

Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP

Faculdade de Ciências Farmacêuticas

Programa de Pós-Graduação em Biociências e Biotecnologia Aplicadas  
à Farmácia

**Análise de polimorfismos do DNA mitocondrial  
em indivíduos residentes na Grande São Paulo  
para aplicação na Identificação Humana.**

**GREICIANE GABURRO PANETO**

ARARAQUARA - SP

2010

GREICIANE GABURRO PANETO

**Análise de polimorfismos do DNA mitocondrial  
em indivíduos residentes na Grande São Paulo  
para aplicação na Identificação Humana.**

---

Tese apresentada ao Programa de Pós-Graduação em Biociências e Biotecnologia Aplicadas à Farmácia da Faculdade de Ciências Farmacêuticas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, como parte dos requisitos para obtenção do Título de Doutor em Biociências e Biotecnologia Aplicadas à Farmácia. Área de concentração: Biologia Molecular.

---

Orientadora: Regina Maria Barretto Cicarelli

ARARAQUARA - SP

2010

### **Ficha Catalográfica**

Elaborada Pelo Serviço Técnico de Biblioteca e Documentação  
Faculdade de Ciências Farmacêuticas  
UNESP – Campus de Araraquara

P191a Paneto, Greiciane Gaburro  
Análise de polimorfismo do DNA mitocondrial em indivíduos residentes na grande São Paulo para aplicação na identificação humana. / Greiciane Gaburro Paneto. – Araraquara, 2010.  
118 f.

Tese (Doutorado) – Universidade Estadual Paulista. “Júlio de Mesquita Filho”. Faculdade de Ciências Farmacêuticas. Programa de Pós Graduação em Biociências e Biotecnologia aplicadas à Farmácia  
Orientador: Regina Maria Barretto Cicarelli

1. Polimorfismo (Genética). 2. Heteroplasmia. 3. DNA Mitocondrial. 4. Identificação humana. I. Cicarelli, Regina Maria Barretto, orient. II. Título.

**CAPES: 40300005**

*Aos meus pais Zenóbia  
(in memoriam) e Cleozadir,  
ao Francisco e aos meus irmãos.*

## AGRADECIMENTOS

À Profa. Dra. Regina Maria Barretto Cicarelli pela orientação e confiança e pela contribuição a minha formação acadêmica e pessoal.

Ao Prof. Dr. Mario Hiroyuki Hirata pela doação das amostras utilizadas neste projeto.

Aos meus pais que sempre me incentivaram nos estudos e sempre me apoiaram na tomada de decisões.

Ao Francisco por estar sempre ao meu lado, pela paciência, compreensão e incentivo.

Aos meus irmãos por terem sempre me apoiado e incentivado.

À Joyce pela disposição em ajudar a qualquer momento, pela troca de idéias, assim como pela amizade.

À Larissa e Aline por terem contribuído substancialmente neste trabalho, assim como pela amizade.

Ao Túlio e à Daniella que sempre estiveram dispostos a ajudar, assim como pela amizade.

A todos os amigos que mesmo longe torceram por mim.

A Lourdes Prieto e Walther Parson pela conferência dos resultados.

Ao Stephan e a Dra Heidi pela ótima experiência na Alemanha.

Ao CNPq pela bolsa de estudo concedida durante o doutorado e ao DAAD pela bolsa na Alemanha.

À FAPESP pelo auxílio financeiro para a realização deste trabalho.

## SUMÁRIO

RESUMO

ABSTRACT

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

LISTA DE SÍMBOLOS

### CAPÍTULO 1

<b>1.</b>	<b>INTRODUÇÃO.....</b>	<b>19</b>
<b>2.</b>	<b>REVISÃO BIBLIOGRÁFICA.....</b>	<b>21</b>
2.1	O DNA MITOCONDRIAL E SUA UTILIZAÇÃO COMO MARCADOR GENÉTICO FORENSE.....	21
2.2	REGIÃO HIPERVARIÁVEL DO DNA MITOCONDRIAL.....	23
2.3	EXAMES FORENSES DE DNA MITOCONDRIAL.....	23
2.4	BANCO DE DADOS DE DNA MITOCONDRIAL.....	25
2.5	CLASSIFICAÇÃO EM HAPLOGRUPOS DO DNA MITOCONDRIAL.....	26
2.6	HETEROPLASMIA.....	27
<b>3.</b>	<b>JUSTIFICATIVA E RELEVÂNCIA DO PROJETO.....</b>	<b>30</b>
<b>4.</b>	<b>OBJETIVOS.....</b>	<b>32</b>
<b>5.</b>	<b>MATERIAL E MÉTODOS.....</b>	<b>33</b>
5.1	CASUÍSTICA.....	33
5.2	SEQUENCIAMENTO DA REGIÃO HIPERVARIÁVEL DO DNA MITOCONDRIAL EM AMOSTRAS DE SANGUE.....	33
<b>5.2.1</b>	<b>Extração do DNA.....</b>	<b>33</b>
<b>5.2.2</b>	<b>Reação de PCR.....</b>	<b>34</b>
<b>5.2.3</b>	<b>Purificação do produto de PCR.....</b>	<b>35</b>
<b>5.2.4</b>	<b>Reação de sequenciamento.....</b>	<b>36</b>
<b>5.2.5</b>	<b>Purificação do produto da reação de sequenciamento.....</b>	<b>38</b>

<b>5.2.6</b>	<b>Eletoforese e detecção.....</b>	<b>38</b>
<b>5.2.7</b>	<b>Análise dos resultados.....</b>	<b>38</b>
5.2.7.1	Ferramentas de bioinformática.....	38
5.2.7.2	Classificação em haplogrupos.....	39
5.2.7.3	Análise estatística.....	39
5.3	PCR EM TEMPO REAL.....	39
<b>5.3.1</b>	<b>Extração do DNA.....</b>	<b>39</b>
<b>5.3.2</b>	<b>Construção dos <i>primers</i> e sondas.....</b>	<b>40</b>
5.4	SNAPSHOT.....	41
<b>5.4.1</b>	<b>Extração do DNA.....</b>	<b>41</b>
<b>5.4.2</b>	<b>Seleção dos SNPs.....</b>	<b>41</b>
<b>5.4.3</b>	<b>Reação de PCR em <i>multiplex</i>.....</b>	<b>43</b>
<b>5.4.4</b>	<b>Reação de SNaPshot.....</b>	<b>44</b>
<b>5.4.5</b>	<b>Sensibilidade.....</b>	<b>45</b>
<b>5.4.6</b>	<b>Classificação em haplogrupos.....</b>	<b>45</b>
5.5	SEQUENCIAMENTO DA REGIÃO HV3 DO DNA MITOCONDRIAL EM AMOSTRAS DE CABELO.....	46
<b>5.5.1</b>	<b>Extração do DNA.....</b>	<b>46</b>
<b>5.5.2</b>	<b>Reação de PCR.....</b>	<b>47</b>
<b>5.5.3</b>	<b>Clonagem.....</b>	<b>48</b>
<b>5.5.4</b>	<b>Análise dos resultados.....</b>	<b>48</b>
5.5.4.1	Ferramentas de bioinformática.....	48
5.5.4.2	Detecção de heteroplasmia.....	48
<b>6.</b>	<b>RESULTADOS E DISCUSSÃO.....</b>	<b>49</b>
6.1	SEQUENCIAMENTO DAS AMOSTRAS DE SANGUE.....	49
6.2	COMPARAÇÃO DOS RESULTADOS COM OUTRAS POPULAÇÕES MUNDIAIS.....	53
6.3	ANÁLISE DO SNP 3010 POR PCR EM TEMPO REAL.....	55
<b>6.3.1</b>	<b>Aumento do poder discriminatório do SNP 3010.....</b>	<b>56</b>
6.4	REAÇÃO DE SNAPSHOT.....	58
6.5	DISTRIBUIÇÃO DOS HAPLOGRUPOS DE ACORDO COM A COR DE PELE DECLARADA PELO INDIVÍDUO.....	64
6.6	FREQUÊNCIA DE HETEROPLASMIA EM HV3 EM AMOSTRAS DE CABELO E SANGUE.....	66
<b>7.</b>	<b>CONCLUSÕES.....</b>	<b>68</b>
<b>8.</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>70</b>

<b>APÊNDICE A.....</b>	<b>79</b>
<b>APÊNDICE B.....</b>	<b>83</b>
<b>APÊNDICE C.....</b>	<b>86</b>
<b>APÊNDICE D.....</b>	<b>96</b>
<b>APÊNDICE E.....</b>	<b>100</b>

## **CAPÍTULO 2**

ARTIGO PUBLICADO

## RESUMO

A identificação humana por meio da análise de DNA utiliza o perfil genético de um indivíduo baseado na combinação de diversos marcadores que são herdados de seus progenitores. Estes marcadores são geralmente diferenças nas sequências de DNA nuclear entre os indivíduos (polimorfismos). Em alguns casos, entretanto, a análise do DNA nuclear não pode ser aplicada. Isso ocorre quando o DNA da amostra apresenta-se degradado ou em casos onde o material biológico não apresenta o DNA nuclear. Nestes casos, a análise do DNA mitocondrial (DNA mt) é o método de escolha. Entretanto, em um mesmo indivíduo podem existir populações de DNA mt diferentes, fenômeno denominado heteroplasmia.

Este trabalho teve como objetivo o estudo de polimorfismos presentes no DNA mt de indivíduos residentes na Grande São Paulo para aplicação na Identificação Humana. Para isso, foi sequenciada toda a região hipervariável do DNA mt de 160 indivíduos. Além disso, o SNP 3010 foi analisado por discriminação alélica (PCR em tempo real) para estudo do aumento do poder discriminatório quando os indivíduos não puderam ser diferenciados pela análise da região hipervariável. Foi desenvolvido, também, uma reação de SNaPshot em *multiplex* contendo 42 SNPs que permitiram a classificação das amostras em haplogrupos do DNA mt. Por fim, foram analisadas 100 amostras de cabelo dos mesmos indivíduos para o estudo da frequência de heteroplasmia na região HV3 do DNA mt.

De um total de 160 amostras, 144 haplótipos diferentes foram encontrados quando analisamos toda a região hipervariável do DNA mt; 131 haplótipos foram únicos. O SNP 3010 foi suficientemente discriminatório para conseguir distinguir indivíduos que apresentavam o mesmo haplótipo em dois dos treze casos que apresentavam mais de um indivíduo com o mesmo haplótipo.

Nossa população foi classificada em 46,6% de origem africana, 27,3% de origem européia, 25,5% de origem nativo-americana e 0,6% de origem asiática utilizando o painel de SNaPshot contendo 42 SNPs em *multiplex*.

Nenhuma diferença entre amostras de cabelo e sangue na região HV3 foi observada nas 100 amostras analisadas. Além disso, nenhuma heteroplasmia de sequência (ou pontual) foi detectada. Entretanto, uma amostra apresentou heteroplasmia de comprimento (tanto na amostra de sangue como na amostra de cabelo) em uma região que apresenta repetições de CA (citosina/adenina), com impossibilidade de leitura após a heteroplasmia.

Os resultados obtidos neste trabalho contribuem de forma singular para a utilização do DNA mt no contexto forense brasileiro.

Palavras-chave: Identificação Humana. DNA Mitocondrial. Polimorfismos. Heteroplasmia

## ABSTRACT

The human identification by DNA analysis uses the genetic profile of an individual based on the combination of diverse markers that are inherited of its ancestors. These markers are generally differences in sequences of nuclear DNA between individuals (polymorphisms). In some cases, however, the analysis of the nuclear DNA cannot be applied. It occurs when the DNA sample is degraded or in cases which the biological material does not content nuclear DNA. In these cases, the mitochondrial DNA (mtDNA) analysis is the choice method. However, inside one individual can exist different mtDNA populations, phenomenon called heteroplasmy.

The aim of this work was to study mtDNA polymorphisms in individuals residents in São Paulo metropolitan area for application in Human Identification. For this, we have sequenced the entire hypervariable region of mtDNA for 160 individuals. Furthermore, the SNP 3010 was analyzed by allelic discrimination (Real-Time PCR) to study the increase of the discriminatory power when individuals could not be differentiated by analysis of the hypervariable region. It was developed also, a SNaPshot multiplex reaction containing 42 SNPs that allowed the classification of samples in mtDNA haplogroups. Finally, 100 hair samples, from the same individuals, were analyzed for the study of heteroplasmy frequency in the HV3 region of mtDNA.

Among 160 samples, 144 different haplotypes were found when the entire hypervariable region of mtDNA was analyzed; 131 haplotypes were unique. The SNP 3010 SNP was able to distinguish individuals who had the same haplotype in two of thirteen cases with more than one individual with the same haplotype.

Our population was classified in 46.6% of African, 27.3% of European, 25.5% of Native American and 0.6% Asian origin using the SNaPshot panel of 42 SNPs in multiplex.

No differences between hair and blood samples in the HV3 region was observed in 100 samples. Furthermore, no sequence heteroplasmy (or point) was detected. However, one sample showed length heteroplasmy (both in the blood sample as the sample of hair) in a region with repetitions of CA (cytosine/adenine), with impossible to read after heteroplasmy.

The results of this study contribute to the singular use of mtDNA in forensic Brazilian context.

*Keywords:* Human Identification. Mitochondrial DNA. Polymorphisms. Heteroplasmy

## LISTA DE FIGURAS

Figura 1 –	DNA mitocondrial.....	22
Figura 2	Árvore filogenética humana do DNA mt simplificada.....	27
Figura 3 –	Heteroplasmia de sequência.....	29
Figura 4 –	Heteroplasmia de comprimento.....	29
Figura 5 –	Esquema da localização dos <i>primers</i> utilizados na PCR e no sequenciamento.....	37
Figura 6 –	Esquema da localização dos 42 polimorfismos utilizados na reação de SNaPshot para classificação em haplogrupos do DNA mt.....	42
Figura 7 –	Esquema indicando todas as etapas realizadas para análise das amostras pela técnica de SNaPshot.....	46
Figura 8 –	Esquema das etapas realizadas para análise de heteroplasmia por sequenciamento nas amostras de sangue e cabelo.....	49
Figura 9 –	Resultado da genotipagem do SNP 3010 gerado pelo software 7500 System.....	56
Figura 10 –	Distribuição dos alelos do SNP 3010 após as amostras serem classificadas em haplogrupos nativo-americanos, africanos e europeus.....	58
Figura 11 –	Árvore filogenética englobando haplogrupos europeus, africanos e nativo-americanos gerada por este painel.....	60
Figura 12 –	Exemplo de eletroferograma mostrando o perfil de análise para os 42 SNPs.....	61
Figura 13 –	Classificação das amostras segundo a origem dos haplogrupos.....	63

Figura 14 – Distribuição dos haplogrupos em brancos e pretos.....	65
Figura 15 – Heteroplasmia de comprimento na região de repetição de CA em HV3 com impossibilidade de leitura após esta região.....	66
Figura 16 – Eletroferograma na região de repetição de CA demonstrando mistura dos alelos e impossibilidade de leitura após esta região.....	67

## LISTA DE TABELAS

Tabela 1 –	<i>Primers</i> utilizados na reação de PCR e sequenciamento.....	35
Tabela 2 –	<i>Primers</i> e sondas TaqMan utilizadas nas reações de PCR em tempo real.....	40
Tabela 3 –	Informações sobre <i>primers</i> da reação de PCR em <i>multiplex</i> .....	79
Tabela 4 –	Informações sobre <i>primers</i> da reação de SNaPshot em <i>multiplex</i> .....	83
Tabela 5 –	Haplótipos e classificação em haplogrupos das 160 amostras analisadas por sequenciamento de toda a região hipervariável.....	86
Tabela 6 –	Haplótipos encontrados em mais de um indivíduo em nossa amostra populacional.....	50
Tabela 7 –	Parâmetros estatísticos calculados a partir da análise de toda a região hipervariável e de cada uma das regiões hipervariáveis em separado e agrupadas nos 160 indivíduos.....	51
Tabela 8 –	Índices de diversidade das populações do Sudeste, Sul, Nordeste e Norte do Brasil para a região HV1.....	52
Tabela 9 –	Parâmetros estatísticos em nossa amostra populacional e em outras populações mundiais.....	53
Tabela 10 –	Valores de <i>Fst</i> e <i>Fst p value</i> resultantes da comparação pareada entre populações.....	54
Tabela 11 –	Valores de <i>Fst</i> e <i>Fst p value</i> resultantes da comparação pareada entre subpopulações.....	55
Tabela 12 –	Fenótipo e resultado da genotipagem do SNP 3010 por PCR em tempo real..	96
Tabela 13 –	Resultado da genotipagem do SNP 3010 em diversas populações.....	57

Tabela 14 – Resultado da análise dos 42 SNPs por SNaPshot e classificação em haplogrupos das 160 amostras.....	79
--	----

## LISTA DE ABREVIATURAS E SIGLAS

A	Base nitrogenada adenina
AMOVA	<i>Analysis of Molecular Variance</i>
C	Base nitrogenada citosina
rCRS	<i>Cambridge Reference Sequence</i> revisada
D	Deleção
D-loop	<i>Displacement loop</i>
dNTP	desorribonucleotídeo trifosfatado
DNA	Ácido desoxirribonucléico
DNA mt	DNA mitocondrial
DTT	Dithiothreitol
EDTA	Ácido etilenodiaminotetracético
EUA	Estados Unidos da América
Exo	Exonuclease I
FBI	<i>Federal Bureau of Investigation</i>
Fig	Figura
<i>Fst</i>	<i>Fixation index</i>
G	Base nitrogenada guanina
HV1	Região Hipervariável 1
HV2	Região Hipervariável 2
HV3	Região Hipervariável 3
IBGE	Instituto Brasileiro de Geografia e Estatística
MgCl <sub>2</sub>	Cloreto de Magnésio
min	Minuto
N	Número
nt	Nucleotídeo
OBS	Observação
PCR	<i>Polymerase Chain Reaction</i>
PCA	Fenol/clorofórmio/álcool isoamílico
qsp	Quantidade suficiente para
RFLP	<i>Restriction Fragment Length Polymorphism</i>
RNA	Ácido ribonucléico

SAP	<i>Shrimp Alkaline Phosphatase</i>
SNP	<i>Single Nucleotide Polymorphism</i>
STRs	<i>Short Tandem Repeats</i>
seg	Segundos
T	Base nitrogenada timina
TE	Tris EDTA
Tris	Tris (hidroximetil) aminometano
VNTRs	<i>Variable Number of Tandem Repeats</i>
WEB	<i>World Wide Web</i>

## LISTA DE SÍMBOLOS

%	porcentagem
°C	grau Celsius
cm	centímetro
ml	mililitro
μl	microlitro
mM	milimolar
M	molar
μg	micrograma
ng	nanograma
mm	milímetro
pb	pares de base
pH	potencial hidrogeniônico
U	unidade
pmol	picomol
X	vezes
∞	infinito
~	aproximadamente
v	versão

## **CAPÍTULO 1**

## 1. INTRODUÇÃO

A identificação humana por meio do DNA constitui um dos produtos mais revolucionários da Genética Moderna. Em menos de 20 anos, tornou-se uma ferramenta indispensável na área forense, sendo aceita rotineiramente em processos judiciais em todo o mundo. Sua importância não está apenas na possibilidade de provar a culpabilidade de uma pessoa ou inocentá-la, mas em estabelecer uma conexão irrefutável entre a pessoa e sua presença na cena do crime. O perfil genético de um indivíduo, comumente utilizado na identificação humana, está baseado na combinação de diversos marcadores que são herdados de seus progenitores. Esses marcadores referem-se geralmente a diferenças nas sequências de DNA entre os indivíduos (polimorfismos).

Os polimorfismos podem apresentar-se de duas formas: o polimorfismo de sequência e o polimorfismo de comprimento. O primeiro, é composto de diferentes nucleotídeos em uma determinada localização do genoma e, geralmente, se baseia em mutações pontuais (WEEDEN; SWARNEN, 1998). O segundo corresponde à sequências de nucleotídeos que se repetem consecutivamente (motivos), conhecidas como VNTRs (*Variable Number of Tandem Repeats*) (MOLLER; BRINKMANN, 1995). Esses últimos podem ser subdivididos em um tipo particular, os STRs (*Short Tandem Repeats*), em que o motivo repetitivo tem cerca de 2 a 16 pares de bases. Os STRs estão espalhados pelos cromossomos, principalmente no DNA nuclear, em regiões não codificadoras entre os genes ou dentre eles (introns), correspondendo a, aproximadamente, 3% do genoma humano (SUBRAMANIAN; MISHRA; SINGH, 2003).

Os marcadores genéticos utilizados em identificação humana podem ser biparentais (metade dos cromossomos herdada da mãe e a outra metade herdada do pai) ou uniparentais (cromossomo Y herdado apenas do pai em indivíduos do sexo masculino e DNA mitocondrial herdado apenas da mãe em todos os indivíduos). Os marcadores biparentais autossômicos

(STRs autossômicos) permitem uma identificação individual porque um determinado perfil genético é específico de um só indivíduo, sendo apenas partilhado por gêmeos monozigóticos. Os marcadores uniparentais não permitem a identificação individual, pois todos os indivíduos aparentados por via masculina partilharão a mesma linhagem de cromossomo Y, enquanto todos os indivíduos aparentados por via feminina partilharão a mesma linhagem de DNA mitocondrial, salvo nos casos de ocorrência de mutações durante a transmissão.

A tecnologia mais recente, utilizada pela comunidade que atua na área de identificação humana, consiste na amplificação de regiões STRs, pela reação de PCR (*Polymerase Chain Reaction*), um processo *in vitro* que aumenta a quantidade de um fragmento pequeno de DNA previamente selecionado (MULLIS; FALOONA, 1987). A análise de STRs do DNA nuclear é utilizada como método padrão em identificação humana vista a grande capacidade de discriminação entre os indivíduos, sendo utilizada na rotina forense.

Em alguns casos, entretanto, a análise do DNA nuclear não pode ser aplicada. Isso ocorre quando o DNA da amostra apresenta-se muito degradado, ou os indivíduos em questão apresentam vínculo genético distantes, ou ainda em situações onde o material biológico não apresenta o DNA nuclear, como em fragmentos de cabelo, por exemplo (SALAS; LAREU; CARRACEDO, 2001). Nestes casos, a única alternativa, ou pelo menos a alternativa de maior sucesso, é a análise do DNA mitocondrial (DNA mt).

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 O DNA MITOCONDRIAL E SUA UTILIZAÇÃO COMO MARCADOR GENÉTICO FORENSE

No contexto da análise forense, o interesse pelo DNA mt surgiu por vários motivos: primeiro esse DNA também contém regiões polimórficas que permitem sua individualização; segundo, os descendentes recebem esse DNA apenas da mãe, o que permite traçar a linhagem materna de uma pessoa; e, terceiro, é mais resistente à degradação do que o DNA nuclear (BUTLER; LEVIN, 1998; FORAN, 2006; SULLIVAN et al., 1992), além de estar presente em várias cópias por célula (ROBIN; WONG, 1988). Assim, em grandes desastres (incêndios, explosões, queda de avião, etc.), quando é mais difícil identificar os corpos, analisa-se o DNA mt (ANJOS et al., 2004), que é extraído dos restos mortais e a sequência de interesse é comparada com sequências obtidas de irmãos ou ascendentes maternos.

Pelo fato do genoma mitocondrial ser haplóide, ou seja, possuir apenas a contribuição maternal, os diferentes DNA mt são ditos haplótipos.

A sequência completa do genoma mitocondrial (mt) humano foi determinada em 1981 por Anderson e colaboradores. Essa sequência foi re-analisada e revisada por Andrews e colaboradores em 1999, sendo modificada e denominada *Cambridge Reference Sequence* (rCRS).

O genoma mt é composto por 16.569 nucleotídeos e apresenta-se como uma dupla fita circular. A maior parte do genoma é formada por genes. Entretanto, uma região não codificadora de aproximadamente 1.200 nucleotídeos flanqueia a posição “0” do genoma e é conhecida como região controle, D-loop ou região hipervariável (Fig. 1).



## 2.2 REGIÃO HIPERVARIÁVEL DO DNA MITOCONDRIAL

A região hipervariável do DNA mt é de interesse para a identificação humana por causa do alto grau de polimorfismos na sequência de nucleotídeos, e compreende três segmentos da região controle: HV1 (da posição 16024 a 16365), HV2 (da posição 73 a 340) e HV3 (da posição 438 a 576) (LUTZ et al. 2000). As duas primeiras regiões são utilizadas na rotina forense, a terceira região, entretanto, vem sendo explorada nos últimos anos (HOONG; LEK, 2005; ZHANG et al., 2005; LEE et al., 2006), além de ser importante para a correta classificação de algumas amostras em seus respectivos haplogrupos (PARSON; BANDELT, 2007).

A amplificação e análise, em separado, de dois ou três segmentos do DNA mt aumenta o risco de troca de amostras, levando a recombinações artificiais. Esses tipos de erros, que só podem ser investigados posteriormente por análise filogenética, resultam em haplótipos recombinantes, consistindo de segmentos que pertencem a diferentes haplogrupos. A fim de solucionar este problema, foi sugerida a amplificação de toda a região hipervariável em uma única reação (PARSON; BANDELT, 2007) e o sequenciamento da mesma (não apenas HV1, HV2 e HV3, mas também das regiões entre esses segmentos) em ambas as fitas do DNA, utilizando para isto uma série de *primers* (ou iniciadores).

## 2.3 EXAMES FORENSES DE DNA MITOCONDRIAL

A maioria dos laboratórios que utiliza a tipagem do DNA mt baseia-se nos polimorfismos presentes na sequência de nucleotídeos nas regiões HV1 e HV2 (SNPs, deleções e inserções) comparando a amostra questionada com a sequência referência (rCRS) para a anotação das diferenças (polimorfismos). A sequência da amostra questionada é então

comparada com a da amostra referência (do indivíduo em questão ou de parentes com vínculo materno). De acordo com as regras estabelecidas pelo FBI, se essas sequências diferirem em dois ou mais polimorfismos, uma exclusão inequívoca pode ser feita. Se as sequências corresponderem uma à outra, existe a possibilidade da amostra questionada pertencer ao indivíduo em questão ou demonstrar vínculo materno. Existe ainda a possibilidade de a amostra pertencer a outro indivíduo não aparentado. Por essa razão, a identidade de duas amostras na região HV1 e HV2 não pode ser considerada uma identificação conclusiva (SWGDM, 2003). Para se determinar a significância dessa identidade é necessário verificar a frequência com que esse conjunto de polimorfismos ocorre na população (frequência do haplótipo), o que gera a necessidade de um banco de dados populacional de DNA mt (HOLLAND; PARSONS, 1999). Haplótipos pouco frequentes no banco de dados indicam maior probabilidade de a amostra questionada ser realmente de mesma origem da amostra referência. Haplótipos muito frequentes indicam menor probabilidade de a amostra questionada ser originária do mesmo indivíduo da amostra referência, já que muitos indivíduos apresentam o mesmo conjunto de polimorfismos naquela população. No caso das amostras diferirem em apenas um nucleotídeo, o resultado é dito inconclusivo e procede-se à verificação de outras regiões polimórficas, como por exemplo HV3 (LUTZ et al., 2000).

