genotipagem do SNP rs5987052. A.** Mulher homocigota para o marcador DXS154088940 (alelo 428 pb). **B.** Após a digestão com a enzima de restrição Fat I o alelo é reduzido para 95 pb, o que indica que a mulher é homocigota para o alelo G do SNP. Assim sendo, os haplótipos dos cromossomos X nessa mulher são (SNPSTR) G/428.

MPX STR HBMA.2 982...M.fsa 4 Yellow MPX STR HBMA.2 982M

**A****B**

Sau3I3-1-12-12-30 PM.fsa 3 Yellow Sau3I



**Figura 17. Eletroferogramas representativos da genotipagem do SNP rs79260793. A.** Mulher 982M é heterozigota para o marcador DXS154103837 (alelos 304 e 313pb). **B.** Após a digestão com a enzima de restrição Sau3AI ambos os alelos do microssatélite são reduzidos para 160 pb, o que indica que a mulher é homozigota para o alelo C do SNP. Assim sendo, os haplótipos dos cromossomos X nessa mulher são SNPSTR C/304 e C/313.



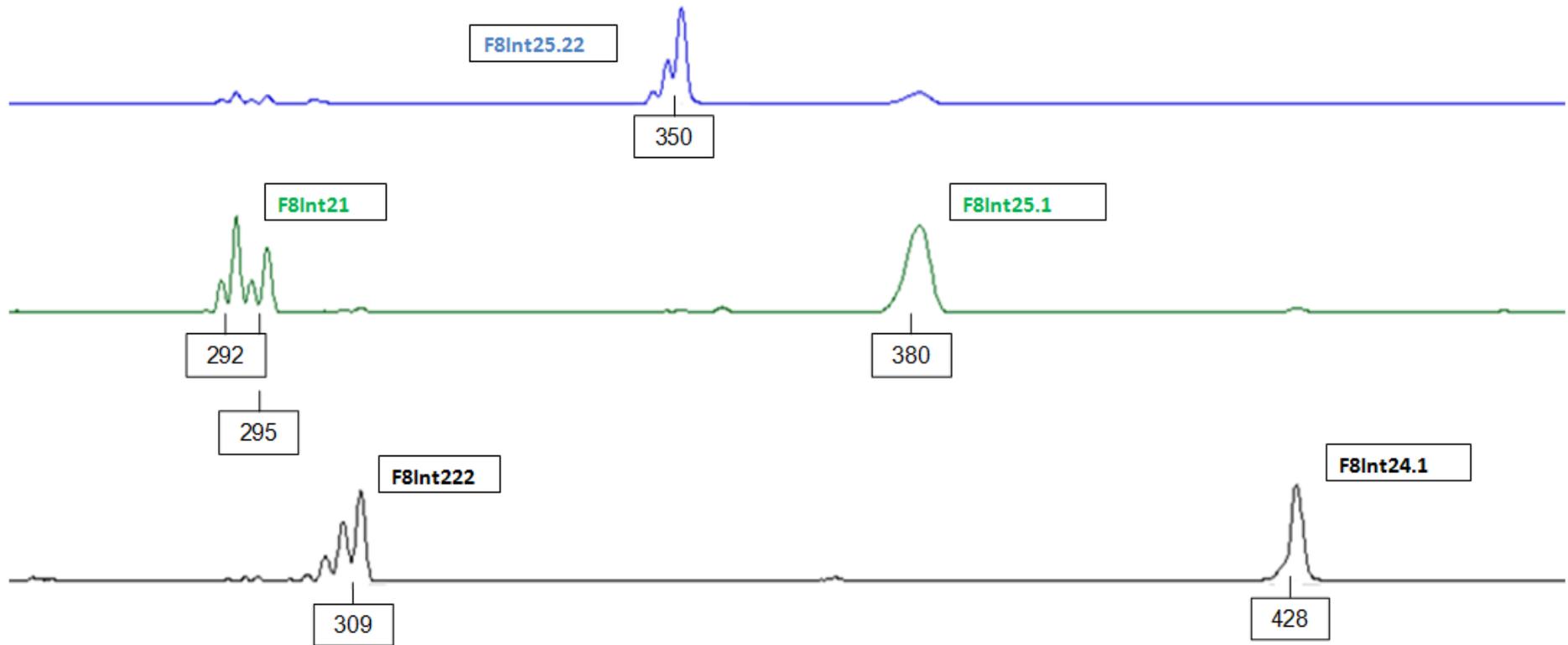
**Figura 18. Eletroferogramas representativos da genotipagem do SNP rs112922881. A.** Mulher homozigota para o marcador DXS154087367 (alelo 379 pb). **B.** Após a digestão com a enzima de restrição Fau I o alelo é para 319 pb, o que indica que a mulher é homozigota para o alelo C do SNP. Assim sendo, o haplótipo dos cromossomos X nessa mulher é SNPSTR G/379.

