



**Universidade do Estado do Rio de Janeiro**  
Centro Biomédico  
Instituto de Biologia Roberto Alcântara Gomes

Suellen Silva Bernardo dos Santos

**Estudo da ancestralidade materna da população do Rio de Janeiro:  
análise do DNA mitocondrial**

Rio de Janeiro  
2012

Suellen Silva Bernardo dos Santos

**Estudo da ancestralidade materna da população do Rio de Janeiro: análise do  
DNA mitocondrial**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-graduação em Biociências, da Universidade do Estado do Rio de Janeiro.

Orientador: Prof. Dr. Elizeu Fagundes de Carvalho

Coorientadora: Dra. Dayse Aparecida da Silva

Rio de Janeiro

2012

CATALOGAÇÃO NA FONTE  
UERJ/REDE SIRIUS/BIBLIOTECA CB-A

S237 Santos, Suellen Silva Bernardo dos.  
Estudo da ancestralidade materna da população do Rio de Janeiro: análise do DNA mitocondrial / Suellen Silva Bernardo dos Santos. – 2012.  
81 f. : il.

Orientador: Elizeu Fagundes de Carvalho  
Coorientadora: Dayse Aparecida da Silva

Dissertação (Mestrado) – Universidade do Estado do Rio de Janeiro, Instituto de Biologia Roberto Alcântara Gomes. Pós-graduação em Biociências.

1. DNA Mitocondrial – Análise – Teses. 2. Haplótipos. 3. Genética da população humana – Teses. 4. População – Aspectos genéticos - Teses. 5. População – Brasil – Teses. I. Carvalho, Elizeu Fagundes de. II. Silva, Dayse Aparecida da. III. Universidade do Estado do Rio de Janeiro. Instituto de Biologia Roberto Alcântara Gomes. IV. Título.

CDU 575.113.1

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

---

Assinatura

---

Data

Suellen Silva Bernardo dos Santos

**Estudo da ancestralidade materna da população do Rio de Janeiro: análise do  
DNA mitocondrial**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-graduação em Biociências, da Universidade do Estado do Rio de Janeiro.

Aprovada em 12 de abril de 2012.

Banca Examinadora:

---

Prof. Dr. Marcelo Aguiar Costa Lima  
Instituto de Biologia Roberto Alcântara Gomes - UERJ

---

Prof. Dr. Paulo Marques Machado Brito  
Instituto de Biologia Roberto Alcântara Gomes - UERJ

---

Prof. Dr. Milton Ozório Moraes  
Instituto Oswaldo Cruz- IOC

Rio de Janeiro

2012

## DEDICATÓRIA

A Deus por tudo que me proporcionou na vida.

Ao meu avô, José Francisco (*in memoriam*) que foi o meu maior incentivador nessa e em outras grandes conquistas.

Aos meus pais, Rosângela e Ubiratan, e a minha avó Jandyra, os quais eu amo muito pelo apoio, educação e exemplos de vida e família.

Ao meu namorado, Gleidson, pelo amor, compreensão e acima de tudo pelo companheirismo.

## **AGRADECIMENTOS**

Ao Dr. Elizeu pela orientação, confiança e grande oportunidade de fazer parte de sua equipe.

A Dra. Dayse Aparecida da Silva, pelos conhecimentos e ensinamentos transmitidos necessários para o desenvolvimento deste trabalho.

À Rose pelos conhecimentos sobre DNA mitocondrial gentilmente transmitidos e apoio fundamentais para a elaboração deste trabalho.

À Fernanda pela amizade e companheirismo envolvendo a vida acadêmica, profissional e pessoal.

Aos amigos do LDD, Alexandre, Juliana, Andréa, Paty, Monique, João, Marcinha, Saulo, Raquel, Cris, Esther, Anna, Silvia e Celso pela amizade cultivada, dia pós dia, e pelo esclarecimento técnico e conceitual nos mais diversos assuntos abordados neste trabalho, obrigada por tudo!

A Agda, Márcia, Helô e Jaque... pela amizade e por todo auxílio na resolução das questões burocráticas!

De um modo geral, a todo corpo técnico e administrativo do LDD, pelo companheirismo de cada dia e auxílio nas mais diversas etapas deste trabalho.

Aos meus amigos, Camila, Delly, Rachel, Carol, Luciana, Thaís, Felipe, Allan, Patrícia, Elisa e Janine que mesmo de longe incentivaram e me ajudaram quando possível nessa e em outras conquistas!

Às amigas que a Uerj me apresentou, companheiras de graduação e mestrado, que dividiram comigo todas as angústias, vitórias, cada passo ensaiado e concretizado dentro e fora da universidade.

A todos que direta ou indiretamente colaboraram para este trabalho, o meu muito obrigada.

De tudo ficaram três coisas: a certeza de que estamos começando, a certeza de que é preciso continuar e a certeza de que podemos ser interrompidos antes de terminar. Portanto, devemos fazer da interrupção um caminho novo, da queda um passo de dança, do medo uma escada, do sonho uma ponte, da procura um encontro.

*Fernando Sabino*

## RESUMO

SANTOS, Suellen Silva Bernardo dos. *Estudo da ancestralidade materna da população do Rio de Janeiro: análise do DNA mitocondrial*. 2012. 81f. Dissertação (Mestrado em Biociências) - Instituto de Biologia Roberto Alcântara Gomes, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2012.

As porções uniparentais do genoma humano, representadas pelo cromossomo Y e pelo DNA mitocondrial (DNAmt), contêm informação genética relacionada às heranças patrilinear e matrilinear, respectivamente. Além da aplicabilidade em genética médica e forense, o DNAmt tem sido utilizado como um importante marcador molecular em estudos sobre evolução para traçar inferências filogenéticas e filogeográficas sobre as populações humanas. A análise de linhagens de DNAmt presentes em diferentes populações mundiais levou à identificação de haplogrupos reunindo diversos haplótipos específicos dos grandes grupos étnicos: africanos, europeus, asiáticos e nativos americanos. A população brasileira é conhecida como uma das mais heterogêneas do mundo, resultado do processo de colonização do país, abrangendo mais de cinco séculos de miscigenação entre povos de diferentes continentes. Este trabalho teve como objetivo estimar a partir da análise do DNA mitocondrial as proporções ancestrais africanas, européias e ameríndias na população do Rio de Janeiro. Para isso foram sequenciadas as regiões hipervariáveis HVI e HVII do DNAmt de 109 indivíduos não relacionados geneticamente residentes no Rio de Janeiro. Os haplogrupos foram classificados de acordo com o conjunto de polimorfismos dos haplótipos individuais. Programas estatísticos foram utilizados para a determinação de parâmetros de diversidade genética e comparações populacionais. A diversidade haplotípica foi estimada em 0,9988. Nossos resultados demonstraram na população do Rio de Janeiro percentuais de cerca de 60%, 25% e 15% de ancestralidades maternas africana, ameríndia e européia, respectivamente. Através da análise de distâncias genéticas, evidenciou-se que a população do Rio de Janeiro está mais próxima das populações brasileiras dos estados de São Paulo e Alagoas. Como descrito nos registros históricos, algumas regiões do país tiveram processos de colonização muito específicos que se refletem nas proporções ancestrais maternas e paternas observadas. Em relação ao DNAmt, não se verificou diferença genética significativa entre as populações do Rio de Janeiro e a de Angola, uma população africana. Os resultados obtidos estão em estreita concordância com os registros históricos e outros estudos genéticos acerca da formação da população brasileira.

Palavras-chave: DNA mitocondrial. População brasileira. Ancestralidade. Haplogrupos. Rio de Janeiro.



## ABSTRACT

The uniparental portions of the human genome, represented by the Y chromosome and mitochondrial DNA (mtDNA), contain genetic information related to patrilineal and matrilineal inheritance, respectively. Besides the application in medical genetics and forensic, the mtDNA has been used as an important molecular marker in studies of evolution to trace phylogenetic and phylogeographic inferences on human populations. The analysis of mtDNA lineages present in different populations worldwide led to the identification of haplogroups combining several specific haplotypes of the major ethnic groups: Africans, Europeans, Asians and Native Americans. The Brazilian population is known as one of the most heterogeneous in the world, a result of the colonization process of the country, covering more than five centuries of miscegenation between peoples of different continents. This study aimed to estimate from the analysis of mitochondrial DNA, the ancestral proportion of African, European and Amerindian in the Rio de Janeiro population. For that reason, the hypervariable regions HVI and HVII of mtDNA from 109 genetically unrelated individuals residing in Rio de Janeiro were sequenced. The haplogroups were classified according to the set of polymorphism of the individual haplotypes. Statistical programs were used for determination of parameters of genetic diversity and population comparisons. The haplotype diversity was estimated to be 0.9988. Our results showed in the population of Rio de Janeiro percentage of about 60%, 25% and 15% of maternal African ancestry, Amerindian and European, respectively. Through the analysis of genetic distances, it was showed that the population of Rio de Janeiro is closer to Brazilian populations of the states of Sao Paulo and Alagoas. As described in historical records, some regions had very specific processes of colonization that are reflected in the maternal and paternal ancestors proportions observed. Regarding mtDNA, there was no significant genetic difference between the populations of Rio de Janeiro and Angola, an African population. The current results are in close agreement with the historical records and other genetic studies on the formation of the Brazilian population.

Keywords: Mitochondrial DNA. Brazilian population. Ancestry. Haplogroups. Rio de Janeiro.

## LISTA DE FIGURAS

Figura 1 –	Esquema representativo do genoma humano (adaptado de Butler, 2009).....	15
Figura 2 –	Genoma mitocondrial humano (adaptado Taylor & Turnbull, 2005).....	18
Figura 3 –	Árvore filogenética simplificada do DNAMt (adaptado de van Oven & Kayser, 2009).....	25
Figura 4 –	Migrações humanas a partir dos estudos do DNA mitocondrial.....	26
Figura 5 –	Produtos de amplificação por PCR das regiões HVI e HVII do DNA mitocondrial.....	38
Figura 6 –	Eletroferograma representativo do sequenciamento da região HVI do DNA mitocondrial (orientação 5´-3´) da amostra RJM63.....	39
Figura 7 –	Eletroferograma representativo do sequenciamento da região HVII do DNA mitocondrial (orientação 5´-3´) da amostra RJM63.....	40
Figura 8 –	Proporção de distintas ancestralidades maternas do DNA mitocondrial na amostra da população do Rio de Janeiro.....	46
Figura 9 –	Eletroferograma da reação de sequenciamento da região HVI do DNA mitocondrial da amostras RJM17. Sequências senso e reverso complemento - orientação 3´-5´ – alinhadas no SeqScape®Software - versão 2.5 (Applied Biosystems).....	50
Figura 10–	Projeção em duas dimensões do MDS das distâncias genéticas ( $F_{ST}$ ) entre as 10 populações analisadas (valor de stress=0,11235).	53

## LISTA DE TABELAS

Tabela 1 – Sequências de <i>primers</i> utilizados para amplificação das regiões HVI e HVII do DNA mitocondrial.....	32
Tabela 2 – Condições termocíclicas utilizadas para amplificação das regiões HVI e HVII do DNA mitocondrial.....	32
Tabela 3 – Haplótipos, frequências haplotípicas, polimorfismos das regiões HVI e HVII do DNA mitocondrial e haplogrupos verificados na amostra da população do Rio de Janeiro.....	42
Tabela 4 – Parâmetros de diversidade a partir da análise das regiões HVI e HVII do DNA mitocondrial em uma amostra populacional do Rio de Janeiro.....	45
Tabela 5 – Haplogrupos africanos e respectivas frequências na população do Rio de Janeiro.....	47
Tabela 6 – Haplogrupos nativos americanos e respectivas frequências na população do Rio de Janeiro.....	48
Tabela 7 – Haplogrupos europeus e respectivas frequências na população do Rio de Janeiro.....	49
Tabela 8 – Matriz das distâncias genéticas ( $F_{ST}$ ) para frequências de haplogrupos de DNAm <sub>t</sub> entre pares de populações. Abaixo da diagonal são mostrados os valores de P e acima da diagonal os valores de $F_{ST}$ .....	52
Tabela 9 – Proporção de ancestralidade materna ameríndia, européia e africana em diferentes estados do Brasil estimados a partir do DNA mitocondrial.....	61
Tabela 10 – Naturalidade dos indivíduos da população do Rio de Janeiro selecionados para o presente estudo.....	66
Tabela 11 – Proporção das distintas ancestralidades maternas em indivíduos residentes e naturais do Rio de Janeiro.....	67

## LISTA DE ABREVIATURAS E SIGLAS

AIMs-	<i>Ancestry Informative Markers</i>
AL-	Alagoas
AN-	Angola
ATP-	Trifosfato de Adenosina
BSA-	<i>Bovine Serum Albumin</i>
CH-	<i>China</i>
COI-	Citocromo oxidase subunidade I
COII-	Citocromo oxidase subunidade II
Cytb-	Citocromo b
DNA-	<i>DeoxyriboNucleic Acid</i> - Ácido desoxirribonucléico
DNAm-t-	DNA mitocondrial
EDTA -	Ácido etilenodiaminotetracético
EMPOP-	<i>European MtDNA Population Database</i>
FBI-	<i>Federal Bureau of Investigation</i>
F <sub>ST</sub> -	Índice de fixação usado como medida de diferenciação genética entre subpopulações
HVI-	Hipervariável I
HVII-	Hipervariável II
IUPAC-	<i>International Union of Pure and Applied Chemistry</i>
LDD-	Laboratório de Diagnósticos por DNA

MDS-	<i>Multidimensional scaling</i>
NLB -	Tampão de lise de membrana nuclear
PA-	Pará
PCR -	<i>Polymerase Chain Reaction</i> - Reação em cadeia da polimerase
POP -	<i>Performance Optimizer Polymer</i>
PT-	Portugal
rCRS-	<i>Revised Cambridge Reference Sequence</i>
RNA-	<i>Ribonucleic Acid</i>
rRNA-	<i>Ribosomal Ribonucleic Acid</i>
RS-	Rio Grande do Sul
SC-	Santa Catarina
SNP-	<i>Single Nucleotide Polymorphism</i>
SP-	São Paulo
STR -	<i>Short Tandem Repeat</i>
Taq -	<i>Thermus aquaticus</i>
TE -	Tampão Tris-EDTA
TLH -	Tampão de lise de hemácias
Tris -	Tris-Hidroxametil-Aminometano
tRNA-	<i>Transfer Ribonucleic Acid</i>
UERJ-	Universidade do Estado do Rio de Janeiro

## LISTA DE SÍMBOLOS

~ -	Aproximadamente
% -	porcentagem
$\geq$ -	maior ou igual
$\mu\text{L}$ -	microlitro
$\mu\text{M}$ -	micromolar
$^{\circ}\text{C}$ -	graus Celsius
A-	Adenine
C -	Citosina
g -	Gravidade
G -	Guanina
M -	molar
Mg -	miligrama
mL -	mililitro
ng -	nanograma
nm-	Nanômetro
p/v-	peso por volume
T -	timina
U -	unidade de enzima
V-	Volt

## SUMÁRIO

	<b>INTRODUÇÃO</b> .....	15
1	<b>OBJETIVOS</b> .....	29
1.1	<b>Objetivo Geral</b> .....	29
1.2	<b>Objetivos Específicos</b> .....	29
2	<b>MATERIAL E MÉTODOS</b> .....	30
2.1	<b>Amostras biológicas</b> .....	30
2.2	<b>Extração de DNA</b> .....	30
2.2.1	<u>Extração por Salting-out</u> .....	30
2.2.2	<u>Extração por Chelex</u> .....	31
2.3	<b>Determinação da concentração de DNA</b> .....	31
2.4	<b>Reação de polimerização em cadeia (PCR) para Regiões Hipervariáveis do DNA mitocondrial</b> .....	31
2.5	<b>Sequenciamento das Regiões HVI e HVII do DNA mitocondrial</b> .....	33
2.6	<b>Análise das sequências e determinação de haplogrupos</b> .....	34
2.7	<b>Análises Estatísticas</b> .....	34
2.8	<b>Cálculo de distâncias genéticas</b> .....	35
2.9	<b>Lista de soluções e tampões utilizados</b> .....	36
3	<b>RESULTADOS</b> .....	37
3.1	<b>Análise dos produtos de PCR e Sequenciamento</b> .....	37
3.1.1	<u>Análise dos produtos de amplificação em gel de agarose 2%</u> .....	37
3.1.2	<u>Sequenciamento das regiões HVI e HVII do DNA mitocondrial</u> ....	38
3.2	<b>Frequências haplotípicas e determinação dos haplogrupos</b> .....	41
3.2.1	<u>Análise de parâmetros de diversidade</u> .....	45
3.3	<b>Proporção dos haplogrupos na população do Rio de Janeiro</b> .....	45
3.3.1	<u>Haplogrupos africanos</u> .....	46
3.3.2	<u>Haplogrupos nativos americanos</u> .....	48
3.3.3	<u>Haplogrupos europeus</u> .....	49
3.4	<b>Proporção de heteroplasma de sequência na população do Rio de Janeiro</b> .....	50
3.5	<b>Estimativas de distâncias genéticas</b> .....	51

3.5.1	<u>Distâncias genéticas em relação a outras populações brasileiras.....</u>	51
3.5.2	<u>Distâncias genéticas em relação a populações de grupos ancestrais..</u>	51
4	<b>DISCUSSÃO.....</b>	55
4.1	<b>Diversidade genética e haplotípica .....</b>	56
4.2	<b>Frequências de Heteroplasmias.....</b>	59
4.3	<b>Frequência e origem geográfica dos haplogrupos verificados na população do Rio de Janeiro.....</b>	60
4.4	<b>Distâncias genéticas.....</b>	68
5	<b>CONCLUSÃO.....</b>	71
	<b>REFERÊNCIAS.....</b>	72



## INTRODUÇÃO

### GENOMA HUMANO

O genoma humano corresponde ao total de informação genética contida nas células humanas, armazenado na molécula de DNA (ácido desoxirribonucléico). Divide-se em um complexo genoma nuclear, que contém quase que a integralidade da informação genética, e um pequeno genoma mitocondrial (Strachan & Read, 2004). É constituído por aproximadamente 3,2 bilhões de pares de bases organizados em 23 pares de cromossomos, sendo 22 pares autossômicos e um par de cromossomos sexuais X e Y (Figura 1) (Butler, 2009).

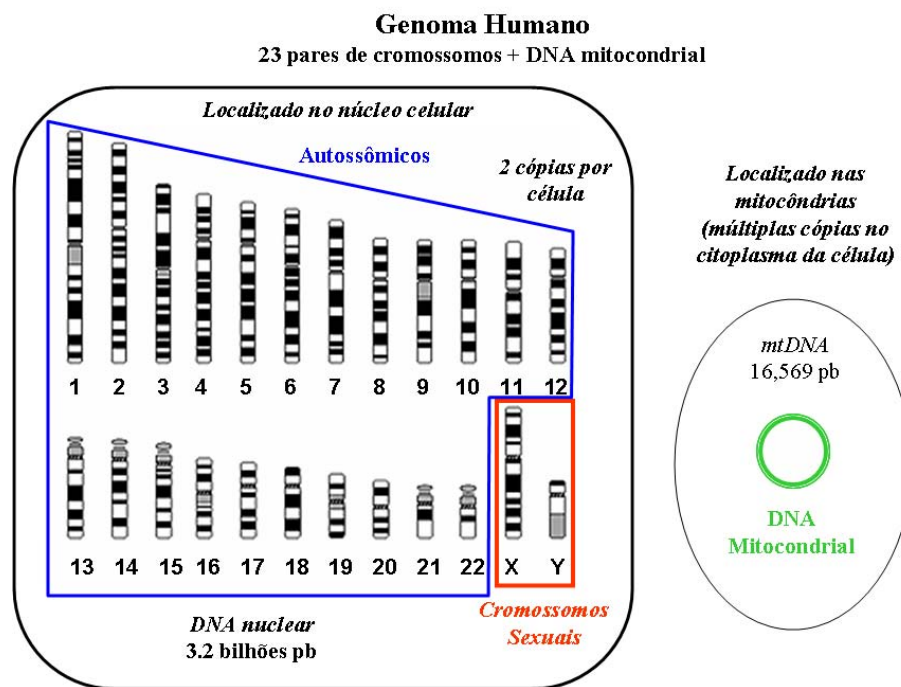


Figura 1 - Esquema representativo do genoma humano (adaptado de Butler, 2009).

A molécula de DNA está presente em todas as células nucleadas do nosso organismo. Cada célula, exceto as da linhagem germinativa, possui dois pares de cromossomos, um herdado do pai e outro da mãe. Os cromossomos de um mesmo par são ditos homólogos. O único par de cromossomos não homólogos é dos cromossomos sexuais em indivíduos do sexo masculino, onde o cromossomo X foi herdado da mãe e o cromossomo Y do pai, não havendo recombinação entre eles, com exceção das regiões pseudoautosômicas (Jobling & Tyler-Smith, 2003).

Desde a descoberta e definição da estrutura do DNA (Franklin & Gosling, 1953, Watson & Crick, 1953), diversas metodologias e ferramentas biotecnológicas foram desenvolvidas com o objetivo de estudar a molécula de forma mais minuciosa. Na década de 1950, foi identificada e purificada a enzima DNA polimerase (Lehman *et al.*, 1958). Em 1970, foi realizado o primeiro isolamento de uma enzima de restrição (Smith, 1970). Nesta mesma década, Sanger desenvolveu uma técnica de sequenciamento do DNA (Sanger *et al.*, 1977).