Com o objetivo de obter informações adicionais, ou seja, aumentar o grau de discriminação do DNA mt, estudos recentes sugerem a tipagem adicional de polimorfismos de nucleotídeos únicos (SNPs) em outras regiões do DNA mt nos casos onde não é possível diferenciar os indivíduos apenas pela análise da região hipervariável (COBLE et al., 2004). Alguns autores já relataram casos na literatura onde houve a necessidade de analisar polimorfismos adicionais, presentes na região codificadora do DNA mt, a fim de solucionar os casos (JUST et al., 2010; KIM et al., 2010).

Não existe, entretanto, um consenso ou padronização mundial sobre a forma correta de se analisar um resultado de DNA mt, apenas recomendações. Muitos laboratórios não adotam as regras sugeridas pelo FBI (SWGDM, 2003) por considerarem uma regra muito simplista, que não leva em consideração a taxa de mutação de cada posição nucleotídica e de cada tecido. (PRIETO et al., 2008).

#### 2.4 BANCO DE DADOS DE DNA MITOCONDRIAL

A criação de bancos de dados de DNA mt para estimar a frequência dos haplótipos em determinada população é indispensável para permitir sua utilização. Entretanto, por estar sujeito a vários tipos de erros, principalmente humanos (na leitura e digitação de dados, por exemplo), existe a necessidade de estabelecer regras de controle de qualidade tanto na geração de resultados como na compilação dos mesmos. Um exemplo de banco de dados que está sendo desenvolvido com rigorosos parâmetros de qualidade é o banco de dados mundial EMPOP (PARSON; DÜR, 2007). A padronização da geração, análise e transferência de dados, além do uso de softwares que permitem a visualização dos dados, interpretação e detecção de erros, garantem a qualidade dos resultados (PARSON; BANDELT, 2007).

Outros bancos de dados de DNA mt estão disponíveis na WEB, como por exemplo: AMBase (Universidade Federal do Pará – Brasil), MITOMAP (Universidade da Califórnia - EUA), mtDB (Universidade de Uppsala – Suécia), Mitosearch (Family Tree DNA - EUA), mtDNAManager (Universidade Yonsei – Coreia) e FBI Software mtDNA Population Database (FBI – EUA). Entretanto cada um apresenta um regulamento próprio para a deposição dos perfis.

## 2.5 CLASSIFICAÇÃO EM HAPLOGRUPOS DO DNA MITOCONDRIAL

O termo "haplogrupo" refere-se a um conjunto de haplótipos derivados por descendência da mesma molécula ancestral, possuindo o mesmo padrão basal de mutações. Os haplogrupos são assim definidos por motivos, constituídos de mutações particulares presentes na região não-codificadora e na região codificadora do DNA mt. A nomenclatura utilizada para haplogrupos é designada por letras do alfabeto e foi inicialmente utilizada por Torroni et al. (1993), o qual diferenciou uma amostra de nativo-americanos em quatro grupos basais (A, B, C, D) baseados na análise de fragmentos de restrição (RFLP). Mais tarde, os haplogrupos de outras populações de diferentes continentes foram caracterizados, seguindo-se a sua designação por letras. Subgrupos dentro de um haplogrupo, sub-haplogrupos, são designados por números, como por exemplo os sub-haplogrupos H1 e H2, pertencentes ao haplogrupo H. Ambos estes sub-haplogrupos partilham os polimorfismos (mutações) que caracterizam H, mas todos os indivíduos H1 partilham um ou vários polimorfismos adicionais, que são diferentes dos polimorfismos partilhados pelos indivíduos pertencentes a H2. A nomenclatura para subagrupamentos adicionais segue a alternativa letra-número-letra-número, em ordem crescente alfabética e numérica a partir do último grupo publicado. O símbolo asterisco é utilizado para denotar membros de um haplogrupo que não pertencem a nenhum dos sub-haplogrupos já denominados, por exemplo, HV\* apresenta polimorfismos que não o permitem classificar nem como HV1, R0 ou qualquer um dos sub-haplogrupos H (BANDEL et al., 2006).

Os haplogrupos são continental ou etnicamente específicos (RICHARDS et al., 2000). Sete deles (L0, L1, L2, L3, L4, L5, L6) formam as linhagens africanas (VAN OVEN; KAYSER, 2009; SCHLEBUSCH et al., 2009), dez (H, HV, I, J, K, T, U, V, W e X) englobam quase todas as linhagens de europeus, norte-africanos e caucasianos do Leste-

asiático (MACAULAY et al., 1999; TORRONI et al., 1994; FINNILA et al., 2001). Os haplogrupos A, B, C, D, E, F, G e M englobam a maioria das linhagens descritas para a Ásia, Oceania e nativo-americanos (KIVISILD et al., 2002; TORRONI et al., 1993; KONG et al., 2003) (Fig. 2).

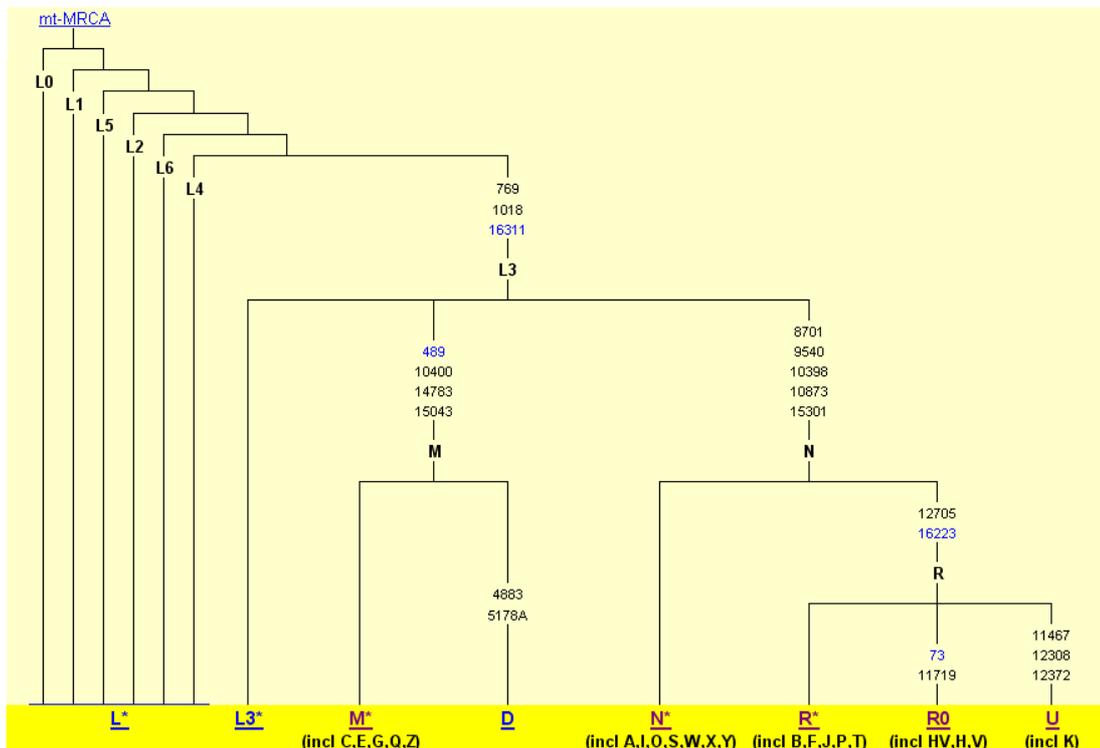


Figura 2 – Árvore filogenética humana do DNA mt simplificada; Mt-MRCA: *More Recent Common Ancestor*; posições nucleotídicas são relativas à CRS revisada; mutações na região codificadora estão descritas em preto e mutações na região não codificadora estão descritas em azul.

Fonte: <http://www.phylotree.org/tree/main.htm>

## 2.6 HETEROPLASMIA

A heteroplasmia, ou seja, a presença de mais de um tipo de DNA mt dentro de um indivíduo é uma situação encontrada mais comumente em amostras de cabelo do que de sangue (PANETO et al., 2007). A natureza clonal dos folículos capilares e a alta energia requerida na queratinização das células são duas características da histogênese que podem

contribuir para o aumento da segregação de variantes heteroplásmicas em amostras de cabelo (TULLY et al. 2004). Em contraste, amostras de sangue periférico consistem de linfócitos produzidos por um largo número de células tronco hematopoiéticas na medula óssea e, se sua proporção diferir em células tronco individuais, a quantidade de heteroplasmia observada será proporcional à média de todas as células tronco presentes.

A heteroplasmia foi primeiro observada em humanos em associação com uma doença mitocondrial (HOLT; HARDING; MORGAN-HUGHER, 1988). Depois, a heteroplasmia foi também demonstrada na região controle do DNA mt (JAZIN et al., 1996). Mutações são acumuladas durante a vida de um indivíduo, logo é esperado que todos os indivíduos possuam heteroplasmia. A heteroplasmia pode ser observada em diferentes situações: 1) indivíduos podem possuir mais do que um tipo de DNA mt em um determinado tecido; 2) indivíduos podem exibir um tipo de DNA mt em um tecido e um tipo diferente em outro tecido; e/ou 3) o indivíduo pode ser heteroplásmico em uma amostra de tecido e homoplásmico em outra amostra do mesmo tecido (ALONSO et al., 2002; HE et al., 2010).

A heteroplasmia pode apresentar-se de duas formas: uma delas dita heteroplasmia de sequência (ou pontual) quando os diferentes tipos de DNA mt diferem em apenas um nucleotídeo (Fig. 3). A heteroplasmia é dita de comprimento quando ocorre a deleção ou inserção de um ou mais nucleotídeos, normalmente citosinas, em regiões de repetições (Fig. 4). Tanto HV1 como HV2 e HV3 possuem regiões poli-C (nt 16184-16193, nt 303-315 e nt 568-573, respectivamente).

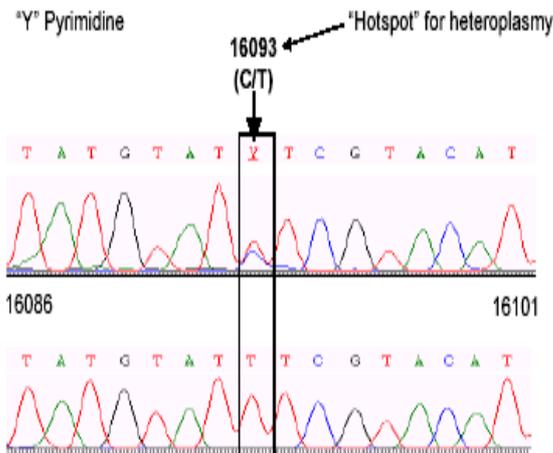


Figura 3 - Heteroplasmia de sequência.

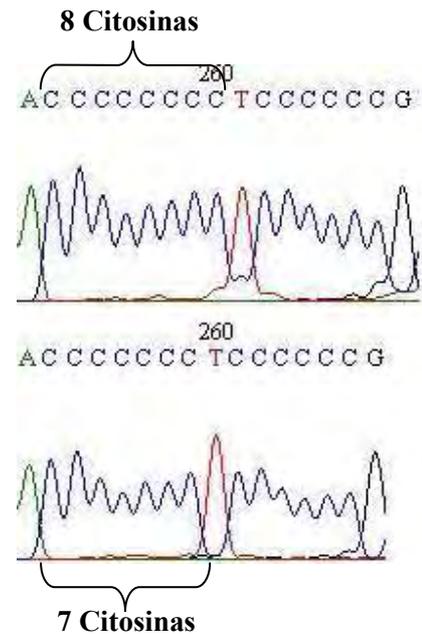


Figura 4 - Heteroplasmia de comprimento.

Fonte: BUTLER, J. M. Forensic DNA Typing.  
London: Elsevier Academic Press, 2005.

A presença de heteroplasmia em um caso forense pode ser relevante porque duas amostras de um mesmo indivíduo podem diferir em níveis de heteroplasmia ou, algumas vezes, na substituição de uma base (SALAS et al., 2001). Os cientistas forenses devem ser preparados para lidar com esse processo e interpretá-lo. Dado que a heteroplasmia possa ocorrer, em algum nível, em um ou dois loci em um indivíduo, isso não é um fato que, *a priori*, possa causar confusão na análise forense, pois pode representar um nível adicional de variação, aumentando o poder discriminatório do teste. Por exemplo, se uma amostra referência e uma amostra questionada possuírem a mesma constelação de polimorfismos, a coocorrência de heteroplasmia em um locus em particular pode prover evidência adicional de que a amostra referência é a fonte da amostra questionada (IVANOV et al., 1996). Entretanto, se a heteroplasmia for encontrada em apenas uma das amostras, referência ou questionada, o resultado pode tornar-se duvidoso.

É importante determinar a frequência e a localização (*hotspots*) das heteroplasmias em cada tecido, de forma a evitar interpretações incorretas de resultados. Devido ao grande número de cópias do DNA mt presentes nas células e a alta sensibilidade das técnicas de detecção, a ocorrência de contaminação entre duas ou mais amostras pode ocorrer, o que pode ser incorretamente interpretado como heteroplasmia. Essa hipótese deve ser considerada, o analista deve verificar na literatura a frequência e localização da mesma naquele tecido, assim como verificar se a amostra foi manipulada corretamente para evitar contaminação com o seu ou outro perfil do laboratório.

### **3. JUSTIFICATIVA E RELEVÂNCIA DO PROJETO**

A utilização do DNA mitocondrial em casos forenses no Brasil apresenta dificuldades, pois poucos laboratórios dispõem do *know-how* necessário para sua análise e interpretação (BARALDI et al., 2008). Entretanto, a atual política de segurança pública equipou, nos últimos anos, 19 laboratórios públicos para análise de DNA pela polícia científica em diferentes Estados brasileiros, os quais utilizarão as diversas técnicas de identificação humana. Assim, faz-se necessário o estudo das frequências dos marcadores moleculares, como as frequências dos haplótipos do DNA mt, com a geração de resultados de alta qualidade e deposição dos mesmos em bancos de dados viabilizando sua aplicação também no contexto brasileiro. Além disso, a análise de SNPs que permitem a discriminação de indivíduos de nossa população, que apresentem o mesmo perfil quando a região hipervariável é analisada, como o SNP 3010, por exemplo, vem adicionar informações importantes para difundir seu uso no Brasil.

A par disso, a classificação das amostras em haplogrupos do DNA mt permite-nos obter a origem materna do indivíduo, além de funcionar como um controle de qualidade dos resultados do sequenciamento. Para isso, são selecionados SNPs presentes na região codificadora do DNA mt e, algumas vezes SNPs da região não-codificadora, que discriminam os diversos haplogrupos possíveis (Fig. 2).

Como a população brasileira é altamente miscigenada, apresentando origem européia, africana e nativo-americana, a análise de um grande número de SNPs é necessária para permitir a classificação em todos os prováveis haplogrupos. Atualmente essa classificação é feita utilizando-se a técnica de SNaPshot que permite a análise de diversos SNPs em uma única reação em *multiplex*. Assim, a proposta deste trabalho foi analisar, em amostras brasileiras, 42 SNPs que permitissem a classificação das mesmas em haplogrupos.

A frequência de heteroplasmia é também de interesse prático nos testes para identificação forense, onde pode ser relevante na solução de um caso (SALAS; LAREU; CARRACEDO, 2001). Novos estudos são necessários para se avaliar a localização e frequência de heteroplasmia em todas as regiões utilizadas (toda a região hipervariável, por exemplo) e nos diversos tecidos, visto que diferenças entre os mesmos já foram relatadas (PANETO et al., 2007; HE et al., 2010). Para isso o método de escolha é o sequenciamento, pois é o método utilizado na rotina forense. Como não existem relatos na literatura da frequência de heteroplasmia em HV3 em amostras de cabelo, esse trabalho também propôs esta análise, cujos resultados poderão ser aplicados nacional e internacionalmente.

#### 4. OBJETIVOS

- Realizar o sequenciamento da região hipervariável do DNA mt em amostras de sangue;
- Calcular os índices de diversidade (diversidade haplotípica, probabilidade de semelhança, número de haplótipos e sítios polimórficos) da região hipervariável do DNA mt entre os indivíduos de nossa amostra populacional e compará-los com resultados de populações já descritas na literatura;
- Analisar, nas amostras de sangue, o SNP 3010 pela técnica de PCR em Tempo Real para estudo do aumento do poder discriminatório do DNA mt quando subtipos comuns do mesmo são encontrados pela análise da região hipervariável comumente utilizada;
- Analisar, nas amostras de sangue, 42 polimorfismos que permitam sua classificação em haplogrupos do DNA mt, utilizando a técnica de SNaPshot;
- Comparar a sequência da terceira região hipervariável (HV3) do DNA mt em amostras de sangue com aquelas obtidas das respectivas amostras de cabelo de indivíduos residentes na Grande São Paulo, para o cálculo da frequência de heteroplasmia e possíveis posições heteroplásmicas de alta frequência (*hotspots*).

## 5. MATERIAL E MÉTODOS

### 5.1 CASUÍSTICA

Foram analisadas 160 amostras de sangue (depositadas em papel de filtro comum) e 100 amostras de cabelo de indivíduos da Grande São Paulo (as amostras de cabelo pertenciam aos mesmos indivíduos que doaram as respectivas amostras de sangue). Os indivíduos foram selecionados a partir de sua autodenominação em brancos ou pretos, segundo critérios do IBGE ([www.ibge.gov.br](http://www.ibge.gov.br)).

OBSERVAÇÃO: As amostras foram doadas pelo Prof. Dr. Mário Hirata do Departamento de Análises Clínicas e Toxicológicas da Faculdade de Ciências Farmacêuticas da Universidade de São Paulo, conforme Resolução nº 347 de 13 de janeiro de 2005 e aprovado pelo CEP da FCF/USP (ofício nº 52, agosto de 2001).

### 5.2 SEQUENCIAMENTO DA REGIÃO HIPERVARIÁVEL DO DNA MITOCONDRIAL EM AMOSTRAS DE SANGUE

#### 5.2.1 Extração do DNA

Para a extração do DNA das amostras de sangue foi utilizado o primeiro protocolo descrito a seguir. Quando não foi possível a amplificação do DNA com este protocolo, foi utilizado o segundo protocolo descrito:

(1) Um disco de 3 mm de papel de filtro contendo a amostra de sangue foi cortado utilizando *Harris Micro Punche e Cutting Mat* (Whatman-Bioscience) e depositado em tubo de 1,5 ml.

Procedeu-se à extração utilizando *FTA Reagent* (Whatman) e eluição com TE (10 mM Tris-HCl; 0,1 mM EDTA; pH 7,5) deixando-se o papel (DNA adsorvido) secar em estufa por 2 horas; ou (2) Seis discos de 3 mm de papel de filtro foram utilizados para extração com o kit *DNA IQ* (Promega), eluindo-se o DNA em 50 µl de tampão (concentração final de 2 ng/µl).

### 5.2.2 Reação de PCR (*Reação em Cadeia da Polimerase*)

A região hipervariável foi amplificada em toda a sua extensão utilizando-se:

2,5 µl de *GoldStar Buffer* (Promega)

0,2 µl (1 U) de *Platinum Taq DNA polymerase* (Invitrogen)

0,2 µM de cada *primer* (HV1-L15997 e HV3-R639) (Tabela 1)

1 disco de 3 mm de papel contendo o DNA ou 4 ng de DNA

Água ultrapura (Milli-Q) autoclavada qsp. 25 µl.

Os fragmentos foram amplificados em termociclador *GeneAmp PCR System 9700*

(Perkin Elmer) conforme ciclagem abaixo:

95°C – 11 min

96°C – 1 min

94°C – 30 seg

56°C – 30 seg

72°C – 90 seg

} 40 X

60°C – 11 min

4°C - ∞

Tabela 1 – *Primers* utilizados na reação de PCR e sequenciamento.

	<i>Primer</i>	Orientação	Sequência
HV1	L15997	<i>forward</i>	5'-CACCATTAGCACCCAAAGCT-3'
	H16401	<i>reverse</i>	5'-TGATTTACGGAGGATGGTG-3'
	L16209	<i>forward</i>	5'-CCCCATGCTTACAAGCAAGT-3'
	R16164 <sup>(*)</sup>	<i>reverse</i>	5'-TTGATGTGGATTGGGT-3'
HV2	L29	<i>forward</i>	5'-GGTCTATCACCCCTATTAACCAC-3'
	H408	<i>reverse</i>	5'-CTGTAAAAGTGCATACCGCCA-3'
	159R	<i>reverse</i>	5'-AAATAATAGGATGAGGCAGGAATC-3'
	H285 <sup>(*)</sup>	<i>reverse</i>	5'-GGGGTTTGGTGGAAATTTTTTG-3'
HV3	F314	<i>forward</i>	5'-CCGCTTCTGGCCACAGCACT-3'
	599R <sup>(*)</sup>	<i>reverse</i>	5'-TTGAGGAGGTAAGCTACATA-3'
	R639	<i>reverse</i>	5'-GGGTGATGTGAGCCCGTCTA-3'

Nota: (\*) *Primers* internos à sequência de interesse que foram utilizados quando não foi possível obter a leitura da sequência devido à presença de heteroplasmia de comprimento ou não anelamento do *primer*.

Após a amplificação dos fragmentos, a avaliação dos produtos da reação e sua quantificação foram realizadas visualmente após eletroforese em gel de agarose 1,0% (Cuba Horizon 58, Model 200 - Gibco) por 15 minutos a 100 volts, utilizando-se 4 µl do produto da PCR e comparando-se a intensidade das bandas com aquelas do marcador *Low DNA Mass Ladder* (Invitrogen) após coloração com Brometo de Etídio (BrEt).

### 5.2.3 Purificação do produto de PCR

O produto de PCR (~1200 nt) foi purificado previamente à realização da reação de sequenciamento para retirada dos nucleotídeos não incorporados e excesso de *primers*. Nessa

etapa, as amostras que não estavam na concentração mínima necessária para o sequenciamento (de 1 a 40 ng de produto de PCR por reação, dependendo do *primer* utilizado) foram concentradas. Foram utilizadas as colunas de purificação *GFX PCR DNA and gel band purification* kit (GE Healthcare), conforme as recomendações do fabricante. Após a purificação das amostras, as mesmas foram quantificadas por fluorimetria utilizando aparelho *Qubit Fluorometer* (Invitrogen).

#### 5.2.4 Reação de sequenciamento

A reação de sequenciamento foi realizada conforme protocolo abaixo:

- 1 µl de *Big Dye Terminator* v 1.0 ou 3.1 (Applied Biosystems by Life Technologies)
- 2 µl do *primer forward* ou *reverse* (1,6 picomolar)
- 2 µl de tampão para *Big Dye*
- produto de PCR purificado suficiente para a reação
- água ultrapura (Milli-Q) autoclavada qsp. 10 µl

Um esquema dos *primers* utilizados na amplificação por PCR e sequenciamento pode ser visualizado na figura 5.

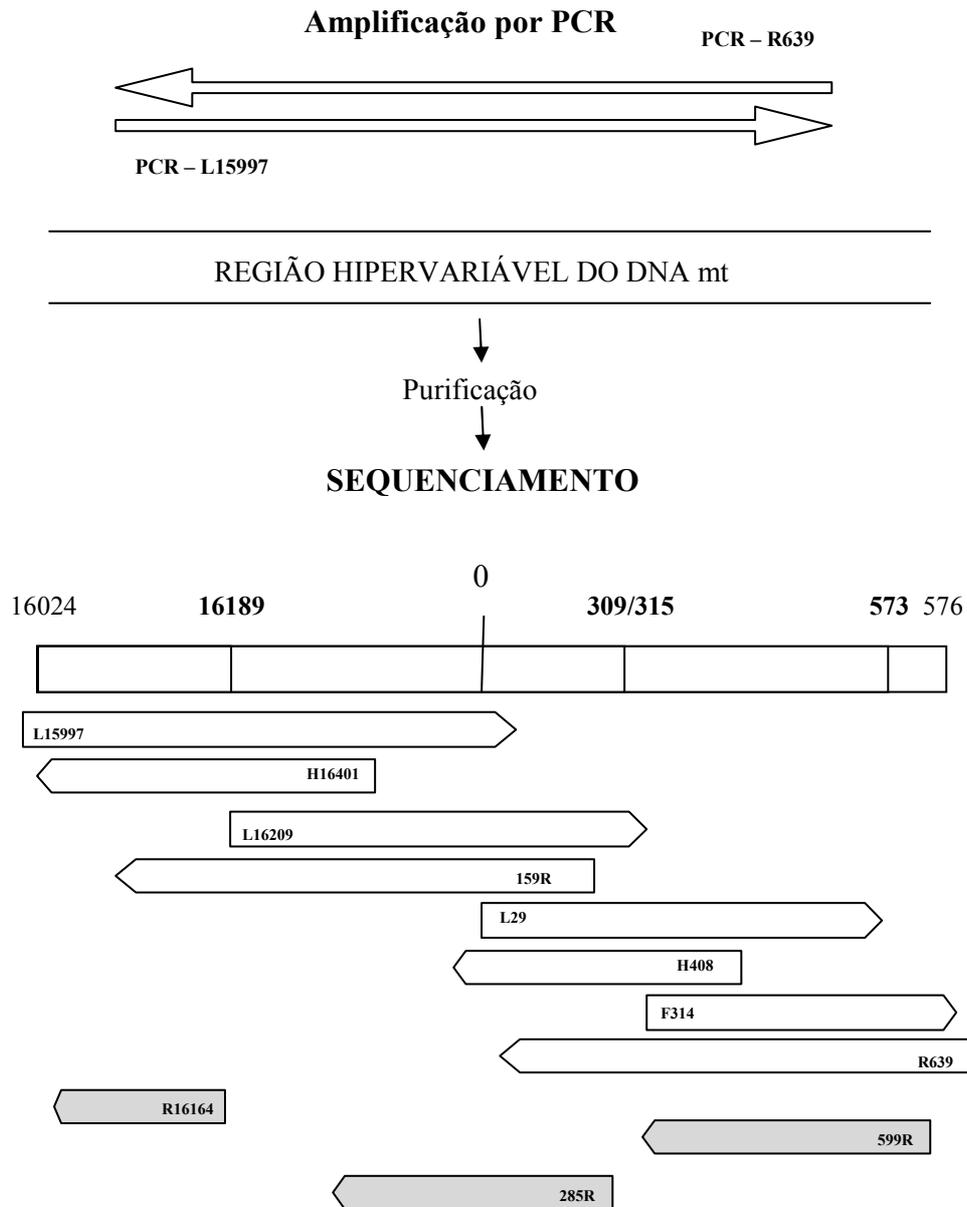


Figura 5 – Esquema da localização dos *primers* utilizados na PCR e sequenciamento.