#### 4.4 Análise dos SNPSTRs

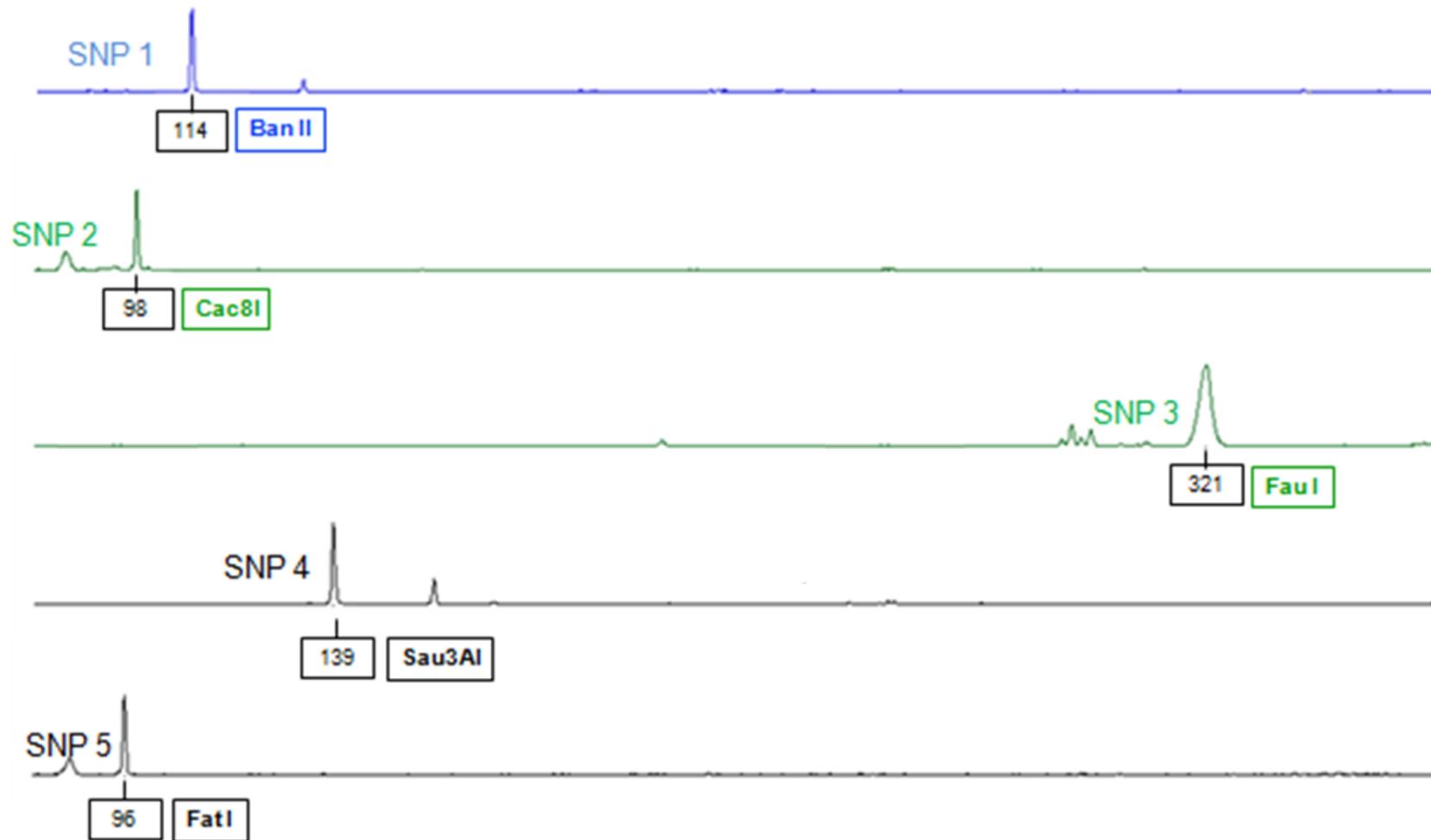
Após o estudo *in silico* e padronização, pode-se determinar a fase gamética através da análise dos marcadores SNPSTRs (Tabela 8) correlacionando os tamanhos dos amplicons em pares de base do STR antes (Figura 18) e após a digestão (Figura 19). Em um dos marcadores, por exemplo, antes da digestão o *amplicon* era de 350pb (Figura 18). Após a digestão com a enzima de restrição BanII o alelo 300pb é reduzido para 114pb (Figura 19), o que indica que o alelo G está em fase gamética com alelo 350pb.