Na década de 80, outra técnica revolucionou a biologia molecular, a reação em cadeia da polimerase (PCR), um processo *in vitro* de amplificação do DNA (Mullis *et al.*, 1986, White *et al.*, 1989), que tornou possível a redução drástica da quantidade de DNA e o tempo necessários para as análises. Além disso, a PCR possibilitou uma utilização generalizada de análise de DNA em casos forenses onde as quantidades da molécula eram diminutas (Decorte *et al.*, 1990, Westwood & Werrett, 1990).

O projeto genoma, através do sequenciamento do DNA humano, revelou que o nosso material genético contém aproximadamente três bilhões de pares de bases, dos quais 1,1% corresponde a éxons, 24% a íntrons e 75% a sequências intergênicas (Venter *et al.*, 2001). Segundo dados do consórcio público responsável pelo projeto genoma, a espécie humana possui aproximadamente 32 mil genes e deste total, apenas 1,5% codifica proteínas (Lander *et al.*, 2001).

O projeto genoma possibilitou um maior conhecimento sobre os novos polimorfismos descritos, como os microsátélites ou *short tandem repeat* (STRs) e os *single nucleotide polymorphisms* (SNPs). A menor distribuição de tamanho dos alelos dos STRs em conjunto com a técnica de PCR possibilitou a análise até mesmo de DNA degradado, o que não era possível com as metodologias anteriores (Holland *et al.*, 1993). Os sistemas automáticos de detecção a laser tornaram

possível a análise de múltiplos *loci* STRs amplificados em uma mesma reação de PCR (Kimpton *et al.*, 1993).

Os polimorfismos de nucleotídeo único (SNPs) ocorrem a cada 100-300 bases no genoma humano, sendo desta maneira, os polimorfismos mais frequentes (Lander *et al.*, 2001). No banco de dados dbSNP (versão 32), já foram validados mais de 17 milhões de SNPs (<http://www.ncbi.nlm.nih.gov/projects/SNP/>). Tais polimorfismos são resultantes de uma alteração, deleção ou inserção de um nucleotídeo em uma determinada posição no genoma humano. A análise de um SNP pode ser feita através do sequenciamento da região de interesse ou pela amplificação de um pequeno segmento em torno da posição SNP seguido de um minisequenciamento (Sobrinho *et al.*, 2005).

Os diversos polimorfismos de DNA são amplamente utilizados em genética forense, podendo ocorrer em cromossomos sexuais, autossômicos e no DNA mitocondrial (DNAMt). O conhecimento da estrutura genética da população possibilita uma aplicação direta na construção de bancos de dados com interesses forenses. As regiões polimórficas do DNA podem ser utilizadas em identificação de indivíduos, estudos de filogenia, genética de populações e mapeamento de genes. Atualmente as sequências STR autossômicas são muito utilizadas na identificação humana, seja na investigação criminal ou de vínculo genético ou na tentativa de localizar pessoas desaparecidas (Carracedo & Sanchez-Diz, 2005).

Assim como no DNA nuclear, no DNAMt ocorrem muitos SNPs. A análise dos polimorfismos verificados na região controle do DNAMt é uma eficiente metodologia para genotipagem de materiais degradados, ossos, dentes e fios de cabelo (Parsons & Coble, 2001, von Wurmb-Schwark *et al.*, 2003). Além desta aplicação, devido a características específicas desta molécula, o DNA mitocondrial é muito empregado em estudos de evolução e ancestralidade (Underhill & Kivisild, 2007).

## DNA MITOCONDRIAL

As mitocôndrias apresentam várias cópias de seu material genético, o DNA mitocondrial (DNAmt). Dependendo do tipo celular e função biológica, uma única célula pode conter de centenas a milhares de mitocôndrias. As principais funções dessa organela no nosso organismo são: produção de adenosina trifosfato (ATP), mediação da morte celular programada (apoptose) e produção de calor (Finsterer, 2004).

Anderson e colaboradores (1981) descreveram a primeira sequência do genoma mitocondrial humano. O DNA mitocondrial (Figura 2) possui 16.569 pares de bases e é composto por uma dupla fita circular, com uma distribuição assimétrica de guaninas e citosinas, o que gera uma cadeia pesada (H) rica em guaninas e uma cadeia leve (L) rica em citosina, no qual estão contidos 37 genes envolvidos na produção de energia e no seu armazenamento na forma de ATP. Deste total, dois codificam RNAs ribossômicos (12S e 16S), 22 codificam RNAs transportadores e 13 codificam subunidades do sistema de fosforilação oxidativa (Anderson *et al.*, 1981, Taylor & Turnbull, 2005).

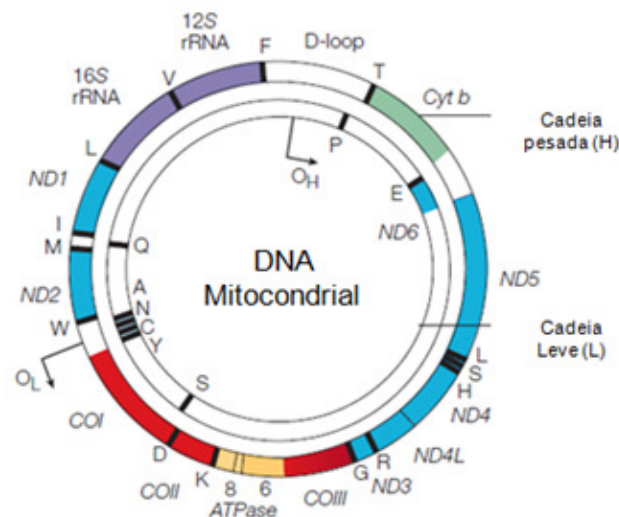


Figura 2 - Genoma mitocondrial humano (adaptado de Taylor & Turnbull, 2005). Os genes que codificam as subunidades do complexo I são mostrados em azul; citocromo c oxidase (COI-COIII) é mostrado em vermelho; citocromo b do complexo III é mostrado em verde e as subunidades da ATP sintase são mostradas em amarelo. Os dois RNAs ribossomais (rRNAs; 12S e 16S) são mostrados em roxo e 22tRNAs são indicado por linhas. Também é destacada a região controle (D-loop).

O genoma mitocondrial humano possui uma região codificante e uma não codificante. A região não codificante, também conhecida como região controle ou alça D, é o local em que se inicia a replicação da cadeia H. Esta região contém sequências que são vitais para o início da replicação e transcrição do DNAm (Taylor & Turnbull, 2005).

O DNAm pode ser utilizado para a identificação taxonômica, genética populacional e em análises forenses por apresentar vantagens em relação ao DNA nuclear. A principal vantagem desta molécula para uso forense é a maior abundância de cópias em relação ao DNA nuclear. Cada célula possui centenas de mitocôndrias e cada mitocôndria possui dezenas de cópias de seu cromossomo circular. Essas características possibilitam que em uma pequena amostra de tecido biológico possa ser encontrada uma quantidade razoável de DNAm (Scheffler, 2002, Budowle *et al.*, 2003).

O DNAm não possui íntrons, não sofre recombinação e é haplóide, o que o torna interessante para a identificação de espécies (Saccone *et al.*, 1999). Além disso, possui certos genes, como o citocromo b (cytb) e o citocromo oxidase I (COI) que possuem regiões bastante conservadas, o que possibilita uma análise a partir de uma metodologia simples em diversos organismos (Folmer *et al.*, 1994).

A molécula do DNAm possui regiões bastante variáveis que podem ser utilizadas para diferenciar indivíduos dentro de uma espécie. Com ocorrência, sobretudo na região controle, tais variações são decorrentes de mutações (Brown *et al.*, 1979, Brown *et al.*, 1982, Pakendorf & Stoneking, 2005). Em mamíferos, o DNAm possui uma taxa de evolução ~10 vezes mais rápida do que aquela encontrada no DNA nuclear (Brown *et al.*, 1979, Brown *et al.*, 1982, Pakendorf & Stoneking, 2005).

A alta taxa de mutação da molécula de DNAm é explicada por uma combinação de fatores. O DNAm está localizado na organela em que ocorre a cadeia respiratória, que é uma fonte endógena de espécies reativas de oxigênio, proporcionando um ambiente favorável a mutações. Associadas a este fator, a ausência de proteínas histonas e uma capacidade limitada de reparo justificam a alta susceptibilidade a danos em relação ao DNA nuclear (Taylor & Turnbull, 2005).

A variabilidade apresentada por uma molécula de DNA é resultante de um balanço entre as mutações que surgem e as restrições, estruturais e funcionais, que elas ocasionam. A hipótese do relógio molecular propõe que os genes e seus

produtos evoluem sob uma taxa constante ao longo do tempo e das linhagens. A principal vantagem que a teoria ofereceu foi a possibilidade de inferir o tempo de divergência entre duas linhagens a partir do número de substituições entre as suas sequências de DNA, independentemente de informações fósseis (Pauling, 1964, Hillis, 1996).

Por essa proposta, se a taxa de substituição for constante entre grupos taxonomicamente próximos, é possível estimar o tempo de divergência dentre os mesmos aplicando um ponto de calibração específico para o grupo. Esta teoria foi difundida rapidamente na comunidade científica e diversos relógios moleculares universais foram propostos. Estes foram empregados em diferentes trabalhos que buscavam elucidar os fundamentos que norteavam as questões evolutivas e biogeográficas (Doolittle *et al.*, 1996, Hedges *et al.*, 1996, Holder & Lewis, 2003).

Através de marcadores moleculares que apresentem alguma variabilidade entre os indivíduos podemos analisar e entender os mais diversos processos evolutivos. O DNAm possui excelentes marcadores para a inferência de filogenias maternas, principalmente entre espécies próximas, devido às peculiaridades de sua herança uniparental clonal e ausência de recombinação. As árvores de evolução ilustram o conceito de descendência com modificação e exibem os processos de deriva e migração que influenciam a estrutura da população. Filogenias do DNAm registraram a acumulação da diversidade genética e são indicadoras únicas da estrutura da população existente (Underhill & Kivisild, 2007).

## **ESPECIFICIDADES DO DNA MITOCONDRIAL**

O DNAm está presente em um grande número de cópias por célula. Em uma célula somática, enquanto ocorrem duas cópias do genoma nuclear podem ser encontradas de 100 a 1000 cópias do genoma mitocondrial (Robin & Wong, 1988). Esta propriedade associada ao seu diminuto tamanho relativo torna a molécula de escolha em investigações forenses, análises de amostras antigas e outras que apresentem DNA degradado ou em pouca quantidade (Pakendorf & Stoneking, 2005).

A existência de múltiplas cópias do DNA mitocondrial por célula pode gerar um fenômeno conhecido como heteroplasmia, caracterizado pela coexistência de

dois ou mais genomas mitocondriais em uma mesma célula ou tecido de um indivíduo (Tuppen *et al.*, 2010). A heteroplasmia pode apresentar-se de duas formas: heteroplasmia de sequência e heteroplasmia de comprimento. A heteroplasmia de sequência consiste na ocorrência de duas populações de DNA mitocondrial diferenciadas em uma determinada posição de nucleotídeos (Stewart *et al.*, 2001). O caso mais famoso desse tipo de heteroplasmia foi verificado quando da análise do DNAmT dos restos mortais da família Romanov (Gill *et al.*, 1994, Ivanov *et al.*, 1996).

A heteroplasmia de comprimento quando comparada a heteroplasmia de sequência ocorre mais comumente e possui menos especificidade populacional. É caracterizada pela deleção ou inserção de citosinas em regiões contendo repetições mononucleotídicas. Na região controle do DNAmT existem quatro regiões que contêm repetições de citosinas consecutivas, conhecidas como *C-tracts*, propensas a heteroplasmia de comprimento do DNAmT. Um indivíduo pode apresentar em suas células distintos comprimentos de *C-tracts* (Bendall & Sykes, 1995). Diversas pesquisas buscam elucidar os mecanismos celulares e circunstâncias em que ocorrem este tipo de heteroplasmia (Irwin *et al.*, 2009, Lacan *et al.*, 2009, Forster *et al.*, 2010).

Existem algumas dificuldades na interpretação das heteroplasmias, são elas: o mecanismo de surgimento desses eventos não é claro; a taxa de mutação em que ocorrem é desconhecida; a notação de heteroplasmia pode ser complexa e até mesmo a sua presença é difícil de detectar se o pesquisador não recorrer, por exemplo, à clonagem de bactérias (Lutz-Bonengel *et al.*, 2004, Forster *et al.*, 2010). A prática usual, tanto em ciência forense, quanto em estudos evolutivos humanos tem sido a de desconsiderar as heteroplasmias de comprimento nas análises de DNA mitocondrial. Já as heteroplasmias pontuais são raras e são importantes em interpretações forenses (Stewart *et al.*, 2001).

A herança do DNAmT se dá exclusivamente por via materna e com ausência de recombinação. A transmissão uniparental do DNA mitocondrial se apresenta como uma ferramenta importante em estudos de ancestralidade em uma população, sendo possível traçar a origem materna de indivíduos (Pakendorf & Stoneking, 2005).

Diversos estudos buscaram elucidar o porquê de as mitocôndrias paternas não serem herdadas. É bem descrito que, no momento da fecundação, as mitocôndrias paternas são seletivamente destruídas no oócito. As mitocôndrias

presentes nos espermatozoides são marcadas para destruição por ubiquitinação no momento da fecundação (Robin & Wong, 1988). Algumas discussões na comunidade científica sugeriram que este mecanismo poderia falhar possibilitando uma herança de mitocôndrias paternas, justificadas pelo estudo da herança em algumas doenças (Schwartz & Vissing, 2003). Mesmo diante de tais discussões a herança materna exclusiva do DNAm<sub>t</sub> observada em seres humanos ainda pode ser considerada como regra (Schwartz & Vissing, 2003, Pakendorf & Stoneking, 2005).

Outras teorias também apontaram para a possibilidade de recombinação do DNAm<sub>t</sub>, seriam elas: recombinação com DNAm<sub>t</sub> paterno; recombinação com pseudogenes do DNAm<sub>t</sub> presentes no DNA nuclear e recombinação entre diferentes moléculas de DNAm<sub>t</sub> de um mesmo indivíduo que apresente heteroplasmia. Cada uma destas vias exigiria que as mitocôndrias possuísem enzimas necessárias para recombinação homóloga, ou que o DNA pudesse ser tomado a partir do citoplasma pela mitocôndria e não existem boas evidências dessa recombinação *in vivo* em mamíferos (Eyre-Walker, 2000).

## **REGIÕES HIPERVARIÁVEIS E OS HAPLOGRUPOS DO DNA MITOCONDRIAL**

A Região controle do DNAm<sub>t</sub> possui duas regiões que são os focos de estudos populacionais devido às variações de sequências dentro desses segmentos (Bonatto & Salzano, 1997, Torroni *et al.*, 2006): a Região Hipervariável I (HVI) que compreende as posições 16.024 à 16.383; e a Região Hipervariável II (HVII) que estende-se da posição 57 até 372 (Stoneking, 2000). Uma terceira região chamada de segmento hipervariável III (HVIII) também está presente na região controle e está situada entre as posições 438 e 574, no entanto é menos polimórfica que HVI e HVII (Lutz *et al.*, 1997).

Por convenção, as posições de nucleotídeos do genoma mitocondrial humano são numeradas de 1-16569 de acordo com a Sequência de Referência Cambridge revisada (rCRS) (Anderson *et al.*, 1981, Anderson *et al.*, 1999, Andrews *et al.*, 1999). Os polimorfismos de base única (transições, transversões, inserções ou deleções) são determinados através das diferenças encontradas entre as sequências geradas e a rCRS. Cada conjunto de polimorfismos verificados em um indivíduo é chamado



de haplótipo. Por sua vez, ocorrem haplogrupos definidos por grupos de sequências (haplótipos) que se relacionam a uma mesma origem ou região geográfica (Pakendorf & Stoneking, 2005).

A maioria dos laboratórios que utiliza a tipagem do DNAMt se baseia nos polimorfismos presentes na sequência de nucleotídeos das regiões HVI e HVII. 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 DNAMt (Holland & Parsons, 1999). Haplótipos pouco frequentes indicam maior probabilidade de amostra questionada idêntica ter a mesma origem da amostra referência. Haplótipos muito frequentes indicam menor probabilidade de coincidência, já que muitos indivíduos apresentam o mesmo conjunto de polimorfismos naquela população. Em 1999, foi criado o EMPOP, um banco de dados *on line* que visa a coleta, controle de qualidade e divulgação de haplótipos de DNAMt mundiais. Em dezembro de 2011, existiam mais de 16 mil sequências em seu acervo digital (<http://empop.org/>). O FBI reúne as informações de DNAMt em seu banco de DNA conhecido como CODIS<sup>mt</sup> (Monson *et al.*, 2002). Em 2001, ano de sua criação, o CODIS<sup>mt</sup> continha 4839 perfis de DNAMt representativos de 14 populações. O banco de dados da Yonsei University mtDNAManager (Lee *et al.*, 2009) continha, até Dezembro de 2010, 9294 sequências da região controle do DNAMt divididas em 5 grupos populacionais.

A taxa de mutação não é constante em toda a molécula de DNAMt. Na região controle é cerca de 10 vezes superior à observada na região codificante (Pakendorf & Stoneking, 2005). Tal característica pode ser explicada pela seleção natural que ocorre na região codificante, eliminando as mutações funcionalmente deletérias. Na região controle, que embora seja importante para a replicação do DNA mitocondrial, não existe a codificação de proteínas, sendo assim esta região está mais propensa a acumular mutações. Além disso, também existem diferenças entre taxas mutacionais na região controle. Em algumas posições desta região, nos *hot spots*, as mutações ocorrem de forma mais frequente, como por exemplo, nas posições 146, 150, 152, 195, 16189, 16311 e 16362 enquanto que outros sítios são bastante estáveis (477, 493, 16108, 16219) (Meyer *et al.*, 1999).

Questões amplas como a origem do homem moderno e os grandes trajetos da sua dispersão pelo planeta, bem como as grandes migrações intercontinentais que

ocorreram ao longo de milhares de anos e que moldaram a atual diversidade dos povos foram tratadas em diversos estudos por meio de análises de polimorfismos de DNA.

Os marcadores uniparentais do genoma humano, aqueles presentes no cromossomo Y e no DNA mitocondrial, têm importantes aplicações em inúmeros campos como a antropologia evolutiva e história das populações (Underhill & Kivisild, 2007). Como já citado, o DNA mitocondrial possui características especiais que incluem a herança matrilinear, a ausência de recombinação, número elevado de cópias e uma taxa evolutiva mais elevada que a do DNA nuclear.

Os estudos acerca da evolução humana indicam que todos os tipos de DNAmt poderiam ser agrupados em uma única filogenia e que as variantes mais polimórficas, situadas na raiz dessa filogenia, estariam presentes nas linhagens mitocondriais africanas. Deste modo, chegou-se à conclusão de que todos os DNA mitocondriais derivam de um ancestral matrilinear comum que viveu a cerca de 200.000 anos na África (Mishmar *et al.*, 2003).

Os conjuntos de variações nas sequências do DNA mitocondrial encontrados em um haplótipo evoluíram como resultado da acumulação de mutações nas matrinhagens herdadas ao longo das gerações. A partir desta inferência foi possível construir uma árvore filogenética que reflete as relações entre as diferentes filogenias em escala global (Figura 3). Um grupo de haplótipos relacionados forma um clado genético definido, denominado haplogrupo. Em 2008, a fim de possibilitar uma melhor compreensão das variações do DNA mitocondrial, foi gerada uma árvore atualizada de variação do DNA mitocondrial considerando todos os dados de sequências disponíveis no NCBI GenBank ([www.ncbi.nlm.nih.gov/GenBank](http://www.ncbi.nlm.nih.gov/GenBank)) (van Oven & Kayser, 2009).

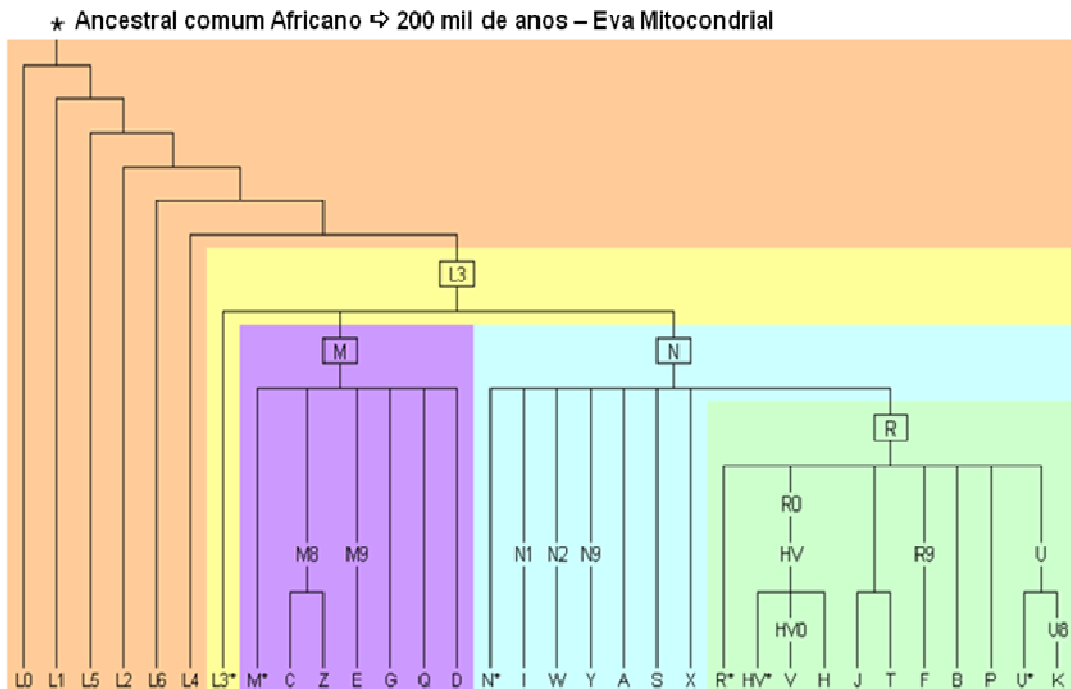


Figura 3 - Árvore filogenética simplificada do DNAMt (adaptado de van Oven & Kayser, 2009).