Nota: em negrito encontramos as regiões poli-C (posição 16189, 309/315 e 573).

*Primers* extras destacados na parte inferior na figura.

Fonte: Figura modificada de Parson; Bandelt, 2007.

A reação de sequenciamento foi realizada em termociclador *GeneAmp PCR System 9700* utilizando-se a ciclagem abaixo:

96°C – 1 min  
96°C – 10 seg  
56°C – 10 seg  
60°C – 4 min  
4°C - ∞

} 25X

### 5.2.5 Purificação do produto da reação de sequenciamento

O produto da reação de sequenciamento foi purificado para retirada de nucleotídeos não incorporados por meio da técnica de precipitação com etanol/EDTA ou resina *Big Dye X Terminator Purification Kit* (no caso de fragmentos pequenos) conforme recomendações do fabricante do *Big Dye* (Applied Biosystems by Life Technologies).

### 5.2.6 Eletroforese e detecção

A amostra foi eluída em formamida e a eletroforese foi realizada em ABI 377 ou 3130 (Applied Biosystems by Life Technologies), conforme recomendações do fabricante.

### 5.2.7 Análise dos resultados

#### 5.2.7.1 Ferramentas de bioinformática

A análise dos resultados do sequenciamento foi feita utilizando o programa BioEdit (HALL, 1999). Os resultados do sequenciamento de ambas as fitas foram comparados para se

chegar a um consenso de sequência. Os polimorfismos encontrados foram anotados em planilha Excel 2000.

#### 5.2.7.2 Classificação em haplogrupos

A partir dos resultados do sequenciamento, a classificação em haplogrupos foi gerada a partir do banco de dados mtDNAManager (LEE et al., 2008) e confirmada em Phylotree.org (VAN OVEN; KAYSER, 2009).

#### 5.2.7.3 Análise estatística

Os cálculos dos índices de diversidade nas amostras de sangue foram realizados utilizando o programa Arlequin v 3.1 (EXCOFFIER et al., 2005). A estrutura genética das populações foi analisada utilizando-se AMOVA (Análise de Variância Molecular) (EXCOFFIER et al., 1992). A sequência, de toda a região hipervariável, para cada indivíduo analisado de cada população utilizada para comparação com a nossa população, foi obtida a partir do banco de dados GenBank e configurada para o programa Arlequin utilizando o programa DNA SP v 5.10 (LIBRADO; ROZAS, 2009).

### 5.3 PCR EM TEMPO REAL

#### 5.3.1 Extração do DNA

Foi utilizado o protocolo de extração de DNA com o kit *DNA IQ*, já descrito neste trabalho.

### 5.3.2 Construção dos *primers* e sondas

Os *primers* e sondas para discriminação alélica por PCR em tempo real foram construídos e padronizados pela empresa fornecedora dos mesmos (Assay by Design, Applied Biosystems by Life Technologies), utilizando o sistema *TaqMan* e podem ser visualizados na tabela 2.

Tabela 2 - *Primers* e sondas *TaqMan* utilizadas nas reações de PCR em tempo real

SNP 3010	Sequência
<i>Primer Forward</i>	TTTACGACCTCGATGTTGGATCAG
<i>Primer Reverse</i>	TCACGTAGGACTTTAATCGTTGAACAA
Sondas	VIC-ACATCCC <u>G</u> ATGGTGC-NFQ
	FAM-ACATCCC <u>A</u> ATGGTGC-NFQ

Nota: VIC e FAM são fluoróforos, NFQ (Non Fluorescent Quencher). Em negrito e sublinhados na sequência das sondas estão os nucleotídeos polimórficos.

As reações de genotipagem foram realizadas em volume final de 25 µl:

- 12,5 µl de 2x *TaqMan Universal PCR Master Mix* (Applied Biosystems by Life Technologies)
- 1,25 µl de 20x *SNP Genotyping Assay Mix* (Applied Biosystems by Life Technologies)
- 2 µl de DNA (4 ng)
- água ultrapura (Milli-Q) autoclavada qsp. 25 µl

Para a genotipagem foi realizada uma leitura pré e pós amplificação (*pre-read and post-read*) utilizando a opção “discriminação alélica” do software *7500 System* em ABI 7500, conforme protocolo do fabricante (*pre e post-read* a 60°C por 1 minuto, com amplificação consistindo de 10 min a 95°C, seguida de 40 ciclos de 15 seg a 92°C e 1 min a 60°C).

Previamente às reações de genotipagem por PCR em tempo real algumas amostras foram genotipadas por sequenciamento para serem utilizadas como controle de ambos os alelos possíveis.

## 5.4 SNAPSHOT

### 5.4.1 Extração do DNA

Para a extração do DNA das amostras de sangue, dois discos de 3 mm de papel de filtro foram cortados utilizando *Harris Micro Punche e Cutting Mat* e o DNA foi extraído utilizando a resina *Chelex* (SINGER-SAM et al., 1989). A quantificação das amostras foi realizada com o kit *Quantifier Human DNA Quantification* (Applied Biosystems by Life Technologies) por PCR em tempo real.

### 5.4.2 Seleção dos SNPs

Quarenta polimorfismos (39 SNPs e 1 deleção) presentes na região codificadora do DNA mt foram selecionados de forma a permitirem a classificação das amostras em haplogrupos europeus, africanos e nativo-americanos. Outros dois SNPs presentes na região não-codificadora (SNP 16519 e 16362) foram escolhidos para o estudo do aumento do poder discriminatório dos mesmos em nossa população. Isto foi feito utilizando uma reação

SNaPshot em *multiplex* previamente publicada, composta apenas por SNPs que permitiam classificação em haplogrupos europeus (KÖHNEMANN et al., 2009), a qual foram adicionados SNPs para a classificação em haplogrupos africanos (L0, L1, L2, L3, L4) e nativo-americanos (A, B, C, D) baseados nos haplogrupos já descritos na literatura para a população brasileira (ALVES-SILVA et al., 2000). Os SNPs escolhidos foram descritos por PhyloTree.org (VAN OVEN; KAYSER, 2009) (Fig. 6).

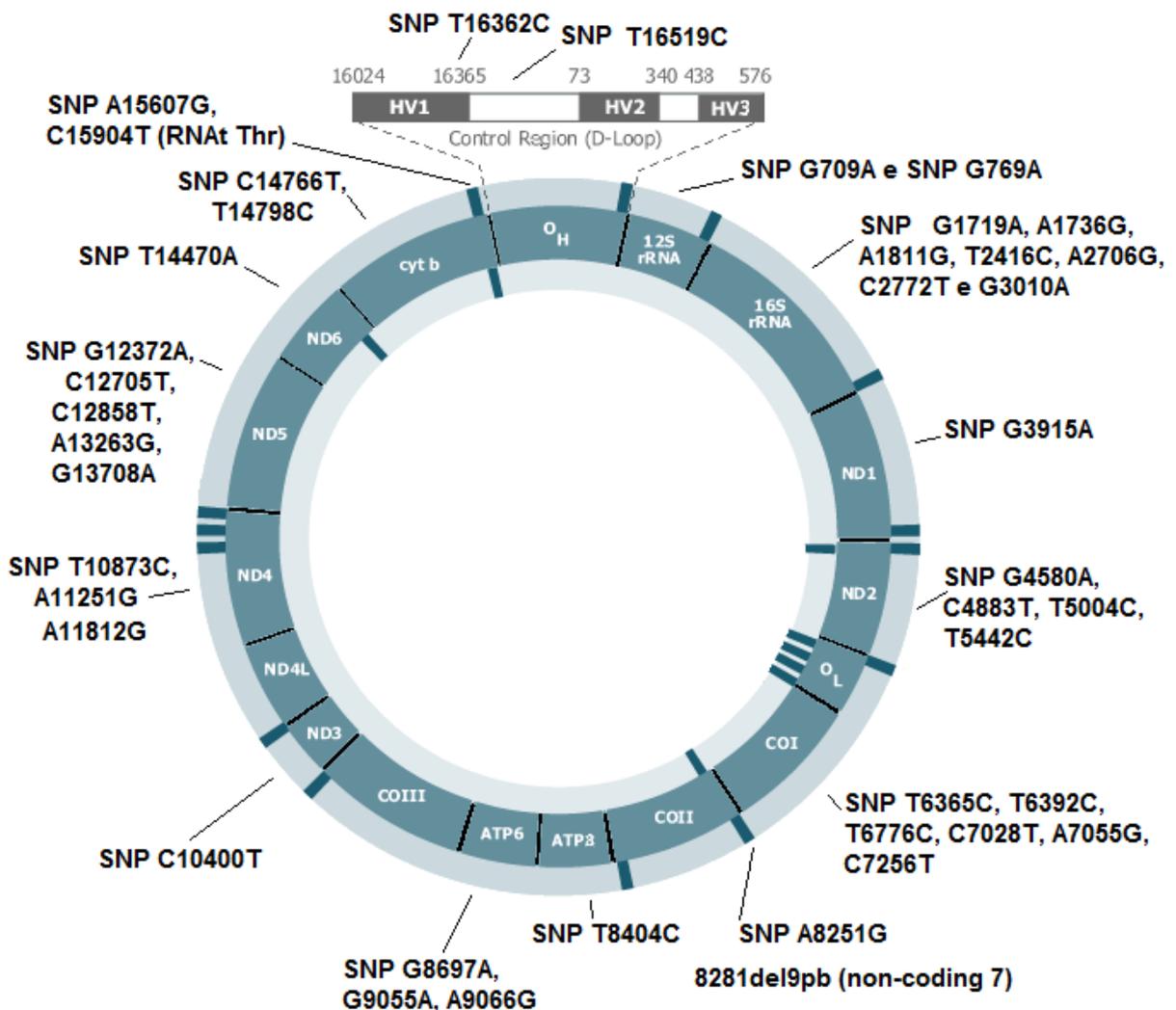


Figura 6: Esquema da localização dos 42 polimorfismos utilizados na reação de SNaPshot para classificação em haplogrupos do DNA mt.

OBS. Todas as etapas da reação de SNaPshot descritas nesta tese foram realizadas durante estágio de doutorado no Instituto de Medicina Legal da Universidade de Münster, Alemanha, sob orientação da Dra. Heidi Pfeiffer e colaboração do doutorando Stephan Köhnemann.

### 5.4.3 Reação de PCR em *multiplex*

Os *primers* para a reação de PCR em *multiplex* foram desenhados utilizando o programa PerlPrimer (MARSHALL, 2004) e a possibilidade de interação entre os mesmos foi descartada utilizando o programa Autodimer (VALLONE; BUTLER, 2004). Trinta e três pares de *primers* foram utilizados para amplificar 42 SNPs. O tamanho dos amplicons variou de 55 a 196 pares de base para permitir a utilização da metodologia para análise de amostras degradadas. A concentração final dos *primers* utilizada na reação e outras informações estão dispostas na Tabela 3 (Apêndice A).

A reação de PCR *multiplex* foi realizada utilizando-se:

- 1,25 µl de MgCl<sub>2</sub> (50 mM)
- 1,25 µl de dNTPs (2,5 mM)
- 1,25 µl de *PCR Gold buffer* (10x) (Applied Biosystems by Life Technologies)
- 0,2 µl de soro albumina bovina (20 mg/ml)
- 0,2 µl de *AmpliTaq Gold* (Applied Biosystems by Life Technologies)
- 100 pg de DNA
- *primers* (Tabela 3)
- água ultrapura (Milli-Q) autoclavada qsp. 12,5 µl

Os fragmentos foram amplificados em termociclador *GeneAmp PCR System 9700* (Perkin Elmer) conforme a ciclagem: Pré-incubação a 95°C por 10 min, seguida de três ciclos de 30 seg a 95°C, 45 seg a 55°C e 30 seg a 72°C, após esta etapa 19 ciclos de 30 seg a 95°C, 45 seg a 55°C com aumento de temperatura em 0,2°C por ciclo, 30 seg a 72°C; seguida de 11 ciclos de 30 seg a 95°C, 45 seg a 60°C, 30 seg a 72°C; e finalmente 7 min a 72°C. Os produtos da reação de PCR foram purificados através de reação enzimática adicionando-se 2,5 µl de *Exo/SAP-IT* (USB) e incubando-se 120 min a 37°C, seguida de desnaturação da enzima a 75°C por 15 min.

#### 5.4.4 Reação de SNaPshot

Os *primers* para a reação de SNaPshot foram desenhados manualmente para se anelarem a um nucleotídeo antes da posição do SNP que se deseja genotipar. Uma cauda nucleotídica, descrita por Sanchez et al. (2005), com diferentes números de bases foi adicionada à região 5' de cada *primer*. Esta cauda, que não se anelava ao produto de PCR previamente amplificado, foi adicionada para permitir a separação dos fragmentos por tamanho em eletroforese capilar. A temperatura de anelamento de cada um dos *primers* foi checada com o programa Perlprimer e a formação de dímeros e *hairpins* foi descartada a partir da análise dos mesmos pelo programa Autodimer. Os *primers* utilizados e outras informações podem ser visualizados na Tabela 4 (Apêndice B).

A reação de SNaPshot foi realizada utilizando-se:

- 2,5 µl *SNaPshot Mastermix*
- 0,5 µl *PCR Gold buffer*
- 0,5 µl do *mix* dos *primers*

- 1,0  $\mu$ l do produto de PCR purificado
- água ultrapura (Milli-Q) autoclavada qsp. 10  $\mu$ l

A reação de sequenciamento foi realizada em termociclador PTC-100 (MJ Research) utilizando-se a ciclagem abaixo:

96°C – 10 seg	}	25X
55°C – 5 seg		
60°C – 30 seg		
4°C - $\infty$		

Os produtos da reação foram purificados por tratamento enzimático adicionando-se 1,25  $\mu$ l de *SAP* (fosfatase alcalina de camarão) e incubação a 37°C por 75 min, seguida de desnaturação da enzima por 15 min a 75°C.

Os fragmentos foram separados por eletroforese em *ABI 3100-Avant Genetic Analyzer*, conforme recomendações do fabricante. O resultado foi analisado utilizando o programa GeneScan (v 3.7, Applied Biosystems by Life Technologies) e Genotyper (v 2.5, Applied Biosystems by Life Technologies) (Fig. 7).

#### 5.4.5 Sensibilidade

A validação da sensibilidade da técnica foi feita utilizando uma série de diluições de um DNA padrão (DNA 9947A - Promega): 10, 1, 0,1, 0,01 e 0,001 ng.

#### 5.4.6 Classificação em haplogrupos

Para os resultados do SNaPshot a classificação em haplogrupos foi baseada em Phylotree.org.

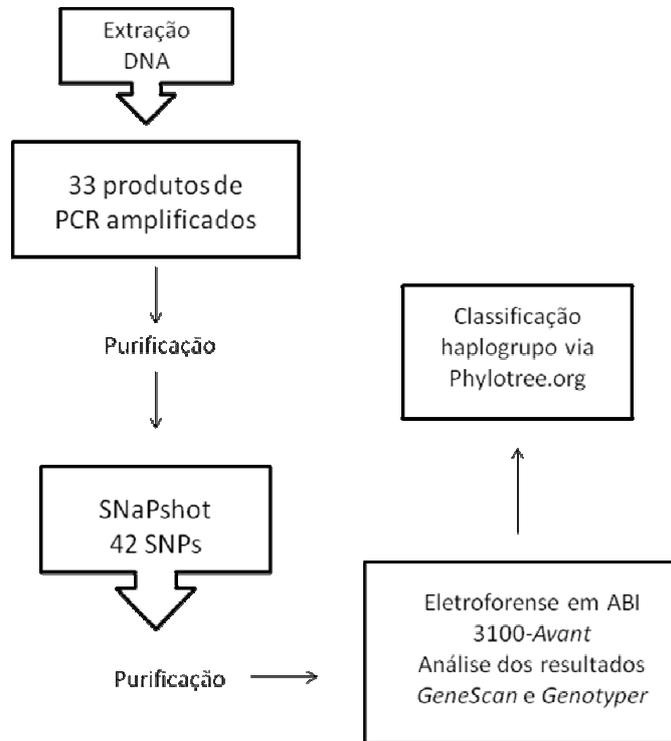


Figura 7: Esquema indicando todas as etapas realizadas para análise das amostras pela técnica de SNaPshot.

## 5.5 SEQUENCIAMENTO DA REGIÃO HV3 DO DNA MITOCONDRIAL EM AMOSTRAS DE CABELO

### 5.5.1 Extração do DNA

Para a extração do DNA das amostras de cabelo, foram usados 4 cm de cabelo (sem raiz) de cada indivíduo. Retirou-se a raiz cortando-se 0,5 cm da extremidade do fio.

OBS: Em indivíduos com fios menores que 4 cm foram utilizados diferentes fios até que totalizassem esse tamanho.

Os fios foram cortados em pedaços de 0,5 cm e colocados em tubo eppendorf de 1,5 ml. Procedeu-se ao protocolo de descontaminação conforme Jehaes et al. (1998). Após esta

etapa, procedeu-se à extração utilizando o kit *Tissue and Hair Extraction* (Promega) com as seguintes modificações:

As amostras foram incubadas por 2 horas e 30 min para digestão com 75 µl de Proteinase-K (18 µg/µl em buffer) e 100 µl de *dithiothreitol* (DTT) 1M (Promega) seguido de extração com fenol:clorofórmio:álcool isoamílico (PCA) 25:24:1 (Invitrogen) e purificação por coluna *GFX PCR DNA and gel band purification kit*, conforme recomendações do fabricante.

### 5.5.2 Reação de PCR (*Reação em Cadeia da Polimerase*)

A reação de PCR foi realizada utilizando-se:

- 5,0 µl de *GoldStar Buffer* (Promega)
- 0,4 µl (2 U) de *AmpliTaqGold polymerase* (Applied Biosystems by Life Technologies)
- 0,2 µM de cada *primer* (HV3-F314 e HV3-R639) (IDT) (Tabela 1)
- 10 µl de DNA (cabelo)
- água ultrapura (Milli-Q) autoclavada qsp. 50 µl.

OBS. A cada conjunto de amostras amplificadas foi realizada uma reação de controle negativo e positivo utilizando água ultrapura autoclavada e 2 ng de um DNA padrão (DNA 9947A – Promega), respectivamente.

As etapas que se sucederam à amplificação são as mesmas já descritas para o sequenciamento de amostras de sangue.

### 5.5.3 Clonagem

Os produtos de PCR contendo heteroplasmia de comprimento na região de repetições de CA em HV3 foram clonados em *pGEM-T Vector* (Promega) para o isolamento de cada um dos alelos. A mini-prep foi realizada conforme protocolo descrito por Sambrook et. al. (1989). Após o isolamento as amostras foram ressequenciadas.

### 5.5.4 Análise dos resultados

#### 5.5.4.1 Ferramentas de bioinformática

A análise dos resultados do sequenciamento foi feita utilizando o programa BioEdit. Os resultados do sequenciamento de ambas as fitas foram comparados para se chegar a um consenso de sequência. Nas amostras de cabelo, os perfis encontrados foram comparados aos obtidos nas respectivas amostras de sangue. A leitura de heteroplasmia foi realizada visualmente em cada eletroferograma (Fig. 8).

#### 5.5.4.2 Detecção de heteroplasmia

Como pré-requisitos para se afirmar a presença de heteroplasmia de sequência nas amostras de cabelo foram considerados os seguintes parâmetros:

- Presença de dois nucleotídeos (dois picos) em uma mesma posição em ambas as fitas do DNA;
- Porcentagem do pico menor deve ser pelo menos 10% da área do pico maior;
- Reprodutibilidade do resultado (todas as etapas foram refeitas utilizando o mesmo DNA extraído).

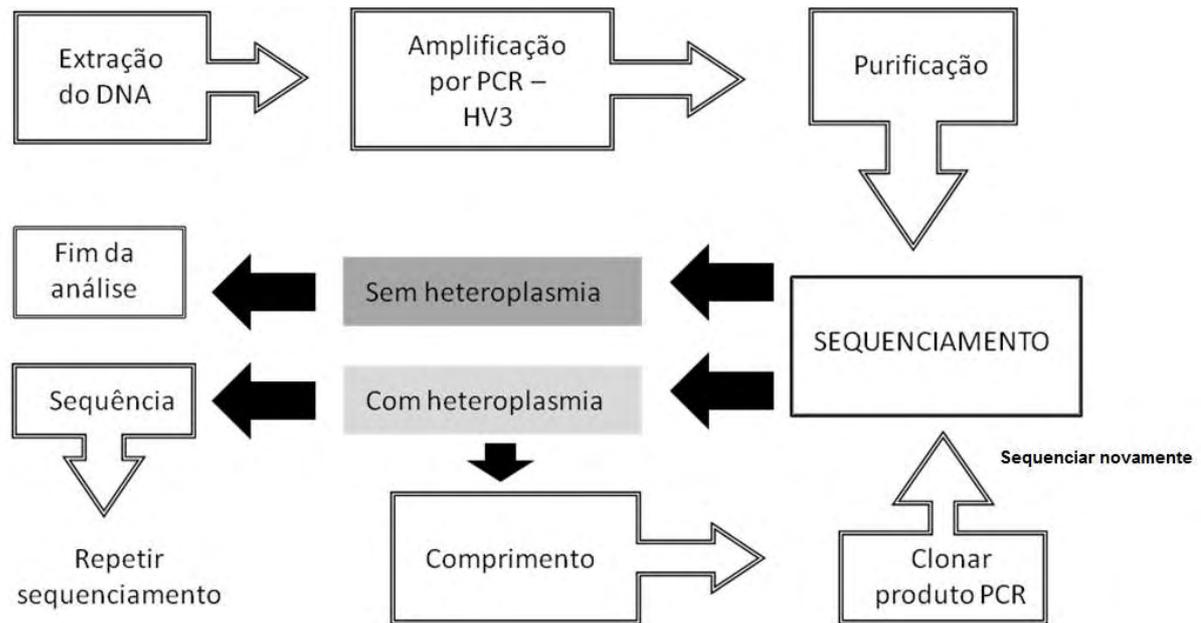


Figura 8: Esquema das etapas realizadas para análise de heteroplasmia por sequenciamento nas amostras de sangue e cabelo.

## 6. RESULTADOS E DISCUSSÃO

### 6.1 SEQUENCIAMENTO DAS AMOSTRAS DE SANGUE

Os haplótipos obtidos neste trabalho podem ser visualizados na tabela 5 (Apêndice C). De um total de 160 amostras, 144 haplótipos diferentes foram encontrados quando analisada toda a região hipervariável do DNA mt (desconsiderando as regiões poli-C presentes nas posições 16193, 309 e 573), dentre os quais um haplótipo foi compartilhado por quatro indivíduos, outro haplótipo compartilhado por três indivíduos e onze haplótipos estavam presentes em dois indivíduos (Tabela 6); 131 haplótipos (81,88% das amostras) foram únicos. Dentre os 1.122 nucleotídeos presentes na região hipervariável (posição 16024 a 576), 190 (16,86%) posições nucleotídicas apresentavam-se polimórficas. As posições nucleotídicas

mais discriminatórias, ou seja, as posições que apresentaram os dois alelos possíveis em proporções semelhantes foram 152 (50,12%), 16519 (47,77%), e 16223 (45,78%).

Tabela 6 - Haplótipos encontrados em mais de um indivíduo em nossa amostra populacional.

Haplogrupo <sup>†</sup>	Haplótipo	N
B2c	73G, 103A, 152C, 263G, 499A, 16182C, 16183C, 16189C, 16217C, 16241G, 16519C	4
H	152C, 263G, 16519C	3
H	263G, 16311C, 16519C	2
H	263G, 16519C	2
C1b	73G, 249DEL, 263G, 290DEL, 291DEL, 489C, 493G, 523DEL, 524DEL, 16223T, 16298C, 16325C, 16327T	2
C1b	73G, 249DEL, 263G, 290DEL, 291DEL, 489C, 493G, 513A, 523DEL, 524DEL, 16223T, 16292T, 16298C, 16325C, 16327T, 16362C	2
C1b	73G, 249DEL, 263G, 290DEL, 291DEL, 489C, 493G, 523DEL, 524DEL, 16126C, 16223T, 16298C, 16325C, 16327T	2
L0a1b	93G, 95C, 185A, 189G, 236C, 247A, 263G, 523DEL, 524DEL, 16129A, 16148T, 16168T, 16172C, 16187T, 16188G, 16189C, 16223T, 16230G, 16278T, 16293G, 16311C, 16320T	2
L2a	73G, 146C, 152C, 195C, 263G, 16182C, 16183C, 16189C, 16223T, 16278T, 16290T, 16294T, 16309G, 16390A	2
L2a1	73G, 146C, 152C, 195C, 263G, 16189C, 16223T, 16278T, 16294T, 16309G, 16390A, 16519C	2
L2b1a	73G, 150T, 152C, 182T, 195C, 198T, 204C, 263G, 418T, 523DEL, 524DEL, 16114A, 16129A, 16213A, 16223T, 16278T, 16355T, 16362C, 16390A	2
L3e1a	73G, 150T, 189G, 200G, 263G, 16185T, 16223T, 16327T, 16519C	2
L3e2b	73G, 150T, 152C, 195C, 16172C, 16183C, 16189C, 16223T, 16320T, 16519C	2

Nota: <sup>†</sup> Classificação baseada nos resultados do sequenciamento da região hipervariável. “N” corresponde ao número de amostras.

A probabilidade de encontrarmos dois indivíduos não relacionados com perfis idênticos em nossa amostra (probabilidade de semelhança) foi de 0,78% e a diversidade

haplotípica calculada foi de 0,9984 +/- 0,0010 (para toda a região hipervariável do DNA mt e desconsiderando as inserções e deleções presentes nas regiões poli-C). As análises estatísticas de toda a região hipervariável (16024-576) e de cada uma das regiões hipervariáveis HV1, HV2 e HV3 em separado e agrupadas podem ser visualizadas na tabela 7.