Tabela 8. Informações sobre os cinco marcadores SNPSTRs utilizados neste estudo.

STR	Produto (pb)	STR	SNP	Enzima	Sítio de restrição	Digestão (pb)
F8Int25.2	350	TG	C/G	Ban II	gggctc	114
F8Int21	292/295	AC	A/G	Cac8I	gcctgc	98
F8Int25.1	380	AG	C/T	FauI	cccgc	321
F8Int22	309	TCCT/CT	C/T	Sau3AI	gatc	139
F8Int24.1	428	GT	A/G	FatI	catg	96

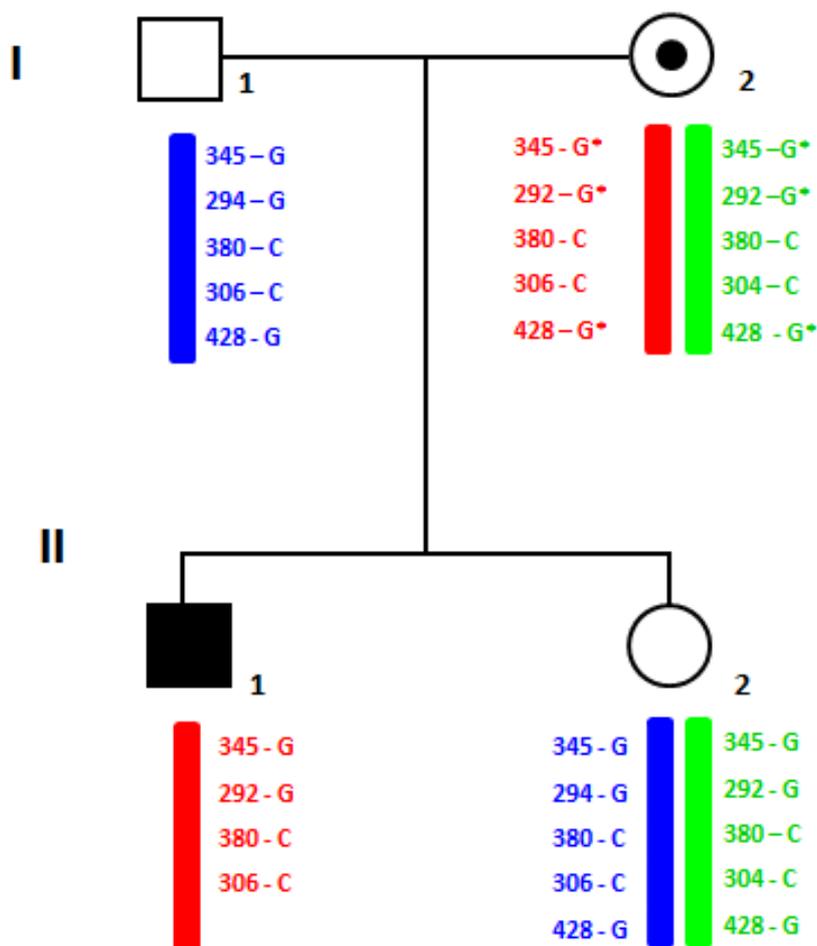


**Figura 19.** Eletroferograma representativo do perfil alélico de uma mulher genotipada com marcadores STR em reação pentaplex. Para os marcadores *F8Int25.2*, *F8Int25.1*, *F8Int222* e *F8Int24.1*, a mulher é homocigota, e para o marcador *F8Int21* é heterocigota.



**Figura 20.** Eletroferograma representativo do perfil alélico gerado por PCR-RFLP. A genotipagem com SNPSTR pode-se determinar a fase gamética do cromossomo.

A figura 20 mostra um exemplo de heredograma montado com um núcleo familiar. Através da tipagem dos marcadores SNPSTR foi possível estabelecer o perfil alélico dos genitores, e através destes marcadores, foi possível diferenciar, nas mulheres das famílias com acometidos, o cromossomo X normal, do cromossomo X que possui a mutação pela identificação no indivíduo acometido, do haplótipo que co-segrega com a mutação.