Cada haplogrupo está relacionado a uma origem geográfica. Os haplogrupos A, B, C, D, E, F, G, M e N estão tipicamente associados a asiáticos enquanto que a maioria dos nativos americanos pertence aos haplogrupos A, B, C, D; os haplogrupos L0, L1, L2, L3 e L4 são encontrados em africanos e os haplogrupos H, I, J, K, T, U, V, W e X estão tipicamente associados às populações europeias (van Oven & Kayser, 2009).

As linhagens mitocondriais mais diversificadas e mais antigas são as africanas que são classificadas em cinco grandes haplogrupos: L0 (mais antigo), L1, L2, L3 e L4 (linhagens mais recentes). Os haplogrupos L0, L1 e L2 são os mais frequentes na África subsaariana apresentando proporções que representam de 70 a 100% da totalidade dos haplogrupos encontrados nesta região (Salas *et al.*, 2002).

Os macro-haplogrupos asiáticos M e N surgiram do haplogrupo L3, característico da África do Norte, há aproximadamente 65.000 anos. O macro-haplogrupo M é subdividido em haplogrupos denominados de C, D, G, E enquanto o macro-haplogrupo N é subdividido nos haplogrupos A, B, F (Torroni & Wallace, 1994, Forster *et al.*, 1996). Devido à grande dimensão de seu território e à

complexidade de sua população, a Ásia apresenta uma elevada variação na frequência haplotípica entre os territórios.

Os haplogrupos A, B, C e D são reconhecidos como os progenitores de todo o DNAmT de nativos americanos, indicando que esses haplogrupos anteciparam a colonização das Américas. Os haplogrupos A, C e D verificados nos ameríndios derivam de populações Siberianas localizadas perto do Estreito de Bering, entretanto há ausência do haplogrupo B nesta população e uma alta frequência deste haplogrupo em populações da costa asiática do Pacífico. Especula-se que o haplogrupo B originou-se no sul da Sibéria e migrou para a América mais tarde do que a migração que carregou os haplogrupos A, C e D (Torrioni *et al.*, 1993, Torrioni *et al.*, 1993).

Em relação à população europeia, os haplogrupos H, I, J, T, U (incluindo o subhaplogrupo K), V, W e X abrangem praticamente toda a variabilidade do DNAmT da Europa (Torrioni *et al.*, 1996). A Figura 4 ilustra a distribuição dos haplogrupos, sugerindo como teriam ocorrido as migrações humanas.

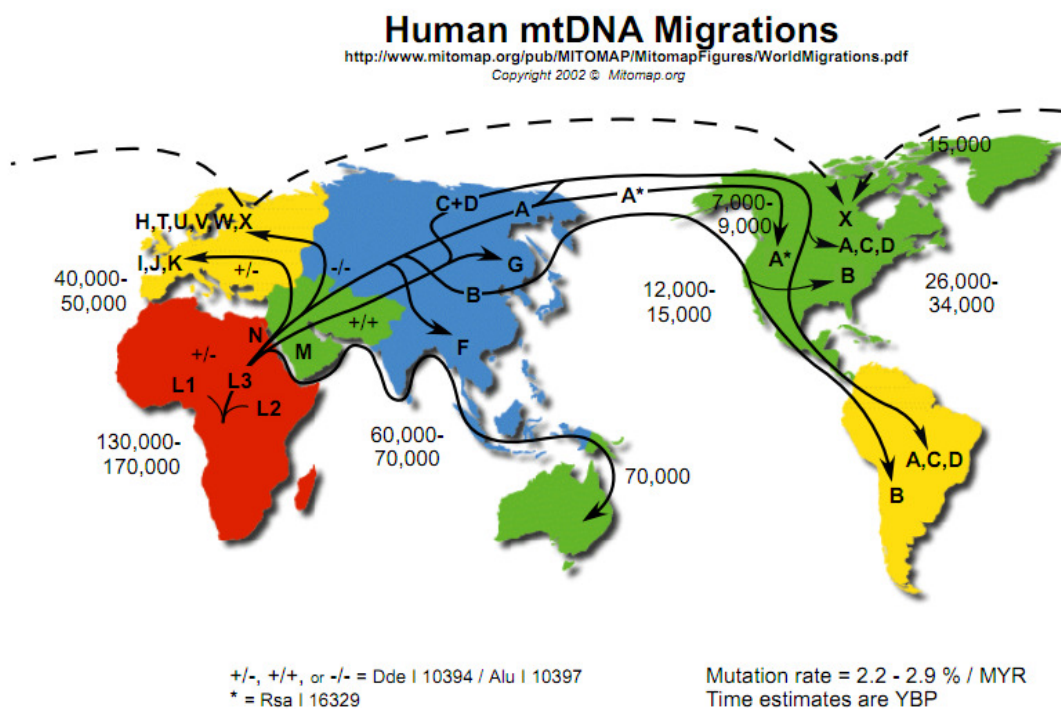


Figura 4 - Migrações Humanas a partir dos estudos do DNA mitocondrial. (Fonte: <http://www.mitomap.org>).

## POPULAÇÃO BRASILEIRA

Populações miscigenadas podem ser definidas como populações com ascendência populacional recente de dois ou mais continentes e que foram formadas a partir de processos de miscigenação entre as populações ancestrais durante centenas de anos. Tais processos foram consequência de acontecimentos históricos, como o tráfico transatlântico de escravos, a colonização das Américas e de outras migrações populacionais (Seldin *et al.*, 2011).

A população brasileira é conhecida como uma das mais heterogêneas do mundo. Tal característica é um reflexo do processo de colonização do Brasil. A nossa população é o resultado de cruzamentos entre diferentes grupos étnicos, tais como nativos americanos, africanos e europeus que vêm ocorrendo há mais de cinco séculos, desde que o Brasil foi descoberto e colonizado por Portugal. Quando em 1500 os portugueses chegaram ao Brasil, a terra era habitada por várias tribos nativas americanas (Bethell, 1984). Nos quatro séculos que se seguiram, cerca de doze milhões de africanos das regiões ao sul do Saara foram transportados para as Américas durante o tráfico de escravos. Destes, quatro milhões foram trazidos para o Brasil, sendo a grande maioria deles originários da costa Atlântica (África central e ocidental). Devido a numerosas questões, incluindo as de natureza demográfica, uma intensa mistura gênica ocorreu entre colonizadores, colonos e escravos desde o início desse contato (Ribeiro, 1995).

Entre os séculos XVI e XIX, em uma primeira grande onda migratória, vieram cerca de 500 mil portugueses para o Brasil enquanto, no mesmo período, cerca de quatro milhões de africanos foram trazidos como escravos. Na primeira metade do século XIX, indivíduos de outras nacionalidades, a exemplo dos italianos, espanhóis e alemães, migraram para diferentes regiões do país. No século XX o país recebeu imigrantes asiáticos e árabes (Callegari-Jacques & Salzano, 1999).

A formação da população brasileira vem sendo estudada há mais de duas décadas pela genética. Polimorfismos de DNA, presentes nos cromossomos autossomos, sexuais e no DNA mitocondrial, vêm sendo utilizados em estudos relacionados à formação da população brasileira e à origem do homem no novo mundo (Silva *et al.*, 2004, de Souza Goes *et al.*, 2005, Silva *et al.*, 2006, Domingues *et al.*, 2007, Barbosa *et al.*, 2008, Manta *et al.*, 2011, Nunes *et al.*, 2011).

O Rio de Janeiro é um estado localizado no sudeste do Brasil. Na época que os primeiros europeus chegaram, as principais tribos que habitavam esta região eram os Goytacazes, Paraíba, Puris, Tamoios, Goiânia, Coropós, Coroado, Kayapó, Guarus, Tupiminós e Temiminós (Peixoto, 1966). O porto do Rio de Janeiro foi um dos maiores pontos de entrada de escravos no Brasil até quase a metade do século XIX, recebendo escravos predominantemente oriundos da África Central e Ocidental do Atlântico.

Já foram realizados na população do Rio de Janeiro diversos estudos buscando estimar a diversidade e formação da população utilizando cromossomos autossômicos e sexuais (de Souza Goes *et al.*, 2005, Domingues *et al.*, 2007, Ferreira da Silva *et al.*, 2010). Não existem estudos que avaliem a ancestralidade materna a partir do DNAMt na população geral do Rio de Janeiro, a exceção de um estudo com uma amostra populacional auto declarada afro descendente (Hunemeier *et al.*, 2007). O objetivo deste trabalho foi, a partir de uma amostra da população atual, analisar os polimorfismos de sequência das regiões hipervariáveis do DNAMt e determinar as linhagens ancestrais maternas envolvidas no processo de povoamento que culminou na formação da população do Rio de Janeiro.

## **1 OBJETIVOS**

### **1.1 Objetivo geral**

Contribuir para o conhecimento sobre a ancestralidade materna na população brasileira através da análise do DNA mitocondrial em uma amostra da população do Rio de Janeiro.

### **1.2 Objetivos específicos**

1. Caracterizar a população do Rio de Janeiro quanto aos polimorfismos das regiões HVI e HVII do DNA mitocondrial.
2. Contribuir para a construção de uma base de dados dos perfis haplotípicos com vista à eventual aplicação na área de identificação humana.
3. Traçar a origem geográfica das matrilineagens presentes na amostra da população do Rio de Janeiro, bem como a proporção de cada grupo ancestral na população atual.
4. Comparar os dados com os de populações ancestrais e outras populações do Brasil.

## 2 MATERIAL E MÉTODOS

### 2.1 Amostras biológicas

A amostra do presente estudo foi obtida de 109 indivíduos não relacionados geneticamente, residentes no Rio de Janeiro, sendo 86 naturais desse estado. Todos participaram de perícias realizadas no Laboratório de Diagnósticos por DNA (LDD) da Universidade do Estado do Rio de Janeiro (UERJ). Através de termo de consentimento livre e esclarecido foi autorizada a utilização de suas amostras biológicas, de forma não identificada, para a realização de estudos científicos.

As amostras de sangue periférico foram coletadas através do sistema Vacutainer contendo EDTA como anticoagulante e mantidas a 4<sup>o</sup> C até a extração do DNA.

### 2.2 Extração de DNA

#### 2.2.1 Extração por *Salting-out* (Miller *et al.*, 1988)

Às alíquotas de 2mL de sangue foram adicionados 10mL de tampão de lise de hemácias (TLH). A preparação foi homogeneizada manualmente e centrifugada por 5 minutos a 100 x g (Centrifuga Excelsa modelo 206BL). Em seguida, o sobrenadante foi descartado. Ao sedimento com células foram adicionados 10mL de tampão TLH, repetindo-se a centrifugação. Após a remoção do sobrenadante, as células foram ressuspensas em 500µL de tampão de lise de membrana nuclear (NLB), seguindo-se a adição de 15µL de SDS 10% e 40µL de proteinase K 20mg/mL. Após homogeneização, a preparação foi mantida por 15 a 18 horas a 37°C. Ao final do período foram adicionados 150µL de NaCl 6M, sendo a preparação agitada manualmente, por aproximadamente 30 segundos, e centrifugada por 10 minutos a 905 x g (Centrifuga Excelsa modelo 206BL). Ao sobrenadante, transferido para um tubo de 10 mL do tipo Falcon, foram adicionados 2,5mL de etanol absoluto sob agitação manual em movimentos circulares com vista à precipitação dos ácidos nucléicos. O álcool foi descartado e o precipitado ressuspensado em 100 a 500µL de tampão TE.



### 2.2.2 Extração por Chelex (Walsh *et al.*, 1991)

Em tubos de micro centrífuga de 1,5mL foi adicionado 1mL de água deionizada autoclavada para cada 10 $\mu$ L de amostra de sangue. Após 30 minutos à temperatura ambiente, o material foi centrifugado por 5 minutos à 12000 x g (Centrifuga Centrimicro 242 - FANEM). O sobrenadante foi descartado e o precipitado ressuspenso em 0,5 mL de água deionizada autoclavada mantido, por 30 minutos, à temperatura ambiente. O material foi centrifugado, por 5 minutos, à 12000 x g e ao precipitado foram adicionados 300 $\mu$ L de solução de Chelex-100 a 5% (p/v). A preparação foi mantida em banho-maria, a 56°C, por 20 minutos e, posteriormente, a 95°C, por 8 minutos. Após a adição de 1mL de TE, o Chelex foi sedimentado através de centrifugação por 3 minutos à 12000 x g, sendo a solução de DNA armazenada a 4°C.

### 2.3 **Determinação da concentração de DNA**

A concentração do DNA extraído por *salting-out* foi determinada por espectrofotometria. O equipamento utilizado para a leitura foi o NanoDrop 2000 (Thermo scientific), que utiliza um sistema de retenção e medição do volume da amostra diretamente em um pedestal ótico com um espectro de 190 a 840 nm.

### 2.4 **Reação de polimerização em cadeia (PCR) para Regiões Hipervariáveis do DNA mitocondrial**

Alíquotas contendo 2 ng de DNA total foram transferidas para tubo estéril apropriado para a reação de PCR. A mistura da reação foi composta por 7,5 $\mu$ L de Tampão gold 3,3x (Applied Biosystems), 0,5 $\mu$ L de BSA (0,5mg/mL), 2,5U de Taq Gold DNA polimerase (0,5 $\mu$ L de 5U/ $\mu$ L) (Applied Biosystems) e 0,5 $\mu$ L de cada iniciador (10  $\mu$ M) em um volume final de 25 $\mu$ L.

Para amplificação das regiões HVI e HVII do DNA mitocondrial foram utilizados os pares de L15997-H16391 e L048-H408, respectivamente

(Hedman *et al.*, 2007). As sequências dos pares de iniciadores encontram-se na Tabela 1.

Tabela 1 - Sequências de iniciadores utilizados para amplificação das regiões HVI e HVII do DNA mitocondrial.

Iniciador	Sequência dos iniciadores (5'- 3')
L15997	CACCATTAGCACCCAAAGCT
H16391	GAGGATGGTGGTCAAGGGAC
L48	CTCACGGGAGCTCTCCATGC
H408	CTGTTAAAAGTGCATACCGCCA

As condições termocíclicas utilizadas nas reações de amplificação estão apresentadas na Tabela 2.

Tabela 2 – Condições termocíclicas utilizadas para amplificação das regiões HVI e HVII do DNA mitocondrial

Temperatura	Tempo	
95° C	11 minutos	
95° C	10 segundos	
60° C	30 segundos	<b>32 ciclos</b>
72° C	30 segundos	

Os produtos de PCR foram separados por eletroforese (tensão média de 110V por 15 minutos) em gel de agarose a concentração de 2%, corados em solução de brometo de etídeo (0,5 µg/mL) por 10 minutos e visualizados em

transiluminador de luz ultravioleta (256nm). Para o registro da imagem foi utilizado o leitor de fluorescência FMBIO II (Hitachi, Co).

## 2.5 Sequenciamento das Regiões HVI e HVII do DNA mitocondrial.

Para a purificação dos produtos de PCR, foram adicionados 2  $\mu\text{L}$  da enzima ExoSAP-IT® (USB Corporation) às alíquotas contendo 5  $\mu\text{L}$  do produto de amplificação. Esta preparação foi incubada a 37°C por 15 minutos e, posteriormente, 85°C por 15 minutos. Após esta etapa foi realizada a reação de sequenciamento utilizando o tampão *Big Dye Terminator Cycle Sequencing Ready Reaction Kit v3.0* (Applied Biosystems). A mistura da reação foi realizada em microtubos, sendo composta por: 3  $\mu\text{L}$  do produto de amplificação purificado, 3  $\mu\text{L}$  do iniciador a 1  $\mu\text{M}$  e 4  $\mu\text{L}$  da solução de reação (kit 2:2) em um volume final de 10  $\mu\text{L}$ . Para cada indivíduo foram realizadas reações de sequenciamento para as fitas: senso e anti-senso das regiões HVI e HVII utilizando os respectivos iniciadores, totalizando 4 reações de sequenciamento distintas por indivíduo. Em seguida as misturas foram submetidas a 25 ciclos térmicos nos termocicladores GeneAmp 2400, GeneAmp 9700 e Veriti (Applied Biosystems) nas seguintes condições: 96°C por 15 segundos, 50°C por 1 segundo, 60°C por 2 minutos, cada ciclo.

Após a reação, o produto de sequenciamento foi purificado em resina Sephadex G-50 (Sambrook *et al.*, 1989). Nesta metodologia, para preparo da resina, 750  $\mu\text{L}$  de Sephadex (solução de 10g em 150 mL) foram aplicados em uma coluna de purificação própria. Esta foi colocada dentro de um tubo de 1,5 mL e posteriormente centrifugada a 5.000 x g por 1 minuto e 30 segundos (Centrifuga Centrimicro 242 - FANEM). O filtrado foi descartado e a coluna centrifugada novamente. Após esta etapa, a resina Sephadex em coluna de purificação própria foi transferida para um novo tubo de 1,5 mL e a ela foi adicionado o volume total do produto de sequenciamento, seguindo para centrifugação a 5.000 x g por 3 minutos.

Ao filtrado foram adicionados 10  $\mu\text{L}$  de formamida (Hi-Di Formamide, Applied Biosystems), e, desta mistura, 10  $\mu\text{L}$  foram aplicados em cada poço da placa do sequenciador. A placa foi centrifugada a 1237 x g por 1 minuto (Laboratory Centrifuge 2-16P Sigma). As amostras foram então submetidas à eletroforese capilar nas condições desnaturantes já descritas acima, utilizando o sequenciador

automático ABI Prism 3130/3500 (Applied Biosystems) preenchido com polímero POP-7™ (Applied Biosystems).

## 2.6 Análise das sequências e determinação de haplogrupos

Após a eletroforese capilar, para determinação da qualidade das sequências geradas, estas foram visualizadas por meio do software Sequencing Analysis 3.7 (Applied Biosystems). A análise das sequências e a determinação dos polimorfismos foram realizadas com o auxílio do *software* SeqScape®Software - versão 2.5 (Applied Biosystems) através de alinhamento e comparação entre as sequências obtidas e a sequência de referência de Cambridge (rCRS) (Anderson *et al.*, 1981).

Os haplogrupos foram classificados com o auxílio da ferramenta mtDNAManager ([mtmanager.yonsei.ac.kr/](http://mtmanager.yonsei.ac.kr/)). Tal ferramenta é um recurso de bioinformática online gratuito que, através de uma interface simples, analisa as sequências de DNA mitocondrial, executando cálculos sobre os polimorfismos encontrados nas regiões hipervariáveis a fim de estimar o haplogrupo mais provável referente a dado conjunto de polimorfismos (Lee *et al.*, 2008).

## 2.7 Análises Estatísticas

Os cálculos referentes aos parâmetros de diversidade: frequências haplotípicas, diversidade haplotípica, número de haplótipos, diversidade nucleotídica, número de transições, transversões, substituições e diferenças aos pares (pair-wise) foram realizados através do *software* Arlequin versão 3.01 (Excoffier *et al.*, 2005).

As frequências dos haplogrupos na amostra da população do Rio de Janeiro foram estimadas através de contagem direta.

## 2.8 Cálculo de distâncias genéticas

Com o propósito de estimar a diferenciação genética entre as populações do Rio de Janeiro, outros estados brasileiros, bem como populações africanas, europeias, ameríndias e asiáticas, os dados obtidos para as frequências dos haplogrupos de DNAmT foram comparados com aqueles das seguintes populações: Alagoas (AL) (Barbosa *et al.*, 2008); Pará (PA) (Feio-Dos-Santos *et al.*, 2006); Rio Grande do Sul (RS) (Marrero *et al.*, 2005); Santa Catarina (SC) (Palencia *et al.*, 2010); São Paulo (SP) (Paneto *et al.*, 2011); Angola (AN) (Plaza *et al.*, 2004); Portugal (PT) (Pereira *et al.*, 2000); Terenas (TE) (SILVA D.A., comunicação pessoal); China (CH) (Yao *et al.*, 2002).

Esta comparação foi realizada através da medida de distância genética ( $F_{ST}$ ) que é calculada a partir das frequências dos haplogrupos verificadas nas populações analisadas. A matriz de  $F_{ST}$  foi construída por meio do *software* Arlequin versão 3.0 (Excoffier *et al.*, 2005). Como alguns haplogrupos apresentaram um número muito reduzido de haplótipos, para a construção da matriz de distância genética, foram considerados níveis hierárquicos superiores, a saber: L0, L1, L2, L3, L4, A, B, C, D, H, K, N, U, M. Os haplogrupos não verificados em todas as populações ainda foram agrupados em uma categoria nomeada O (outros).