Tabela 7 - Parâmetros estatísticos calculados a partir da análise de toda a região hipervariável e de cada uma das regiões hipervariáveis em separado e agrupadas nos 160 indivíduos.

<i>Parâmetros estatísticos</i>	<i>HV1</i>	<i>HV2</i>	<i>HV3</i>	<i>HV1+2</i>	<i>HV1+2+3</i>	<i>16024-576</i>
Número de haplótipos	121	75	26	136	140	144
Diversidade haplotípica	0,9939 +/- 0,0021	0,9642 +/- 0,0078	0,7844 +/- 0,0257	0,9972 +/- 0,0012	0,9978 +/- 0,0011	0,9984 +/- 0,0010
Probabilidade de semelhança	1,23%	4,19%	22,05%	0,9%	0,84%	0,78%
Transições observadas	95	43	13	138	151	174
Transversões observadas	11	4	0	15	15	15
Substituições observadas	106	47	13	153	166	189
Deleções observadas	1	7	7	8	15	16
Comparação pareada entre haplótipos	8,050550 +/- 3,758206	4,979245 +/- 2,434421	2,030582 +/- 1,147573	13,312107 +/- 6,018481	15,342689 +/- 6,889882	16,668947 +/- 7,458895
Diversidade nucleotídica	0,023540 +/- 0,012163	0,018649 +/- 0,010092	0,014200 +/- 0,008882	0,021823 +/- 0,010920	0,020375 +/- 0,010127	0,014791 +/- 0,007325
Sítios polimórficos	94	52	20	146	166	190
Nucleotídeos editados	342	248	139	590	729	1122

Dentre as três regiões HV1, HV2 e HV3, a região HV1 foi a que sozinha apresentou maior diversidade haplotípica, seguida de HV2 e HV3. A análise conjunta de HV1+2, que comumente é utilizada na prática forense, resultou em um incremento significativo da

diversidade. A adição da análise de HV3 e das outras regiões intermediárias (totalizando a análise de toda a região hipervariável) dentre as mesmas, apesar de não incrementarem tanto a diversidade haplotípica, adicionam informações complementares que podem ser importantes na solução de casos forenses (LUTZ et al., 2000).

Quando os resultados obtidos neste trabalho (apenas HV1, visto que nem todos os autores analisaram todas as regiões hipervariáveis) foram comparados àqueles obtidos em outras regiões do Brasil, verificamos que os mesmos são semelhantes aos resultados publicados para a região Nordeste do país (Estado do Alagoas). Os resultados da região Norte do país (Estado do Pará) apresentaram uma diversidade inferior à da região Sudeste e Nordeste, com maior proporção de haplogrupos nativo-americanos em sua constituição (Tabela 8). Os resultados da região Sul (Santa Catarina) foram os que apresentaram a menor diversidade. Isso se deve, provavelmente, a grande proporção de indivíduos de origem europeia, os quais apresentam menor diversidade haplotípica (SALAS et al., 1998).

Tabela 8 – Índices de diversidade das populações do Sudeste, Sul, Nordeste e Norte do Brasil para a região HV1.

<i>Estatística populacional</i>	<i>Sudeste (SP)</i>	<i>Sul(SC)</i>	<i>Nordeste(AL)</i>	<i>Norte(PA)</i>
N	<b>160</b>	80	167	158
Referência	<b>Este trabalho</b>	Palencia et al., 2010	Barbosa et al., 2008	Feio-dos-Santos et al., 2006
Número de haplótipos	<b>121</b>	ND	124	49
Diversidade haplotípica	<b>0,9939 +/- 0,0021</b>	0,9684+/- 0,0124	0,9934+/- 0,0020	0,989+/- 0,0067
Probabilidade de semelhança	<b>1,23%</b>	4,37%	1,25%	2,8%

Nota: “N” corresponde ao número de indivíduos analisados. ND: Não Declarado. SP (São Paulo), SC (Santa Catarina), AL (Alagoas), PA (Pará).

## 6.2 COMPARAÇÃO DOS RESULTADOS COM OUTRAS POPULAÇÕES MUNDIAIS

Quando os resultados estatísticos de nossa amostra populacional foram comparados com resultados de outras populações mundiais observou-se que, mesmo não possuindo um número de tão grande de amostras genotipadas como o da Suécia, Egito e EUA (*Afri*), os mesmos apresentaram uma altíssima diversidade haplotípica, resultante da grande miscigenação da população brasileira (Tabela 9).

Tabela 9: Parâmetros estatísticos em nossa amostra populacional e em outras populações mundiais.

<i>Estatística populacional</i>	<i>Brasil</i>	<i>Argentina (Cent)</i>	<i>EUA (Afri)</i>	<i>EUA (Hisp)</i>	<i>Suécia</i>	<i>Egito</i>
N	<b>160</b>	193	248	128	335	277
Referência	<b>Este trabalho</b>	Bobillo et al., 2009	Diegoli et al., 2009	Saunier et al., 2008	Tillmar et al., 2010	Saunier et al., 2009
Número de haplótipos	<b>144</b>	176	219	115	224	238
Diversidade haplotípica	<b>0,9984</b>	0,937	0,9985	0,9980	0,996	0,9986
Probabilidade de semelhança	<b>0,78%</b>	6,3%	0,14%	0,23%	0,5%	0,49%

Nota: N: número de amostras; Cent (Central); Afri (Pretos americanos); Hisp (hispânicos). Toda região hipervariável (16024-576) foi levada em consideração.

A fim de analisar e comparar a estrutura genética de nossa amostra populacional com outras populações mundiais nós selecionamos duas populações (USA – composto de populações classificadas em pretos e hispânicos - e Kênia), cujos haplótipos de toda a região hipervariável estavam disponíveis no GenBank (BRANDSTÄTTER et al., 2004a; DIEGOLI et al., 2009; SAUNIER et al., 2008). A variabilidade genética dentro e entre as populações (Brasil, EUA e Kênia) foi avaliada utilizando o teste AMOVA (*Analysis of Molecular*

*Variance*). Observou-se que a maior parte da variabilidade genética ocorre dentro de cada uma das populações (97,27%) e não entre as mesmas (2,73%), confirmando que existem mais semelhanças genéticas entre populações humanas geograficamente distintas do que diferenças. As estimativas par a par de *Fst* foram significativas para todas as comparações entre as três populações (*Fst p value* = 0) (Tabela 10), demonstrando uma diferenciação genética, no que se refere ao DNA mt, entre as mesmas. Isso reforça a importância da criação e utilização de bancos de dados de frequências dos haplótipos do DNA mt específicas para cada população.

Tabela 10: Valores de *Fst* e *Fst p value* resultantes da comparação pareada entre populações.

	<b>Brasil/São Paulo</b>	<b>EUA/Pretos e Hispanicos</b>	<b>Kênia</b>
<b>Brasil/São Paulo</b>	*	0,00000	0,00000
<b>EUA/Pretos e Hispanicos</b>	0,01141	*	0,00000
<b>Kênia</b>	0,04872	0,04069	*

Nota: Abaixo do eixo transversal estão representados os valores de *Fst*; acima do eixo transversal estão representados os valores de *Fst p value*.

Visto que nossa amostragem, assim como a da população americana, poderia ser dividida em dois grupos de acordo com a cor de pele dos indivíduos (brancos e pretos na população estudada) foi realizada uma nova análise separando os subgrupos (Tabela 11). Dentre os 160 indivíduos analisados neste trabalho, 89 se declararam como brancos e 71 se declararam como pretos segundo critério do IBGE (branco, preto, pardo, amarelo ou indígena) (Tabela 12 - Apêndice D). Observou-se que, quando comparamos par a par os valores de *Fst* das diversas subpopulações, todas apresentaram valores estatisticamente significantes (*Fst p value* < 0,05), ou seja, todas diferiram em sua estrutura genética. No entanto, quando comparadas as subpopulações deste trabalho, cujos indivíduos foram

classificados como brancos ou pretos, as mesmas estavam geneticamente mais próximas ( $F_{st}$   $p$  value = 0,03604), apesar deste valor ainda demonstrar uma diferenciação significativa.

Tabela 11: Valores de  $F_{st}$  e  $F_{st}$   $p$  value resultantes da comparação pareada entre subpopulações.

	<b>BR/SP- pretos</b>	<b>BR/SP- brancos</b>	<b>EUA- pretos</b>	<b>EUA- hispanicos</b>	<b>Kenianos</b>
<b>BR/SP-pretos</b>	*	<b>0,03604</b>	0,00000	0,00000	0,00000
<b>BR/SP-brancos</b>	0,01308	*	0,00000	0,00000	0,00000
<b>EUA- pretos</b>	0,02451	0,06334	*	0,00000	0,00000
<b>EUA-hispanicos</b>	0,05169	0,03442	0,12238	*	0,00000
<b>Kenianos</b>	0,04507	0,05416	0,05381	0,09628	*

Nota: Abaixo do eixo transversal estão representados os valores de  $F_{st}$ ; acima do eixo transversal estão representados os valores de  $F_{st}$   $p$  value. BR/SP (Brasil/São Paulo), EUA (Estados Unidos da América). Para valores estatisticamente significantes  $F_{st}$   $p$  value <0,05.

Os resultados do sequenciamento de toda a região hipervariável do DNA mt foram enviados para o *Grupo de Habla Española y Portuguesa*, da Sociedade Internacional de Genética Forense (GHEP-ISFG), sob coordenação da Dra. Lourdes Prieto Solla, para conferência e atualmente estão sendo depositados no banco de dados mundial EMPOP.

### 6.3 ANÁLISE DO SNP 3010 POR PCR EM TEMPO REAL

Os resultados da genotipagem do SNP 3010 podem ser visualizados na tabela 12 (Apêndice D). Das 160 amostras analisadas, 140 apresentaram o alelo G, 17 apresentaram o alelo A e 3 amostras não puderam ser genotipadas pela técnica (resultado indeterminado), provavelmente por apresentarem uma concentração de DNA menor do que o recomendado para a análise. Estas últimas foram sequenciadas e todas apresentavam o alelo G. O gráfico da

figura 9 apresenta como foram gerados os resultados pelo software de análise do PCR em tempo real (*7500 System*, modo Discriminação Alélica).

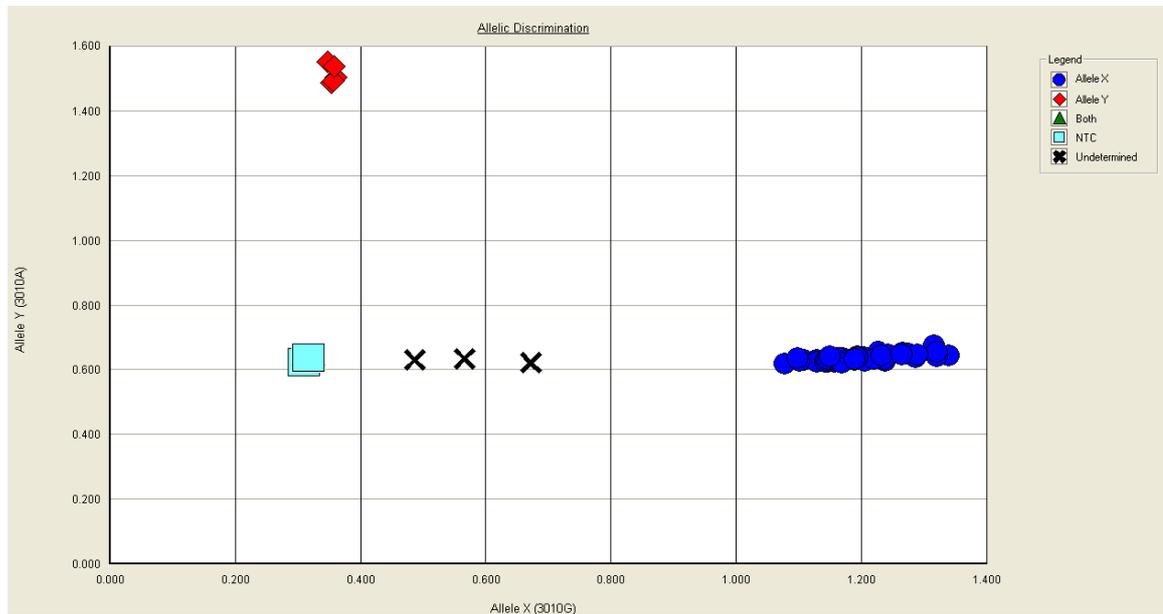


Figura 9: Resultado de uma reação de genotipagem do SNP 3010 gerado pelo software 7500 System. Quadrados correspondem ao controle negativo; losangos vermelhos correspondem às amostras com alelo A; bolas azuis correspondem às amostras com alelo G; X correspondem às amostras com resultado indeterminado.

### 6.3.1 Aumento do poder discriminatório do SNP 3010

Os dados da tabela 13 apresentam o resultado estatístico da genotipagem do SNP 3010 na população estudada e em populações descritas no banco de dados mtDB - *Human Mitochondrial Genome Database* (INGMAN; GYLLENSTEN, 2006).

Observou-se que este SNP apresenta-se mais polimórfico em populações asiáticas, europeias e nativo-americanas. Coble et al. (2004) encontrou uma distribuição ainda mais polimórfica quando estudou 241 indivíduos classificados como europeus e residentes nos EUA (32,36% genotipados para alelo A e 67,64% para alelo G) e afirma que este SNP é

extremamente útil em casos forenses, especialmente para diferenciar indivíduos em populações europeias classificadas com o haplogrupo H.

Tabela 13 - Resultado da genotipagem do SNP 3010 em diversas populações.

<i>População</i>	<i>Alelo A</i>	<i>Alelo G</i>
<b>Este trabalho</b>	<b>10,63%</b>	<b>89,37%</b>
Africanos	4,41%	95,59%
Nativo-americanos	20,96%	79,07%
Asiáticos	27,70%	72,30%
Australianos	8,82%	91,18%
Europeus	21,90%	78,10%

A população brasileira, tratando-se de uma população miscigenada, apresentou uma distribuição abaixo da média total (Alelo A presente em 16,76% dos indivíduos e Alelo G em 83,24% dos indivíduos, excluindo-se do cálculo da média os resultados de nossa população).

Entretanto, este SNP foi suficientemente discriminatório para conseguir distinguir indivíduos que apresentavam o mesmo haplótipo em dois dos treze grupos com mais de um indivíduo com o mesmo haplótipo (Tabela 6). Em ambos os casos, os indivíduos apresentavam o haplogrupo europeu H.

Quando classificamos nossas amostras em haplogrupos africanos, nativo-americanos e europeus, a distribuição dos alelos foi claramente alterada, estando o alelo “A” presente em 4,9% dos nativo-americanos (presente somente no haplogrupo D), 4,0% dos africanos (presente apenas nos haplogrupos L2 e L4) e 27,3% dos europeus (presente apenas nos haplogrupos H, J e U) (Fig. 10).

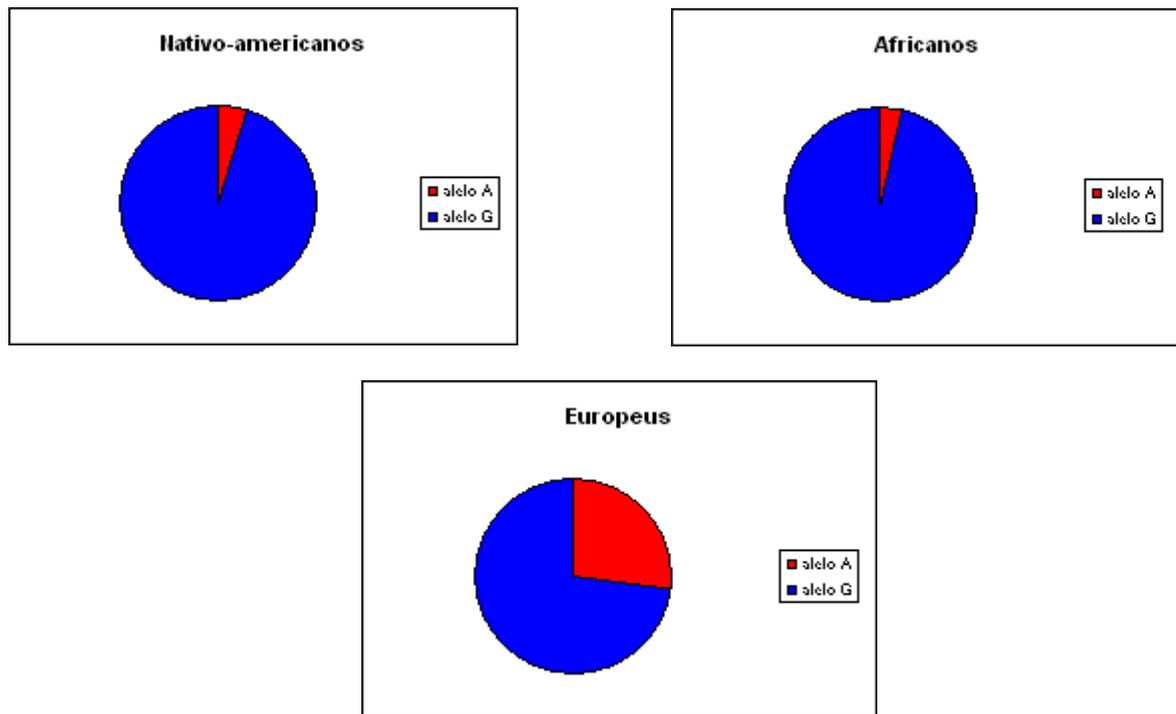


Figura 10: Distribuição dos alelos do SNP 3010 após as amostras serem classificadas em haplogrupos nativo-americanos, africanos e europeus.

#### 6.4 REAÇÃO DE SNAPSHOT

Neste projeto foram analisados 42 SNPs do DNA mt em uma única reação de PCR e SNaPshot a fim de permitir a classificação em haplogrupos da população em estudo.

Os fatores mais importantes observados antes da construção deste painel foram: (1) a seleção de *primers* para a reação de PCR e SNaPshot para SNPs que permitissem a classificação de diferentes haplogrupos já descritos para nossa população; (2) seleção criteriosa de *primers* que não apresentassem interações entre eles e (3) seleção de *primers* para a PCR que amplificassem fragmentos menores que 200 pb a fim de permitir a aplicação deste painel também para amostras degradadas. Até o momento, poucas reações em *multiplex* com um número tão grande de SNPs haviam sido relatadas, e até o momento, nenhum *multiplex* com 42 SNPs em uma única reação de PCR e SNaPshot foram publicados.

Quarenta polimorfismos (39 SNPs e 1 deleção) da região codificadora do DNA mt e dois SNPs da região não-codificadora (também conhecida como região hipervariável – SNPs 16519 e 16362) foram estrategicamente escolhidos baseado na sua habilidade de classificar os haplogrupos presentes em nossa população e a fim de estudar o aumento do poder discriminatório do painel (Fig. 11).

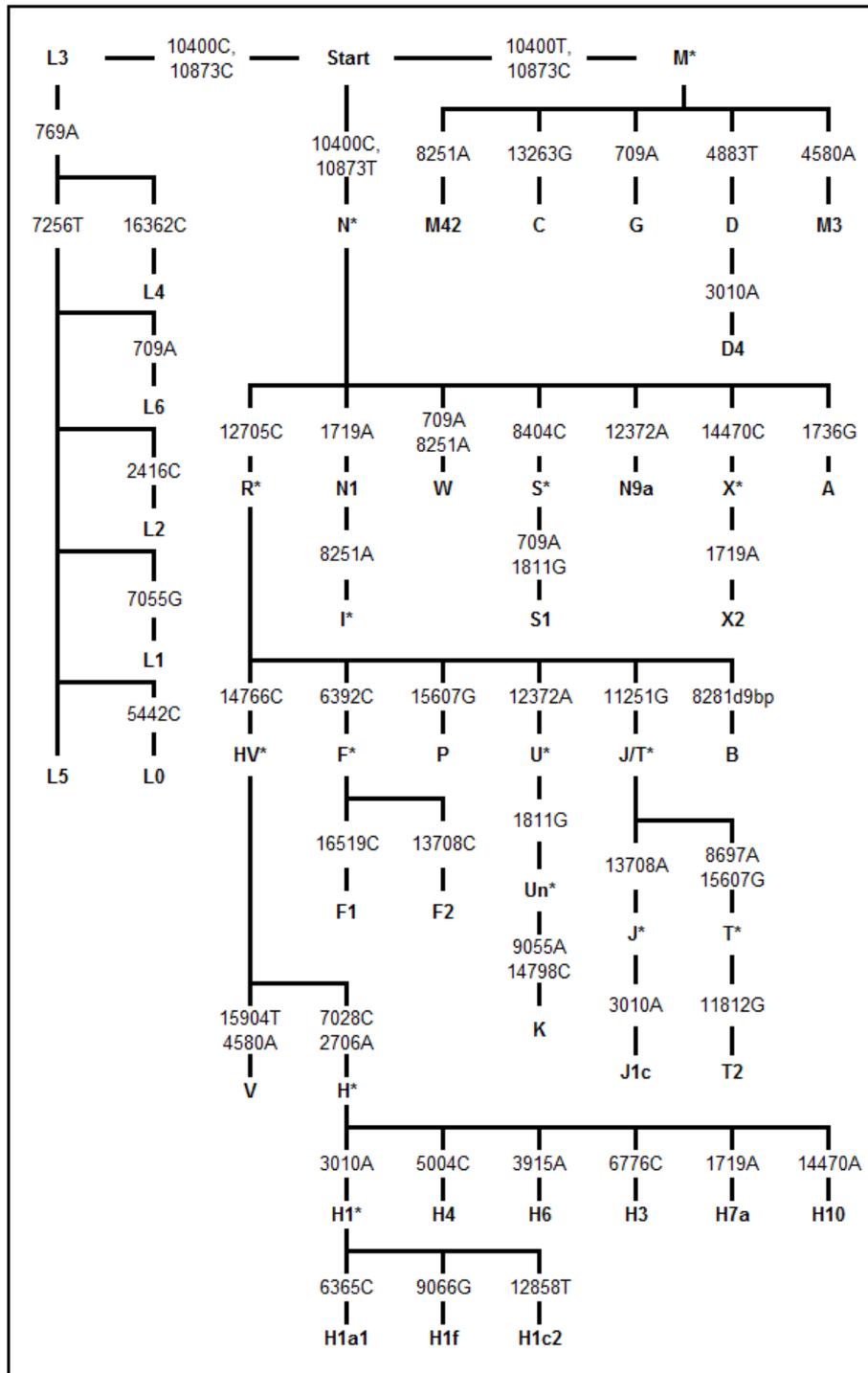


Figura 11: Árvore filogenética englobando haplogrupos europeus, africanos e nativo-americanos gerada por este painel de 42 SNPs. Para classificação de uma amostra no respectivo haplogrupo iniciamos a análise pelos SNPs 10400 e 10873 (posição *Start* na região central/superior da figura) e seguimos, baseados nos alelos de cada SNP naquela amostra até o haplogrupo correspondente.

Usando este painel, pudemos classificar em haplogrupos do DNA mt as 160 amostras analisadas neste projeto (Fig. 12).

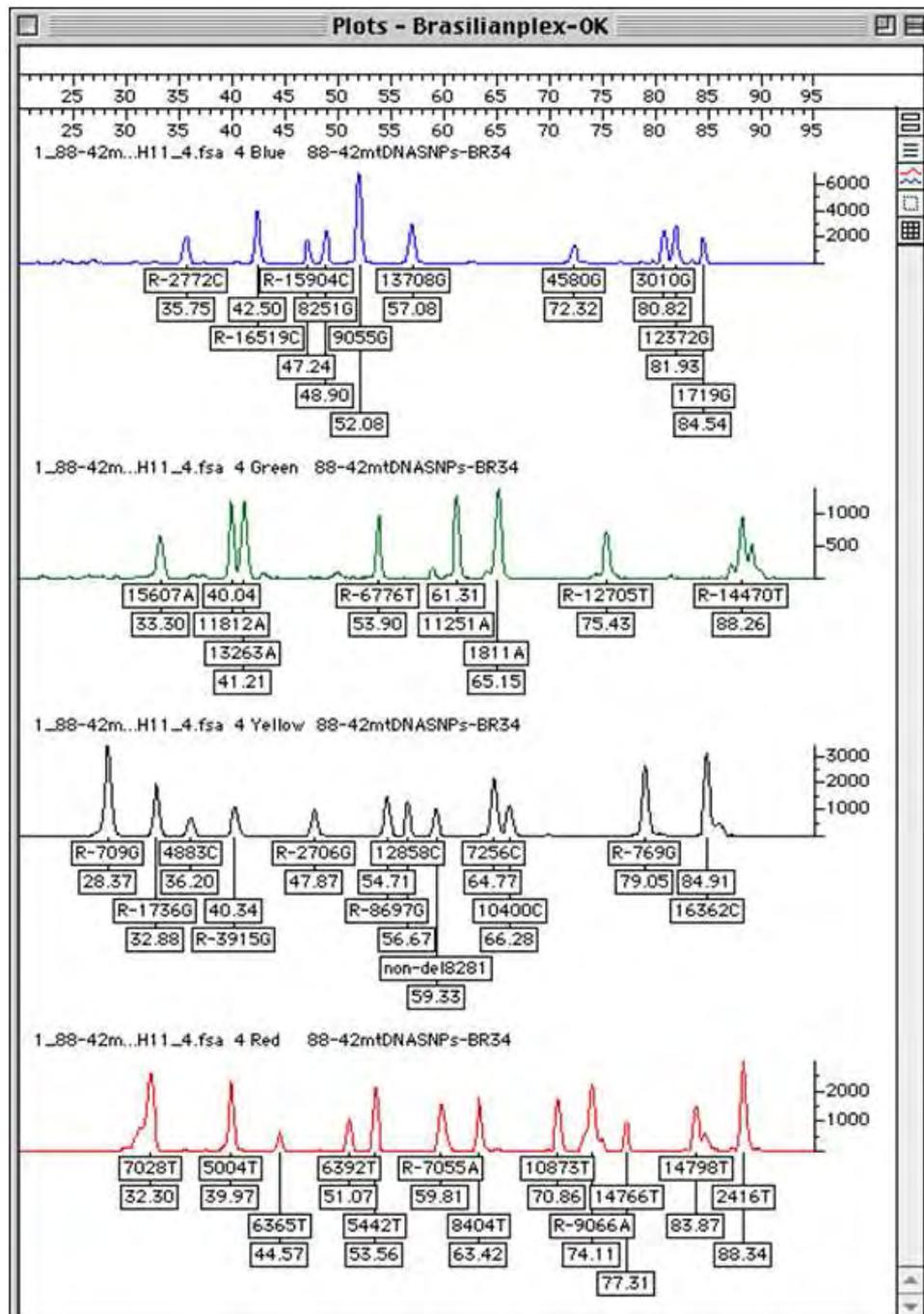


Figura 12: Exemplo de eletroferograma mostrando o perfil de análise de uma amostra para os 42 SNPs; Nos quadrados abaixo dos picos os números seguidos de letras indicam a posição do SNP e a variação do nucleotídeo; números sozinhos indicam tamanho do fragmento após corrida eletroforética; Eixo-y indica intensidade de fluorescência RFU (*relative fluorescence units*).