**Figura 21. Heredograma da segregação de haplótipos em núcleo familiar com um acometido pela Hemofilia A.** Os quadrados representam os homens e os círculos, as mulheres. Indivíduos acometidos são representados por sombreamento total. Nomes dos alelos dos marcadores STR são o comprimento do produto em pares de bases e o nome dos alelos dos marcadores SNP, por letras (A – adenina; C – citosina; T – timina; G – guanina) A irmã (indivíduo II:2) do acometido (indivíduo II:1) não compartilha o mesmo haplótipo com o irmão. Ou seja, não há possibilidade desta mulher (II:2) vir a ter um filho acometido pela doença. Os alelos dos marcadores SNPs com asterisco são aqueles que não amplificaram mas puderam ser inferidos de acordo com os alelos dos outros indivíduos que tiveram seus alelos determinados.



#### 4.5. Dados populacionais relativos ao ensaio pentaplex de Hemofilia A

Após padronização dos ensaios, estes foram validados através da utilização de amostras controle, referentes a indivíduos normais não aparentados. Os ensaios foram satisfatórios e confirmaram o perfil genético esperado para as amostras.

Para os cinco marcadores STR desenvolvidos foram determinados parâmetros populacionais pela tipagem de DNA em 73 mulheres não relacionadas geneticamente. Os alelos foram nomeados conforme o tamanho (em pares de bases) do produto de amplificação observado. As variações em comprimento para um determinado alelo foram sempre múltiplas a uma unidade da repetição do marcador. Os alelos observados por marcador STR, e suas respectivas frequências foram computados. Os dados foram computados e são apresentados na tabela 9.

A chance de erros em diagnóstico indireto está relacionada à taxa de recombinação esperada entre os marcadores mais distantes (*F8Int25.2* e *F8Int24.1*). Como a distância genética interpolada entre os marcadores do ensaio pentaplex mais distantes, é de 0,01 cM, a possibilidade de haver erros no diagnóstico indireto, diminui para 0,01%, pois os marcadores irão segregar em haplótipos, facilitando a análise.

Os marcadores *F8Int25.2*, *F8Int22*, *F8Int21* exibiram uma alta informatividade por serem polimórficos e exibirem taxa de heterozigose maior 50%. Já os marcadores *F8Int25.1* e *F8Int24.1* possuiu perfil monomórfico o que compromete sua informatividade (Tabela 10). O gráfico 1 mostra a pontuação média *in silico* versus heterozigose observada. As maiores pontuações rendem maiores taxas de heterozigose.

Dos cinco marcadores SNPs, quatro marcadores rs28890032, rs112922881, rs5987052, rs79260793 foram considerados monomórficos na pequena amostragem testada. Seria necessário uma amostragem maior de indivíduos testados para analisar com mais acurácia a informatividade desses marcadores. Apesar da pequena amostragem o marcador SNP rs60186903 demonstrou ter significativa taxa de heterozigose, sendo informativo (Tabela 11).

Cinco novos loci SNPSTR foram identificados e desenvolveu-se um ensaio de tipagem de um único tubo de Pentaplex. Usando o ensaio de Pentaplex a fase gamética dos cromossomos X estudados foram determinados. As frequências dos alelos e haplótipos também foram estimadas em um pequeno número de mulheres independentes. Estes resultados mostraram que a genotipagem com loci SNPSTR é

um método rápido e preciso de determinação da fase gamética de cromossomas X. Este método é um ativo em diagnóstico molecular em que confere maior resolução para o aconselhamento genético de heterozigotas portadoras de hemofilia A.

**Tabela 9. Frequências alélicas utilizando o ensaio pentaplex desenvolvido (n=73).**

Alelo	F8Int25.2	F8Int25.1	F8Int24.1	F8Int22	F8Int21
290				0,007	0,041
292					0,349
294					0,158
296					0,425
298				0,007	0,014
300					0,014
302				0,103	
304				0,068	
306				0,315	
308				0,452	
310				0,027	
312				0,021	
345	0,466				
347	0,130				
349	0,404				
380		1,000			
428			1,000		

**Tabela 10. Taxas de heterozigose e número de alelos de 5 microssatélites utilizados na análise de ligação para mutações no gene F8.**

Microssatélite	Heterozigose	Número de alelos identificados.
<i>F8Int25.2</i>	61,6%	3
<i>F8Int25.1</i>	0	1
<i>F8Int24.1</i>	0	1
<i>F8Int22</i>	65,8%	8
<i>F8Int21</i>	69,9%	6