Para obter uma representação gráfica das afinidades entre as distâncias das várias populações comparadas, foi realizada a análise do escalonamento multidimensional (MDS) das distâncias genéticas ( $F_{ST}$ ) entre as diversas populações e o cálculo do valor de stress, utilizando o programa STATISTICA versão 8.0 (StatSoft, 2007; [www.statsoft.com](http://www.statsoft.com)).

## 2.9 Lista de soluções e tampões utilizados

### Extração de DNA Genômico

---

#### **Tampão de lise de hemáceas (TLH)**

---

NH <sub>4</sub> CL	155 mM
KHCO <sub>3</sub>	10 mM
EDTA	1 mM

---

---

#### **Tampão TE (pH 7,6)**

---

Tris-Cl pH 7,6	10 mM
EDTA pH 8,0	1 mM

---

---

#### **Tampão de Lise de núcleo (NLB)**

---

Tris-Cl pH 8	10 mM
NaCl	0,4 mM
EDTA	2 mM

---

### **3 RESULTADOS**

#### **3.1 Análise dos produtos de PCR e Sequenciamento**

Os polimorfismos das regiões hipervariáveis HVI e HVII do DNA mitocondrial, localizados entre as posições 16024-16365 e 70-340, respectivamente, foram analisados. Para isto, tais regiões foram amplificadas por PCR e posteriormente sequenciadas, conforme descrito em material e métodos. Obteve-se amplificação por PCR de todas as amostras do presente estudo.

##### **3.1.1 Análise dos produtos de amplificação em gel de agarose 2%**

Os produtos relativos às amplificações das regiões HVI e HVII do DNA mitocondrial foram visualizados em gel de agarose, a 2%, corado com brometo de etídeo, conforme ilustrado na Figura 5.

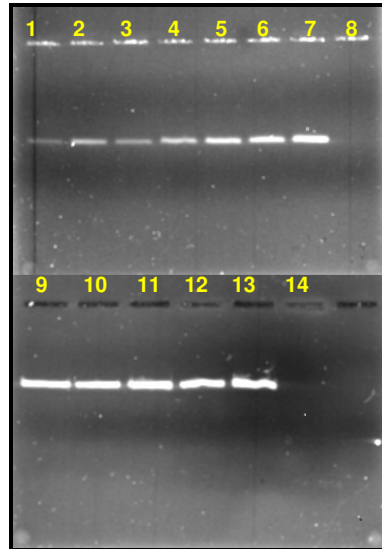


Figura 5 - Produtos de amplificação por PCR das regiões HVI e HVII do DNA mitocondrial.

As regiões HVI e HVII do DNA mitocondrial foram amplificadas com os pares de iniciadores L15997-H16391 e L048-H408, respectivamente. Após eletroforese em gel de agarose 2%, o gel foi corado com brometo de etídeo (0,5 µg/mL) e os produtos de amplificação visualizados em transiluminador de luz ultravioleta. Amostras de 1 a 6: produtos de amplificação da região HVI de 6 indivíduos; Amostras 7 e 8: controles positivo (DNA K562) e negativo da reação de amplificação da região HVI, respectivamente; Amostras de 9 a 12: produtos de amplificação da região HVII de 4 indivíduos; Amostras 13 e 14: controles positivo (DNA K562) e negativo da reação de amplificação da região HVII, respectivamente.

### 3.1.2 Sequenciamento das regiões HVI e HVII do DNA mitocondrial

Resultados relativos aos sequenciamentos de produtos de amplificação por PCR das regiões HVI e HVII do DNA mitocondrial são mostrados nas Figuras 6 e 7, respectivamente.



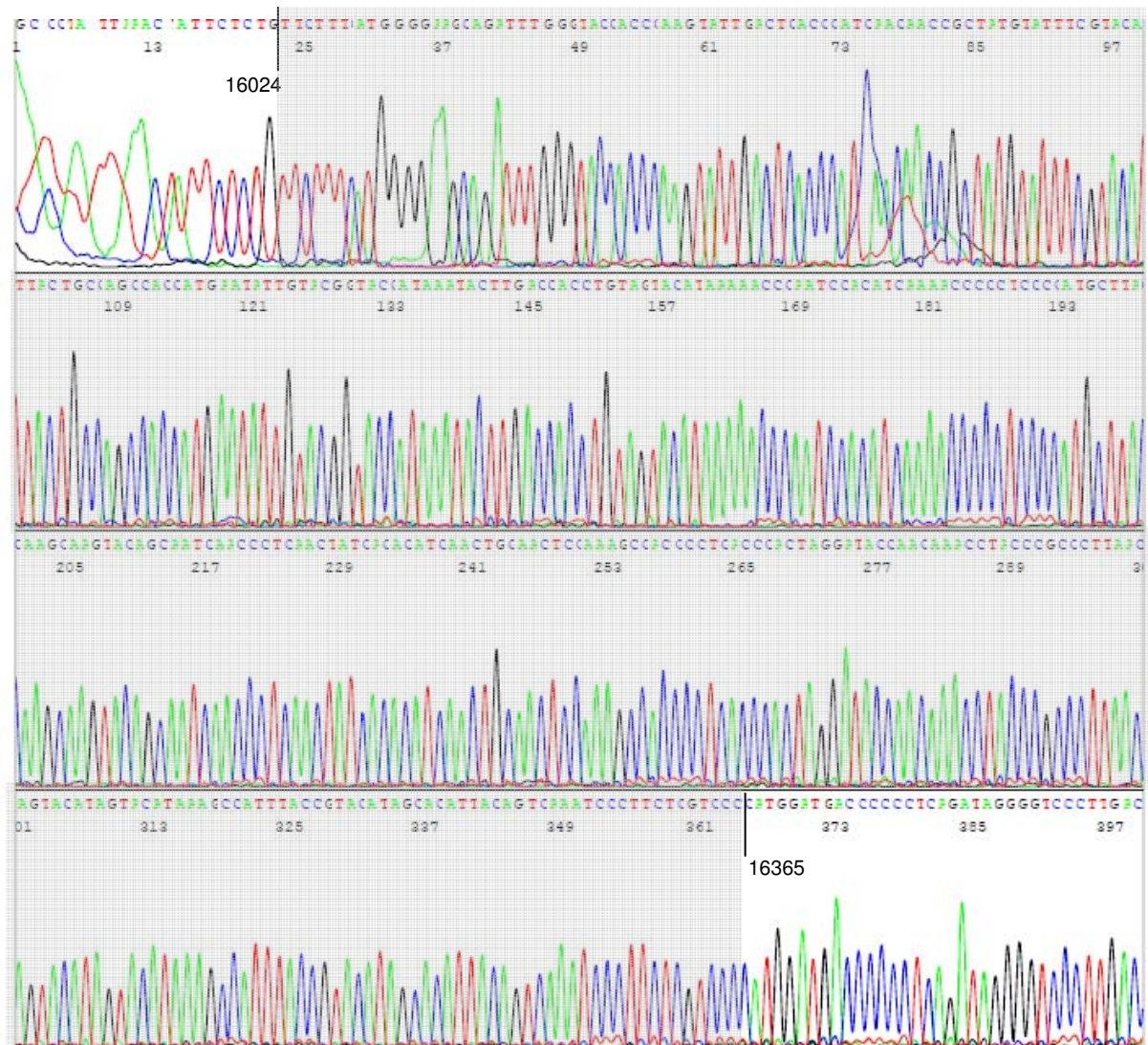


Figura 6 - Eletroferograma representativo do sequenciamento da região HVI do DNA mitocondrial (orientação 5'-3') da amostra RJM63.

Alíquotas das preparações de DNA foram amplificadas com os pares de iniciadores L15997-H16391 para a região HVI do DNA mitocondrial. Após a purificação dos produtos de PCR foram realizadas reações individuais de sequenciamento para produtos gerados por cada iniciador. A sequência de bases foi determinada em leitores de fluorescência ABI 3130/3500 durante eletroforese capilar e visualizada com o software Sequencing Analysis 3.7 (Applied Biosystems). A sequência HVI, compreendida entre as posições 16024 e 16365, encontra-se delimitada na Figura.

Legenda: T = Timina; A = Adenina; G = Guanina; C = Citosina.

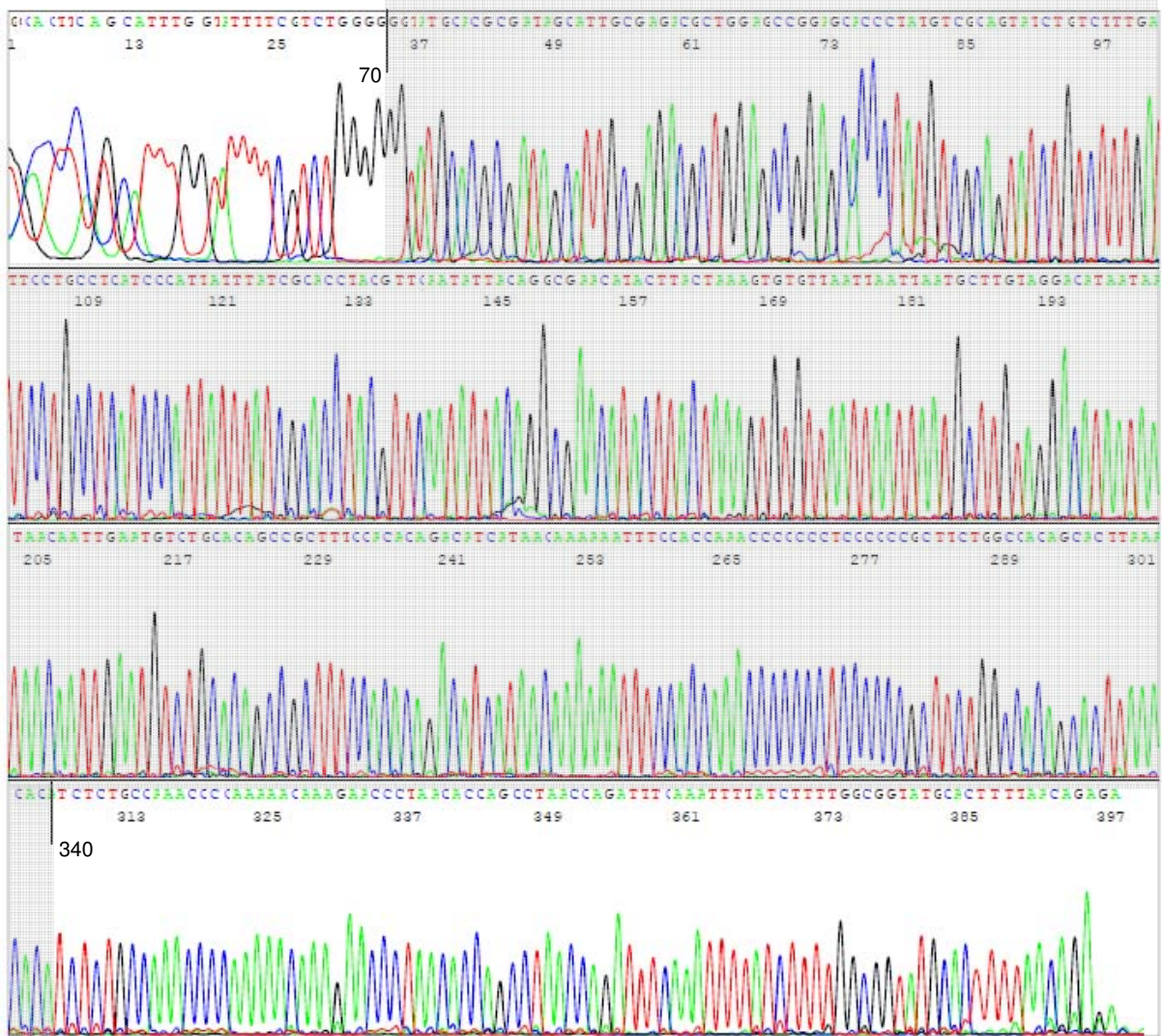


Figura 7 - Eletroferograma representativo do sequenciamento da região HVII do DNA mitocondrial (orientação 5'-3') da amostra RJM63.

Alíquotas das preparações de DNA foram amplificadas com os pares de iniciadores L048-H408 para a região HVII do DNA mitocondrial. Após a purificação dos produtos de PCR foram realizadas reações individuais de sequenciamento para produtos gerados por cada iniciador. A sequência de bases foi determinada em leitores de fluorescência ABI 3130/3500 durante eletroforese capilar e visualizada com o software Sequencing Analysis 3.7 (Applied Biosystems). A sequência HVII, compreendida entre as posições 70 e 340, encontra-se delimitada na Figura.

Legenda: **T** = Timina; **A** = Adenina; **G** = Guanina; **C** = Citosina.

### 3.2 Frequências haplotípicas e determinação dos haplogrupos

Para a genotipagem dos indivíduos, os polimorfismos foram identificados através da comparação entre as sequências HVI e HVII do DNA mitocondrial de cada indivíduo e a Sequência de Referência de Cambridge (rCRS), utilizando-se para isto o software SeqScape® - versão 2.5 (Applied Biosystems). Os haplótipos caracterizados pelo conjunto de polimorfismos verificados em um indivíduo e suas respectivas frequências foram então determinados. Os haplogrupos foram classificados com o auxílio da ferramenta de bioinformática mtDNAManager (mtmanager.yonsei.ac.kr/).

Um total de 142 diferentes *loci* polimórficos foram verificados nas regiões HVI e HVII do DNA mitocondrial, sendo identificados 102 haplótipos distintos em 109 indivíduos. A análise desses resultados mostrou que sete haplótipos foram compartilhados por 2 indivíduos e 95 ocorreram apenas uma vez. A frequência dos haplótipos únicos foi de 0.00917 e a diversidade haplotípica foi estimada em  $0.9988 \pm 0.0014$ .

Os diferentes haplótipos com suas respectivas frequências, os polimorfismos que os caracterizam e a classificação em haplogrupos são mostrados na Tabela 3.



Tabela 3 – Haplótipos, frequências haplotípicas, polimorfismos das regiões HVI e HVII do DNA mitocondrial e haplogrupos verificados na amostra da população do Rio de Janeiro.

	Haplótipo	Frequência	Polimorfismos HVI	Polimorfismos HVII	Haplogrupo
1	RJM42	0.00917	16109T 16180G 16223T 16290T 16319A	73G 146C 153G 235G 263G 309.1C 309.2C 315.1C	A
2	RJM109	0.00917	16097C 16098G 16109T 16189C 16223T 16290T 16319A 16320T	73G 146C 153G 235G 263G 309.1C 315.1C	A
3	RJM53	0.00917	16104T 16109T 16223T 16290T 16319A 16362C	73G 146C 153G 235G 263G 309.1C 315.1C	A2
4	RJM109	0.00917	16109T 16126C 16223T 16259T 16290T 16319A 16327T 16362C	73G 146C 153G 235G 263G 315.1C	A2
5	RJM112	0.00917	16109T 16215 G 16223T 16290T 16319A 16362C	73G 146C 153G 235G 263G 309.1C 315.1C	A2
6	RJM65	0.0183	16109T 16223T 16290T 16319A 16362C	73G 146C 153G 235G 263G 315.1C	A2
7	RJM3	0.00917	16182C 16183C 16189C 16217C 16241G	73G 103A 152C 263G 309.1C 315.1C	B4
8	RJM21	0.00917	16182C 16183C 16189C 16217C 16249C 16312G 16344T	73G 152C 263G 271T 315.1C	B4
9	RJM25	0.00917	16168T 16182C 16183C 16189C 16217C 16249C 16312G	73G 152C 263G 271T 309.1C 315.1C	B4
10	RJM48	0.00917	16182C 16183C 16186T 16189C 16217C	73G 263G 315.1C	B4
11	RJM54	0.00917	16057A 16182C 16183C 16189C 16217C 16249C 16312G 16344T	73G 152C 263G 271T 309.1C 309.2C 315.1C	B4
12	RJM64	0.00917	16183C 16189C 16217C 16241G	73G 103A 152C 263G 309.1C 309.2C 315.1C	B4
13	RJM72	0.00917	16178C 16183C 16189C 16217C	73G 263G 309.1C 309.2C 315.1C	B4
14	RJM80	0.0183	16092C 16183C 16189C 16217C	73G 152C 263G 309.1C 315.1C	B4
15	RJM88	0.00917	16168T 16182C 16183C 16189C 16217C 16249C 16312G	73G 152C 263G 271T 309.1C 309.2C 315.1C	B4
16	RJM89	0.00917	16183C 16189C 16217C 16241G 16258C	73G 103A 152C 263G 309.1C 315.1C	B4
17	RJM60	0.00917	16126C 16223T 16258C 16298C 16325C 16327T	73G 150T 200G 204C 249D 263G 283C 285T 287.1C 309.1C 315.1C	C
18	RJM38	0.00917	16223T 16298C 16325C 16327T 16362C	73G 249D 263G 290D 291D 309.1C 315.1C	C1
19	RJM55	0.00917	16154C 16223T 16298C 16325C 16327T	73G 249D 290D 291D 309.1C 309.2C 315.1C	C1
20	RJM62	0.00917	16223T 16239T 16298C 16311C 16325C 16327T	73G 146C 249D 263G 290D 291D 309.1C 315.1C	C1
21	RJM108	0.00917	16223T 16298C 16325C 16327T	73G 194T 249D 263G 290D 291D 309.1C 315.1C	C1
22	RJM45	0.00917	16051G 16172C 16223T 16298C 16325C 16327T	73G 194T 249D 263G 290D 291D 309.1C 315.1C	C1d
23	RJM76	0.00917	16051G 16184T 16223T 16287T 16298C 16311C 16325C 16327T	73G 146C 194T 249D 263G 290D 291D 309.1C 315.1C	C1d
24	RJM87	0.00917	16051G 16184T 16223T 16287T 16298C 16311C 16325C 16327T	73G 146C 194T 249D 263G 290D 291D 309.1C 309.2C 315.1C	C1d
25	RJM2	0.00917	16124C 16223T 16362C	73G 152C 195C 263G 315.1C	D4/G
26	RJM22	0.00917	16037G 16180G	263G 309.1C 315.1C	H
27	RJM28	0.00917	16271C	263G 315.1C	H
28	RJM29	0.00917		263G 309.1C 309.2C 315.1C	H
29	RJM44	0.00917		263G 309.1C 315.1C	H
30	RJM56	0.00917	16311C	263G 309.1C 309.2C 315.1C	H
31	RJM74	0.00917	16109T	263G 309.1C 309.2C 315.1C	H
32	RJM82	0.00917	16293G	263G 309.1C 315.1C	H
33	RJM63	0.00917	16293G	152C 263G 315.1C	H
34	RJM19	0.00917	16162G 16209C	73G 153G 263G 309.1C 315.1C	H1a
35	RJM52	0.00917	16162G 16209C	73G 263G 315.1C	H1a
36	RJM61	0.00917	16288C 16362C	146C 152C 195C 263G 309.1C 315.1C	H8
37	RJM15	0.00917	16086C 16224C 16311C	73G 152C 315.1C	K
38	RJM9	0.00917	16172C 16213A 16223T 16224C 16278T	73G 146C 152C 263G 309.1C 315.1C	K2a
39	RJM6	0.00917	16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16293G 16311C	93G 95C 185A 189G 236C 247A 263G 309.1C 309.2C 315.1C	L0a1b
40	RJM86	0.00917	16093C 16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16293G	93G 95C 185A 189G 236C 247A 263G 315.1C	L0a1b