Como a classificação em haplogrupos baseada apenas na análise da região hipervariável é, em alguns casos, deficiente e pode não expressar a correta classificação em haplogrupos, a classificação obtida a partir da análise da região hipervariável em nossa amostragem (realização com a técnica de sequenciamento) foi comparada com a classificação gerada por SNaPshot. Algumas amostras que não puderam ser corretamente classificadas pelo sequenciamento, obtiveram sucesso pelo SNaPshot (principalmente na correta definição dos haplogrupos HV, H e V – amostras 04, 35, 114 – tabela 14). Em algumas amostras, entretanto, o painel completo (42 SNPs) não foi obtido, o que, entretanto, não inviabilizou sua classificação em haplogrupo. Isso aconteceu mais comumente em amostras classificadas em haplogrupos africanos e pode ser explicada por mutações nos sítios de ligação dos *primers*. Um exemplo claro ocorreu com amostras que apresentavam mutação no nucleotídeo 16360, o que impedia o anelamento do *primer* responsável pela genotipagem do SNP 16362 na reação de SNaPshot, causando uma falha na detecção deste SNP (Amostras 54, 122, 124, 130, 187, 189 - tabela 14)

A população estudada foi, então, classificada em 46,6% de origem africana, 27,3% européia, 25,5% nativo-americana e 0,6% de origem asiática (Fig. 13). Os haplogrupos mais comumente encontrados foram L3 (19,4%), L2 (14,4%) e H (13,7%).

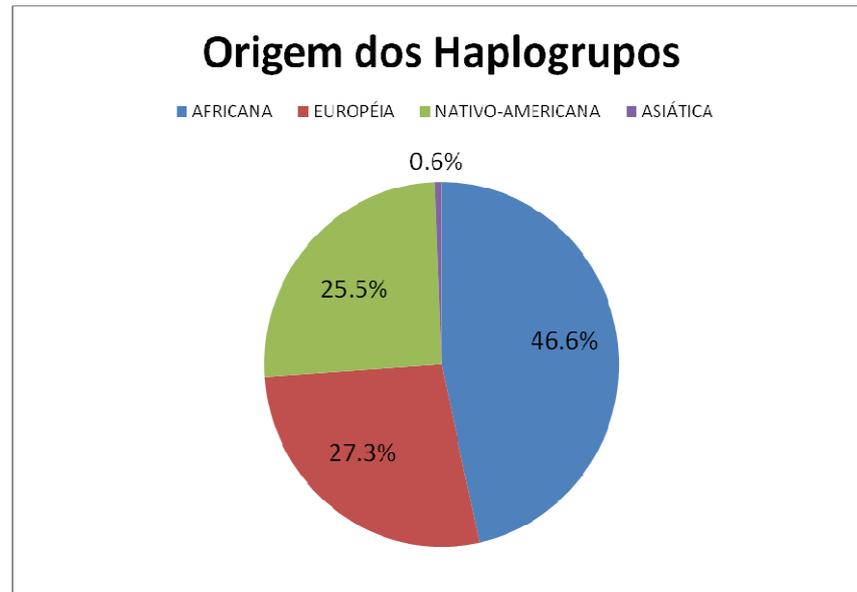


Figura 13: Classificação das amostras segundo a origem dos haplogrupos.

Dentre as amostras de origem africana, o haplogrupo mais comumente encontrado foi L3 (41,3%), seguido por L2 (30,7%), L1 (12,0%), L0 (8,0%), L4 (2,7%) e L\* (4,0%). Resultados similares foram descritos por outros autores (GONÇALVES et al., 2008; HÜNEMEIER et al., 2007).

Dentre as amostras de origem européia, o haplogrupo mais comumente encontrado foi H (50,0%), seguido por J (11,3%), K, U e T (9,1% cada), X (4,5%), W, HV e V (2,3% cada). A frequência encontrada para o haplogrupo predominante H foi similar àquela observada em países da Europa ocidental (PEREIRA et al., 2000; ROOSTALU et al., 2007; TURCHI et al., 2008).

Os haplogrupos nativo-americanos mais comumente encontrados foram: Haplogrupos B e C (33,3% cada), seguido por A (26,2%) e D (4,8%). Duas amostras foram classificadas em haplogrupo M, de origem asiática. Entretanto, uma das amostras foi corretamente classificada como haplogrupo M1a1 africano (2,4% da população africana) somente quando o sequenciamento da região hipervariável foi analisado (OLIVIERI et al.,

2006). A outra amostra classificada como haplogrupo M parece representar a população asiática que recentemente imigrou para o Brasil.

Os resultados para a genotipagem dos 42 SNPs em nossa população podem ser visualizados na tabela 14 (Apêndice E). A diversidade haplotípica encontrada foi de 96,04% utilizando o painel com 42 SNPs em nossa amostra populacional. Um achado interessante, os SNPs 16519 e 16362 que são considerados altamente variáveis, apresentaram diferentes poderes de discriminação dentre os vários haplogrupos:

O SNP 16519 foi altamente informativo (presença do alelo mais raro em mais de 30% das amostras) nos haplogrupos A, D, L0, L2, L3 e J. Nos haplogrupos europeus H e U, o alelo mais raro estava presente em aproximadamente 25% das amostras. Uma pequena variação foi encontrada nos haplogrupos C e L1. O SNP 16362 não foi tão informativo quanto o SNP 16519, mas variações foram encontradas nos haplogrupos U, H, L3, L2 e C. Estes SNPs podem ser importantes para o uso forense, por exemplo, a fim de escolher o melhor SNP a ser tipado para discriminar dois indivíduos com o mesmo haplogrupo.

No que se refere à sensibilidade da técnica, o painel completo foi obtido com 10 picogramas de DNA. Esse resultado, aliado à amplificação de tamanhos pequenos gerados na PCR, permitem sugerir o uso deste painel em casos forenses, especialmente para amostras contendo pequena quantidade ou com degradação do DNA.

## 6.5 DISTRIBUIÇÃO DOS HAPLOGRUPOS DE ACORDO COM A COR DE PELE DECLARADA PELO INDIVÍDUO

Quando foram analisados os haplogrupos presentes nos indivíduos brancos e pretos foi obtido uma distribuição desigual entre os mesmos (Fig.14). Dentre os indivíduos classificados como brancos houve uma distribuição equivalente entre haplogrupos europeus, africanos e

nativo-americanos, com ligeira predominância de haplogrupos europeus. Entretanto, dentre os indivíduos classificados como pretos houve uma predominância de haplogrupos africanos, seguido de haplogrupos nativo-americanos e europeus.

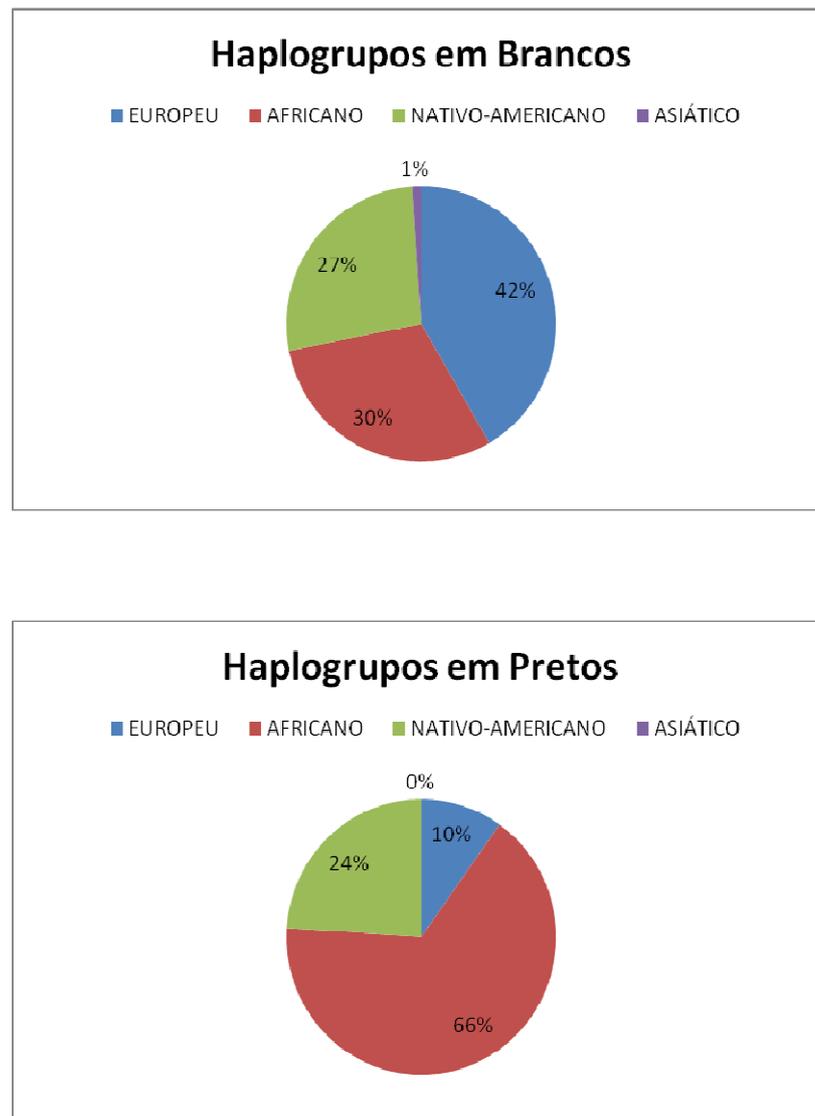


Figura 14: Distribuição dos haplogrupos em brancos e pretos.

Nossos dados confirmam os resultados já relatados na literatura de que não é possível prever o perfil genético de um indivíduo em nossa população, em especial a classificação em

haplogrupos do DNA mt, baseado no fenótipo do mesmo e nem o oposto é verdadeiro (PARRA et al., 2003; PENA et al., 2009).

## 6.6 FREQUÊNCIA DE HETEROPLASMIA EM HV3 EM AMOSTRAS DE CABELO E SANGUE

Nenhuma diferença entre amostras de cabelo e sangue na região HV3 foi observada nas 100 amostras analisadas. Além disso, nenhuma heteroplasmia de sequência (ou pontual) foi detectada. Entretanto, uma amostra apresentou heteroplasmia de comprimento (tanto na amostra de sangue como na amostra de cabelo) em uma região que apresenta repetições de CA, com impossibilidade de leitura após a heteroplasmia (Fig. 15). Este resultado (1,0%), no que se refere às amostras de sangue, está em acordo com aqueles relatados por outros autores. Szibor et al. (2007) encontrou 1,38% de heteroplasmia na região de repetições de CA na população Germânica e Chung et al. (2005) encontrou 0,6% na população Coreana. Heteroplasmia de comprimento também foi observada na posição 573 (região poli-C) em 3 indivíduos. Não há relatos na literatura sobre a frequência de heteroplasmia em HV3 em amostras de cabelo.

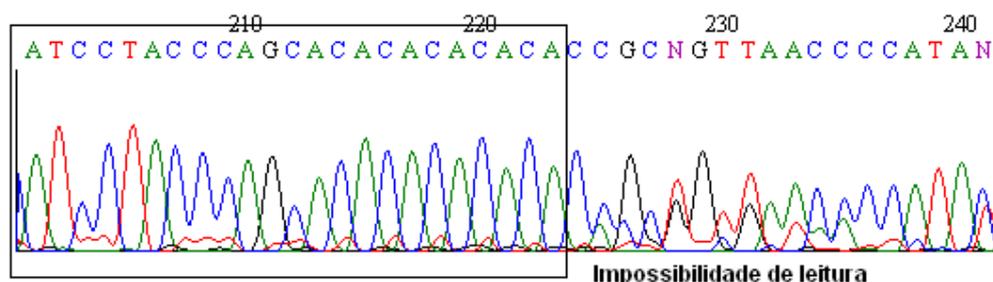


Figura 15: Heteroplasmia de comprimento na região de repetição de CA em HV3 com impossibilidade de leitura após esta região.

A fim de observar quais variantes correspondentes às repetições de CA estavam presentes em ambos os tecidos (cabelo e sangue), os produtos de PCR das amostras que apresentaram heteroplasmia na região de repetição de CA foram clonados e ressequenciados. A clonagem permite o isolamento de cada um dos variantes e o seu subsequente sequenciamento permite a observação de cada um dos nucleotídeos presentes. Após o sequenciamento, encontrou-se uma mistura de “alelos”, com presença de DNA mt com seis repetições de CA e com sete repetições de CA, em ambas as amostras (Fig. 16). A presença de uma mistura de “alelos” sugere uma possibilidade de variação intraindividual e possibilita a seleção de um deles através das gerações. Recentemente, alguns autores relataram herança instável de variantes heteroplásmicas em HV3 quando analisavam heredogramas. Szibor et al. (2007) demonstrou herança instável na região de repetição de CA de HV3 em dois heredogramas. Brandstätter et al. (2004b) relatou um caso de divergência de repetições de CA em 135 famílias (pares de mães e filhos).

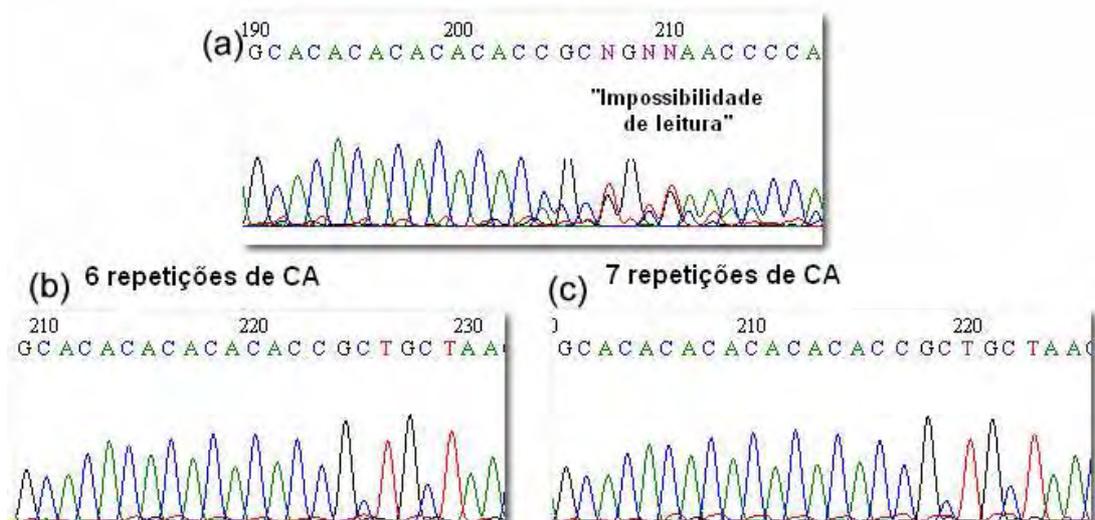


Figura 16: (a) Eletroferograma na região de repetição de CA demonstrando mistura dos nucleotídeos e impossibilidade de leitura após esta região; (b) e (c) Eletroferogramas após clonagem mostrando um dos alelos com seis repetições e outro alelo com sete repetições de CA.

## 7. CONCLUSÕES

- As amostras de indivíduos residentes na Grande São Paulo apresentam grande diversidade de haplótipos do DNA mt e confirmam a miscigenação brasileira, apresentando perfis de africanos, europeus e nativo-americanos, em sua grande maioria;
- Não é possível prever a classificação genética em haplogrupos do DNA mt baseado na cor de pele do indivíduo em nossa população devido a grande miscigenação;
- O SNP 3010 parece ser interessante para discriminar indivíduos que apresentem o mesmo perfil quando a região hipervariável é analisada especialmente aqueles classificados nos haplogrupos H, J e U;
- Podemos concluir que foi possível estabelecer reações de PCR e SNaPshot em *multiplex* sensíveis e confiáveis para a detecção de 42 polimorfismos do DNA mt. O painel obtido proporcionou a classificação das amostras na grande maioria dos haplogrupos africanos, europeus, nativo-americanos, possibilitando seu uso no contexto forense brasileiro;
- O painel completo (detecção dos 42 polimorfismos) foi obtido com 10 picogramas de DNA. Esse resultado, aliado à amplificação de tamanhos pequenos gerados na PCR, permite-nos sugerir o uso deste painel em casos forenses, especialmente para amostras contendo pequena quantidade ou com degradação do DNA;
- Em relação à heteroplasmia na região HV3, podemos concluir que sua presença não é tão frequente como em HV1 e HV2;

- A análise de HV3, entretanto, pode ser importante para a solução de um caso forense por permitir um aumento do poder discriminatório e classificação filogenética;
  
- A grande diversidade genética na região de repetição de CA pode contribuir para a ratificação do seu uso, devendo ser, no entanto, realizada com cautela quando avaliada em tecidos diferentes.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

ALONSO, A.; SALAS, A.; ALBARRÁN, C. et al. Results of the collaborative exercise and proficiency testing program on mitochondrial DNA of the GEP-ISFG: an inter-laboratory study of the observed variability in the heteroplasmy level of hair from the same donor. **Forensic Sci. Int.**, v. 125, p. 1-7, 2002.

ALVES-SILVA, J.; DA SILVA SANTOS, M.; GUIMARÃES, P.E. et al. The ancestry of Brazilian mtDNA lineages. **Am. J. Hum. Genet.**, v. 67, p. 444-61, 2000.

ANDERSON, S.; BANKIER, A.T.; BARRELL, B.G. et al. Sequence and organization of the human mitochondrial genome. **Nature**, v. 290, p. 457-465, 1981.

ANDREWS, R.; KUBACKA, I.; CHINNERY, P.F. et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. **Nat. Genet.**, v. 23, p. 147, 1999.

ANJOS, M. J.; CARVALHO, M.; ANDRADE, L. et al. Individual genetic identification of biological samples: a case of an aircraft accident. **Forensic Sci. Int.**, v. 146S, p. S115–S117, 2004.

BANDELT, H-J; MACAULAY, V.; RICHARDS, M. **Human mitochondrial DNA and the evolution of *Homo sapiens***. 1st edition, Leipzig: Springer-Verlag Berlin Heidelberg, 2006, p.153.

BARALDI, A.M.; MENEZES, L.M.B.; SILVA JR, J.P. et al. Panorama nacional do uso da técnica de identificação genética nos serviços oficiais de identificação e a participação do cirurgião-dentista. **Revista de Pós-Graduação (USP)**, v. 15, p. 261-265, 2008.

BARBOSA, A.B.; DA SILVA, L.A.; AZEVEDO, D.A. et al. Mitochondrial DNA control region polymorphism in the population of Alagoas state, north-eastern Brazil. **J. Forensic Sci.**, v. 53, p. 142-6, 2008.

BOBILLO, M.C.; ZIMMERMANN, B.; SALA, A. et al. Amerindian mitochondrial DNA haplogroups predominate in the population of Argentina: towards a first nationwide forensic mitochondrial DNA sequence database. **Int. J. Legal Med.**, 2009. In press.

BRANDSTÄTTER, A.; PETERSON, C.T.; IRWIN, J.A. et al. Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring phylogenetics parameters for the establishment of a forensic database. **Int. J. Legal Med.**, v. 118, p. 294-306, 2004a.

BRANDSTÄTTER, A.; NIEDERSTÄTTER, H.; PARSON, W. Monitoring the inheritance of heteroplasmy by computer-assisted detection of mixed basecalls in the entire human mitochondrial DNA control region. **Int. J. Legal Med.**, v. 118, p. 47-54, 2004b.

BUTLER, J.M. **Forensic DNA typing**. 2nd edition, London: Elsevier Academic Press, 2005, p.243.

BUTLER, J.M.; LEVIN, B.C. Forensic applications of mitochondrial DNA, **Trends Biotechnol.**, v. 16, p. 158–162, 1998.

CHUNG, U.; LEE, H.Y.; YOO, J.E. et al. Mitochondrial DNA CA dinucleotide repeats in Koreans: the presence of length heteroplasmy. **Int. J. Legal Med.**, v.119, p. 50-53, 2005.

COBLE, M. D.; JUST, R. S.; O'CALLAGHAN, J. E. et al. Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians. **Int. J. Legal Med.**, v.118, p. 137-146, 2004.

DIEGOLI, T.M.; IRWIN, J.A; JUST, R.S. et al. Mitochondrial control region sequences from an African American population sample. **Forensic Sci. Int. Genet.**, v. 4, p. e45-52, 2009.

EXCOFFIER, L.; SMOUSE, P.E.; QUATTRO, J.M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. **Genetics**, v. 131, p. 479-491, 1992.

EXCOFFIER, L.; LAVAL, G.; SCHNEIDER, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. **Evol. Bioinform. Online**, v.1, p.47-50, 2005.

FEIO-DOS-SANTOS, A.C.; CARVALHO, B.M.; BATISTA DOS SANTOS, S.E. et al. Nucleotide variability of HV-I in admixed population of the Brazilian Amazon Region. **Forensic Sci. Int.**, v. 164, p. 276-7, 2006.

FINNILÄ, S.; LEHTONEN, M.S.; MAJAMAA, K. Phylogenetic network for European mtDNA. **Am. J. Hum. Genet.**, v. 68, p. 1475–1484, 2001.

FORAN, D. Relative degradation of nuclear and mitochondrial DNA: an experimental approach. **J. Forensic Sci.**, v. 51, p. 766-770, 2006.

GONÇALVES, V.F.; CARVALHO, C.M.; BORTOLINI, M.C. et al. The phylogeography of African Brazilians. **Hum. Hered.**, v. 65, p. 23-32, 2008.

HALL, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41, p. 95-98, 1999.

HE, Y.; DRESSMAN, D.C.; IACOBUZIO-DONAHUE et al. Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. **Nature**, v. 464,p. 610-4, 2010.

HOLLAND, M. M.; PARSONS, T. J. Mitochondrial DNA sequence analysis-validation and use for forensic casework. **Forensic Sci. Rev.**, v. 11, p. 22-50, 1999.

HOLT, I. J; HARDING, A. E.; MORGAN-HUGHER, J. A. Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. **Nature**, v. 25, p. 717–719, 1988.

HOONG, L. L.; LEK, K. C. Genetic polymorphisms in mitochondrial DNA hypervariable regions I, II and III of the Malaysian population. **AsPac J. Mol. Biol. Biotechnol.**, v. 13, n. 2, p.79-85, 2005.

HÜNEMEIER, T.; CARVALHO, C.; MARRERO, A.R. et al. Niger-Congo speaking populations and the formation of the Brazilian gene pool: mtDNA and Y-chromosome data. **Am. J. Phys. Anthropol.**, v.133, p. 854-67, 2007.

INGMAN, M.; GYLLENSTEN, U. mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. **Nucleic Acids Res.**, v. 34, p. D749-51, 2006.

IVANOV, P.L.; WADHAMS, M.J.; ROBY, R.K. et al. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. **Nat. Genet.**, v. 12, p. 417-20, 1996

JAZIN, E. E. et al. Human brain contains high level of heteroplasmy in the noncoding regions of mitochondrial DNA. **Proc. Natl. Acad. Sci. USA.**, v. 93, p. 12382-7, 1996.

JEHAES, E.; GILISSEN, A.; CASSIMAN, J.J. et al. Evaluation of a decontamination protocol for hair shafts before mtDNA sequencing. **Forensic Sci. Int.**, v. 94, p. 65-71, 1998.

JUST, R.S.; LOREILLE, O.M.; MOLTO, J.E. et al. Titanic's unknown child: The critical role of the mitochondrial DNA coding region in a re-identification effort. **Forensic Sci. Int. Genet.**, 2010, *in press*.

KIM, N.Y.; LEE, H.Y.; PARK, M.J et al. A genetic investigation of Korean mummies from the Joseon Dynasty. **Mol. Biol. Rep.**, 2010, *in press*.

KIVISILD, T.; TOLK, H-V.; PARIK, J. et al. The emerging limbs and twigs of the East Asian mtDNA tree. **Mol. Biol. Evol.**, v. 19, p. 1737-1751, 2002.

KÖHNEMANN, S.; HOHOFF, C.; PFEIFFER, H. An economical mtDNA SNP assay detecting different mitochondrial haplogroups in identical HVR 1 samples of Caucasian ancestry. **Mitochondrion**, v. 9, p. 370-5, 2009.

KONG, Q.P.; YAO, Y.G.; SUN, C. et al. Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. **Am. J. Hum. Genet.**, v. 73, p. 671–676, 2003.

LEE, H.R.; JOHNSON, K.A. Fidelity of the human mitochondrial DNA polymerase. **J. Biol. Chem.**, v. 281, p. 36236-40, 2006.

LEE, H.Y., YOO, J.E.; PARK, M.J. et al. Mitochondrial DNA control region sequences in Koreans: identification of useful variable sites and phylogenetic analysis for mtDNA data quality control. **Int. J. Legal Med.**, v. 120, p. 5–14, 2006.

LEE, H.Y.; SONG, I.; HA, E. et al. mtDNAMAN: a Web-based tool for the management and quality analysis of mitochondrial DNA control-region sequences. **BMC Bioinformatics**, v. 9, p. 483, 2008.

LIBRADO, P.; ROZAS, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**, v. 25, p. 1451-1452, 2009.

LUTZ, S.; WITTIG, H.; WEISSER, H.J. et al. Is it possible to differentiate mtDNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? **Forensic Sci. Int.**, v. 113, p. 97-101, 2000.

MACAULAY, V.; RICHARDS, M.; HICKEY, E. et al. The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. **Am. J. Hum. Genet.**, v. 64, p. 232–249, 1999.

MARSHALL O.J. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. **Bioinformatics**, v. 20, p. 2471-2472, 2004.

MOLLER, A.; BRINKMANN, B. PCR-VNTRs (PCR-variable number of tandem repeats) in forensic science. **Cell. Mol. Biol.**, v.41, n.5, p.715-724, 1995.

MULLIS, K.B.; FALOONA, F.A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. **Methods Enzymol.**, v.155, p.335-350, 1987.

OLIVIERI, A.; ACHILLI, A. PALA, M. et al. The mtDNA legacy of the Levantine early Upper Palaeolithic in África. **Science**, v. 314, p. 1767-70, 2006.