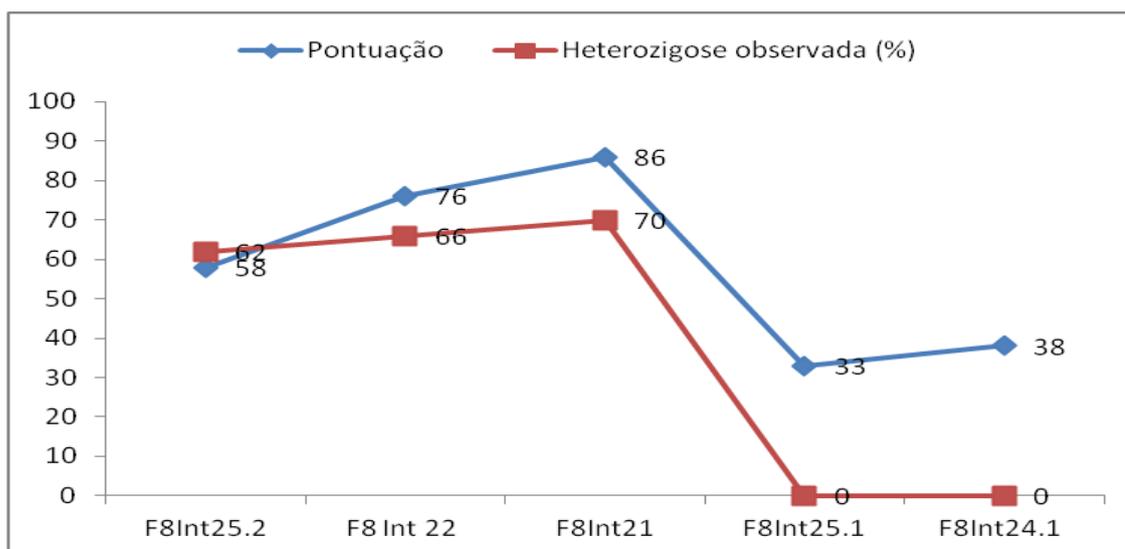


Gráfico 1. Relação entre a pontuação e a heterozigose observada

Tabela 11. Frequência dos marcadores SNPs.

	rs28890032 BanII (n=22)	rs60186903 Cac8I (n=14)	rs79260793 Sau3AI (n=16)	rs5987052 Fat I (n=10)	rs112922881 Fau I (n=19)
Homozigoto (GG)	1,000	0,286		1,000	
Heterozigoto(GA)		0,500			
Homozigoto (AA)		0,214			
Homozigoto(CC)			1,000		1,000

## 5. DISCUSSÃO

Observa-se que há um número limitado de trabalhos de apenas quatro artigos que descrevem marcadores SNPSTRs utilizados para análise de ligação (Mountain *et al.*, 2002; Ramakrishnan *et al.*, 2004; Agrafioti *et al.*, 2007; Odriozola *et al.*, 2009).

Na primeira parte deste estudo foram encontradas evidências de 14 locos SNPSTRs no gene *F8*. Dos 14 marcadores, cinco foram testados como critério para escolha dos locos SNP e STR próximos, com distância menor que 0,01 cM, para não haver risco de recombinação. A chance de erros em diagnóstico indireto utilizando os marcadores identificados neste estudo está relacionada à taxa de recombinação que é esperada entre os marcadores mais distantes, ou seja, de 0,01%. Pelos marcadores SNP estarem próximos fisicamente aos marcadores STR, segregam em haplótipos.

Os marcadores STR *F8Int25.2*, *F8Int22* e *F8Int21* exibiram polimorfismo, ao contrário dos marcadores *F8Int25.1* e *F8Int24.1* que foram monomórficos. As taxas de heterozigose observadas para esses três STR foram comparáveis às relatadas na literatura (Machado, Alves *et al.*, 2009; Machado *et al.*, 2009). Os dados experimentais confirmaram que o marcador *F8Int25.2* exibe baixo número de alelos e os marcadores *F8Int22*, *F8Int21* apresentaram uma maior número de alelos. Houve uma relação direta entre a pontuação e a taxa de heterozigose observada para cada marcador, em que locos com maiores pontuações exibem as maiores taxas de heterozigose observada.

O resultado para os marcadores *F8Int25.1* e *F8Int24.1* era esperado, pois a validação *in silico* e a baixa pontuação das repetições sugeriam que não haviapolimorfismo (HEMApSTR) (Medina-Acosta *et al.*, 2008).

Os marcadores SNPs rs28890032, rs112922881, rs5987052, rs79260793 apresentaram apenas um alelo na pequena amostragem testada. Apenas o SNP rs60186903 apresentou dois alelos, com heterozigose de 0,5, demonstrando ser altamente polimórfico. Interessantemente, esse SNP já foi tipado em outras populações (Projeto 1000 Genomes (Herz *et al.*), dbSNP (Sherry *et al.*, 2001)) e a heterozigose observada foi bem menor, variando de 0,011 a 0,066. Embora o número de amostras aqui genotipadas seja pequena, essa diferença na heterozigose é importante. Contudo, seria necessária uma amostragem maior de indivíduos testados para analisar com mais acurácia esses marcadores. É importante ressaltar que a restrição por digestão é sensível a impurezas e assim

necessita de amostras com DNA de alta pureza para serem analisadas, para eliminar falsos negativos.