	Haplótipo	Frequência	Polimorfismos HVI	Polimorfismos HVII	Haplogrupo
41	RJM102	0.00917	16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16293G 16311C 16320T	93G 95C 185A 189G 236C 247A 263G 315.1C	L0a1b
42	RJM4	0.00917	16148T 16172C 16187T 16188G 16189C 16223T 16230G 16311C 16320T	93G 152C 189G 204C 207A 236C 247A 263G 315.1C	L0a2
43	RJM69	0.00917	16148T 16172C 16187T 16188G 16189C 16223T 16230G 16311C 16320T	93G 152C 189G 204C 207A 236C 247A 263G 309.1C 315.1C	L0a2
44	RJM98	0.00917	16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C	73G 152C 182T 185T 189G 195C 247A 263G 309.1C 315.1C	L1b1
45	RJM43	0.00917	16126C 16187T 16189C 16193T 16223T 16264T 16270T 16278T 16311C	73G 152C 182T 185T 195C 203A 247A 263G 315.1C	L1b1*
46	RJM7	0.00917	16037G 16187T 16189C 16223T 16271C 16274A 16278T 16294T 16311C 16360T	73G 151T 152C 182T 186A 189C 195C 204C 247A 263G 297G 315.1C 316A	L1c
47	RJM26	0.00917	16129A 16145A 16187T 16189C 16213A 16223T 16265C 16271C 16278T 16286G 16294T 16311C	73G 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A	L1c
48	RJM57	0.00917	16093C 16129A 16183C 16189C 16215G 16223T 16278T 16294T 16311C 16319A	73G 151T 152C 182T 186A 189C 247A 263G 315.1C 316A	L1c
49	RJM58	0.00917	16093C 16129A 16144A 16183C 16189C 16215G 16223T 16278T 16294T 16311C	73G 151T 152C 186A 189C 247A 263G 315.1C 316A	L1c
50	RJM91	0.00917	16038G 16129A 16187T 16189C 16223T 16278T 16293G 16294T 16311C 16360T	73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A	L1c1
51	RJM16	0.00917	16129A 16187T 16189C 16214T 16223T 16265C 16278T 16286A 16291T 16294T 16311C 16360T	73G 151T 152C 182T 186A 189C 195C 198T 204C 247A 263G 297G 315.1C 316A	L1c2
52	RJM30	0.00917	16129A 16134T 16145A 16187T 16189C 16213A 16223T 16265C 16274A 16278T 16286G 16294T	73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A	L1c2
53	RJM39	0.00917	16071T 16129A 16145A 16187T 16189C 16213A 16223T 16234T 16265C 16278T 16286G 16294T	73G 151T 152C 182T 186A 189C 195C 198T 234G 247A 263G 297G 315.1C 316A	L1c2
54	RJM27	0.00917	16129A 16163G 16187T 16189C 16209C 16223T 16278T 16293G 16294T 16311C 16360T	73G 151T 152C 182T 186A 189C 247A 263G 309.1C 315.1C 316A	L1c3b
55	RJM70	0.00917	16129A 16163G 16187T 16189C 16223T 16249C 16278T 16293G 16294T 16311C 16360T	73G 151T 152C 182T 186A 189C 247A 263G 315.1C 316A	L1c3b
56	RJM50	0.00917	16223T 16278T	73G 89C 93G 146C 150T 182T 195C 198T 263G 315.1C 325T	L2
57	RJM8	0.00917	16223T 16278T 16294T 16311C 16355T	73G 143A 146C 152C 195C 263G 315.1C	L2a
58	RJM73	0.00917	16093C 16189C 16223T 16278T 16294T	73G 143A 146C 152C 195C 263G 315.1C	L2a
59	RJM23	0.00917	16172C 16223T 16278T 16294T 16309G	73G 143A 146C 152C 195C 263G 315.1C	L2a1
60	RJM31	0.00917	16092C 16223T 16278T 16294T 16309G	73G 146C 152C 195C 198T 263G 315.1C	L2a1
61	RJM32	0.00917	16223T 16278T 16294T 16309G	73G 143A 146C 152C 195C 263G 315.1C	L2a1
62	RJM96	0.00917	16093C 16129A 16189C 16223T 16278T 16294T 16309G	73G 143A 146C 152C 195C 263G 309.1C 315.1C	L2a1
63	RJM100	0.00917	16093C 16223T 16256T 16278T 16294T 16309G	73G 143A 146C 152C 195C 263G 309.1C 309.2C 315.1C	L2a1
64	RJM114	0.00917	16189C 16223T 16278T 16294T 16309G 16342C	73G 146C 150T 152C 195C 263G 309.1C 315.1C	L2a1
65	RJM110	0.00917	16189C 16192T 16223T 16274A 16278T 16294T 16309G	73G 109C 146C 152C 195C 263G 309.1C 309.2C 315.1C	L2a1*
66	RJM94	0.00917	16150T 16223T 16278T 16286T 16294T 16309G	73G 146C 152C 195C 263G 309.1C 315.1C	L2a1a
67	RJM81	0.00917	16175C 16182C 16183C 16189C 16192T 16223T 16278T 16290T 16294T 16309G	73G 146C 152C 195C 263G 315.1C	L2a1b
68	RJM41	0.00917	16114A 16129A 16213A 16223T 16278T 16354T	73G 146C 150T 152C 182T 195C 198T 204C 263G 315.1C	L2b
69	RJM10	0.00917	16114A 16129A 16213A 16223T 16278T 16355T 16362C	73G 150T 152C 182T 195C 198T 204C 263G 315.1C	L2b1
70	RJM101	0.00917	16114A 16129A 16172C 16174T 16213A 16223T 16278T 16355T 16362C	73G 150T 152C 182T 189G 195C 198T 204C 263G 309.1C 315.1C	L2b1
71	RJM18	0.0183	16124C 16223T 16278T 16362C	73G 263G 315.1C	L3b
72	RJM97	0.00917	16124C 16223T 16278T 16362C	73G 263G 309.1C 315.1C	L3b
73	RJM99	0.00917	16124C 16223T 16259T 16278T 16362C	73G 263G 315.1C	L3b
74	RJM13	0.00917	16124C 16223T 16278T 16311C 16362C	73G 150T 263G 315.1C	L3b1*
75	RJM68	0.00917	16124C 16223T 16278T 16311C 16362C	73G 263G 315.1C	L3b1*
76	RJM34	0.00917	16145A 16223T 16278T 16362C	73G 152C 195C 247A 263G 279C 315.1C	L3b1a
77	RJM77	0.00917	16223T 16278T 16362C	73G 263G 309.1C 315.1C	L3b1a
78	RJM84	0.00917	16124C 16223T 16278T 16362C	73G 152C 263G 315.1C	L3b1b
79	RJM17	0.00917	16124C 16223T 16257T	73G 151T 152C 263G 315.1C	L3d
80	RJM1	0.0183	16124C 16166G 16223T	73G 152C 263G 315.1C	L3d1c
81	RJM24	0.00917	16124C 16223T 16256T	73G 152C 263G 315.1C	L3d1d
82	RJM14	0.00917	16109T 16124C 16223T	73G 152C 199C 263G 315.1C	L3d3
83	RJM79	0.00917	16207G 16223T 16327T	73G 150T 183G 189G 195C 200G 263G 315.1C	L3e1
84	RJM93	0.00917	16223T 16327T	73G 150T 152C 189G 200G 263G 309.1C 315.1C	L3e1

	Haplótipo	Frequência	Polimorfismos HVI	Polimorfismos HVII	Haplogrupo
85	RJM35	0.00917	16185T 16223T 16327T	73G 150T 189G 200G 263G 309.1C 315.1C	L3e1a
86	RJM106	0.00917	16185T 16223T 16311C 16327T	73G 150T 185A 189G 200G 263G 315.1C	L3e1a
87	RJM20	0.00917	16176T 16223T 16256T 16327T	73G 150T 152C 189G 263G 315.1C	L3e1d
88	RJM11	0.00917	16223T 16311C 16320T	73G 150T 195C 198T 263G 309.1C 315.1C	L3e2a
89	RJM46	0.00917	16164G 16172C 16189C 16223T 16320T	73G 146C 150T 195C 263G 315.1C	L3e2b
90	RJM51	0.00917	16172C 16189C 16223T 16278T 16320T	73G 150T 195C 263G 315.1C	L3e2b
91	RJM90	0.00917	16076T 16172C 16183C 16189C 16223T 16320T	73G 150T 195C 263G 315.1C	L3e2b
92	RJM95	0.00917	16172C 16183C 16189C 16223T 16320T	73G 150T 152C 195C 263G 309.1C 309.2C 315.1C	L3e2b*
93	RJM47	0.0183	16223T 16265T	73G 150T 195C 263G 315.1C	L3e3
94	RJM75	0.00917	16223T 16265T	73G 150T 195C 198T 263G 315.1C	L3e3
95	RJM105	0.00917	16093C 16223T 16265T 16316G	73G 150T 195C 263G 315.1C	L3e3
96	RJM113	0.00917	16051G 16223T 16264T	73G 150T 263G 309.1C 315.1C	L3e4
97	RJM33	0.00917	16209C 16223T 16245T 16311C	73G 150T 189G 200G 263G 309.1C 315.1C	L3f2
98	RJM12	0.0183	16129A 16172C 16174T 16192T 16218T 16223T 16256A 16311C 16362C	73G 151T 152C 189C 195C 263G 294C 315.1C	L3h1b
99	RJM92	0.0183	16093G 16223T 16287A 16293T 16301T 16311C 16355T 16362C	73G 146C 150T 152C 195C 200G 244G 263G 315.1C	L4b2
100	RJM49	0.00917	16086C 16147A 16164G 16172C 16223T 16248T 16320T 16355T	73G 152C 199C 204C 207A 263G 315.1C	N1a
101	RJM104	0.00917	16080T 16189C 16270T	73G 150T 263G 309.1C 315.1C	U5b1b
102	RJM78	0.00917	16189C 16223T 16325C	73G 263G 315.1C	U5b2a1

As frequências haplotípicas foram calculadas dividindo-se o número de ocorrência de cada haplótipo pelo número total de indivíduos. Os haplótipos em destaque foram verificados duas vezes.

### 3.2.1 Análise de parâmetros de diversidade

Parâmetros de diversidade, tais como número de distintos haplótipos; diversidade genética; diversidade nucleotídica; número de transições, de transversões e de substituições e diferenças aos pares (pair-wise), computados separadamente e em conjunto para as regiões HVI e HVII do DNA mitocondrial dos indivíduos estudados, são apresentados na Tabela 4.

Tabela 4 - Parâmetros de diversidade a partir da análise das regiões HVI e HVII do DNA mitocondrial em uma amostra populacional do Rio de Janeiro.

Parâmetros	HVI	HVII	HVI + II
Número de <i>Loci</i> polimórficos	96	46	142
Número de Haplótipos	91	73	102
Diversidade Haplótipica	0.9963 ± 0.0019	0.9886 ± 0.0035	0.9988 ± 0.0014
Número de Transições	86	35	121
Número de Transversões	17	6	23
Número de Substituições	103	41	144
*Número de diferenças aos pares	8.734455 ± 4.062186	6.408767±3.0358784	15.143221 ± 6.820880
.Diversidade Nucleotídica	0.0759525 ± 0.039118	0.103367 ± 0.054635	0.085074 ± 0.042436

### 3.3 **Proporção dos haplogrupos na população do Rio de Janeiro**

Os haplótipos observados na amostra populacional do Rio de Janeiro foram associados aos respectivos haplogrupos. As proporções das distintas ancestralidades maternas foram determinadas por contagem direta sendo tal distribuição ilustrada na Figura 8.

Na amostra populacional do Rio de Janeiro composta por 109 indivíduos, verificou-se a ocorrência das ancestralidades maternas africana, nativa americana e europeia em 66, 27 e 16 indivíduos, respectivamente.

**Proporção de distintas ancestralidades maternas na população do Rio de Janeiro**

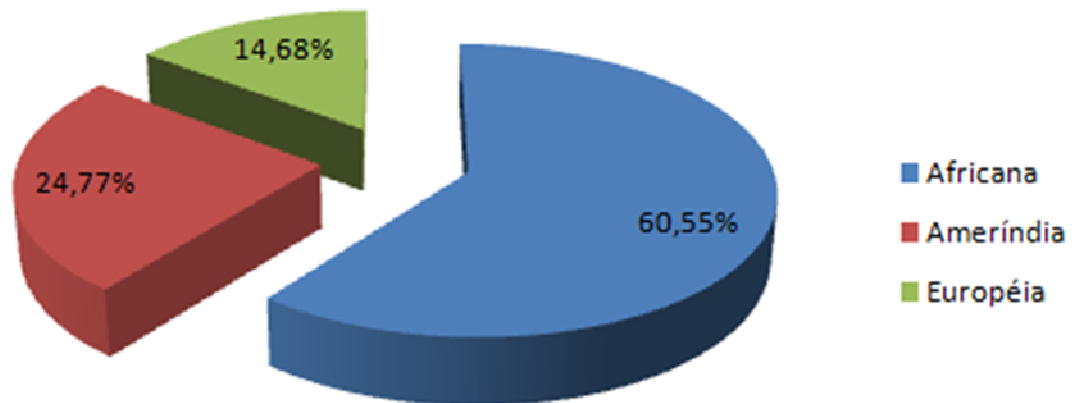


Figura 8 - Proporção de distintas ancestralidades maternas do DNA mitocondrial na amostra da população do Rio de Janeiro.

### 3.3.1 Haplogrupos africanos

Na amostra da população do Rio de Janeiro foram verificados 66 indivíduos (60,55%) com ancestralidade materna africana. Dentre os 35 diferentes haplogrupos verificados neste grupo o L2a1 foi o mais frequente e corresponde a 5,50% da amostra populacional estudada. Cada um dos haplogrupos L1c, L3b e L3e3 foi verificado com frequência de 3,67% enquanto para os haplogrupos L0a1b, L1c2 e L3e2b esta foi de 2,75%. Os demais haplogrupos apresentaram frequências iguais ou inferiores a 1,83%. Os distintos haplogrupos africanos verificados na amostra populacional do Rio de Janeiro, e suas respectivas frequências, são mostrados na Tabela 5.



Tabela 5 – Haplogrupos africanos e respectivas frequências na população do Rio de Janeiro

Haplogrupos africanos	n	Frequência
L0a1b	3	2,75%
L0a2	2	1,83%
L1b1	1	0,92%
L1b1*	1	0,92%
L1c	4	3,67%
L1c1	1	0,92%
L1c2	3	2,75%
L1c3b	2	1,83%
L2	1	0,92%
L2a	2	1,83%
L2a1	6	5,50%
L2a1*	1	0,92%
L2a1a	1	0,92%
L2a1b	1	0,92%
L2b	1	0,92%
L2b1	2	1,83%
L3b	4	3,67%
L3b1*	2	1,83%
L3b1a	2	1,83%
L3b1b	1	0,92%
L3d	1	0,92%
L3d1c	2	1,83%
L3d1d	1	0,92%
L3d3	1	0,92%
L3e1	2	1,83%
L3e1a	2	1,83%
L3e1d	1	0,92%
L3e2a	1	0,92%
L3e2b	3	2,75%
L3e2b*	1	0,92%
L3e3	4	3,67%
L3e4	1	0,92%
L3f2	1	0,92%
L3h1b	2	1,83%
L4b2	2	1,83%

Os haplogrupos foram determinados com o auxílio do software mtDNAManager (mtmanager.yonsei.ac.kr). As frequências foram calculadas dividindo-se o número de ocorrência de cada haplogrupo pelo número total de indivíduos (109) da amostra populacional do Rio de Janeiro estudada.

### 3.3.2 Haplogrupos nativos americanos

Na amostra da população do Rio de Janeiro foram verificados 27 indivíduos (24,77%) com ancestralidade materna nativa americana. Dentre os sete distintos haplogrupos verificados neste grupo o B4 foi o mais frequente e corresponde a 10,09% da amostra populacional estudada. Os distintos haplogrupos nativos americanos verificados na amostra populacional do Rio de Janeiro, e suas respectivas frequências, são mostrados na Tabela 6.

Tabela 6 – Haplogrupos nativos americanos e respectivas frequências na população do Rio de Janeiro

Haplogrupos nativos americanos	n	Frequência.
A	2	1,83%
A2	5	4,59%
B4	11	10,09%
C	1	0,92%
C1	4	3,67%
C1d	3	2,75%
D4/G	1	0,92%

Os haplogrupos foram determinados com o auxílio do software mtDNAManager (mtmanager.yonsei.ac.kr). As frequências foram calculadas dividindo-se o número de ocorrência de cada haplogrupo pelo número total de indivíduos (109) da amostra populacional do Rio de Janeiro estudada.

### 3.3.3 Haplogrupos europeus

Na amostra da população do Rio de Janeiro foram verificados dezesseis indivíduos (14,68%) com ancestralidade materna europeia. Dentre os 8 diferentes haplogrupos verificados neste grupo o H foi o mais frequente e corresponde a 7,34% da amostra populacional estudada. Os distintos haplogrupos europeus verificados na amostra populacional do Rio de Janeiro, e suas respectivas frequências, são mostrados na Tabela 7.

Tabela 7 – Haplogrupos europeus e respectivas frequências na população do Rio de Janeiro

Haplogrupos europeus	n	Frequência
H	8	7,34%
H1a	2	1,83%
H8	1	0,92%
K	1	0,92%
K2a	1	0,92%
N1a	1	0,92%
U5b1b	1	0,92%
U5b2a1	1	0,92%

Os haplogrupos foram determinados com o auxílio do software mtDNAManager (mtmanager.yonsei.ac.kr). As frequências foram calculadas dividindo-se o número de ocorrência de cada haplogrupo pelo número total de indivíduos (109) da amostra populacional do Rio de Janeiro estudada.

### 3.4 Proporção de heteroplasmia de sequência na população do Rio de Janeiro

A frequência de heteroplasmia de sequência nos 109 indivíduos analisados foi de 2,75%. A Figura 9 ilustra uma heteroplasmia de sequência caracterizada pela ocorrência de dois diferentes polimorfismos na posição 16257 da região HVI da sequência de DNA mitocondrial.

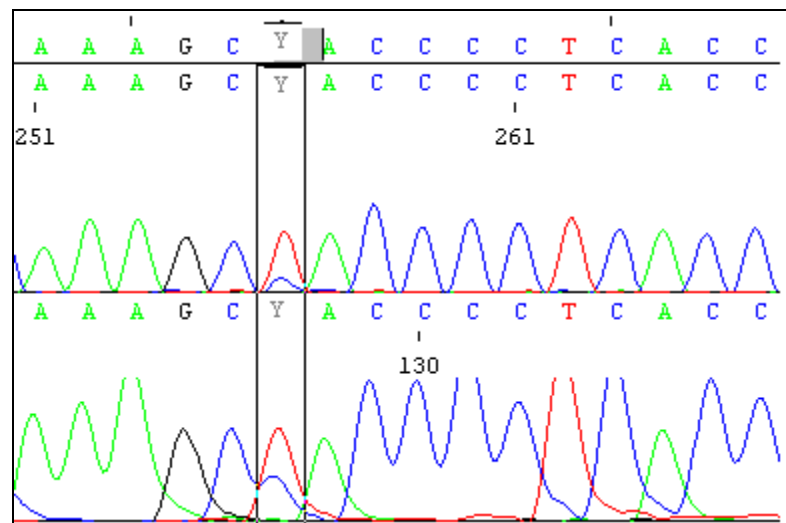


Figura 9 - Eletroferograma da reação de sequenciamento da região HVI do DNA mitocondrial da amostra RJM17. Sequências senso e reverso complemento - orientação 3'-5' - alinhadas no SeqScape®Software - versão 2.5 (Applied Biosystems).

Legenda: **T** = Timina; **A** = Adenina; **G** = Guanina; **C** = Citosina; **Y** = Heteroplasmia  
 No exemplo acima a letra Y representa a coexistência das bases C e T na posição 16257 da região HVI, segundo a norma de notação de heteroplasmia do *International Union of Pure and Applied Chemistry* (IUPAC) (FBI, 2003).

### 3.5 Estimativas de distâncias genéticas

#### 3.5.1 Distâncias genéticas em relação a outras populações brasileiras

A matriz das distâncias genéticas ( $F_{ST}$ ) entre as populações e os respectivos valores de significância ( $p$ ) são mostrados na Tabela 8. Aos valores de  $p$  foi aplicada a correção sequencial de Holm-Bonferroni (Holm, 1979, Rice, 1989). Tal correção consiste na divisão do valor de  $p$  ( $p \leq 0,05$ ) pelo número de haplogrupos observados menos um ( $0,05 / n-1$ ), obtendo-se um valor de  $p$  corrigido. A correção resultou na aceitação da hipótese nula (não diferenciação entre as populações) para valores de  $p$  maiores do que 0,00357.

Nas comparações com populações de outras regiões brasileiras, a população do Rio de Janeiro não apresentou diferenciação genética significativa em relação a Alagoas ( $F_{ST} = 0,01048$ ;  $P = 0,02441$ ), Pará ( $F_{ST} = 0,02378$ ;  $P = 0,01367$ ) e São Paulo ( $F_{ST} = 0,00773$ ;  $P = 0,05469$ ). Por outro lado, as populações de Santa Catarina ( $F_{ST} = 0,08377$ ;  $P = 0,00000$ ) e Rio Grande do Sul ( $F_{ST} = 0,06399$ ;  $P = 0,00000$ ) apresentaram diferenciação genética significativa em relação ao Rio de Janeiro (Tabela 8).

#### 3.5.2 Distâncias genéticas em relação a populações de grupos ancestrais

Quando da comparação com populações atuais de países pertencentes aos grupos ancestrais que contribuíram para a formação da população brasileira, em relação à herança materna, o Rio de Janeiro está mais proximamente relacionado à população Africana (AN), representada neste estudo por Angola, não havendo diferenciação genética significativa entre elas ( $F_{ST} = 0,03159$ ;  $P = 0,01465$ ). Foi observada diferenciação genética significativa entre Rio de Janeiro e as populações ameríndias (TE), representada neste estudo pela população dos índios brasileiros Terenas ( $F_{ST} = 0,14945$ ;  $P = 0,00000$ ) e europeias (PT), representada neste estudo por Portugal ( $F_{ST} = 0,22524$ ;  $P = 0,00000$ ). Também houve diferenciação genética significativa entre o Rio de Janeiro e a população asiática (CH), representada neste estudo pela China ( $F_{ST} = 0,14945$ ;  $P = 0,00000$ ) (Tabela 8).

Tabela 8 - Matriz das distâncias genéticas ( $F_{ST}$ ) para frequências de haplogrupos de DNAmT entre pares de populações. Abaixo da diagonal são mostrados os valores de P e acima da diagonal os valores de  $F_{ST}$ .