PALENCIA, L.; VALVERDE, L. ALVAREZ, A. et al. Mitochondrial DNA diversity in a population from Santa Catarina (Brazil): predominance of the European input. **Int. J. Legal Med.**, 2010. In press.

PANETO, G.G.; MARTINS, J.A.; LONGO, L.V.G. et al. Heteroplasmy in hairs: Differences among hair and blood from the same individuals are still a matter of debate. **Forensic Sci. Int.**, v. 173, p. 117-121, 2007.

PARRA, F.C.; AMADO, R.C.; LAMBERTUCCI, J.R. et al. Color and genomic ancestry in Brazilians. **Proc. Natl. Acad. Sci. U.S.A.**, v. 100, p. 177-182, 2003

PARSON, W.; BANDELT, H.J. Extended guidelines for mtDNA typing of population data in forensic science. **Forensic Sci. Int. Genet.**, v. 1, p. 13-9, 2007.

PARSON, W.; DÜR, A. EMPOP – a forensic mtDNA database. **Forensic Sci. Int. Genet.**, v. 1, p. 88-92, 2007.

PENA, S.D.; BASTOS-RODRIGUES, L.; PIMENTA, J.R. et al. DNA tests probe the genomic ancestry of Brazilians. **Braz. J. Med. Biol. Res.**, v. 42, p. 870-876, 2009.

PEREIRA, L.; PRATA, M.J.; AMORIM, A. Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. **Ann. Hum. Genet.**, v. 64, p. 491-506, 2000.

PRIETO, L.; ALONSO, A.; ALVES, C. et al. 2006 GEP-ISFG collaborative exercise on mtDNA: reflections about interpretation, artefacts, and DNA mixtures. **Forensic Sci. Int. Genet.**, v. 2, p. 126-133, 2008.

RICHARDS, M., MACAULAY, V., HICKEY, E. et al. Tracing European founder lineages in the Near Eastern mtDNA pool. **Am. J. Hum. Genet.**, v. 67, p. 1251-1276, 2000.

ROBIN, E.D.; WONG, R. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. **J. Cell. Physiol.**, v. 136, p. 507–513, 1988.

ROOSTALU, U.; KUTUEV, I.; LOOGVÄLI, E.L. et al. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. **Mol. Biol. Evol.**, v. 24, p. 436-48, 2007.

SALAS, A.; COMAS, D.; LAREU, M.V. et al. mtDNA analysis of the Galician population: a genetic edge of European variation. **Eur. J. Hum. Genet.**, v. 6, p. 365-375, 1998.

SALAS, A.; LAREU, M. V.; CARRACEDO, A. Heteroplasmy in mtDNA and the weight of evidence in forensic mtDNA analysis: a case report. **Int. J. Legal Med.**, v. 114, p. 186-190, 2001.

SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. **Molecular cloning: a laboratory manual**. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989.

SANCHEZ, J.J.; BORSTING, C.; MORLING, N. Typing of Y chromosome SNPs with multiplex PCR methods. **Methods Mol. Biol.**, v. 297, p. 209-228, 2005.

SAUNIER, J.L.; IRWIN, J.A.; JUST, R.S. et al. Mitochondrial control region sequences from a U.S. “Hispanic” population sample. **Forensic Sci. Int. Genet.**, v. 2, p. e19-23, 2008.

SAUNIER, J.L.; IRWIN, J.A.; STROUSS, K.M. et al. Mitochondrial control region sequences from an Egyptian population sample. **Forensic Sci. Int. Genet.**, v. 3, p. e97-103, 2009.

SCHLEBUSCH, C.M., NAIDOO, T.; SOODYALL, H. SNaPshot minisequencing to resolve mitochondrial macro-haplogroups found in Africa. **Electrophoresis**, v. 30, p. 3657-64, 2009.

SINGER-SAM, J.; TANGUAY, R.; RIGGS, A.D. Use of chelex to improve the PCR signal from a small number of cells. **Amplifications: A Forum for PCR Users**, v. 3, p. 11, 1989.

SUBRAMANIAN, S.; MISHRA, R.K.; SINGH, L. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. **Genome Biology**, v.4, R13, 2003.

SULLIVAN, K.M.; HOPGOOD, R.; GILL, P. Identification of human remains by amplification and automated sequencing of mitochondrial DNA. **Int. J. Legal Med.**, v. 105, p. 83-86, 1992.

SWGDM - Scientific Working Group on DNA Analysis Methods. Guidelines for Mitochondrial DNA (mtDNA) Nucleotide Sequence Interpretation. **Forensic Science Communications**, v. 5, n. 2, 2003.

Disponível em: < <http://www.fbi.gov/hq/lab/fsc/backissu/april2003/swgdammitodna.htm> >

Acesso em: 23/03/2010

SZIBOR, R.; PLATE, I.; HEINRICH, M. et al. Mitochondrial D-loop (CA)<sub>n</sub> repeat length heteroplasmy: frequency in a German population sample and inheritance studies in two pedigrees. **Int. J. Legal Med.**, v. 121, p. 207-13, 2007.

TILLMAR, A.O.; COBLE, M.D.; WALLERSTRÖM, T. et al. Homogeneity in mitochondrial DNA control region sequences in Swedish subpopulations. **Int. J. Legal Med.**, v. 124, p. 91-98, 2010.

TORRONI, A.; SCHURR, T.G.; CABELL, M.F. et al. Asian affinities and continental radiation of the four founding Native American mtDNAs. **Am. J. Hum. Genet.**, v. 53, p. 563–590, 1993.

TORRONI, A.; LOTT, M.T.; CABELL, M.F. et al. mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. **Am. J. Hum. Genet.**, v. 55, p. 760–776, 1994.

TULLY, G.; BARRITT, S. M.; BENDER, K. et al. Results of a collaborative study of the EDNAP group regarding mitochondrial DNA heteroplasmy and segregation in hair shafts. **Forensic Sci. Int.**, v. 140, p. 1-11, 2004.

TURCHI, C.; BUSCEMI, L.; PREVIDERÈ, C. et al. Italian mitochondrial DNA database: results of a collaborative exercise and proficiency testing. **Int. J. Legal Med.**, v. 122, p. 199-204, 2008.

VAN OVEN, M.; KAYSER, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. **Hum. Mutat.**, v. 30, p. e386-394, 2009.

VALLONE, P.M.; BUTLER, J.M. AutoDimer: a screening tool for primer-dimer and hairpin structures. **Biotechniques**, v. 37, p. 226-231, 2004.

WEEDEN, V.W.; SWARNEN, S.L. Exames forenses de identificação por análises do DNA. In: Henry, J.B. **Diagnósticos clínicos e tratamento por métodos laboratoriais**. 19. ed. São Paulo: Manole, cap.63, 1998, p.1427-1438.

ZANG, Y.J.; XU, Q.S.; ZHENG, Z.J. et al. Haplotype diversity in mitochondrial DNA hypervariable region I, II and III in northeast China Han. **Forensic Sci. Int.**, v. 149, p. 267-269, 2005.

## APÊNDICE A -

Tabela 3 – Informações sobre *primers* da reação de PCR em *multiplex*.

Posição do SNP	<i>Primer</i>	Seqüência do <i>primer</i> (5'-3')	Tamanho do amplicon (pb)	Concentração final ( $\mu$ M)
G709A e G769A	mtG709A-FWD mtG709A-REV	AAGATTACACATGCAAGCATCC CTAGGCTAAGCGTTTTGAGCTG	112	0,178
G1719A; A1811G e A1736G	mtG1719 + 1811A-FWD mtG1719 + 1811A-REV	CTTGACCCGCTCTGAGCTAAA GCAGAAAGGTATAGGGGTTAGTCC	196	0,178
T2416C	mtT2416C-FWD mtT2416C-R	CAGCCCAATATCTACAAATCAACC GTAAAGATTTGCCGAGTTCCCT	114	0,178
A2706G e C2772T	mtA2706G + C2772T-FWD mtA2706G + C2772T-REV	ACTTTTAAACCAGTGAATTTGACCT GCCCCAACCCGAAATTTTAA	157	0,128
G3010A	mtG3010A-FWD mtG3010A-REV	ACCTCGATGTTGGATCAGGACA CGAACCTTTAATAGGGGCTGCACCAT	55	0,027
G3915A	mtG3915A-FWD mtG3915A-REV	TAGCAGAGACCAACCCGAACC GGCGTATTCGATGTTGAAGC	80	0,134
G4580A	mtG4580A-FWD mtG4580A-REV	CGCTAAGCTCGCACTGATTT TTTTGGTTAGAAGCTGGAATAAAAAGC	130	0,125
C4883T	mtC4883T-FWD mtC4883T-R	CCTTTCACCTTCTGAGTCCCA GCTTACGTYTAGTGAGGGGAG	128	0,036
T5004C	mtT5004C-FWD mtT5004C-REV	TCCATCATAGCAGGCAGTTG TGGTTATGTTAGGGTTGTACGG	124	0,107

Posição do SNP	Primer	Sequência do primer (5'-3')	Tamanho do amplicon (pb)	Concentração final ( $\mu$ M)
T5442C	mtT5442C-FWD mtT5442C-R	AATCTACTCCACCTCAATCAC GTAGGAGTAGCGTGGTRAGG	125	0,178
T6365C e T6392C	mtT6365C + T6392C-FWD mtT6365C + T6392C-REV	TGGAGCCTCCGTAGACCTAA TGGCAGGGGGTTTTATATTG	110	0,045
R-T6776C	mtT6776C-FWD mtT6776C-REV	TGGCTTCCTAGGGTTTATCG CGGAGGTGAAATATGCTCGT	83	0,032
C7028T e A7055G	mtC7028T-FWD mtC7028T-REV	AGCAAACTCACTACTAGACATCG GATGGCAAATACAGCTCCTATTGA	94	0,071
C7256T	mtC7256T-FWD mtC7256T-R	TCTTCCCACAACACTTTCTC GAGAAAATGAATGAAGCCTACAG	109	0,089
A8251G e 8281 del9bp	mtG8251A-FWD mtDEL8281-9-R	CATGCCCATCGTCCTAGAAAT TGCTAAGTTAGCTTTACAGTGG	113	0,089
T8404C	mtT8404C-FWD mtT8404C-REV	CAGTGAAAATGCCCCCAACTAAA TTTTGGTGAGGGAGGTAGGTG	126	0,107
G8697A	mtG8697A-FWD mtG8697A-REV	CCAAACAATGACTAATCAAACCTAACC GTATAAGAGATCAGGTTCCGTCCTT	80	0,036
G9055A e A9066G	mtG9055A + A9066G-FWD mtG9055A + A9066G-REV	CGCCTAACCCGTAACATTACT TGATAAGTGTAGAGGGAAGGTTAATG	100	0,045
C10400T	mtC10400T-FWD mtC10400T-REV	GCCCTAAGTCTGGCCTATGA TGAGTCGAAAATCATTCGTTTTTG	90	0,178
T10873C	mtT10873C-FWD mtT10873C-REV	AATCAACACAACCCACCCACA GGGGAACACGCTAAAATAGGTTG	90	0,178

Posição do SNP	Primer	Sequência do primer (5'-3')	Tamanho do amplicon (pb)	Concentração final ( $\mu$ M)
A11251G	mtA11251G-FWD mtA11251G-REV	CGCAGGCACATACTTCCTATTC AGTGAGCCTAGGGTGTGTGAG	90	0,032
A11812G	mtA11812G-FWD mtA11812G-REV	TCCTCTCTCAAGGACTTCAAACCT GGCTTGCTAGAAAGTCATCAAAAA	63	0,071
G12372A	mtG12372A-FWD mtG12372A-REV	AAAAATTTTGGTGCAACTCCA AACGAGGGTGGTAAAGGATGG	107	0,089
C12705T	mtC12705T-FWD mtC12705T-REV	CCCAACATTAAATCAGTTCCTCAA GAACAGTTGGAAATAGGTTGTTAGC	91	0,107
C12858T	mtC12858T-FWD mtC12858T-REV	ATGATACGCCCGGAGCAGA TGTGGGTCTCATGAGTTGGA	126	0,045
A13263G	mtA13263G-FWD mtA13263G-R	CCCACAAATCCAAACTCTAACAC GCCGATTGTAACATAATATGAGTCC	145	0,045
G13708A	mtG13708A-FWD mtG13708A-REV	CGCTTCCCCACCCCTTACTA GCGGGGGAAATGTTGTTAG	119	0,196
T14470A	mtT14470A-FWD mtT14470A-REV	CCCCATGCCCTCAGGATACTC TTTAGGGGGAATGATGGTTG	126	0,268
C14766T e T14798C	mtC14766T-FWD mtC14798C-REV	CAACTACAAGAAACACCAAATG ATCATGCCGGAGATGTTGGAT	82	0,125
A15607G	mtA15607G-FWD mtA15607G-REV	TCTCCGATCCGTCCCTAAC GATAGTAATAGGGCAAGGAC	55	0,02
C15904T	mtC15904T-FWD mtC15904T-REV	TACTCAAATGGGCCTGTCTT TGATTTGTCCCTTGGAAAAAAGGT	86	0,089

<b>Posição do SNP</b>	<b>Primer</b>	<b>Sequência do <i>primer</i> (5'-3')</b>	<b>Tamanho do amplicon (pb)</b>	<b>Concentração final (<math>\mu</math>M)</b>
T16362C	mfT16362C-FWD mfT16362C-R	TAGCACATTACAGTCAAATCCC GGATATTGATTTCACGGAGGA	151	0,071
T16519C	mfT16519C-FWD mfT16519C-REV	CCATAACACTTGGGGGTAGC TTATTTAAGGGGAACGTTGTTGG	99	0,089

## APÊNDICE B -

Tabela 4 – Informações sobre *primers* da reação de SNaPshot em *multiplex*.

Nome do <i>primer</i>	SNP	Sequência do <i>primer</i>	Orientação	Concentração final ( $\mu\text{M}$ )
R-G709A-SNP	G709A	TTTAGAGGGTGAACCTCACTGGAA	R	0,023
SP2-R-G769A	G769A	CGTCGTGAAAAGTCTGACAAAACTGACTAACTAGGTGCCACGTCTGAAAAGTCTGACAAAAGCGTTTTGAGCTGCATTG	R	0,06
G1719A-SNP	G1719A	AAC TGACTAACTAGGTGCCACGTCTGAAAAGTCTGACAAAACGTCTGAAAAGTCTGACAAACCCACTTACTACCAGACAAACCTTA	F	0,08
SP-R-A1736G	A1736G	CAATTTCAATTTCTATCGCCTACTTTAT	R	0,05
A1811G-SNP	A1811G	AACTGACTAACTAGGTGCCACGTCTGAAAAGTCTGACAAAGCAAGGGAAGATGAAAAATTATA	F	0,1
SP3-T2416C	T2416C	TGCCACGTCTGAAAAGTCTGACAAAAC T GACTAACTAGGTGCCACGTCTGAAAAGTCTGACAAACCAACCAAGTCATTATTACCC	F	0,1
R-A2706G-SNP	A2706G	GCCACGTCTGAAAAGTCTGACAAAGTCTTCTCGTCTGCTGTGT	R	0,1
R-C2772T-SNP	C2772T	GTCTGACAAAGGTTTGGTAGTTTAGGACCTGT	R	0,075
R-G3010A-SNP	G3010A	CGTGAAAAGTCTGACAAAAC T GACTAAAAC T AGGTGCCACGTCTGAAAAGTCTGACAACTCGATGTTGGATCAGGACATCCC	F	0,024
R-G3915A-SNP	G3915A	GTATTTCGATGTTGAAAGCCCTGAGACTAGTTCGGACTC	R	0,06
G4580A-SNP	G4580A	ACAAAAC T GACTAAAAC T AGGTGCCACGTCTGAAAAGTCTGACAAATTTACCTGAGTAGGCCCTAGAATAAACAT	F	0,025
SP2-C4883T	C4883T	GTGAAAAGTCTGACAAAAC T GACAAAAC T AGCCCC	F	0,03
T5004C-SNP	T5004C	CGTGAAAAGTCTGACAAACCAACCCCA GCTACGCCAAAATC	F	0,035
SP2-T5442C	T5442C	TAAAAC T AGGTGCCACGTCTGAAAAGTCTGACAAAAC T GACAAAAC T GACAAAACCAACCCCA	F	0,034
T6365C-SNP	T6365C	ACGTCGTGAAAAGTCTGACAAAATCTTCTCTACACCTAGCAGG	F	0,1
T6392C-SNP	T6392C	GTGCCACGTCTGAAAAGTCTGACAAACCTCTATCTTAGGGCCATCAA	F	0,1
T6776C-SNP	T6776C	TGCCACGTCTGAAAAGTCTGACAAAGTCTAGGTCATTTCTCTACTGTAAAATAT	R	0,012

Nome do primer	SNP	Sequência do primer	Orientação	Concentração final (µM)
C7028T-SNP	C7028T	ACTACACGACCGTACTACGTTGTAGC	F	0,025
SP2-R-A7055G	A7055G	CTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAACACTATGATGGCAAATACAGC	R	0,02
SP2-C7256T	C7256T	CAAAAACCTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAACGATGCATACACCACATGAAA	F	0,06
SP-A8251G new	A8251G	CCACGTCGTGAAAGTCTGACAATTTCCCTAAAAATCTTTTGAAATAGG	F	0,05
SP2-DEL8281-9	8281del9pb	GTCGTGAAAGTCTGACAACCCCTATARCACCCCTCTA	F	0,07
T8404C-SNP	T8404C	ACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAACACTACCCGTATGCCCCACCATAAT	F	0,03
R-G8697A-SNP	G8697A	GTGCCAGTCGTGAAAGTCTGACAACGTCCTTTAGTGTGTTAT	R	0,028
G9055A-SNP	G9055A	TAGGTGCCACGTCGTGAAAGTCTGACAACACTCAATGCACCTAAATTTGGAAGC	F	0,01
R-A9066G-SNP	A9066G	GACAAAACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAACAGAGGAAAGTTAATGGTTGATAT	R	0,125
C10400T-SNP	C10400T	ACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAATGACTACAAAAAGGATTTAGACTGARC	F	0,075
T10873C-SNP	T10873C	GACAAAACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAACAGCCTAATTTAGCATCATCCC	F	0,07
A11251G-SNP	A11251G	AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAACACTTCCCCTACTCATCGCACT	F	0,007
A11812G-SNP	A11812G	TGAAAGTCTGACAACCTCTCAAGGACTTCAAACCTCT	F	0,006
G12372A-SNP	G12372A	CGTCGTGAAAGTCTGACAAAACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAACACTACTATAACCCCTAACCCCT	F	0,1
R-C12705T-SNP	C12705T	CAAAAACCTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAAGTAACTAAGATTAGTATGGTAATTAGGAA	R	0,05
C12858T-SNP	C12858T	AAACTAGGTGCCACGTCGTGAAAGTCTGACAAGCAGCCATTCAAGCAATCCTATA	F	0,023
SP-A13263G	A13263G	GTCGTGAAAGTCTGACAAGCCCTTCTCCACTTCAAGTCA	F	0,03
G13708A-SNP	G13708A	GACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAATAAACCCCATTAACCGCCTG	F	0,16
R-T14470A-SNP	T14470A	TCTGACAAAACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAACAAAAATTTTAGGGGGAATGATGGTTGCTTTGG	R	0,15
C14766T-SNP	C14766T	GAAAAGTCTGACAAAACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAACCAATGACCCCAATACGCCAAAA	F	0,04

Nome do <i>primer</i>	SNP	Sequência do <i>primer</i>	Orientação	Concentração final ( $\mu\text{M}$ )
T14798C-SNP	T14798C	CGTGAAGTCTGACAAAACGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAAACCCCTAATAAAAAATAATTAACCACTCA	F	0,28
A15607G-SNP	A15607	AGTCTGACAACTCCGATCCCGTCCCTAAACA	F	0,018
R-C15904T-SNP	C15904T	ACGTCGTGAAAAGTCTGACAAACCGGTTTACAAGACTGGGTGATTA	R	0,007
SP2-T16362C	T16362C	CCACGTCGTGAAAAGTCTGACAAAACGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAAAATTACAGTCAAAATCCCCTTCTCG	F	0,07
R-T16519C-SNP	T16519C	CGTGAAGTCTGACAAATGTGGGCTATTTAGGCTTTATG	R	0,03

## APÊNDICE C -

Tabela 5: Haplótipos e classificação em haplogrupos das 160 amostras analisadas por sequenciamento de toda a região hipervariável.

<i>Amostra</i>	<i>HG</i>	Polimorfismos da região hipervariável															
<b>BR-SP002</b>	<b><i>C1d</i></b>	73G	152C	189G	194T	249DEL	263G	290DEL	291DEL	309.1C	315.1C	489C	523DEL	524DEL			
		16051G	16093C	16223T	16298C	16325C	16327T	16519C									
<b>BR-SP004</b>	<b><i>HV*</i></b>	152C	263G	315.1C	523DEL	524DEL	16221T	16519C									
<b>BR-SP007</b>	<b><i>L3e1a</i></b>	73G	150T	189G	200G	263G	309.1C	315.1C	16185T	16223T	16327T	16519C					
<b>BR-SP008</b>	<b><i>XI</i></b>	73G	146C	153G	256T	263G	315.1C	16182C	16183C	16189C	16223T	16278T	16519C				
<b>BR-SP009</b>	<b><i>J2</i></b>	73G	150T	152C	263G	295T	315.1C	489C	16069T	16126C	16193T	16278T					
<b>BR-SP010</b>	<b><i>U3a</i></b>	73G	150T	263G	315.1C	16343G	16356C	16390A	16519C								
<b>BR-SP011</b>	<b><i>B4b</i></b>	73G	263G	315.1C	499A	16126C	16183C	16189C	16217C	16372C	16519C						
<b>BR-SP015</b>	<b><i>K2a</i></b>	73G	146C	152C	263G	315.1C	16224C	16311C	16519C								
<b>BR-SP016</b>	<b><i>M*</i></b>	73G	263G	309.1C	315.1C	489C	16192T	16519C									
<b>BR-SP017</b>	<b><i>H3a</i></b>	152C	263G	309.1C	315.1C	16239G	16256T	16519C									
<b>BR-SP018</b>	<b><i>L3e2b</i></b>	73G	150T	195C	263G	309.1C	315.1C	16172C	16183C	16189C	16193.1C	16223T	16320T				
		16519C															
<b>BR-SP019</b>	<b><i>L1c</i></b>	73G	151T	152C	182T	186A	189C	195C	198T	204C	247A	263G	297G	309.1C			
		315.1C	316A	459DEL	16037G	16187T	16189C	16223T	16271C	16274A	16278T	16291T	16294T	16311C			
		16360T	16519C														
<b>BR-SP020</b>	<b><i>L1b</i></b>	73G	152C	182T	185T	195C	228A	247A	263G	315.1C	523DEL	524DEL	16114G	16126C			
		16187T	16189C	16223T	16264T	16270T	16274A	16278T	16293G	16311C	16519C						
<b>BR-SP021</b>	<b><i>L3e2b</i></b>	73G	150T	152C	195C	263G	315.1C	16172C	16183C	16189C	16223T	16239T	16320T	16519C			
<b>BR-SP022</b>	<b><i>C1b</i></b>	71A	73G	249DEL	263G	290DEL	291DEL	309.1C	315.1C	489C	493G	523DEL	524DEL	16126C			
		16223T	16270T	16298C	16325C	16327T	16438A										
<b>BR-SP023</b>	<b><i>B4b</i></b>	73G	103A	152C	263G	309.1C	315.1C	499A	16183C	16189C	16217C	16241G	16519C				







<i>Amostra</i>	<i>HG</i>	<b>Polimorfismos da região hipervariável</b>												
<i>BR-SP086</i>	<i>C1b</i>	73G	152C	249DEL	263G	290DEL	291DEL	309.1C	315.1C	489C	493G	523DEL	524DEL	16037G
			16223T	16294T	16298C	16325C	16327T							
<i>BR-SP087</i>	<i>K1a1a</i>	73G	263G	315.1C	497T	524.1A	524.2C	16093C	16224C	16311C	16519C			
<i>BR-SP089</i>	<i>L3f1b4a</i>	73G	150T	189G	200G	204C	263G	309.1C	315.1C	16209C	16223T	16292T	16311C	16519C
<i>BR-SP090</i>	<i>A2</i>	64T	73G	146C	153G	210G	235G	263G	309.1C	315.1C	523DEL	524DEL	16111T	16209C
		16223T	16290T	16319A	16362C									
<i>BR-SP091</i>	<i>B2c</i>	73G	103A	152C	263G	315.1C	499A	16182C	16183C	16189C	16217C	16241G	16519C	
<i>BR-SP092</i>	<i>B4b</i>	73G	263G	309.1C	315.1C	499A	16183C	16189C	16217C	16519C				
<i>BR-SP094</i>	<i>L3b</i>	73G	195C	263G	279C	309.1C	315.1C	523DEL	524DEL	16145A	16223T	16278T	16362C	16519C
<i>BR-SP095</i>	<i>L3k</i>	73G	146C	150T	152C	235G	263G	315.1C	494T	16223T	16355T			
<i>BR-SP098</i>	<i>B2c</i>	73G	103A	152C	263G	315.1C	499A	16182C	16183C	16189C	16217C	16241G	16519C	
<i>BR-SP099</i>	<i>H</i>	263G	309.1C	309.2C	315.1C	16519C								
<i>BR-SP100</i>	<i>L3d</i>	73G	152C	263G	315.1C	523DEL	524DEL	16124C	16183C	16189C	16223T	16278T	16304C	16311C
<i>BR-SP101</i>	<i>B2c</i>	73G	103A	152C	263G	315.1C	499A	16182C	16183C	16189C	16217C	16241G	16519C	
<i>BR-SP102</i>	<i>D1</i>	73G	263G	315.1C	489C	16223T	16325C	16362C						
<i>BR-SP103</i>	<i>L3e2</i>	73G	150T	195C	198T	263G	315.1C	499A	16223T	16269G	16320T	16399G	16519C	
<i>BR-SP106</i>	<i>B2c</i>	73G	103A	152C	263G	499A	16182C	16183C	16189C	16217C	16241G	16519C		
<i>BR-SP109</i>	<i>T2</i>	73G	263G	315.1C	16126C	16140C	16189C	16294T	16296T	16311C	16519C			
<i>BR-SP110</i>	<i>L4b2</i>	73G	146C	152C	195C	244G	263G	309.1C	315.1C	340T	523DEL	524DEL	16051G	16114T
		16189C	16192T	16223T	16293T	16311C	16316G	16355T	16362C	16399G	16519C			
<i>BR-SP111</i>	<i>L3e1</i>	73G	150T	189G	200G	263G	315.1C	16172C	16223T	16327T	16399G			
<i>BR-SP112</i>	<i>L2a</i>	73G	143A	146C	152C	263G	309.1C	315.1C	16189C	16193T	16193.1C	16223T	16278T	16294T
		16362C	16390A											
<i>BR-SP113</i>	<i>L3e1b</i>	73G	150T	185A	189G	263G	309.1C	315.1C	16223T	16325DEL	16327T			
<i>BR-SP114</i>	<i>HV0?</i>	72C	263G	309.1C	315.1C	16124C	16298C	16319A						
<i>BR-SP117</i>	<i>L2a1a</i>	73G	146C	195C	263G	309.1C	315.1C	16189C	16192T	16223T	16278T	16294T	16309G	16390A