Concluindo, *F8Int25.2* , *F8Int22*, *F8Int21* e SNP rs60186903 devem ser considerados os principais locos a serem testados por terem apresentado as maiores taxas de heterozigose, e por isso são os mais informativos para aconselhamento genético.

Neste estudo foi possível a amplificação simultânea dos cinco marcadores STR selecionados em uma única reação, pois os cinco se anelarem na temperatura de 63°C. Já a análise dos marcadores SNPs, através da digestão com enzimas de restrição foi feita em reações separadas. A utilização dos marcadores em reação multiplex reduz o tempo de entrega do resultado do diagnóstico indireto.

As portadoras podem ser identificadas por métodos moleculares indiretos como a análise de ligação de marcadores polimórficos em núcleos familiares com pelo menos um acometido. A inclusão dos marcadores SNPSTR combinados torna o diagnóstico indireto para Hemofilia A mais robusto, além de aumentarem o poder de resolução do teste, e facilitar a determinação da fase gamética em mulheres heterozigotas.

Para o único núcleo familiar testado que contém um filho acometido por Hemofilia A, os marcadores SNPs não foram informativos para determinar a fase gamética. Neste caso, os marcadores STRs foram suficientes para determinar e para concluir que a irmã do acometido não era portadora do gene *F8* com a mutação e por isso ela não apresenta chances de ter um filho acometido pela Hemofilia A.

Os resultados mostraram que a genotipagem com locos SNPSTR é um método rápido e acurado para determinação da fase gamética do cromossomo X. Esta técnica em diagnóstico molecular confere maior resolução para o aconselhamento genético para investigação de mulheres com membros familiares acometidos por Hemofilia A.

## 6. CONCLUSÕES

- A predição do polimorfismo *in silico* foi validada experimentalmente pela tipagem dos novos marcadores em amostras de DNA, demonstrando o potencial informativo em diagnóstico indireto por gerarem dados de mais fácil interpretação, garantindo uma maior rapidez na entrega dos resultados.
- Dos cinco marcadores STR, três marcadores foram polimórficos *F8Int25.2*, *F8Int 22* e *F8Int21* e informativos para aconselhamento genético, e dois marcadores *F8Int25.1*, *F8Int24.1* foram considerados monomórficos não sendo considerados marcadores informativos para análise de ligação.
- Dos cinco marcadores SNPs, apenas o rs60186903 foi polimórfico sendo informativo para a determinação da fase gamética. Os demais marcadores não apresentaram variação alélica na amostragem testada.
- Os novos marcadores SNPSTR representam um aumento no número de marcadores já descritos, tornando o teste molecular para diagnóstico indireto de portadoras de Hemofilia A um teste mais robusto além de permitir a determinação da fase gamética em mulheres duplo heterozigotas.
- A identificação de portadoras de mutações causadoras da Hemofilia A é importante para aconselhamento genético. A identificação pode ser realizada por métodos moleculares indiretos como a análise de ligação de marcadores polimórficos SNPSTR em núcleos familiares com pelo menos um acometido.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

Agrafioti, I.eM. P. Stumpf. SNPSTR: a database of compound microsatellite-SNP markers. Nucleic Acids Res, v. 35, n.Database issue, Jan, p.D71-5. 2007.

Barbosa De Souza, T., L. Pessanha Duarte, J. T. D. Souto Filho, J. Hernandez FernandezE. Medina-Acosta. Farmacogenômica do desenvolvimento de anticorpos inibidores do fator VIII na hemofilia A. Revista Científica da Faculdade de Medicina de Campos, v. 6, n.1, p.7-13. 2011.

Butler, J. M. Forensic DNA Typing, v. 2005. p.,

Davies, K. E., B. D. Young, R. G. Elles, M. E. HilleR. Williamson. Cloning of a representative genomic library of the human X chromosome after sorting by flow cytometry. Nature, v. 293, n.5831, Oct 1, p.374-6. 1983.

Ellegren. Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics. v. 5, p.435-444. 2004.

Goodeve, A. C.el. R. Peake. The molecular basis of hemophilia A: genotype-phenotype relationships and inhibitor development. Semin Thromb Hemost, v. 29, n.1, Feb, p.23-30. 2003.

Hering, S. A., C.; Edelmann, J.; Heidel, M.; Dressler, J.; Rodig, H.; Kuhlisch, E.; Szibor, R. DXS10079, DXS10074 and DXS10075 are STRs located within a 280-kb region of Xq12 and provide stable haplotypes useful for complex kinship cases., v., p. 337-345. 2006.