	RJ	NA	TE	PT	CH	AL	PA	SC	RS	SP
RJ	*	0.03159	0.14945	0.22524	0.14945	0.01048	0.02378	0.08377	0.06399	0.00773
AN	♦ 0.01465+- 0.0034	*	0.26905	0.31720	0.22406	0.06162	0.10851	0.16654	0.14865	0.06146
TE	0.00000+- 0.0000	0.00000+- 0.0000	*	0.34375	0.14021	0.10817	0.07073	0.17058	0.12400	0.13073
PT	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	*	0.20979	0.17225	0.25653	0.05694	0.13116	0.14578
CH	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	*	0.10594	0.10403	0.10514	0.08633	0.10745
AL	♦ 0.02441+- 0.0047	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	*	0.02276	0.04528	0.02948	0.00311
PA	♦ 0.01367+- 0.0034	0.00000+- 0.0000	0.00098+- 0.0010	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	*	0.09485	0.05193	0.02428
SC	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	*	0.01141	0.03183
RS	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.04688+- 0.0079	*	0.02628
SP	♦ 0.05469+- 0.0076	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.14062+- 0.0142	0.00293+- 0.0016	0.00098+- 0.0010	0.00000+- 0.0000	*

Foi gerada uma matriz de distâncias genéticas ( $F_{ST}$ ) entre as populações por meio do programa Arlequin.

A população do Rio de Janeiro é representada pela sigla RJ. AL- Alagoas; AN- Angola (África); SP- São Paulo; RS- Rio Grande do Sul; CH – China (Ásia); TE - Terenas (Ameríndios); PA - Pará; SC - Santa Catarina; PT – Portugal (Europa).

♦ Valores de p não significativos após correção de Bonferroni (aceitação da hipótese nula, de não diferenciação,  $p \geq 0,00357$ ).

Para melhor observação da posição relativa das populações no espaço genético em duas dimensões, foi feito o escalonamento multidimensional (MDS) (Figura 10) da matriz de distâncias entre pares de populações (Tabela 8). A medida de stress, utilizada para mensurar a qualidade do ajuste da análise do MDS, foi de 0,11235.

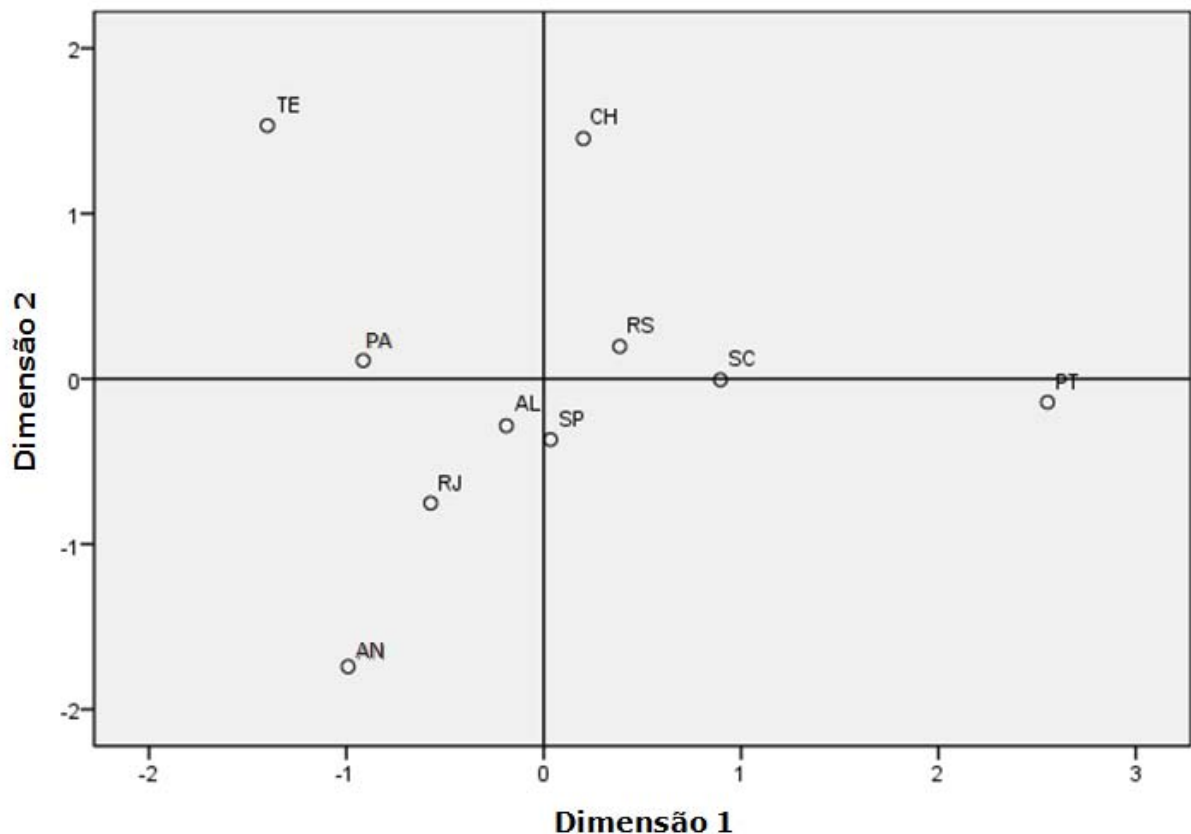


Figura 10 - Projeção em duas dimensões do MDS das distâncias genéticas ( $F_{ST}$ ) entre as 10 populações analisadas (valor de stress = 0,11235).

Foi gerada uma matriz de distâncias genéticas ( $F_{ST}$ ) entre as populações, por meio do programa Arlequin. Esta matriz foi baseada nas frequências dos haplogrupos verificados nas populações estudadas. A partir desta matriz foi realizada uma análise multidimensional (MDS), com auxílio do software STATISTICA v.8.0 e gerado um gráfico em duas dimensões. Os círculos representam as populações. A população do Rio de Janeiro é representada pela sigla RJ. AL- Alagoas; AN- Angola (África); SP- São Paulo; RS- Rio Grande do Sul; CH – China (Ásia); TE - Terenas (Ameríndios); PA - Pará; SC - Santa Catarina; PT – Portugal (Europa).

Os resultados da análise por MDS (mostrados na Figura 10) possibilitam observar um agrupamento formado pelas populações brasileiras Pará, Rio de Janeiro, Alagoas, São Paulo, Rio Grande do Sul e Santa Catarina localizado entre as populações ancestrais (Portugal, Angola e Terenas). Em relação a este agrupamento, a população do Rio de Janeiro situa-se mais próxima das populações brasileiras de Alagoas e de São Paulo e da população ancestral de Angola. As populações ancestrais: Portugal, Angola, China e Terenas mostram-se distantes uma das outras.



## 4 DISCUSSÃO

Os diferentes polimorfismos existentes no DNA, classificados de acordo com a sua natureza molecular e localização no genoma, podem ser utilizados com diferentes objetivos na genética forense e populacional. Aqueles presentes em cromossomos autossômicos são marcadores de escolha para discriminação de indivíduos, apesar de também serem úteis como marcadores informativos de ancestralidade (AIMs) (Shriver *et al.*, 2003).

Os polimorfismos uniparentais maternos (DNA mitocondrial) e paternos (região não recombinante do cromossomo Y) localizam-se em segmentos do DNA que não sofrem recombinação, caracterizando-se como marcadores de linhagens. Este material genético é transmitido às gerações seguintes permanecendo inalterado nas matrilineagens e patrilineagens na dependência exclusiva da ocorrência de mutação (Underhill & Kivisild, 2007).

Características específicas do DNA mitocondrial, tais como o elevado número de cópias por célula, herança exclusivamente materna, ausência de recombinação e uma taxa mutacional geralmente maior do que a encontrada no DNA nuclear, fizeram com que, nos últimos 30 anos, (Brown, 1980) esta molécula se tornasse fundamental para a realização de estudos em identificação forense, genética de populações e estudos filogenéticos (Cann *et al.*, 1987, Krings *et al.*, 1997, Ingman *et al.*, 2000, Underhill & Kivisild, 2007).

A partir dos estudos sobre evolução humana realizados através da análise do DNAm, identificou-se que as linhagens matrilineares estão relacionadas com regiões geográficas específicas, o que possibilita a reconstrução da história demográfica, migrações e determinação da ancestralidade genética de indivíduos e grupos humanos (Hunemeier *et al.*, 2007, Barbosa *et al.*, 2008, Bobillo *et al.*, 2010, Paneto *et al.*, 2011).

Assim como ocorre em outros estados brasileiros, do ponto de vista de sua formação, o Rio de Janeiro possui uma população bastante heterogênea. Na história do estado, que abriga a cidade que permaneceu por quase 200 anos como a capital do Brasil, pode-se identificar diversos fluxos migratórios, os quais contribuíram para o elevado grau de miscigenação da sua população atual.

Até meados do século XVI, a região permaneceu habitada exclusivamente pelos nativos. Em 1565, foi fundada a cidade de São Sebastião do Rio de Janeiro,

como estratégia para efetivar a presença portuguesa na região e neutralizar a ameaça de invasores franceses. A partir de então, destacou-se economicamente abrigando atividades de alto valor econômico, na época, como a cultura da cana e do café (Peixoto, 1966).

Em 1763, a capital do Brasil foi transferida de Salvador para o Rio de Janeiro, o que permaneceu até 1960. A descoberta e a exploração de ouro em Minas Gerais aumentaram a importância do Rio de Janeiro, principalmente no tocante a atividade portuária. A cidade abrigava um porto que funcionava como tradicional entrada de escravos africanos para dar suporte às atividades econômicas da região e do país. Assim como em outras cidades do país, a partir de 1870, o Rio de Janeiro começou a receber diversos outros fluxos de imigrantes, a exemplo dos italianos, espanhóis, árabes e asiáticos (Peixoto, 1966, Reis, 2000).

#### 4.1 Diversidade genética e haplotípica

O presente estudo avaliou a diversidade das linhagens matrilineares da população atual do Rio de Janeiro através da análise dos polimorfismos verificados nas sequências HVI e HVII da região controle do DNA mitocondrial de 109 indivíduos não relacionados geneticamente.

Na população do Rio de Janeiro, os parâmetros de diversidade das regiões hipervariáveis HVI e HVII foram analisados tanto separadamente como em conjunto e são apresentados na Tabela 4. Tais parâmetros de diversidade do DNA mitocondrial são importantes no tocante à utilização da molécula em genética forense e populacional (Holland & Huffine, 2001).

Em nosso estudo, a partir de um total de 142 diferentes *loci* polimórficos observados em HVI e HVII foram identificados 102 haplótipos distintos, conforme pode ser observado na Tabela 3. Quatorze indivíduos compartilharam, aos pares, sete distintos haplótipos enquanto 95 haplótipos ocorreram apenas uma vez. A frequência de cada haplótipo único foi de 0,00917.

A região HVI apresentou 67,6% dos polimorfismos observados e o maior número médio de diferenças aos pares, resultados esses muito provavelmente relacionados ao maior número de posições polimórficas verificadas na região em

relação à HVII. Na região HVI, foram verificados 96 polimorfismos enquanto que na região HVII foram verificados 46.

Verificou-se 91 haplótipos distintos em HVI e 73 em HVII, sendo as diversidades haplotípicas  $0.9963 \pm 0.0019$  e  $0.9886 \pm 0.0035$ , respectivamente. Por outro lado, a diversidade nucleotídica foi maior em HVII, refletindo uma maior diversidade de bases em seus *loci* polimórficos. Valores semelhantes aos observados para a população do Rio de Janeiro foram verificados em estudos envolvendo outras populações brasileiras (Barbosa *et al.*, 2008, Palencia *et al.*, 2010, Paneto *et al.*, 2011).

A diversidade haplotípica total (HVI + HVII) foi de  $0.9988 \pm 0.0014$ , demonstrando que, na população do Rio de Janeiro, o DNA mitocondrial é muito informativo para fins de identificação forense face à elevada probabilidade de se diferenciar dois indivíduos pelas suas linhagens maternas. Em populações nativas americanas da América do Sul, valores menores de diversidade haplotípica são observados, a exemplo do grupo indígena Cayapa do Equador (0,860) (Vona *et al.*, 2005). Esses grupos não apresentam importante miscigenação, o que justifica valores menores de diversidade haplotípica em relação aos encontrados em populações abertas como a do Rio de Janeiro, alvo desse estudo.

Em relação às linhagens maternas, quando a diversidade haplotípica da população do Rio de Janeiro (0,9988) é comparada às de outras populações que também apresentam níveis elevados de miscigenação, são encontrados valores semelhantes àqueles verificados para tais populações. Os valores observados em estudos com as populações da região central da Argentina, da cidade de Caracas, na Venezuela, dos estados de Santa Catarina e Alagoas, no Brasil foram 0,937, 0,9863, 0,9930 e 0,9975, respectivamente (Vona *et al.*, 2005, Barbosa *et al.*, 2008, Lander *et al.*, 2008, Palencia *et al.*, 2010).

Em relação às linhagens de DNA mitocondrial, estes achados sugerem que a população do Rio de Janeiro recebeu um maior fluxo genético originário da África, fixando maior número de distintas linhagens até os dias atuais. Em contrapartida, em outras regiões, a menor diversidade haplotípica pode ser reflexo da não entrada ou eliminação de certos haplótipos naquelas regiões. Cabe ainda ressaltar que a variabilidade das linhagens maternas tende a ser menor em populações humanas tradicionalmente isoladas, provavelmente por causa do fluxo gênico restrito e altos níveis de endogamia durante longos períodos. O poder de discriminação do DNA

mitocondrial em procedimentos forenses é limitado neste tipo de população comparado às populações cujos fluxos gênicos não foram, ou são, restritivos (Alfonso-Sanchez *et al.*, 2006).

A alta diversidade haplotípica do DNA mitocondrial pode ser explicada pelo fato de a população atual do Rio de Janeiro ter sido formada por grupos de origens geográficas diversas e ainda distintas quanto às suas ancestralidades maternas, os quais não sendo reprodutivamente discriminados originaram uma população altamente miscigenada, geneticamente diversa. Quanto mais intensa for a miscigenação ocorrida no processo de formação de uma população mais elevado será o valor de diversidade haplotípica encontrado nessa população.

A diversidade do genoma mitocondrial humano é o resultado da acumulação de mutações sucessivas nas linhagens matrilineares enquanto, durante e após a saída dos humanos modernos da África. Com o tempo, o processo de divergência molecular do DNA mitocondrial deu origem a grupos monofiléticos, conhecidos como haplogrupos (Torrioni *et al.*, 1993), que estão diretamente relacionados com grupos específicos da população e/ou áreas geográficas (Achilli *et al.*, 2005).

É sabido que ocorrem regiões do planeta colonizadas por grupos humanos de diferentes origens geográficas, as quais foram, portanto, povoadas, desde as suas bases, por distintas linhagens genéticas femininas. É importante notar que uma vez não tendo ocorrido discriminação positiva ou negativa daquelas linhagens a população atual refletirá as proporções do DNAm<sub>t</sub> na população base. Em relação à formação da população brasileira, há de se considerar a sua origem tri híbrida no que se refere ao fluxo gênico restrito entre indivíduos de origem nativa americana, europeia (Portugal) e africana. O Brasil é conhecido por ter uma das populações mais heterogêneas do mundo, resultado de um processo de colonização caracterizado pela contribuição étnica de colonos europeus, escravos africanos e nativos americanos ao longo dos últimos cinco séculos (Ribeiro, 1995, Reis, 2000).

## 4.2 Frequências de Heteroplasmias

A heteroplasmia é caracterizada pela presença de dois ou mais tipos de DNA mitocondrial em células e tecidos de um mesmo indivíduo, tendo sido descritas duas formas de heteroplasmias: as de sequência e as de comprimento (Pai *et al.*, 2006). A estratégia adotada no presente estudo, para confirmar ou não a ocorrência de heteroplasmias, se baseia em recomendações internacionais a fim de se evitar erros nas análises de sequências de DNA mitocondrial e, em consequência, na construção de bancos internacionais de haplótipos de DNA mitocondrial (Prieto *et al.*, 2008).

Foi verificada uma frequência de 2,75% de heteroplasmia de sequência na amostra populacional do Rio de Janeiro. A frequência verificada na literatura para este tipo de heteroplasmia é de 2 a 8% (Holland & Parsons, 1999). Foi verificado que a frequência de heteroplasmia de sequência pode variar entre diferentes tecidos de um mesmo indivíduo (Calloway *et al.*, 2000); A frequência de heteroplasmia foi verificada ser maior em pêlo do que no sangue (Tully *et al.*, 2004). Uma frequência de 1,13 % do mesmo tipo de heteroplasmia foi observada em amostras de sangue de 88 indivíduos da cidade de São Paulo enquanto que em amostras de cabelo dos mesmos indivíduos a frequência foi de 10,5% (Paneto *et al.*, 2007). Este tipo de heteroplasmia é um evento raro sendo significativo em interpretações forenses (Stewart *et al.*, 2001).

As heteroplasmias de comprimento são observadas frequentemente no DNA mitocondrial, sendo as múltiplas cópias da macromolécula diferenciadas entre si devido ao tamanho da região Poli C, geradas por falhas da DNA polimerase. Na comunidade científica, as mesmas não são consideradas em análises forenses, pois não existem estudos claros sobre a sua taxa de ocorrência (Lutz-Bonengel *et al.*, 2004, Forster *et al.*, 2010). A frequência de heteroplasmia de comprimento não foi avaliada no presente estudo, uma vez que, a técnica de sequenciamento utilizada não é indicada para a avaliação do número de citosinas nas regiões poli C devido à dificuldade de interpretação dos resultados (Lutz-Bonengel *et al.*, 2004).

### 4.3 Frequência e origem geográfica dos haplogrupos verificados na população do Rio de Janeiro

A população brasileira apresenta ancestralidade tri-híbrida formada a partir de cruzamento entre brancos (europeus), negros (africanos) e índios (nativos americanos), os quais ocorreram de maneira assimétrica e diferenciada de acordo com a região do país. Diversos estudos acerca da ancestralidade materna da população brasileira realizados pela análise do DNAmT corroboram os registros históricos que relatam diferenças nos padrões de miscigenação levada a termo no processo, de colonização e povoamento do país. Estas diferenças foram motivadas principalmente pelas diferentes atividades econômicas desenvolvidas nas diversas regiões do país. A descoberta de ouro e de diamante na região sudeste nos séculos XVII e XVIII, por exemplo, intensificou o comércio e a utilização de mão de obra escrava nesta região. No nordeste, devido ao cultivo de cana de açúcar, também se verificou um grande aporte de escravos provenientes da África. Já na região sul do país, atividades de cultivo de trigo e criação de gado eram predominantes e nessa região não se registrou uma entrada maciça de escravos conforme ocorreu nas regiões sudeste e nordeste (Gomes, 2007).

Estudo pioneiro realizado no ano de 2000 com indivíduos brancos de diferentes regiões do Brasil buscou identificar as diferentes proporções de ancestralidades maternas nessas regiões. Os pesquisadores verificaram que as proporções das ancestralidades maternas se alteravam de acordo com a região estudada. Na região norte foi verificada uma ocorrência mais elevada de linhagens mitocondriais ameríndias, nas regiões nordeste e sudeste, as linhagens ancestrais africanas eram as mais frequentes enquanto na região sul as matrilineagens europeias predominavam (Alves-Silva *et al.*, 2000).

No que se refere à origem geográfica dos haplogrupos do DNAmT verificados no presente estudo nos 109 indivíduos analisados, observou-se 66, 27 e 16 matrilineagens associadas às ancestralidades africana, nativa americana e europeia, respectivamente. Portanto, na amostra da população do Rio de Janeiro 60,55% dos indivíduos apresentaram ancestralidade africana, 24,77% ameríndia e 14,68% europeia (Figura 8). Na Tabela 9 são listados estudos, mais recentes, e seus respectivos resultados, acerca da ancestralidade materna em diferentes estados do território brasileiro.

Tabela 9 – Proporção de ancestralidade materna ameríndia, europeia e africana em diferentes estados do Brasil estimados a partir do DNA mitocondrial.

Estado	Ancestralidade			Referências
	Ameríndia	Europeia	Africana	
Rio de Janeiro	24,77%	14,68%	60,55%	Presente estudo
São Paulo	26,1%	27,3%	46,6%	(Paneto <i>et al.</i> , 2011)
Santa Catarina	21%	64%	15%	(Palencia <i>et al.</i> , 2010)
Rio Grande do Sul	36%	48%	16%	(Marrero <i>et al.</i> , 2005)
Alagoas	33%	21%	44%	(Barbosa <i>et al.</i> , 2008)
Pará	66%	5%	29%	(Feio-Dos-Santos <i>et al.</i> , 2006)

Na comparação dos nossos resultados com os de outras regiões do Brasil, é possível observar que as proporções das diferentes ancestralidades maternas se alteram, fato este que, aparentemente, reflete a história da colonização regional. Como verificado na amostra do Rio de Janeiro, a predominância do componente materno africano é também observada para uma amostra de 160 indivíduos residentes na grande São Paulo. Dentre estes verificou-se 46,6% de indivíduos com matrilineagens africanas, enquanto 27,3% e 26,1% apresentavam ancestralidades europeia e ameríndia, respectivamente (Paneto *et al.*, 2011). Conquanto sejam estados da mesma região geográfica, é observada diferença no que se refere à distribuição de origem das linhagens maternas. Embora a ancestralidade africana seja mais elevada em ambas populações, verifica-se em São Paulo uma maior proporção de linhagens maternas europeias enquanto a frequência de indivíduos com ancestralidade materna ameríndia se mantém em percentuais muito próximos em ambos os estados.