<i>Amostra</i>	<i>HG</i>	<b>Polimorfismos da região hipervariável</b>													
<b>BR-SPI33</b>	A2	64T	73G	146C	235G	263G	309.1C	309.2C	315.1C	523DEL	524DEL	16111T	16223T	16290T	
		16319A	16362C												
<b>BR-SPI34</b>	<b>CIb</b>	73G	152C	249DEL	263G	290DEL	291DEL	309.1C	309.2C	315.1C	489C	493G	523DEL	524DEL	
		16223T	16298C	16325C	16327T	16526A									
<b>BR-SPI36</b>	<b>CIb</b>	73G	249DEL	263G	290DEL	291DEL	309.1C	315.1C	489C	493G	523DEL	524DEL	16223T	16298C	
		16325C	16327T	16526A											
<b>BR-SPI37</b>	A2	73G	146C	153G	235G	263G	315.1C	523DEL	524DEL	16092C	16111T	16223T	16290T	16319A	
		16362C													
<b>BR-SPI38</b>	<b>H</b>	93G	263G	309.1C	315.1C	16519C									
<b>BR-SPI39</b>	<b>L3a2b</b>	73G	150T	152C	195C	263G	309.1C	315.1C	16172C	16183C	16189C	16223T	16320T	16519C	
<b>BR-SPI40</b>	<b>CIb</b>	73G	249DEL	263G	290DEL	291DEL	309.1C	309.2C	315.1C	489C	493G	523DEL	524DEL	16223T	
		16298C	16325C	16327T											
<b>BR-SPI42</b>	<b>H</b>	263G	309.1C	315.1C	16311C	16519C									
<b>BR-SPI43</b>	<b>L0a1b</b>	93G	95C	185A	189G	236C	247A	263G	315.1C	523DEL	524DEL	16129A	16148T	16168T	
		16172C	16187T	16188G	16189C	16223T	16230G	16278T	16293G	16311C	16320T				
<b>BR-SPI45</b>	<b>L2a1</b>	73G	143A	146C	152C	195C	263G	309.1C	315.1C	16223T	16256T	16278T	16294T	16309G	
		16390A													
<b>BR-SPI46</b>	<b>H</b>	152C	263G	309.1C	309.2C	315.1C	16519C								
<b>BR-SPI47</b>	<b>N9b?</b>	73G	150T	200G	263G	309.1C	309.2C	315.1C	16129A	16183C	16189C	16193.1C	16223T	16260T	
		16327T													
<b>BR-SPI48</b>	<b>H5</b>	93G	263G	309.1C	315.1C	456T	16294T	16304C							
<b>BR-SPI49</b>	<b>H</b>	263G	309.1C	315.1C	523DEL	524DEL	16519C								
<b>BR-SPI50</b>	<b>H5</b>	263G	309.1C	315.1C	456T	16294T	16304C	16320T							
<b>BR-SPI51</b>	<b>K1a4</b>	73G	263G	315.1C	497T	523.1A	524.1C	16129A	16224C	16311C	16519C				
<b>BR-SPI52</b>	<b>T1a</b>	73G	195C	263G	315.1C	16126C	16163G	16186T	16189C	16294T	16519C				
<b>BR-SPI54</b>	<b>U4a1</b>	73G	152C	195C	263G	309.1C	315.1C	499A	523.1A	524.1C	16134T	16240G	16286A	16356C	



<i>Amostra</i>	<i>HG</i>	<b>Polimorfismos da região hipervariável</b>												
<b>BR-SPI170</b>	<b>M1a1</b>	73G	152C	195C	204C	315.1C	489C	16129A	16182C	16183C	16189C	16223T	16249C	16311C
		16359C	16519C	16527T										
<b>BR-SPI171</b>	<b>B2c</b>	73G	263G	309.1C	315.1C	499A	523DEL	524DEL	16182C	16183C	16186T	16189C	16217C	16519C
<b>BR-SPI172</b>	<b>J1</b>	73G	256T	263G	295T	309.1C	315.1C	462T	482C	489C	16069T	16126C		
<b>BR-SPI173</b>	<b>L0a2</b>	64T	93G	152C	189G	236C	247A	263G	309.1C	315.1C	523DEL	524DEL	16148T	16172C
		16187T	16188G	16189C	16223T	16230G	16311C	16320T	16519C					
<b>BR-SPI174</b>	<b>H6</b>	239C	263G	309.1C	315.1C	16192T	16362C	16482G						
<b>BR-SPI177</b>	<b>H</b>	263G	315.1C	16311C	16519C									
<b>BR-SPI178</b>	<b>L2a1</b>	73G	146C	152C	195C	198T	263G	315.1C	16093C	16129A	16189C	16192T	16223T	16278T
		16294T	16309G	16390A										
<b>BR-SPI180</b>	<b>L2a1b</b>	73G	146C	152C	195C	263G	315.1C	16182C	16183C	16189C	16223T	16278T	16290T	16294T
		16309G	16390A											
<b>BR-SPI181</b>	<b>H</b>	263G	315.1C	524.1A	524.2C	16148T	16519C							
<b>BR-SPI182</b>	<b>C1b</b>	73G	249DEL	263G	290DEL	291DEL	309.1C	315.1C	489C	493G	523DEL	524DEL	16126C	16223T
		16298C	16325C	16327T										
<b>BR-SPI183</b>	<b>U4a3</b>	73G	195C	247A	263G	315.1C	499A	523.1A	524.1C	525.1A	526.1C	16265G	16356C	16362C
		16519C												
<b>BR-SPI184</b>	<b>L1b</b>	73G	152C	182T	185T	195C	247A	263G	315.1C	357G	523DEL	524DEL	16126C	16187T
		16189C	16223T	16264T	16270T	16278T	16293G	16311C	16317G	16519C				
<b>BR-SPI185</b>	<b>W4</b>	73G	143A	189G	194T	195C	196C	204C	207A	263G	309.1C	315.1C	16223T	16286T
		16292T	16519C											
<b>BR-SPI186</b>	<b>L2b</b>	73G	150T	152C	182T	195C	198T	204C	263G	315.1C	418T	523DEL	524DEL	16114A
		16129A	16213A	16223T	16278T	16355T	16390A							
<b>BR-SPI187</b>	<b>L1c3</b>	73G	151T	152C	182T	186A	189C	195C	247A	248G	263G	315.1C	316A	523DEL
		524DEL	16129A	16187T	16189C	16223T	16278T	16293G	16294T	16311C	16360T	16519C		
<b>BR-SPI189</b>	<b>L1c2</b>	73G	151T	152C	182T	186A	189C	195C	198T	247A	263G	297G	309.1C	315.1C

<i>Amostra</i>	<i>HG</i>	Polimorfismos da região hipervariável															
		316A	471C	523DEL	524DEL	16129A	16169T	16187T	16189C	16223T	16265C	16278T	16286G	16294T			
		16311C	16360T	16519C	16527T												
<b>BR-SP190</b>	<b>L2a1b</b>	73G	146C	152C	195C	263G	315.1C	16182C	16183C	16189C	16223T	16269C	16278T	16290T			
		16294T	16309G	16390A													
<b>BR-SP192</b>	<b>B2c</b>	73G	103A	152C	263G	309.1C	309.2C	315.1C	499A	16182C	16183C	16189C	16217C	16241G			
		16301T	16519C														
<b>BR-SP193</b>	<b>L2b1</b>	73G	150T	152C	182T	195C	198T	204C	263G	315.1C	418T	523DEL	524DEL	16114A			
		16129A	16213A	16223T	16278T	16355T	16362C	16390A									
<b>BR-SP194</b>	<b>H</b>	263G	309.1C	315.1C	16184T	16263C	16265G	16519C									
<b>BR-SP195</b>	<b>L3a1a</b>	73G	150T	189G	200G	263G	309.1C	315.1C	16185T	16223T	16327T	16519C					
<b>BR-SP196</b>	<b>C1b</b>	73G	249DEL	263G	290DEL	291DEL	309.1C	315.1C	489C	493G	523DEL	524DEL	16126C	16223T			
		16298C	16325C	16327T													
<b>BR-SP197</b>	<b>A2</b>	73G	146C	153G	235G	263G	309.1C	315.1C	16111T	16223T	16290T	16319A	16362C				
<b>BR-SP199</b>	<b>L0a1b</b>	89C	93G	95C	185A	189G	236C	247A	263G	309.1C	315.1C	523DEL	524DEL	16129A			
		16148T	16168T	16172C	16187T	16188G	16189C	16223T	16230G	16278T	16293G	16311C	16320T	16519C			
<b>BR-SP201</b>	<b>B4b</b>	73G	146C	152C	199C	263G	315.1C	499A	16183C	16189C	16217C	16519C					

Nota: HG (haplogrupo); “?”(ou); “?” dúvida na classificação.

## APÊNDICE D -

Tabela 12: Fenótipo e resultado da genotipagem do SNP 3010 por PCR em tempo real.

<i>Amostra</i>	<b>Fenótipo</b>	<b>SNP 3010</b>
BR-SP002	BRANCO	G
BR-SP004	PRETO	A
BR-SP007	PRETO	G
BR-SP008	BRANCO	G
BR-SP009	BRANCO	G
BR-SP010	BRANCO	A
BR-SP011	PRETO	G
BR-SP015	BRANCO	G
BR-SP016	BRANCO	G
BR-SP017	BRANCO	G
BR-SP018	PRETO	G
BR-SP019	PRETO	G
BR-SP020	BRANCO	G
BR-SP021	BRANCO	G
BR-SP022	BRANCO	G
BR-SP023	PRETO	G
BR-SP024	BRANCO	G
BR-SP026	BRANCO	G
BR-SP027	BRANCO	G
BR-SP028	PRETO	G
BR-SP029	PRETO	G
BR-SP030	PRETO	G
BR-SP032	PRETO	G
BR-SP034	BRANCO	G
BR-SP035	BRANCO	G
BR-SP036	BRANCO	G
BR-SP037	PRETO	G
BR-SP038	BRANCO	G
BR-SP039	BRANCO	G
BR-SP040	BRANCO	G
BR-SP041	BRANCO	G
BR-SP042	BRANCO	G
BR-SP044	PRETO	G
BR-SP045	BRANCO	G
BR-SP046	PRETO	G
BR-SP047	BRANCO	G
BR-SP050	PRETO	A
BR-SP051	BRANCO	G
BR-SP052	BRANCO	A
BR-SP053	PRETO	G

<i>Amostra</i>	<b>Fenótipo</b>	<b>SNP 3010</b>
BR-SP054	PRETO	G
BR-SP057	PRETO	G
BR-SP058	BRANCO	A
BR-SP059	BRANCO	G
BR-SP060	PRETO	G
BR-SP061	PRETO	G
BR-SP063	PRETO	G
BR-SP064	PRETO	G
BR-SP065	BRANCO	G
BR-SP066	BRANCO	G
BR-SP067	BRANCO	G
BR-SP068	BRANCO	G
BR-SP070	BRANCO	G
BR-SP071	BRANCO	G
BR-SP072	PRETO	G
BR-SP073	BRANCO	A
BR-SP074	BRANCO	G
BR-SP076	BRANCO	G
BR-SP077	BRANCO	G
BR-SP078	BRANCO	G
BR-SP079	BRANCO	G
BR-SP080	BRANCO	G
BR-SP082	PRETO	G
BR-SP083	PRETO	G
BR-SP084	BRANCO	G
BR-SP085	PRETO	G
BR-SP086	BRANCO	G
BR-SP087	BRANCO	G
BR-SP089	PRETO	G
BR-SP090	BRANCO	G
BR-SP091	BRANCO	G
BR-SP092	BRANCO	G
BR-SP094	PRETO	G
BR-SP095	PRETO	G
BR-SP098	BRANCO	G
BR-SP099	BRANCO	G
BR-SP100	BRANCO	G
BR-SP101	BRANCO	G
BR-SP102	PRETO	A
BR-SP103	PRETO	G
BR-SP106	BRANCO	G
BR-SP109	BRANCO	G
BR-SP110	PRETO	G
BR-SP111	PRETO	G

<i>Amostra</i>	<b>Fenótipo</b>	<b>SNP 3010</b>
BR-SP113	BRANCO	G
BR-SP114	BRANCO	G
BR-SP117	PRETO	G
BR-SP118	PRETO	G
BR-SP119	BRANCO	G
BR-SP121	PRETO	G
BR-SP122	PRETO	G
BR-SP123	PRETO	G
BR-SP124	PRETO	G
BR-SP125	BRANCO	G
BR-SP126	PRETO	G
BR-SP127	PRETO	G
BR-SP128	PRETO	A
BR-SP129	PRETO	G
BR-SP130	PRETO	G
BR-SP131	PRETO	G
BR-SP132	PRETO	G
BR-SP133	BRANCO	G
BR-SP134	PRETO	G
BR-SP136	PRETO	G
BR-SP137	BRANCO	G
BR-SP138	PRETO	A
BR-SP139	PRETO	G
BR-SP140	PRETO	G
BR-SP142	BRANCO	G
BR-SP143	BRANCO	G
BR-SP145	PRETO	G
BR-SP146	BRANCO	G
BR-SP147	PRETO	G
BR-SP148	BRANCO	G
BR-SP149	BRANCO	A
BR-SP150	BRANCO	G
BR-SP151	BRANCO	G
BR-SP152	BRANCO	G
BR-SP154	PRETO	G
BR-SP155	BRANCO	G
BR-SP156	PRETO	G
BR-SP157	BRANCO	A
BR-SP158	BRANCO	G
BR-SP159	PRETO	G
BR-SP160	PRETO	G
BR-SP161	BRANCO	G
BR-SP162	BRANCO	A
BR-SP163	BRANCO	G

---

<i>Amostra</i>	<b>Fenótipo</b>	<b>SNP 3010</b>
BR-SP165	BRANCO	G
BR-SP166	PRETO	G
BR-SP167	PRETO	A
BR-SP168	BRANCO	A
BR-SP169	PRETO	A
BR-SP170	BRANCO	G
BR-SP171	PRETO	G
BR-SP172	BRANCO	A
BR-SP173	BRANCO	G
BR-SP174	BRANCO	G
BR-SP177	BRANCO	G
BR-SP178	BRANCO	G
BR-SP180	PRETO	G
BR-SP181	BRANCO	G
BR-SP182	PRETO	G
BR-SP183	BRANCO	G
BR-SP184	BRANCO	G
BR-SP185	BRANCO	G
BR-SP186	PRETO	G
BR-SP187	PRETO	G
BR-SP189	PRETO	G
BR-SP190	PRETO	G
BR-SP192	BRANCO	G
BR-SP193	BRANCO	G
BR-SP194	BRANCO	G
BR-SP195	PRETO	G
BR-SP196	BRANCO	G
BR-SP197	PRETO	G
BR-SP199	PRETO	G
BR-SP201	BRANCO	G

---

APÊNDICE E -

Tabela 14 – Resultado da análise dos 42 SNPs por SNaPshot e classificação em haplogrupos das 160 amostras.

<i>Amostra</i>	<i>HG</i>	2772	16519	15904	8251	9055	13708	4580	3010	12372	1719	15607	11812	13263	6776	11251	1811	12705	14470	709	4883
<i>BR-SP002</i>	<i>C</i>	C	C	C	G	G	G	G	G	G	G	A	A	G	T	A	A	T	T	G	C
<i>BR-SP004</i>	<i>HI*</i>	C	C	C	G	G	G	G	A	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP007</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP008</i>	<i>X*</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	C	G	C
<i>BR-SP009</i>	<i>J</i>	C	T	C	G	G	A	G	G	G	G	A	A	A	T	G	A	C	T	G	C
<i>BR-SP010</i>	<i>Un*</i>	C	C	C	G	G	G	G	A	A	G	A	A	A	T	A	G	C	T	G	C
<i>BR-SP011</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP015</i>	<i>K</i>	C	C	C	G	A	G	G	G	A	G	A	A	A	T	A	G	C	T	G	C
<i>BR-SP016</i>	<i>M*</i>	C	C	C	G	A	G	A	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP017</i>	<i>H3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	C	A	A	C	T	G	C
<i>BR-SP018</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP019</i>	<i>L1</i>	C	C	C	G	G	G	G	G	G	A	A	A	A	T	A	A	T	T	G	C
<i>BR-SP020</i>	<i>L1</i>	C	C	C	?	G	G	G	G	G	G	A	A	A	T	A	A	T	T	C	C
<i>BR-SP021</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP022</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	G	T	A	A	T	T	G	C
<i>BR-SP023</i>	<i>B</i>	?	C	C	G	G	G	G	G	G	?	?	A	A	T	A	?	C	C	G	?
<i>BR-SP024</i>	<i>L0a2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	A	C
<i>BR-SP026</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	?	A	A	A	T	A	A	T	T	G	C
<i>BR-SP027</i>	<i>A</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP028</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP029</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP030</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP032</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C

<i>Amostra</i>	3915	2706	8697	12858	del8281	10400	10873	2416	7028	1736	5004	6365	6392	5442	7055	8404	7256	9066	14766	769	14798	16362
<i>BR-SP002</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP004</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP007</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP008</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP009</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP010</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP011</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP015</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	C	T
<i>BR-SP016</i>	A	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP017</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP018</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP019</i>	G	G	G	C	C	C	C	?	T	A	T	T	T	T	G	T	T	A	T	A	T	T
<i>BR-SP020</i>	G	G	G	C	C	C	C	T	T	?	T	T	T	T	G	T	T	A	T	A	T	T
<i>BR-SP021</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP022</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP023</i>	G	?	G	C	G	C	T	T	T	?	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP024</i>	G	G	G	C	G	C	C	T	T	A	T	T	T	C	A	T	T	A	T	G	T	T
<i>BR-SP026</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SP027</i>	G	G	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP028</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP029</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP030</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP032</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T

<i>Amostra</i>	<i>HG</i>	2772	16519	15904	8251	9055	13708	4580	3010	12372	1719	15607	11812	13263	6776	11251	1811	12705	14470	709	4883
<i>BR-SP034</i>	<i>A</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP035</i>	<i>HV</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP036</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP037</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP038</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP039</i>	<i>C</i>	C	?	C	G	G	G	G	G	G	G	A	A	G	T	A	A	T	T	G	C
<i>BR-SP040</i>	<i>H</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP041</i>	<i>L2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP042</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP044</i>	<i>L2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP045</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP046</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP047</i>	<i>H3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	C	A	A	C	T	G	C
<i>BR-SP050</i>	<i>D</i>	C	C	C	G	G	G	G	A	G	G	A	A	A	T	A	A	T	T	G	T
<i>BR-SP051</i>	<i>X2g</i>	C	C	C	G	G	G	G	G	G	A	A	A	A	T	A	A	T	C	G	C
<i>BR-SP052</i>	<i>J1c</i>	C	C	C	G	G	A	G	A	G	G	A	A	A	T	G	A	C	T	G	C
<i>BR-SP053</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP054</i>	<i>L1</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP057</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP058</i>	<i>H1*</i>	C	C	C	G	G	G	G	A	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP059</i>	<i>C</i>	C	C	C	G	G	G	G	G	G	G	A	A	G	T	A	A	T	T	G	C
<i>BR-SP060</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	G	T	A	A	T	T	G	C
<i>BR-SP061</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP063</i>	<i>L1</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C

<i>Amostra</i>	3915	2706	8697	12858	del8281	10400	10873	2416	7028	1736	5004	6365	6392	5442	7055	8404	7256	9066	14766	769	14798	16362
<i>BR-SP034</i>	G	G	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP035</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP036</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP037</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP038</i>	G	G	?	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP039</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP040</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP041</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SP042</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP044</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SP045</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SP046</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP047</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP050</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP051</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	A	T	T
<i>BR-SP052</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	C	T
<i>BR-SP053</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP054</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	G	T	T	A	T	A	T	?
<i>BR-SP057</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SP058</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP059</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP060</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP061</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP063</i>	G	G	G	C	C	C	C	T	T	?	T	T	T	T	G	T	T	A	T	A	T	T

<i>Amostra</i>	<i>HG</i>	2772	16519	15904	8251	9055	13708	4580	3010	12372	1719	15607	11812	13263	6776	11251	1811	12705	14470	709	4883
<i>BR-SP064</i>	<i>T2</i>	C	C	C	G	G	G	G	G	G	G	G	G	A	T	G	A	C	T	?	C
<i>BR-SP065</i>	<i>K</i>	C	C	C	G	A	G	G	G	A	G	A	A	A	T	A	G	C	T	G	C
<i>BR-SP066</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP067</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP068</i>	<i>A</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP070</i>	<i>T*</i>	C	C	C	G	G	G	G	G	G	G	G	A	A	T	G	A	C	T	A	C
<i>BR-SP071</i>	<i>A</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP072</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP073</i>	<i>HI*</i>	C	T	C	G	G	G	G	A	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP074</i>	<i>U*</i>	C	T	C	G	G	G	G	G	A	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP076</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP077</i>	<i>L2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP078</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP079</i>	<i>L0a2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP080</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP082</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP083</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP084</i>	<i>L1</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP085</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP086</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP087</i>	<i>K</i>	C	C	C	G	A	G	G	G	A	G	A	A	A	T	A	G	C	T	G	C
<i>BR-SP089</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP090</i>	<i>A</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP091</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	C	G	C
<i>BR-SP092</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP094</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C

<i>Amostra</i>	3915	2706	8697	12858	del8281	10400	10873	2416	7028	1736	5004	6365	6392	5442	7055	8404	7256	9066	14766	769	14798	16362
<i>BR-SP064</i>	G	G	A	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP065</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	C	T
<i>BR-SP066</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP067</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP068</i>	G	G	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP070</i>	G	G	A	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP071</i>	G	G	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP072</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP073</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP074</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP076</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP077</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SP078</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP079</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	C	A	T	T	A	T	A	T	T
<i>BR-SP080</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP082</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP083</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP084</i>	G	G	G	C	C	C	C	?	T	A	T	T	T	T	G	T	T	A	T	A	T	?
<i>BR-SP085</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP086</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP087</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	C	T
<i>BR-SP089</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP090</i>	G	G	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP091</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP092</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP094</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	C

<i>Amostra</i>	<i>HG</i>	2772	16519	15904	8251	9055	13708	4580	3010	12372	1719	15607	11812	13263	6776	11251	1811	12705	14470	709	4883
<i>BR-SP095</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP098</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	C	G	C
<i>BR-SP099</i>	<i>H*</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP100</i>	<i>L3</i>	C	T	C	A	G	G	G	G	G	A	A	A	A	T	A	A	T	T	G	C
<i>BR-SP101</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	C	G	C
<i>BR-SP102</i>	<i>D</i>	C	T	C	G	G	G	G	A	G	G	A	A	A	T	A	A	T	T	G	T
<i>BR-SP103</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP106</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	C	G	C
<i>BR-SP109</i>	<i>T2</i>	C	C	C	G	G	G	G	G	G	G	G	G	A	T	G	A	C	T	A	C
<i>BR-SP110</i>	<i>L4</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	A	C
<i>BR-SP111</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP112</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP113</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP114</i>	<i>V</i>	C	T	T	G	G	G	A	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP117</i>	<i>L2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP118</i>	<i>A</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP119</i>	<i>H*</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP121</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP122</i>	<i>L*</i>	C	C	?	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP123</i>	<i>A</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP124</i>	<i>L1</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP125</i>	<i>A</i>	C	C	C	G	G	G	G	G	G	?	A	A	A	T	A	?	T	T	G	?
<i>BR-SP126</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP127</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	G	T	A	A	T	T	G	C
<i>BR-SP128</i>	<i>HI*</i>	C	C	C	G	G	G	G	A	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SP129</i>	<i>L2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C

<i>Amostra</i>	3915	2706	8697	12858	del8281	10400	10873	2416	7028	1736	5004	6365	6392	5442	7055	8404	7256	9066	14766	769	14798	16362
<i>BR-SP095</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP098</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP099</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP100</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP101</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP102</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP103</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP106</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP109</i>	G	G	A	C	?	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP110</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	A	T	C
<i>BR-SP111</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP112</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	C
<i>BR-SP113</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP114</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP117</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SP118</i>	G	G	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP119</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP121</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP122</i>	G	G	G	C	C	C	C	?	T	A	T	T	T	T	A	T	T	A	T	A	T	?
<i>BR-SP123</i>	G	G	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP124</i>	G	G	G	C	C	C	C	?	T	A	T	T	T	T	G	T	T	A	T	A	T	?
<i>BR-SP125</i>	G	?	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP126</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP127</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP128</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP129</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T

<i>Amostra</i>	<i>HG</i>	2772	16519	15904	8251	9055	13708	4580	3010	12372	1719	15607	11812	13263	6776	11251	1811	12705	14470	709	4883
<i>BR-SPI30</i>	<i>L*</i>	C	C	?	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI31</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI32</i>	<i>L2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI33</i>	<i>A</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI34</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI36</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI37</i>	<i>A</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI38</i>	<i>H1*</i>	C	C	C	G	G	G	G	A	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI39</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI40</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI42</i>	<i>H*</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI43</i>	<i>L0</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI45</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI46</i>	<i>H3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	C	A	A	C	T	G	C
<i>BR-SPI47</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI48</i>	<i>H*</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI49</i>	<i>H1*</i>	C	C	C	G	G	G	G	A	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI50</i>	<i>H*</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI51</i>	<i>K</i>	C	C	C	G	A	G	G	G	A	G	A	A	A	T	A	G	C	T	G	C
<i>BR-SPI52</i>	<i>T*</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	G	A	C	T	A	C
<i>BR-SPI54</i>	<i>Un*</i>	C	C	C	G	G	G	G	G	A	G	A	A	A	T	A	G	C	T	G	C
<i>BR-SPI55</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI56</i>	<i>J*</i>	C	C	C	G	G	A	G	G	G	G	A	A	A	T	G	A	C	T	G	C
<i>BR-SPI57</i>	<i>J1c</i>	C	T	C	G	G	A	G	A	G	G	A	A	A	T	G	A	C	T	G	C
<i>BR-SPI58</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C