Herz, H. M., L. D. Madden, Z. Chen, C. Bolduc, E. Buff, R. Gupta, R. Davuluri, A. Shilatifard, I. K. HariharaneA. Bergmann. The H3K27me3 demethylase dUTX is a suppressor of Notch- and Rb-dependent tumors in Drosophila. Mol Cell Biol, v. 30, n.10, May, p.2485-97.

Hong, S. S., U. K. Kim, J. J. Chae, S. H. Goh, Y. S. Kim, N. Yong, J. E. Park, K. S. Kim, S. E. HongC. C. Lee. Lack of correlation between DNA polymorphisms of the apoB gene and plasma lipid levels in Korean population. Molecules and Cells, v. 5, p.235-242. 1995.

Kemball-Cook, G. The Hemophilia A Mutation, Structure, Test and Resource Site: HAMSTeRS. London, <http://europium.csc.mrc.ac.uk/WebPages/Main/main.htm>. 2005.

Kim, U. K., J. J. Chae, S. H. Lee, C. C. LeeY. Namkoong. Molecular diagnosis of Duchenne/Becker muscular dystrophy by polymerase chain reaction and microsatellite analysis. Mol Cells, v. 13, n.3, Jun 30, p.385-8. 2002.

Kong, A., D. F. Gudbjartsson, J. Sainz, G. M. Jonsdottir, S. A. Gudjonsson, B. Richardsson, S. Sigurdardottir, J. Barnard, B. Hallbeck, G. Masson, A. Shlien, S. T. Palsson, M. L. Frigge, T. E. Thorgeirsson, J. R. GulchereK. Stefansson. A high-resolution recombination map of the human genome. Nat Genet, v. 31, n.3, Jul, p.241-7. 2002.

Leclercq, S., E. RivalseP. Jarne. Detecting microsatellites within genomes: significant variation among algorithms. BMC Bioinformatics, v. 8, p.125-131. 2007.

Machado, F. B. Análise compreensiva de marcadores STR no cromossomo X humano: aplicações na genética médica e forense. (Dissertação de Conclusão de mestrado em Biociências e Biotecnologia). Laboratório de Biotecnologia, Universidade estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, RJ, 2008.103 p.

Machado, F. B., A. F. Alves , L. C. Rossetti, C. D. De BrasieE. Medina-Acosta. Informativeness of a novel multiallelic marker-set comprising an F8 intron 21 and three tightly linked loci for haemophilia A carriership analysis. Haemophilia, v. 17, n.2, Mar, p.257-66. 2009.

Machado, F. B., L. P. DuarteE. Medina-Acosta. Improved criterion-referenced assessment in indirect tracking of haemophilia A using a 0.23 cM-resolution dense polymorphic marker set. Haemophilia, v. 15, n.5, Sep, p.1135-42. 2009a.

\_\_\_\_\_. A novel informative dinucleotide microsatellite marker located on human factor VIII intron 25. Haemophilia, v. 15, n.2, Jan 24, p.613-614. 2009b.

Machado, F. B., D. L.P.eE. Medina Acosta. A novel informative dinucleotide microsatellite marker located on human factor VIII intron 25. v. 2009.

Machado, F. B.eE. Medina-Acosta. High-resolution combined linkage physical map of short tandem repeat loci on human chromosome band Xq28 for indirect haemophilia A carrier detection. Haemophilia, v. 15, n.1, Jan, p.297-308. 2009.

Medina-Acosta, E. Evidence of partial and weak gametic disequilibrium across clusters of pericentromeric short tandem repeats loci on human X chromosome: Proceed with caution in forensic genetics. Forensic Sci Int Genet, v., Jan 13. 2010.

Medina-Acosta, E.eF. B. Machado. HEMApSTR. 2010p. 2008.

Mountain, J. L., A. Knight, M. Jobin, C. Gignoux, A. Miller, A. A. LineP. A. Underhill. SNPSTRs: empirically derived, rapidly typed, autosomal haplotypes for inference of population history and mutational processes. Genome Res, v. 12, n.11, Nov, p.1766-72. 2002.



Nothnagel, M.eK. Rohde. The effect of single-nucleotide polymorphism marker selection on patterns of haplotype blocks and haplotype frequency estimates. Am J Hum Genet, v. 77, n.6, Dec, p.988-98. 2005.