O estado de Alagoas apresenta uma distribuição de matrilineagens semelhante àquela encontrada no Rio de Janeiro. Em um estudo envolvendo 167 indivíduos, as proporções verificadas de linhagens maternas africana, nativa americana e europeia, foram 44%, 33% e 21%, respectivamente (Barbosa *et al.*, 2008). A predominância de linhagens ancestrais maternas africanas nas regiões sudeste e nordeste é justificada pela semelhança dos processos de ocupação e povoamento dessas regiões, que tiveram uma forte influência do tráfico de escravos por mais de 250 anos. Os escravos eram a principal mão de obra utilizada nas lavouras de algodão, fumo e cana de açúcar, e também nas minas de ouro e prata. O sudeste e nordeste foram os grandes centros destas atividades econômicas;

dados do século XIX mostram que os negros representavam mais de dois terços da população brasileira (Gomes, 2007).

No tocante às ancestralidades maternas africana e europeia, uma situação diametralmente oposta é observada em estados da região sul do país. Em Santa Catarina, foram verificadas proporções da ordem de 63,7%, 21,3% e 15% para as ancestralidades maternas europeia, nativa americana e africana, respectivamente (Palencia *et al.*, 2010). No Rio Grande do Sul, tais proporções são da ordem de 48%, 36% e 16% (Marrero *et al.*, 2005). O maior aporte de linhagens mitocondriais europeias na região Sul é justificado pela colonização desta região ter sido predominantemente europeia, com um menor emprego da mão de obra escrava africana. O percentual, de linhagens ancestrais maternas nativas americanas, observado no Rio Grande do Sul é justificado pela contribuição significativa dos ameríndios na formação da população da região durante o período colonial (Kern *et al.*, 1993).

No estado do Pará foram verificadas proporções de 66%, 5% e 29% para as ancestralidades maternas ameríndia, europeia e africana, respectivamente (Feio-Dos-Santos *et al.*, 2006). Durante o período colonial, os índios do norte do país buscavam se refugiar dos colonos em terras do interior. No século XVII, os bandeirantes avançaram para o interior do território brasileiro provocando alterações das fronteiras e, em parte, foram responsáveis pelo deslocamento dos índios para regiões que ainda preservavam a mata nativa, longe do litoral (Ribeiro, 1995). A predominância de linhagens maternas mitocondriais ameríndias na região norte é justificada pelo fato de a densa floresta amazônica ter dificultado o processo de colonização portuguesa e, em consequência, ter possibilitado uma maior resistência e isolamento dos povos nativos americanos da região.

Os dados referentes à ancestralidade paterna do estado do Rio de Janeiro são bem distintos daqueles verificados neste nosso estudo para a ancestralidade materna. Estudo realizado através da análise de marcadores SNP dos cromossomos Y em 127 indivíduos nascidos no Rio de Janeiro, mostrou que cerca de 88% dos indivíduos portavam haplogrupos de origem europeia, 8,7% ancestralidade paterna africana enquanto que 1,6% se associavam à herança paterna nativa americana (Silva *et al.*, 2006).

Quando os dados relativos às ancestralidades maternas e paternas no estado do Rio de Janeiro são analisados em conjunto, evidencia-se um padrão de



povoamento calçado na reprodução assimétrica representada pela forte predominância da paternidade do homem europeu e maternidade de mulheres africanas e indígenas, que se encontra em concordância com a história da colonização dessa região do país no período posterior ao “descobrimento” do Brasil pelos portugueses. Em diversos registros históricos, é possível encontrar relatos a respeito das motivações que favoreceram este padrão reprodutivo dos nossos ascendentes. Um dos principais motivos foi o significativamente menor número de homens portugueses vindos para o Brasil em relação ao de mulheres portuguesas. Em 1552, em carta ao rei D. João III, o padre Manuel da Nóbrega relata a falta de mulheres brancas no país e pede que elas sejam enviadas para que os homens “casem e vivam [...] apartados dos pecados em que agora vivem” (Leite, 1954).

Outro fator crucial que contribuiu para o padrão de cruzamentos assimétricos foram os esforços relacionados com a expansão territorial e a revogação em 1750 do Tratado de Tordesilhas (1494) com a instituição do Tratado de Madri. A partir de então, a posse da terra passava a ser de quem de fato a ocupasse. Deste modo, Portugal possuía interesses em elevar a taxa de povoamento do Brasil, garantindo uma maior ocupação territorial e, em consequência, a posse da terra (Gomes, 2007). Portanto, a coroa portuguesa, que tolerava relacionamentos entre portugueses e índias desde o início da colonização do Brasil, passou a estimular ativamente “casamentos” desse tipo, o que foi inclusive, expresso por meio de um Alvará de Lei promulgado em 4 de abril de 1755, pelo Marquês do Pombal. Os relacionamentos entre homens portugueses e mulheres africanas também persistiram em níveis elevados conforme relatou o historiador Charles Boxer apud Gomes (2007), “a frequência de prostituição de escravas e [...] o duplo padrão de castidade como o que existia entre marido e esposas, concorriam para uma grande quantidade de miscigenação entre homens brancos e mulheres de cor”.

Entre os séculos XVI e XIX, cerca de 10 milhões de escravos africanos foram vendidos para as Américas, 40% destes vieram para o Brasil entrando no país, principalmente, pelo porto do Rio de Janeiro que foi o maior importador de escravos. O historiador Manolo Garcia Florentino estima que durante o século XVIII 850.000 escravos, o equivalente à metade de todos os negros trazidos para o Brasil nesse período, desembarcaram no porto do Rio de Janeiro (apud Gomes, 2007). Predominantemente, os escravos eram oriundos das regiões centro-ocidental e central-Atlântico da África, provavelmente devido à posição geográfica dessas

regiões. A partir de dados relativos aos anos de 1822 a 1833 em relação a origem da população negra do Rio de Janeiro verifica-se que, 86% dos escravos eram recém chegados do continente africano e 14% nascidos ou já residentes na América portuguesa, o que mostra a forte entrada de escravos africanos em nosso país mesmo após o regresso de D. João VI para Portugal (Klein, 2002).

Portanto, o intenso aporte de africanos no estado do Rio de Janeiro e as políticas voltadas para a promoção de um rápido aumento da população brasileira durante o período colonial podem explicar o fato de cerca de 60% das matrilineagens verificadas neste estudo serem de origem Africana.

Entre as linhagens de DNA mitocondrial africano, os subgrupos mais comuns ao longo do grupo L são característicos da África Subsaariana (Salas *et al.*, 2002). Em nossos dados, os haplogrupos originados do L3 (L3b, L3b1\*, L3b1a, L3b1b, L3d, L3d1c, L3d1d, L3d3, L3e1, L3e1a, L3e1d, L3e2a, L3e2b, L3e2b\*, L3e3, L3e4, L3f2, L3h1b) foram os mais frequentes entre as linhagens maternas estando presentes em 29,35% dos indivíduos da amostra. Haplogrupos do L2 (L2a, L2a1, L2a1\*, L2a1a, L2a1b, L2b, L2b1), foram verificados em 13,76% da amostra estudada, haplogrupos do grupo L1 (L1b1, L1b1\*, L1c, L1c1, L1c2, L1c3b) em 11,01%, enquanto que os dos grupos L0 (L0a1b e L0a2) e L4 (Lab2) foram observados em 4,58% e 1,83% dos indivíduos da amostra da população do Rio de Janeiro, respectivamente.

Todos os haplogrupos africanos observados na população do Rio de Janeiro são característicos de regiões que eram tradicionais exportadoras de escravos para a América colonial. O haplogrupo L2a é o mais comum e mais amplamente distribuído na África sub-saariana. A sua ampla distribuição no continente africano torna difícil a identificação da origem geográfica destas linhagens embora, recentes estudos tenham demonstrado que quase todos os haplogrupos derivados do L2a são oriundos do oeste (Salas *et al.*, 2004).

Os haplogrupos relacionados ao L3a encontrados nas Américas têm suas origens geográficas nas regiões centro-oeste e sudeste do continente africano. Os relacionados ao L3e são comuns em Moçambique, Angola, São Tomé e Bioko. As linhagens do haplogrupo L3e1 encontradas nas Américas possuem como provável origem a costa do centro-oeste da África, entre Camarões e Angola, tradicionais portos de exportação de escravos para o Brasil. Em um estudo anterior acerca da origem dos africanos do estado do Rio de Janeiro foi verificado que as linhagens de

DNA mitocondrial tem origem na região centro-ocidental da África (Alves-Silva *et al.*, 2000), os haplogrupos são similares aos verificados no presente estudo.

As proporções de haplogrupos africanos (detalhadas na Tabela 5) verificadas neste presente estudo são concordantes com as observadas para outras populações da América do Sul, incluindo as relativas a do Brasil (Alves-Silva *et al.*, 2000, Salas *et al.*, 2004, Hunemeier *et al.*, 2007, Paneto *et al.*, 2011).

A alta diversidade de haplogrupos africanos é explicada por esta ser reconhecidamente a mais ancestral das populações humanas. A diversidade genética é um parâmetro importante que fornece informações sobre a riqueza genética da população sendo também indicativo de idade populacional, observando-se para as mais antigas o acúmulo de um maior número de mutações (Ingman *et al.*, 2000).

Em relação às matrilineagens nativas americanas, na amostra da população do Rio de Janeiro, o haplogrupo mais frequente foi o B4 verificado em 10,09% dos indivíduos, seguido do A2, C1 e pelo conjunto A, C, C1d e D4/G com 4,59, 3,67%, e 6,42%, respectivamente (Tabela 6). Uma distribuição semelhante de haplogrupos ameríndios também foi verificada em estudos realizados nas regiões sul e sudeste do Brasil (Alves-Silva *et al.*, 2000, Gonçalves *et al.*, 2010). Os haplogrupos verificados no presente estudo (A, B, C e D) são característicos das comunidades ancestrais nativas americanas (Forster *et al.*, 1996).

O haplogrupo B é encontrado com uma frequência de 90% no grupo indígena Aché no Paraguai (Schmitt *et al.*, 2004). Em relação ao haplogrupo A foram verificadas proporções de 82%, 92% e 50%, grupos indígenas Nandeva, Kaiowá e M'byá, respectivamente, habitantes das regiões central e sul do Brasil (Marrero *et al.*, 2007). O haplogrupo A2 foi verificado em 40,5% de uma amostra de 121 indivíduos da comunidade Guarani-M'byá, da Argentina. Os grupos indígenas M'byá, Nandeva, Kaiowá e Aché pertencem à família linguística Tupi-Guarani. Segundo dados do Museu do Índio (FUNAI, 2004), comunidades remanescente dos índios M'byá são encontradas na região de Paraty-Mirim no Rio de Janeiro. O haplogrupo C foi verificado com uma frequência de 53% em índios Kaingang do Rio Grande do Sul (Marrero *et al.*, 2007); registros históricos descrevem um intenso movimento migratório desse grupo indígena, da família Jê, entre as regiões sul e sudeste do Brasil devido ao contato com os colonizadores europeus (Jacobus & Kern, 1991).

A proporção dos haplogrupos nativos americanos observados na amostra populacional do Rio de Janeiro sugere uma forte correlação com aquela relatada em estudos sobre a ancestralidade materna de indivíduos das comunidades indígenas do tronco linguístico Tupi-Guarani (Marrero *et al.*, 2007). No início do século XVI, os principais grupos indígenas que habitavam a região atualmente designada como Rio de Janeiro pertenciam ao tronco linguístico Tupi e Jê. Entre os povos pertencentes ao tronco linguístico Tupi, a família denominada Tupi-Guarani era a mais numerosa (Marconi & Presotto, 1998).

Em relação à ancestralidade materna europeia foram verificados que 7,34% dos indivíduos pertencem ao haplogrupo H (Tabela 7), que juntamente com os demais haplogrupos verificados são representativos dos principais grupos familiares europeus, observando-se ainda uma forte correlação entre os haplogrupos europeus encontrados na amostra da população do Rio de Janeiro e aqueles observados na Península Ibérica (Torrioni *et al.*, 1996, Pereira *et al.*, 2000).

Na realização do presente estudo, as naturalidades de todos os 109 indivíduos selecionados ao acaso e residentes no Rio de Janeiro foram verificadas. A Tabela 10 ilustra as proporções de indivíduos de diferentes naturalidades observadas na amostra populacional.

Tabela 10 – Naturalidade dos indivíduos da população do Rio de Janeiro selecionados para o presente estudo.

Naturalidade	Número de Indivíduos	%
Rio de Janeiro	86	78,90%
Minas gerais	4	3,67%
São Paulo	1	0,92%
Bahia	5	4,59%
Pernambuco	2	1,83%
Rio Grande do Norte	1	0,92%
Paraíba	4	3,67%
Alagoas	1	0,92%
Ceará	1	0,92%
Sergipe	1	0,92%
Maranhão	1	0,92%
Pará	1	0,92%
Paraná	1	0,92%

Entre os 23 indivíduos não naturais do Rio de Janeiro, componentes da região nordeste, com cerca de 70%, representam a maior parcela dos imigrantes presentes na amostra seguidos de 17,5% de nativos de Minas Gerais, estado também da região sudeste, fato justificável pela recente forte atração que o estado do Rio de Janeiro exerceu sobre estas populações. A partir de meados do século XX, verificou-se um aumento significativo da migração de componentes de diversos estados para o Rio de Janeiro, impulsionados pela desvantagem econômica e social da sua região de nascimento em relação às regiões sudeste e sul, observando-se a transferência de mão de obra barata para atender às necessidades de desenvolvimento dos grandes centros urbanos localizados nessas regiões do país.

Com o objetivo de verificar se a presença de indivíduos não nascidos no Rio de Janeiro estariam a influenciar significativamente as proporções das ancestralidades maternas verificadas na amostra da população do Rio de Janeiro, estas foram recalculadas após os 23 indivíduos não naturais do Rio de Janeiro terem sido excluídos da amostra, sendo os dados apresentados na Tabela 11. Os resultados mostram a não ocorrência de diferenças significativas entre as proporções dos grupos ancestrais observadas para as populações residente e natural do Rio de Janeiro. Tal resultado, portanto, sugere que as recentes ondas migratórias não alteraram as proporções de linhagens matrilineares muito provavelmente devido a semelhança entre os processos de colonização verificados no sudeste e nordeste brasileiros.

Tabela 11 – Proporções das distintas ancestralidades maternas em indivíduos residentes e naturais do Rio de Janeiro.

Ancestralidade	RJ Residentes	Rj Naturais
Nativa Americana	24,77%	22,09%
Europeus	14,68%	13,95%
Africanos	60,55%	63,95%

#### 4.4 Distâncias genéticas

A análise de distâncias genéticas envolveu a comparação par a par de 10 populações (Tabela 8). Em termos metodológicos, quando ocorrem múltiplos testes de significância independentes, é recomendável utilizar algum tipo de correção do nível de significância para diminuir a chance de rejeição da hipótese nula (não diferenciação) quando ela é verdadeira (erro tipo I). No presente estudo, foi utilizada a correção sequencial de Holm-Bonferroni, pois este método não aumenta os erros do tipo II (aceitar como verdadeira uma hipótese nula falsa) (Holm, 1979, Rice, 1989).

Em relação a ancestralidade materna, os resultados das análises de distância genética,  $F_{ST}$ , entre pares de populações mostram que a população do Rio de Janeiro diferencia-se geneticamente de forma significativa das populações brasileiras de Santa Catarina ( $F_{ST} = 0,08377$ ,  $P = 0,00000$ ) e Rio Grande do Sul ( $F_{ST} = 0,06399$ ,  $P = 0,00000$ ). A diferenciação genética significativa entre populações do Rio de Janeiro e as de Santa Catarina e Rio Grande do Sul pode ser atribuída aos diferentes processos de colonização que ocorreram durante o povoamento dessas regiões. Em relação ao Rio de Janeiro, a região sul do país, onde estão localizados os estados de Santa Catarina e Rio Grande do Sul, recebeu um fluxo diferenciado de colonizadores. A população europeia predominou nesta região enquanto que o Rio de Janeiro recebeu um intenso contingente de populações africanas (Gomes, 2007).

Para a população do Rio de Janeiro não se observou diferenciação genética em relação às populações de São Paulo ( $F_{ST} = 0,00773$ ,  $P = 0,05469$ ), Alagoas ( $F_{ST} = 0,01048$ ,  $P = 0,02441$ ) e Pará ( $F_{ST} = 0,02378$ ,  $P = 0,01367$ ). A distribuição dos haplogrupos do DNAmT no Rio de Janeiro se assemelha àquelas encontradas nas populações de São Paulo e Alagoas, principalmente no tocante a predominância de haplogrupos de DNAmT de grupos ancestrais africanos. Tal qual o observado para o Rio de Janeiro, estes estados receberam um intenso fluxo migratório de escravos africanos durante o período colonial (Gomes, 2007). A não diferenciação genética entre Rio de Janeiro e Pará, provavelmente se justifica pela semelhança no padrão de distribuição dos haplogrupos europeus. Em ambas as populações os haplogrupos europeus representam a menor proporção de ancestralidade observada nas populações, 14,68% e 5 %, nas populações do Rio de Janeiro e Pará,

respectivamente. Nessas regiões, o padrão assimétrico de cruzamentos entre homens europeus e mulheres ameríndias e africanas ocorrido no processo de colonização do Brasil fica evidenciado observando-se a pequena contribuição ancestral de linhagens maternas europeias (Feio-Dos-Santos *et al.*, 2006, Gomes, 2007).

Como consequência da maior proporção de linhagens africanas a população do Rio de Janeiro se assemelha mais à população africana, representada por Angola, do que às outras duas populações ancestrais, não ocorrendo diferenciação genética entre elas ( $F_{ST} = 0,03159$ ,  $P = 0,01465$ ) (Tabela 8). A análise do MDS (Figura 10) mostra a população do Rio de Janeiro mais próxima da população africana (AN) do que das demais populações portuguesa (PT) e Terena (TE).

Estes resultados sugerem que, no que diz respeito às linhagens femininas, o processo de formação da população brasileira apresentou algumas diferenças entre os estados e regiões do país, estando de acordo com o descrito pela literatura (Alves-Silva *et al.*, 2000). Na região sul (Rio Grande do Sul e Santa Catarina) observa-se predominância de linhagens europeias, com uma menor contribuição de linhagens maternas africanas. Nas regiões sudeste e nordeste, representadas pelas populações Rio de Janeiro, São Paulo e Alagoas, as linhagens africanas correspondem a maior parcela de ancestralidade verificada enquanto que na região norte, representada pela população do Pará, ocorreu uma maior contribuição genética de linhagens maternas ameríndias.

O gráfico do MDS (Figura 10) construído com base na matriz de distâncias genéticas ( $F_{ST}$ ) possui valor de stress estimado em 0,11235 indicando uma boa aproximação dos resultados do MDS aos dados da matriz de distâncias (Sturrock & Rocha, 2000). A medida stress é utilizada para mensurar a qualidade do ajuste da análise indicando a proporção da variação das distâncias originais em relação às distâncias previstas pelo MDS (Knoke & Yang, 2008). A projeção mostra que as populações ancestrais encontram-se destacadas entre si, representando um menor fluxo gênico entre elas. É possível observar um agrupamento formado pelas populações do Brasil (Rio de Janeiro, São Paulo, Alagoas, Pará, Rio Grande do Sul e Santa Catarina) entre as populações ancestrais, evidenciando a contribuição daqueles grupos no processo de formação populacional do país. As populações brasileiras deslocam-se no espaço bidimensional em função das contribuições ancestrais mais significativas. Rio Grande do Sul e Santa Catarina estão mais

próximas da população europeia, representada por Portugal, enquanto Rio de Janeiro, Alagoas e São Paulo estão mais próximas da população africana, representada por Angola e a população do Pará mais próxima da população ameríndia, representada pelos Terenas. Tais resultados estão de acordo com o que foi observado para as amostras populacionais brasileiras em relação a proporção de matrilineagens de DNAm (Alves-Silva *et al.*, 2000, Marrero *et al.*, 2005, Feio-Dos-Santos *et al.*, 2006, Barbosa *et al.*, 2008, Palencia *et al.*, 2010, Paneto *et al.*, 2011).



## 5 CONCLUSÃO

- Um total de 142 *loci* polimórficos da região controle do DNA mitocondrial nos segmentos HVI e HVI-II foi contabilizado, tendo sido identificados 102 haplótipos distintos. A diversidade haplotípica foi de  $0.9988 \pm 0.0014$ , sugerindo que a população do Rio de Janeiro foi formada a partir de povos de distintas origens que contribuíram para o estado de miscigenação verificado atualmente.
- A variedade de polimorfismos do DNA mitocondrial se relacionou com uma elevada probabilidade de se discriminar indivíduos em função de suas ancestralidades.
- A população do Rio de Janeiro apresentou 60,55% da sua ancestralidade materna relacionada com haplogrupos africanos, 24,77% associada com haplogrupos nativos americanos e 14,68% associada a haplogrupos europeus.
- O presente estudo e outros realizados em diversas populações de distintas regiões do Brasil confirmam os relatos históricos sobre as diferenças regionais que ocorreram no processo de colonização do Brasil, assim como a formação tri-híbrida da nossa população.
- Através da análise de distâncias genéticas em relação à ancestralidade materna entre as diversas populações comparadas, verificou-se ser a do Rio de Janeiro a mais próxima da população ancestral africana.