<i>Amostra</i>	3915	2706	8697	12858	del8281	10400	10873	2416	7028	1736	5004	6365	6392	5442	7055	8404	7256	9066	14766	769	14798	16362
<i>BR-SPI30</i>	G	G	G	C	C	C	C	?	T	A	T	T	T	T	A	T	T	A	T	A	T	?
<i>BR-SPI31</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	C
<i>BR-SPI32</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SPI33</i>	G	G	G	C	C	C	C	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SPI34</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI36</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI37</i>	G	G	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SPI38</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI39</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI40</i>	G	G	G	C	C	T	C	T	T	A	C	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI42</i>	G	A	G	C	C	C	C	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI43</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	C	A	T	T	A	T	A	T	T
<i>BR-SPI45</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SPI46</i>	G	?	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI47</i>	G	G	G	C	G	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI48</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI49</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI50</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI51</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	C	T
<i>BR-SPI52</i>	G	G	A	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI54</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI55</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI56</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI57</i>	G	G	G	C	C	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	C	T
<i>BR-SPI58</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	C

<i>Amostra</i>	<i>HG</i>	2772	16519	15904	8251	9055	13708	4580	3010	12372	1719	15607	11812	13263	6776	11251	1811	12705	14470	709	4883
<i>BR-SPI59</i>	<i>L3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI60</i>	<i>L2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI61</i>	<i>L3</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI62</i>	<i>H1*</i>	C	C	C	G	G	G	G	A	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI63</i>	<i>L2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI64</i>	<i>L4</i>	C	C	C	G	G	G	G	A	G	G	A	A	A	T	A	?	T	T	A	C
<i>BR-SPI65</i>	<i>L0</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI66</i>	<i>H3</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	C	A	A	C	T	G	C
<i>BR-SPI67</i>	<i>L2</i>	C	C	C	?	G	G	G	A	G	G	A	A	A	T	A	A	T	T	A	C
<i>BR-SPI68</i>	<i>H1*</i>	C	C	C	G	G	G	G	A	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI69</i>	<i>L2</i>	C	T	C	G	G	G	G	A	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI70</i>	<i>M*</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI71</i>	<i>B</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI72</i>	<i>J1c</i>	C	T	C	G	G	A	G	A	G	G	A	A	A	T	G	A	C	T	G	C
<i>BR-SPI73</i>	<i>L0a2</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI74</i>	<i>H6</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI77</i>	<i>H*</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI78</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI80</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI81</i>	<i>H*</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI82</i>	<i>C</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI83</i>	<i>Un*</i>	C	C	C	G	G	G	G	G	A	G	A	A	A	T	A	G	C	T	G	C
<i>BR-SPI84</i>	<i>L1</i>	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	?	C
<i>BR-SPI85</i>	<i>W</i>	C	C	C	A	G	G	G	G	G	G	A	A	A	T	A	A	T	T	A	C
<i>BR-SPI86</i>	<i>L2</i>	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C

<i>Amostra</i>	3915	2706	8697	12858	del8281	10400	10873	2416	7028	1736	5004	6365	6392	5442	7055	8404	7256	9066	14766	769	14798	16362
<i>BR-SPI59</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI60</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SPI61</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI62</i>	G	A	G	C	C	C	C	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI63</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SPI64</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	A	T	C
<i>BR-SPI65</i>	G	G	G	C	C	C	C	T	T	A	T	T	C	C	A	T	T	A	T	A	T	T
<i>BR-SPI66</i>	G	A	G	C	C	C	C	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI67</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	G	T	T
<i>BR-SPI68</i>	G	A	G	C	C	C	C	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI69</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SPI70</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI71</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI72</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	C	T
<i>BR-SPI73</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	C	A	T	T	A	T	A	T	T
<i>BR-SPI74</i>	A	A	G	C	C	C	C	T	C	A	T	T	T	T	A	T	C	A	C	G	T	C
<i>BR-SPI77</i>	G	A	G	C	C	C	C	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI78</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SPI80</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SPI81</i>	G	A	G	C	C	C	C	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SPI82</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI83</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SPI84</i>	G	G	G	C	C	C	C	T	T	?	T	T	T	T	G	T	T	A	T	A	T	T
<i>BR-SPI85</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SPI86</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T

<i>Amostra</i>	<i>HG</i>	2772	16519	15904	8251	9055	13708	4580	3010	12372	1719	15607	11812	13263	6776	11251	1811	12705	14470	709	4883
<i>BR-SPI87</i>	L*	C	C	?	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI89</i>	L1	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI90</i>	L2	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI92</i>	B	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	C	G	C
<i>BR-SPI93</i>	L2	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI94</i>	H*	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C
<i>BR-SPI95</i>	L3	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI96</i>	C	C	T	C	G	G	G	G	G	G	G	A	A	G	T	A	A	T	T	G	C
<i>BR-SPI97</i>	A	C	T	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SPI99</i>	L0	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	T	T	G	C
<i>BR-SP201</i>	B	C	C	C	G	G	G	G	G	G	G	A	A	A	T	A	A	C	T	G	C

<i>Amostra</i>	3915	2706	8697	12858	de18281	10400	10873	2416	7028	1736	5004	6365	6392	5442	7055	8404	7256	9066	14766	769	14798	16362
<i>BR-SP187</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	?
<i>BR-SP189</i>	G	G	G	C	C	C	C	?	T	A	T	T	T	T	G	T	T	A	T	A	T	?
<i>BR-SP190</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	T
<i>BR-SP192</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP193</i>	G	G	G	C	C	C	C	C	T	A	T	T	T	T	A	T	T	A	T	A	T	C
<i>BR-SP194</i>	G	A	G	C	C	C	T	T	C	A	T	T	T	T	A	T	C	A	C	G	T	T
<i>BR-SP195</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP196</i>	G	G	G	C	C	T	C	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T
<i>BR-SP197</i>	G	G	G	C	C	C	T	T	T	G	T	T	T	T	A	T	C	A	T	G	T	C
<i>BR-SP199</i>	G	G	G	C	C	C	C	T	T	A	T	T	T	C	A	T	T	A	T	A	T	T
<i>BR-SP201</i>	G	G	G	C	G	C	T	T	T	A	T	T	T	T	A	T	C	A	T	G	T	T

Nota: HG (haplogrupo); ? (alelo não foi obtido).

## **CAPÍTULO 2**

**TECHNICAL NOTE**  
**CRIMINALISTICS**

Greiciane G. Paneto,<sup>1</sup> M.Sc.; Larissa V.G. Longo,<sup>1</sup> B.S.; Joyce A. Martins,<sup>1</sup> M.Sc.;  
Maria Angélica de Camargo,<sup>1</sup> B.S.; Jeane C. Costa,<sup>1</sup> B.S.; Aline C.O. de Mello,<sup>1</sup> B.S.; Bety Chen,<sup>2</sup> M.Sc.;  
Rogério N. Oliveira,<sup>3</sup> Ph.D.; Mário H. Hirata,<sup>2</sup> Ph.D.; and Regina M.B. Cicarelli,<sup>1</sup> Ph.D.

## Heteroplasmy in Hair: Study of Mitochondrial DNA Third Hypervariable Region in Hair and Blood Samples\*†

**ABSTRACT:** Mitochondrial DNA (mtDNA) analysis has proved useful for forensic identification especially in cases where nuclear DNA is not available, such as with hair evidence. Heteroplasmy, the presence of more than one type of mtDNA in one individual, is a common situation often reported in the first and second mtDNA hypervariable regions (HV1/HV2), particularly in hair samples. However, there is no data about heteroplasmy frequency in the third mtDNA hypervariable region (HV3). To investigate possible heteroplasmy hotspots, HV3 from hair and blood samples of 100 individuals were sequenced and compared. No point heteroplasmy was observed, but length heteroplasmy was, both in C-stretch and CA repeat. To observe which CA "alleles" were present in each tissue, PCR products were cloned and re-sequenced. However, no variation among CA alleles was observed. Regarding forensic practice, we conclude that point heteroplasmy in HV3 is not as frequent as in the HV1/HV2.

**KEYWORDS:** forensic science, mitochondrial DNA, heteroplasmy, hair, blood, HV3

Heteroplasmy, the presence of more than one type of mitochondrial DNA (mtDNA) within an individual, is a situation found more commonly in hair than in blood samples (1–4). The clonal nature of hair follicles and the high-energy requirements of keratinizing hair shaft cells are two features of hair histogenesis that could contribute to the high observation of the segregation of heteroplasmic variants in mtDNA from hair shafts (5). In contrast, peripheral-blood samples consist of lymphocytes produced by a very large number of hemopoietic stem cells in the bone marrow, and if proportions differ in individual stem cells, the heteroplasmy measured is an averaged value of the proportions of all stem cells. The high mutation rate of the mtDNA, when compared to nuclear DNA, appears to occur because it is more exposed to mutagenic events, owing to its lack of protective histones, proximity to the

respiratory chain, and its relative deficiency in DNA repair mechanisms (6).

Heteroplasmy can be differentiated into two types: point and length heteroplasmy. A sequence point, or site heteroplasmy, is defined as sequences containing different bases at the same position. The sequence electropherogram usually shows two different bases superimposed (7). A length heteroplasmy is represented by multiple populations of mtDNA containing repetitive stretches (typically C-stretches or CA repeats) of various lengths.

The presence of heteroplasmy in a forensic case can be relevant because two samples from the same individual may differ in levels of heteroplasmy or sometimes in a homoplasmic substitution (8). Therefore, it is important to determine the frequency of heteroplasmy and hotspots, to avoid incorrect interpretations.

While a number of studies have been published on heteroplasmy in the first and second hypervariable regions (HV1 and HV2), little is known about the third hypervariable region (HV3). Its importance to forensic science was described by Lutz et al. (9), and it is now incorporated into forensic analysis, mainly because of an informative dinucleotide repeat (CA)<sub>n</sub> (10–15). Analysis of the entire mitochondrial DNA control region, including HV3, is also being proposed, for the best determination of the haplogroup affiliation of populations. Phylogenetic analysis is being used for quality assurance of sequences (16). At this time, there are no data about possible hotspots of heteroplasmic positions and their frequencies in the HV3 region, comparing hair and blood samples.

The aim of our work was to find differences in the HV3 (nt 438–574) sequences between hair and blood from the same

<sup>1</sup>UNESP, São Paulo State University, School of Pharmaceutical Sciences, Rodovia Araraquã-Jaú, Km 01, Campus, 14801-902 Araraquara, SP, Brazil.

<sup>2</sup>USP, University of São Paulo, School of Pharmaceutical Sciences, Av. Prof. Lineu Prestes, 580 Bl. 17, Cidade Universitária, 05508900 São Paulo, SP, Brazil.

<sup>3</sup>USP, University of São Paulo, School of Dentistry, Av. Prof. Lineu Prestes, 2227, Cidade Universitária, 05508900 São Paulo, SP, Brazil.

\*Presented at the DNA in Forensics 2008 Conference, May 27–30, 2008, in Ancona, Italy.

†Funded by a grant from FAPESP - Fundação de Apoio a Pesquisa do Estado de São Paulo (2005/03984-5). G.G. Paneto receives a fellowship from CNPq - Conselho Nacional de Pesquisa (2007-0/140052), and L.V.G. Longo received a fellowship from FAPESP (2006/06772-1). CNPq is the Brazilian National Research Council and FAPESP is the São Paulo State Research Aid Foundation.

Received 11 May 2008; and in revised form 1 May 2009; accepted 3 May 2009.

individuals to determine possible hotspots of heteroplasmic positions and the frequencies in those tissues.

## Materials and Methods

### DNA Samples and Extraction

One hundred unrelated Brazilian individuals were analyzed (all the participants gave their written informed consent prior to their inclusion in this study). Samples from head hair and blood (spotted onto filter paper and air-dried) were collected from the same individuals (3–6 years previously), and stored at 8°C (in refrigerator) until analyzed.

Hair shafts were cut off 0.5 cm above the skin surface, and 4 cm of the hair was cut into pieces of 0.5 cm each (more than one hair was used when the length of hair was smaller than 4 cm). Hair samples were decontaminated as described by Jehaes et al. (17) prior to DNA extraction. DNA was extracted using Tissue and Hair extraction kit (Promega, Madison, WI), with the following modifications: incubation for 2 h 30 min, followed by phenyl/chloroform/isoamyl alcohol (PCIA) extraction (USB-GE GE, Chalfont St. Giles, UK) and purification with columns of illustra GFX™ PCR DNA and gel band purification kit (GE). DNA was eluted in 50 µL double-distilled H<sub>2</sub>O (ddH<sub>2</sub>O).

DNA was extracted from blood with the DNA IQ™ System kit (Promega), following the manufacturer's instructions to obtain a final concentration of 2 ng DNA/µL.

### PCR Amplification

Amplification was carried out on 10 µL of hair extraction product or 4 ng of blood DNA template in a 50-µL reaction containing each of the following primers at 0.2 mM: HV3-F314 (5'-CCG CTT CTG GCC ACA GCA CT-3') and HV3-R639 (5'-GGG TGA TGT GAG CCC GTC TA-3') (IDT, Coralville, IA), 5 µL of GoldSTAR® Buffer (Promega), and 0.4 µL of AmpliTaqGold® DNA polymerase (Applied Biosystems, Foster City, CA), made up to the volume with ddH<sub>2</sub>O. Thermal cycling was performed in a GeneAmp® PCR System 9700 (Perkin-Elmer, Wellesley, MA) starting at 95°C, for 11 min, then 1 min at 96°C, followed by 40 cycles at 94°C for 30 sec, 30 sec at 56°C, and 90 sec at 72°C, ending with 11 min at 60°C. Amplicon yield was evaluated by running 5 µL of all the PCR product reaction and comparing it with 2 µL of Low DNA Mass™ Ladder (Invitrogen, Carlsbad, CA) on 1% agarose gel, stained with ethidium bromide, to estimate the PCR product concentration. Negative controls were included in the extraction and PCR reactions, to detect any possible contamination. PCR products were cleaned and concentrated (to at least 20 ng/µL) with illustra GFX™ PCR DNA and gel band purification kit (GE).

### Sequencing

Cycle sequencing of both strands was carried out using BigDye® Terminator Cycle Sequencing Ready Reaction Kit v1.0 (Applied Biosystems), using the same primers as in the PCR reaction. Reaction products were ethanol precipitated and electrophoresed in an ABI PRISM® 377 Genetic Analyzer (Applied Biosystems).

### Data Analysis

Data analysis was carried out with BioEdit software (18). Only high quality sequences with little or no background signal were used. Sequences were aligned with the Revised Cambridge

Reference Sequence. Point heteroplasmy was only accepted if a secondary peak of more than about 10% of the primary peak was present and confirmed in the reverse sequencing reaction. The proportion of the secondary peak was estimated by measuring the heights of the primary and secondary peaks of mixed basecalls, as depicted in the electropherograms, with at least two independent PCR reactions, to rule out sequencing artifacts or mixtures.

### Cloning

PCR products containing CA heteroplasmy (electropherograms "out of phase") were cloned into the pGEM®-T Vector (Promega), to isolate each allele. Plasmid DNA was isolated by using the protocol described by Sambrook et al. (19).

## Results and Discussion

### Point Heteroplasmy

No difference among hair and blood from the same individuals was observed by direct PCR product sequencing and no point heteroplasmy was observed. Irwin et al. (20) and Zimmermann et al. (21) reported point heteroplasmy at positions 448R and 498Y, respectively, in blood samples. By contrast, no data has been published to date regarding heteroplasmy in the HV3 region in hair samples. Heteroplasmy had been shown to be frequent in hair samples in the HV1 and HV2 regions (3,22) and we expected that it would also occur, to some extent, in HV3. In general, the sites at which heteroplasmy is most commonly observed are correlated with reported control region mutational hotspots. According to a population study of 200 unrelated individuals from Germany, HV1 (np 16,024–16,365) showed 88 variable positions in a total length of 342 bp (26%) and HV2 (np 73–340) displayed 65 mutable sites in 268 bp (24%). HV3 (np 438–574) exhibited a slightly lower variability, with 25 polymorphic sites within 137 bp (18%), but contrasted clearly with the intervening regions, which, respectively, showed variability rates of only 7% (np 16,366–16,569, 1–72) and 3% (np 341–437) (9). In a Japanese study, in HV1, HV2, and HV3, 80, 37, and 14 polymorphic sites were identified, respectively, representing 23, 14, and 10% of the total size of each region, excluding all C-stretches (23). Our population showed 22 polymorphic positions in HV3 (data not shown), excluding all C-stretches, representing 16% of the total size of the region. Apparently, because HV3 is a less polymorphic region than HV1/HV2, it appears to be less prone to mutation (9,24). At some sites, however, the observation of heteroplasmy is not consistent with established mutation rate data, suggesting the action of other mechanisms, both selective and neutral, as occur at position 16,093. However, we did not find any vestige of point heteroplasmy in our samples.

### Length Heteroplasmy

Concerning length heteroplasmy, however, one individual showed heteroplasmy in the CA repeat (in both blood and hair samples) with "out of phase" electropherogram (Fig. 1). This result, regarding blood samples, is consistent with earlier studies. Szibor et al. (13) found 1.38% of CA heteroplasmy in a German population and Chung et al. (25) found 0.6% in a Korean population. No information about its frequency in hair samples has been published to date. As for the mechanism of heteroplasmy formation in the CA dinucleotide repeats, their propensity for insertion/deletion mutation of multiples of the repeating unit during replication

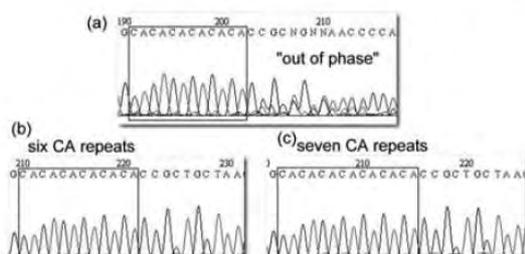


FIG. 1—(a) CA heteroplasmy found in hair and blood of the same individual with electropherogram "out of phase" after CA repeat. (b–c) Two different alleles that compound CA heteroplasmy separated after cloning.

could explain the phenomenon, as could the low fidelity of mitochondrial DNA polymerase. A high slippage rate of dinucleotide repeats can be observed in STRs, and the resultant stutter production is correlated with the length of repeat stretches consisting of uniform repeats (26). Length heteroplasmy was also observed at position 573 (poly-C stretch) in three individuals.

To observe which CA "alleles" were present in each tissue containing CA heteroplasmy, PCR products were cloned and re-sequenced. The presence of mixed alleles would suggest the possibility of an intra-individual difference in CA repeats between hair and blood samples. However, no variation among CA alleles was observed in hair and blood, even after cloning and re-sequencing. Recently, some authors reported unstable inheritance of HV3 heteroplasmy variants over a few generations. Szibor et al. (13) showed unstable inheritance at the CA repeat locus when analyzing two pedigrees. Brandstätter et al. (27) reported one case of CA divergence in 135 families (mother-child pairs). In general, this could be because of a marked reduction in the number of mtDNA molecules being transmitted from mother to offspring (the mitochondrial genetic bottleneck), to the partitioning of mtDNA into homoplasmic segregating units, or to the selection of a group of mtDNA molecules to populate the next generation. Cree et al. (28) recently suggested that the partitioning of mtDNA molecules into different cells before and after embryo implantation, followed by the segregation of replicating mtDNA between proliferating primordial germ cells, is responsible for the varying levels of heteroplasmy seen among the offspring of heteroplasmic females, at least in mice.

### Conclusion

Regarding forensic practice, we have seen that point heteroplasmy in HV3 is not as frequent as in the HV1 and HV2 regions. The contribution of HV3 analysis, however, could play a role in a forensic case. This region can increase discrimination and help in phylogenetic analysis for quality assurance of sequences. Also, the great genetic diversity in the CA repeat seems to increase its potential contribution to the forensic analysis. Nevertheless, because there have been few studies concerning variations of CA among different tissues within the same individual, its analysis should be used with caution.

### References

- Alonso A, Salas A, Albarrán C, Arroyo E, Castro A, Crespillo M, et al. Results of the 1999–2000 collaborative exercise and proficiency testing program on mitochondrial DNA of the GEP-ISFG: an inter-laboratory study of the observed variability in the heteroplasmy level of hair from the same donor. *Forensic Sci Int* 2002;125(1):1–7.

- Calloway CD, Reynolds RL, Herrin GL Jr, Anderson WW. The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age. *Am J Hum Genet* 2000;66(4):1384–97.
- Paneto GG, Martins JA, Longo LV, Pereira GA, Freschi A, Alvarenga VL, et al. Heteroplasmy in hair: differences among hair and blood are still a matter of debate. *Forensic Sci Int* 2007;3:117–22.
- Wang Q, Boles RG. Individual human hair mitochondrial DNA control region heteroplasmy proportions in mothers and children. *Mitochondrion* 2006;6 (1):37–42.
- Linch CA, Whiting MM, Holland MM. Human hair histogenesis for the mitochondrial DNA forensic scientist. *J Forensic Sci* 2001;46(4):844–53.
- Lee HR, Johnson K. Fidelity of the human mitochondrial DNA polymerase. *J Biol Chem* 2007;281 (47):36236–40.
- Lutz-Bonengel S, Sängler T, Parson W, Müller H, Ellwart JW, Follo M, et al. Single lymphocytes from two healthy individuals with mitochondrial point heteroplasmy are mainly homoplasmic. *Int J Legal Med* 2007;122(3):189–97.
- Salas A, Lareu MV, Carracedo A. Heteroplasmy in mtDNA and the weight of evidence in forensic mtDNA analysis: a case report. *Int J Legal Med* 2001;114(3):186–90.
- Lutz S, Wittig H, Weisser HJ, Heizmann J, Junge A, Dimo-Simonin N, et al. Is it possible to differentiate mtDNA by means of HVII in samples that cannot be distinguished by sequencing the HVI and HVII regions? *Forensic Sci Int* 2000;113(1-3):97–101.
- Lee HY, Yoo JE, Park MJ, Chung U, Shin KJ. Mitochondrial DNA control region sequences in Koreans: identification of useful variable sites and phylogenetic analysis for mtDNA data quality control. *Int J Legal Med* 2006;120(1):5–14.
- Szibor R, Michael M, Spitsyn VA, Plate I, Ginter EK, Krause D. Mitochondrial D-loop 3 (CA) repeat polymorphism: optimization of analysis and population data. *Electrophoresis* 1997;18(15):2857–60.
- Szibor R, Plate I, Schmitter H, Wittig H, Krause D. Forensic mass screening using mt DNA. *Int J Legal Med* 2006;120(6):372–6.
- Szibor R, Plate I, Heinrich M, Michael M, Schöning R, Wittig H, et al. Mitochondrial D-loop (CA)<sub>n</sub> repeat length heteroplasmy: frequency in a German population sample and inheritance studies in two pedigrees. *Int J Legal Med* 2007;121(3):207–13.
- Vanecek T, Trubac P, Vorel F, Miroslav S. Population genetics of 11 nuclear and 1 mitochondrial short tandem repeat loci in a population of South Bohemia, Czech Republic. *J Appl Biomed* 2005;3:129–31.
- Yan J, Tang H, Liu Y, Ma W, Liu Y, Zhang Q, et al. Allele frequencies of mitochondrial D-loop (CA)<sub>n</sub> repeat polymorphism in six Chinese ethnic groups. *Legal Med* 2007;9(6):330–1.
- Parson W, Bandelt H-J. Extended guidelines for mtDNA typing of population data in forensic science. *Forensic Sci Int: Genet* 2007;1(1):13–9.
- Jehaes E, Gilissen A, Cassiman JJ, Decorte R. Evaluation of a decontamination protocol for hair shafts before mtDNA sequencing. *Forensic Sci Int* 1998; 94(1-2):65–71.
- <http://www.mbio.ncsu.edu/BioEdit/BioEdit.html>.
- Sambrook J, Fritsch EF, Maniatis T. *Molecular cloning: a laboratory manual*, 2nd edn. New York, NY: Cold Spring Harbor Laboratory Press, 1989.
- Irwin JA, Saunier JL, Strouss KM, Diegoli TM, Sturk KA, O'Callaghan JE, et al. Mitochondrial control region sequences from a Vietnamese population sample. *Int J Legal Med* 2007;122(3):257–9.
- Zimmermann A, Brandstätter A, Dufner N, Niederwieser D, Spiroski M, Arsov T, et al. Mitochondrial DNA control region population data from Macedonia. *Forensic Sci Int: Genet* 2007;4:e4–9.
- Melton T, Dimick G, Higgins B, Lindstrom L, Nelson K. Forensic mitochondrial DNA analysis of 691 casework hairs. *J Forensic Sci* 2005; 50(1):73–80.
- Mabuchi T, Susukida R, Kido A, Oya M. Typing the 1.1 kb control region of human mitochondrial DNA in Japanese individuals. *J Forensic Sci* 2007;52 (2):355–63.
- Bini C, Ceccardi S, Luiselli D, Ferri G, Pelotti S, Colalongo C, et al. Different informativeness of the three hypervariable mitochondrial DNA regions in the population of Bologna (Italy). *Forensic Sci Int* 2003; 135:48–52.
- Chung U, Lee HY, Yoo JE, Park MJ, Shin KJ. Mitochondrial DNA CA dinucleotide repeats in Koreans: the presence of length heteroplasmy. *Int J Legal Med* 2005;119:50–3.
- Shinde D, Lai Y, Sun F, Arnheim N. Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis:

## 718 JOURNAL OF FORENSIC SCIENCES

- (CA/GT) $n$  and (A/T) $n$  microsatellites. *Nucleic Acids Res* 2003;31:974–80.
27. Brandstätter A, Niedstätter H, Parson W. Monitoring the inheritance of heteroplasmy by computer-assisted detection of mixed basecalls in the entire human mitochondrial DNA control region. *Int J Legal Med* 2004;118:47–54.
28. Cree LM, Samuels DC, de Sousa Lopes SC, Rajasimha HK, Wonnapijit P, Mann JR, et al. A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nat Genet* 2008;40:249–54.

Additional information and reprint requests:  
Greiciane Gaburro Paneto, M.Sc.  
Ph.D. Student  
School of Pharmaceutical Sciences  
São Paulo State University  
Rodovia Araraquara-Jaú, Km 01, Campus  
14801-902 Araraquara, SP  
Brazil  
E-mail: greiciane\_gp@hotmail.com