Odriozola, A., J. M. Aznar, L. Valverde, S. Cardoso, M. L. Bravo, J. J. Builes, B. Martinez, D. Sanchez, F. Gonzalez-Andrade, E. Sarasola, M. C. Gonzalez-Fernandez, B. Martinez JarretaeM. M. De Pancorbo. SNPSTR rs59186128\_D7S820 polymorphism distribution in European Caucasoid, Hispanic, and Afro-American populations. Int J Legal Med, v. 123, n.6, Nov, p.527-33. 2009.

Oldenburg, J., N. M. AnanyevaeE. L. Saenko. Molecular basis of haemophilia A. Haemophilia, v. 10 Suppl 4, Oct, p.133-9. 2004.

Oudet, C., R. Heilig, A. HanauereJ. L. Mandel. Nonradioactive assay for new microsatellite polymorphisms at the 5' end of the dystrophin gene, and estimation of intragenic recombination. Am J Hum Genet, v. 49, n.2, Aug, p.311-9. 1991.

Payseur., B. A., M. Place.eJ. L. Weber. Linkage Disequilibrium between STRPs and SNPs across the Human Genome. v. 2008.

Ramakrishnan, U.eJ. L. Mountain. Precision and accuracy of divergence time estimates from STR and SNPSTR variation. Mol Biol Evol, v. 21, n.10, Oct, p.1960-71. 2004.

Ren, Z., C. Zhou, Y. Xu, J. Deng, H. ZengeY. Zeng. Mutation and haplotype analysis for Duchenne muscular dystrophy by single cell multiple displacement amplification. Mol Hum Reprod, v. 13, n.6, Jun, p.431-6. 2007.

Saiki, F. M., Jd; Pimentel, Al; Simões, Al (2010). SNP F8Int18Bcll do gene F8 e sua aplicação no diagnóstico de portadora da hemofilia a no Brasil.

Schwartz, L. S., J. Tarleton, B. Popovich, W. K. SeltzereE. P. Hoffman. Fluorescent multiplex linkage analysis and carrier detection for Duchenne/Becker muscular dystrophy. Am J Hum Genet, v. 51, n.4, Oct, p.721-9. 1992.

Sherry, S. T., M. H. Ward, M. Kholodov, J. Baker, L. Phan, E. M. SmigielskieK. Sirotkin. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res, v. 29, n.1, Jan 1, p.308-11. 2001.

Szibor, R. X-chromosomal markers: past, present and future. Forensic Sci Int Genet, v. 1, n.2, Jun, p.93-9. 2007.

Tasleem Raza, S., N. HusaineA. Kumar. Screening for hemophilia A carriers: utility of PCR-RFLP--based polymorphism analysis. Clin Appl Thromb Hemost, v. 15, n.1, Feb, p.78-83. 2009.

**ANEXO****TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**

Eu \_\_\_\_\_, tendo sido satisfatoriamente informado(a) sobre o estudo supracitado, e seus objetivos gerais, declaro por livre e espontânea vontade que concordo em participar do mesmo. Autorizo a utilização da amostra biológica de mim coletada apenas para este projeto, podendo o DNA extraído dessa amostra ser armazenado para utilização em novo projeto de pesquisa, para o qual deverei ser contatado(a) a fim de conceder ou não autorização para seu uso. Declaro estar ciente das seguintes medidas de proteção contra qualquer tipo de discriminação e/ou estigmatização individual ou coletiva, que serão utilizadas durante a pesquisa: a minha identidade será mantida sob sigilo absoluto; todas as dúvidas que porventura ocorram durante o processo desta pesquisa serão devidamente esclarecidas; não haverá nenhum tipo de constrangimento ou ação em contra, caso decida por revogar a autorização aqui concedida para utilização das informações obtidas a partir dos estudos da amostra biológica e dos dados, de mim coletados, a qualquer momento. Declaro estar ciente que terei a opção de tomar ou não conhecimento dos resultados da pesquisa na forma de parecer técnico individual quanto ao meu perfil genético determinado para os marcadores supracitados. Declaro ainda não ter nenhum tipo de interesse econômico, em qualquer época, sobre os possíveis resultados alcançados durante as etapas subsequentes do andamento deste estudo, deixando claro ser o meu interesse particular somente o de colaborador (a) anônimo (a).

Campos dos Goytacazes, RJ, \_\_\_\_ de \_\_\_\_\_ de 2013

---

Assinatura do voluntário (ou mãe, pai e/ou responsável legal)

---

Enrique Medina-Acosta, MSc., PhD.

Coordenador do Núcleo de Diagnóstico e Investigação Molecular – NUDIM

Universidade Estadual do Norte Fluminense Darcy Ribeiro

Matrícula funcional nº. 0000549-6

Responsável pela pesquisa laboratorial