## REFERÊNCIAS

- Achilli A, Rengo C, Battaglia V, Pala M, Olivieri A, Fornarino S, et al. Saami and Berbers--an unexpected mitochondrial DNA link. *Am J Hum Genet.* 2005 May;76(5):883-6.
- Alfonso-Sanchez MA, Martinez-Bouzas C, Castro A, Pena JA, Fernandez-Fernandez I, Herrera RJ, et al. Sequence polymorphisms of the mtDNA control region in a human isolate: the Georgians from Swantia. *J Hum Genet.* 2006;51(5):429-39.
- Alves-Silva J, da Silva Santos M, Guimaraes PE, Ferreira AC, Bandelt HJ, Pena SD, et al. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet.* 2000 Aug;67(2):444-61.
- Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. *Nature.* 1981 Apr 9;290(5806):457-65.
- Anderson TD, Ross JP, Roby RK, Lee DA, Holland MM. A validation study for the extraction and analysis of DNA from human nail material and its application to forensic casework. *J Forensic Sci.* 1999 Sep;44(5):1053-6.
- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature genetics.* 1999 Oct;23(2):147.
- Barbosa AB, da Silva LA, Azevedo DA, Balbino VQ, Mauricio-da-Silva L. Mitochondrial DNA control region polymorphism in the population of Alagoas state, north-eastern Brazil. *J Forensic Sci.* 2008 Jan;53(1):142-6.
- Bendall KE, Sykes BC. Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. *Am J Hum Genet.* 1995 Aug;57(2):248-56.
- Bethell L. *Colonial Brazil*: Cambridge Univ Pr 1984.
- Bobillo MC, Zimmermann B, Sala A, Huber G, Rock A, Bandelt HJ, 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.* 2010 Jul;124(4):263-8.
- Bonatto SL, Salzano FM. A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. *Proc Natl Acad Sci U S A.* 1997 Mar 4;94(5):1866-71.
- Brown WM, George M, Jr., Wilson AC. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci U S A.* 1979 Apr;76(4):1967-71.
- Brown WM. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc Natl Acad Sci U S A.* 1980 Jun;77(6):3605-9.

Brown WM, Prager EM, Wang A, Wilson AC. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *Journal of molecular evolution*. 1982;18(4):225-39.

Budowle B, Allard MW, Wilson MR, Chakraborty R. Forensics and mitochondrial DNA: applications, debates, and foundations. *Annu Rev Genomics Hum Genet*. 2003;4:119-41.

Butler JM. *Fundamentals of Forensic DNA Typing*: Academic Press/Elsevier 2009.

Callegari-Jacques SM, Salzano FM. Brazilian Indian/non Indian interactions and their effects. *Ciência e Cultura*. 1999;51:166-74.

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 Apr;66(4):1384-97.

Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. *Nature*. 1987 Jan 1-7;325(6099):31-6.

Carracedo A, Sanchez-Diz P. Forensic DNA-typing technologies: a review. *Methods Mol Biol*. 2005;297:1-12.

de Souza Goes AC, de Carvalho EF, Gomes I, da Silva DA, Gil EH, Amorim A, et al. Population and mutation analysis of 17 Y-STR loci from Rio de Janeiro (Brazil). *Int J Legal Med*. 2005 Mar;119(2):70-6.

Decorte R, Cuppens H, Marynen P, Cassiman JJ. Rapid detection of hypervariable regions by the polymerase chain reaction technique. *DNA and cell biology*. 1990 Jul-Aug;9(6):461-9.

Domingues PM, Gusmao L, da Silva DA, Amorim A, Pereira RW, de Carvalho EF. Sub-Saharan Africa descendents in Rio de Janeiro (Brazil): population and mutational data for 12 Y-STR loci. *Int J Legal Med*. 2007 May;121(3):238-41.

Doolittle RF, Feng DF, Tsang S, Cho G, Little E. Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science*. 1996 Jan 26;271(5248):470-7.

Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online*. 2005;1:47-50.

Eyre-Walker A. Do mitochondria recombine in humans? *Philosophical transactions of the Royal Society of London Series B, Biological sciences*. 2000 Nov 29;355(1403):1573-80.

FBI. FEDERAL BUREAU OF INVESTIGATION - 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: 22/09/11.

Feio-Dos-Santos AC, Carvalho BM, Batista dos Santos SE, Ribeiro-dos-Santos AK. Nucleotide variability of HV-I in admixed population of the Brazilian Amazon Region. *Forensic science international*. 2006 Dec 20;164(2-3):276-7.

Ferreira da Silva IH, Barbosa AG, Azevedo DA, Sanchez-Diz P, Gusmao L, Tavares CC, et al. An X-chromosome pentaplex in two linkage groups: haplotype data in Alagoas and Rio de Janeiro populations from Brazil. *Forensic Sci Int Genet*. 2010 Jul;4(4):e95-100.

Finsterer J. Mitochondriopathies. *European journal of neurology : the official journal of the European Federation of Neurological Societies*. 2004 Mar;11(3):163-86.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular marine biology and biotechnology*. 1994 Oct;3(5):294-9.

Forster L, Forster P, Gurney SM, Spencer M, Huang C, Rohl A, et al. Evaluating length heteroplasmy in the human mitochondrial DNA control region. *Int J Legal Med*. 2010 Mar;124(2):133-42.

Forster P, Harding R, Torroni A, Bandelt HJ. Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet*. 1996 Oct;59(4):935-45.

Franklin RE, Gosling RG. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature*. 1953 Jul 25;172(4369):156-7.

FUNAI. Museu do Índio. Os Guarani no Rio de Janeiro. 2004 [cited 2012; Available from: [http://www.museudoindio.org.br/template\\_01/default.asp?ID\\_S=33&ID\\_M=112](http://www.museudoindio.org.br/template_01/default.asp?ID_S=33&ID_M=112).

Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, et al. Identification of the remains of the Romanov family by DNA analysis. *Nature genetics*. 1994 Feb;6(2):130-5.

Gomes L. 1808: como uma rainha louca, um príncipe medroso e uma corte corrupta enganaram Napoleão e mudaram a história de Portugal e do Brasil: Editora Planeta do Brasil 2007.

Gonçalves VF, Parra FC, Gonçalves-Dornelas H, Rodrigues-Carvalho C, Silva HP, Pena S. Recovering mitochondrial DNA lineages of extinct Amerindian nations in extant homopatric Brazilian populations. *Investig Genet*. 2010;1(1):13.

Hedges SB, Parker PH, Sibley CG, Kumar S. Continental breakup and the ordinal diversification of birds and mammals. *Nature*. 1996 May 16;381(6579):226-9.

Hedman M, Brandstatter A, Pimenoff V, Sistonen P, Palo JU, Parson W, et al. Finnish mitochondrial DNA HVS-I and HVS-II population data. *Forensic science international*. 2007 Oct 25;172(2-3):171-8.

Hillis DM. Inferring complex phylogenies. *Nature*. 1996 Sep 12;383(6596):130-1.

Holder M, Lewis PO. Phylogeny estimation: traditional and Bayesian approaches. *Nature reviews Genetics*. 2003 Apr;4(4):275-84.

Holland MM, Fisher DL, Lee DA, Bryson CK, Weedn VW. Short tandem repeat loci: application to forensic and human remains identification. *Exs*. 1993;67:267-74.

Holland MM, Parsons TJ. Mitochondrial DNA sequence analysis-validation and use for forensic casework. *Forensic Science Review* 1999 1999;11(21).

Holland MM, Huffine EF. Molecular analysis of the human mitochondrial DNA control region for forensic identity testing. *Curr Protoc Hum Genet*. 2001 May;Chapter 14:Unit 14 7.

Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. 1979;6:65-70.

Hunemeier T, Carvalho C, Marrero AR, Salzano FM, Junho Pena SD, Bortolini MC. Niger-Congo speaking populations and the formation of the Brazilian gene pool: mtDNA and Y-chromosome data. *Am J Phys Anthropol*. 2007 Jun;133(2):854-67.

Ingman M, Kaessmann H, Paabo S, Gyllensten U. Mitochondrial genome variation and the origin of modern humans. *Nature*. 2000 Dec 7;408(6813):708-13.

Irwin JA, Saunier JL, Niederstatter H, Strouss KM, Sturk KA, Diegoli TM, et al. Investigation of heteroplasmy in the human mitochondrial DNA control region: a synthesis of observations from more than 5000 global population samples. *Journal of molecular evolution*. 2009 May;68(5):516-27.

Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parsons TJ. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nature genetics*. 1996 Apr;12(4):417-20.

Jacobus AL, Kern AA. *Arqueologia pré-histórica do Rio Grande do Sul: Porto Alegre. Mercado Aberto* 1991.

Jobling MA, Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. *Nature reviews Genetics*. 2003 Aug;4(8):598-612.

Jobling MA, Gill P. Encoded evidence: DNA in forensic analysis. *Nature reviews Genetics*. 2004 Oct;5(10):739-51.

Kern AA, Bern Z, Cohen VRA, Avancini EG, Kemel CLA, Gick PW. *Rio Grande do Sul: continente múltiplo. Porto Alegre, Brazil: Riocell Marprom* 1993.

Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M. Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR methods and applications. 1993 Aug;3(1):13-22.

Klein H. As origens africanas dos escravos brasileiros. São Paulo 2002.

Knoke D, Yang S. Social network analysis. 2nd ed Los Angeles: SAGE Publications. 2008.

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Paabo S. Neandertal DNA sequences and the origin of modern humans. Cell. 1997 Jul 11;90(1):19-30.

Lacan M, Theves C, Amory S, Keyser C, Crubezy E, Salles JP, et al. Detection of the A189G mtDNA heteroplasmic mutation in relation to age in modern and ancient bones. Int J Legal Med. 2009 Mar;123(2):161-7.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860-921.

Lander N, Rojas MG, Chiurillo MA, Ramirez JL. Haplotype diversity in human mitochondrial DNA hypervariable regions I-III in the city of Caracas (Venezuela). Forensic Sci Int Genet. 2008 Sep;2(4):e61-4.

Lee HY, Song I, Ha E, Cho SB, Yang WI, Shin KJ. mtDNAManager: a Web-based tool for the management and quality analysis of mitochondrial DNA control-region sequences. BMC bioinformatics. 2008;9:483.

Lee YS, Kim WY, Ji M, Kim JH, Bhak J. MitoVariome: a variome database of human mitochondrial DNA. BMC Genomics. 2009;10 Suppl 3:S12.

Lehman IR, Zimmerman SB, Adler J, Bessman MJ, Simms ES, Kornberg A. Enzymatic Synthesis of Deoxyribonucleic Acid. V. Chemical Composition of Enzymatically Synthesized Deoxyribonucleic Acid. Proc Natl Acad Sci U S A. 1958 Dec 15;44(12):1191-6.

Leite S. Cartas dos primeiros Jesuítas do Brasil: Comissão do IV Centenário da Cidade de São Paulo 1954.

Lutz-Bonengel S, Sanger T, Pollak S, Szibor R. Different methods to determine length heteroplasmy within the mitochondrial control region. Int J Legal Med. 2004 Oct;118(5):274-81.

Lutz S, Weisser HJ, Heizmann J, Pollak S. A third hypervariable region in the human mitochondrial D-loop. Human genetics. 1997 Dec;101(3):384.

Manta FS, Silva DA, Sanchez-Diz P, Aguiar JI, Cabello PH, Carvalho EF. Terena Amerindian group autosomal STR data: comparison studies with other Brazilian populations. Molecular biology reports. 2011 Sep 24.

Marconi MA, Presotto ZMN. Antropología: uma introdução: 4. ed. Editora Atlas 1998.

Marrero AR, Das Neves Leite FP, De Almeida Carvalho B, Peres LM, Kommers TC, Da Cruz IM, et al. Heterogeneity of the genome ancestry of individuals classified as White in the state of Rio Grande do Sul, Brazil. *Am J Hum Biol.* 2005 Jul-Aug;17(4):496-506.

Marrero AR, Silva-Junior WA, Bravi CM, Hutz MH, Petzl-Erler ML, Ruiz-Linares A, et al. Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. *Am J Phys Anthropol.* 2007 Feb;132(2):301-10.

Meyer S, Weiss G, von Haeseler A. Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. *Genetics.* 1999 Jul;152(3):1103-10.

Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988 Feb 11;16(3):1215.

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, et al. Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci U S A.* 2003 Jan 7;100(1):171-6.

Monson KL, Miller KWP, Wilson MR, DiZinno JA, Budowle B. The mtDNA Population Database: An Integrated Software and Database Resource for Forensic Comparison. *Forensic Science Communications.* 2002;4(2). Available at: <http://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/april2002/miller1.htm>.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor symposia on quantitative biology.* 1986;51 Pt 1:263-73.

Nunes AC, Silva DA, Teixeira MA, Nunes DD, Lopes CM, Netto OR, et al. Y chromosome comparative analysis of Rondonia with other Brazilian populations. *Leg Med (Tokyo).* 2011 May;13(3):161-3.

Pai CY, Hsieh LL, Lee TC, Yang SB, Linville J, Chou SL, et al. Mitochondrial DNA sequence alterations observed between blood and buccal cells within the same individuals having betel quid (BQ)-chewing habit. *Forensic science international.* 2006 Jan 27;156(2-3):124-30.

Pakendorf B, Stoneking M. Mitochondrial DNA and human evolution. *Annu Rev Genomics Hum Genet.* 2005;6:165-83.

Palencia L, Valverde L, Alvarez A, Caine LM, Cardoso S, Alfonso-Sanchez MA, et al. Mitochondrial DNA diversity in a population from Santa Catarina (Brazil): predominance of the European input. *Int J Legal Med.* 2010 Jul;124(4):331-6.

Paneto GG, Martins JA, Longo LV, Pereira GA, Freschi A, Alvarenga VL, et al. Heteroplasmy in hair: differences among hair and blood from the same individuals

are still a matter of debate. *Forensic science international*. 2007 Dec 20;173(2-3):117-21.

Paneto GG, Kohnemann S, Martins JA, Cicarelli RM, Pfeiffer H. A single multiplex PCR and SNaPshot minisequencing reaction of 42 SNPs to classify admixture populations into mitochondrial DNA haplogroups. *Mitochondrion*. 2011 Mar;11(2):296-302.

Parsons TJ, Coble MD. Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome. *Croatian medical journal*. 2001 Jun;42(3):304-9.

Pauling L. *Molecular Disease and Evolution*. Bulletin of the New York Academy of Medicine. 1964 May;40:334-42.

Peixoto DC. *História Fluminense*. Niterói: Lumak Gráfica e Editora Ltda 1966.

Pereira L, Prata MJ, Amorim A. Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. *Annals of human genetics*. 2000 Nov;64(Pt 6):491-506.

Plaza S, Salas A, Calafell F, Corte-Real F, Bertranpetit J, Carracedo A, et al. Insights into the western Bantu dispersal: mtDNA lineage analysis in Angola. *Human genetics*. 2004 Oct;115(5):439-47.

Prieto L, Alonso A, Alves C, Crespillo M, Montesino M, Picornell A, et al. 2006 GEP-ISFG collaborative exercise on mtDNA: reflections about interpretation, artefacts, and DNA mixtures. *Forensic Sci Int Genet*. 2008 Mar;2(2):126-33.

Reis JJ. *BRASIL: 500 ANOS DE POVOAMENTO*. Rio de Janeiro: IBGE 2000.

Ribeiro D. *O povo brasileiro: a formação e o sentido do Brasil*. São Paulo 1995.

Rice WR. Analyzing tables of statistical tests. *Evolution* 1989;43:223-5.

Robin ED, Wong R. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *Journal of cellular physiology*. 1988 Sep;136(3):507-13.

Saccone C, De Giorgi C, Gissi C, Pesole G, Reyes A. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. *Gene*. 1999 Sep 30;238(1):195-209.

Salas A, Richards M, De la Fe T, Lareu MV, Sobrino B, Sanchez-Diz P, et al. The making of the African mtDNA landscape. *Am J Hum Genet*. 2002 Nov;71(5):1082-111.

Salas A, Richards M, Lareu MV, Scozzari R, Coppa A, Torroni A, et al. The African diaspora: mitochondrial DNA and the Atlantic slave trade. *Am J Hum Genet*. 2004 Mar;74(3):454-65.



Sambrook M, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual 1989.

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5463-7.

Scheffler IE. Structure and Morphology. *Mitochondria*: John Wiley & Sons, Inc. 2002:15-47.

Schmitt R, Bonatto SL, Freitas LB, Muschner VC, Hill K, Hurtado AM, et al. Extremely limited mitochondrial DNA variability among the Ache Natives of Paraguay. Annals of human biology. 2004 Jan-Feb;31(1):87-94.

Schwartz M, Vissing J. New patterns of inheritance in mitochondrial disease. Biochem Biophys Res Commun. 2003 Oct 17;310(2):247-51.

Schwartz M, Vissing J. [Paternal inheritance of mitochondrial DNA]. Ugeskrift for laeger. 2003 Sep 15;165(38):3627-30.

Seldin MF, Pasaniuc B, Price AL. New approaches to disease mapping in admixed populations. Nature reviews Genetics. 2011 Aug;12(8):523-8.

Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, et al. Skin pigmentation, biogeographical ancestry and admixture mapping. Human genetics. 2003 Apr;112(4):387-99.

Silva DA, Crouse CA, Chakraborty R, Goes AC, Carvalho EF. Statistical analyses of 14 short tandem repeat loci in Brazilian populations from Rio de Janeiro and Mato Grosso do Sul states for forensic and identity testing purposes. Forensic science international. 2004 Jan 28;139(2-3):173-6.

Silva DA, Carvalho E, Costa G, Tavares L, Amorim A, Gusmao L. Y-chromosome genetic variation in Rio de Janeiro population. Am J Hum Biol. 2006 Nov-Dec;18(6):829-37.

Smith HW. Incidence of river water of Escherichia coli containing R factors. Nature. 1970 Dec 26;228(5278):1286-8.

Sobrino B, Brion M, Carracedo A. SNPs in forensic genetics: a review on SNP typing methodologies. Forensic science international. 2005 Nov 25;154(2-3):181-94.

Stewart JE, Fisher CL, Aagaard PJ, Wilson MR, Isenberg AR, Polansky D, et al. Length variation in HV2 of the human mitochondrial DNA control region. J Forensic Sci. 2001 Jul;46(4):862-70.

Stoneking M. Hypervariable sites in the mtDNA control region are mutational hotspots. Am J Hum Genet. 2000 Oct;67(4):1029-32.

Strachan T, Read AP. Human molecular genetics 3: Garland Science 2004.

Sturrock K, Rocha J. A Multidimensional Scaling Stress Evaluation Table

Field Methods. 2000;12:49-60.

Taylor RW, Turnbull DM. Mitochondrial DNA mutations in human disease. *Nature reviews Genetics*. 2005 May;6(5):389-402.

Torrioni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, et al. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet*. 1993 Sep;53(3):563-90.

Torrioni A, Sukernik RI, Schurr TG, Starikorskaya YB, Cabell MF, Crawford MH, et al. mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *Am J Hum Genet*. 1993 Sep;53(3):591-608.

Torrioni A, Wallace DC. Mitochondrial DNA variation in human populations and implications for detection of mitochondrial DNA mutations of pathological significance. *J Bioenerg Biomembr*. 1994 Jun;26(3):261-71.

Torrioni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, et al. Classification of European mtDNAs from an analysis of three European populations. *Genetics*. 1996 Dec;144(4):1835-50.

Torrioni A, Achilli A, Macaulay V, Richards M, Bandelt HJ. Harvesting the fruit of the human mtDNA tree. *Trends Genet*. 2006 Jun;22(6):339-45.

Tully G, Barritt SM, Bender K, Brignon E, Capelli C, Dimo-Simonin N, et al. Results of a collaborative study of the EDNAP group regarding mitochondrial DNA heteroplasmy and segregation in hair shafts. *Forensic science international*. 2004 Feb 10;140(1):1-11.

Tuppen HA, Blakely EL, Turnbull DM, Taylor RW. Mitochondrial DNA mutations and human disease. *Biochim Biophys Acta*. 2010 Feb;1797(2):113-28.

Underhill PA, Kivisild T. Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annual review of genetics*. 2007;41:539-64.

van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human mutation*. 2009 Feb;30(2):E386-94.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science*. 2001 Feb 16;291(5507):1304-51.

von Wurmb-Schwark N, Harbeck M, Wiesbrock U, Schroeder I, Ritz-Timme S, Oehmichen M. Extraction and amplification of nuclear and mitochondrial DNA from ancient and artificially aged bones. *Leg Med (Tokyo)*. 2003 Mar;5 Suppl 1:S169-72.

Vona G, Falchi A, Moral P, Calo CM, Varesi L. Mitochondrial sequence variation in the Guahibo Amerindian population from Venezuela. *Am J Phys Anthropol*. 2005 Jul;127(3):361-9.

- Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques*. 1991 Apr;10(4):506-13.
- Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*. 1953 Apr 25;171(4356):737-8.
- Westwood SA, Werrett DJ. An evaluation of the polymerase chain reaction method for forensic applications. *Forensic science international*. 1990 Apr;45(3):201-15.
- White TJ, Arnheim N, Erlich HA. The polymerase chain reaction. *Trends Genet*. 1989 Jun;5(6):185-9.
- Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet*. 2002 Mar;70(3):635-51.