



**Universidade Federal de Pernambuco**  
**Centro de Ciências Biológicas**  
**Laboratório de Imunopatologia Keizo Asami**  
**Programa de Pós-Graduação em Biologia Aplicada à Saúde**

***Identificação Humana como Ferramenta de Investigações Criminais: Estudo de  
Frequências Alélicas de Marcadores de Interesse Forense no Estado de  
Pernambuco***

**SANDRA MARIA DOS SANTOS**

**Setembro, 2014**  
**Recife – PE**



**Universidade Federal de Pernambuco**  
**Centro de Ciências Biológicas**  
**Laboratório de Imunopatologia Keizo Asami**  
**Programa de Pós-Graduação em Biologia Aplicada à Saúde**

***Identificação Humana como Ferramenta de Investigações Criminais: Estudo de  
Frequências Alélicas de Marcadores de Interesse Forense no Estado de  
Pernambuco***

Tese submetida ao Programa de Pós-Graduação em Biologia Aplicada à Saúde do Laboratório de Imunopatologia Keizo Asami, do Centro de Ciências Biológicas, da Universidade Federal de Pernambuco, para defesa de doutorado.

**Doutoranda: Sandra Maria dos Santos**  
**Orientador: Prof. Dr. Sergio Crovella**  
**Co-orientador: Prof. Dr. Paulo Roberto Eleutério de Souza**

**Recife - PE**  
**2014**

Catálogo na Fonte:

Bibliotecário Bruno Márcio Gouveia, CRB-4/1788

Santos, Sandra Maria dos

Identificação humana como ferramenta de investigações criminais: estudo de frequências alélicas de marcadores de interesse forense no Estado de Pernambuco / Sandra Maria dos Santos. – Recife: O Autor, 2014.

157 f.: il.

Orientadores: Sergio Crovella, Paulo Roberto Eleutério de Souza

Tese (doutorado) – Universidade Federal de Pernambuco. Centro de Ciências Biológicas. Pós-graduação em Biologia aplicada à Saúde, 2014.

Inclui referências e anexos

1. Medicina legal 2. Genética 2. Genética legal 3. Polimorfismos (Genética) I. Crovella, Sergio (orient.) II Souza, Paulo Roberto Eleutério de (co-orient.) III. Título.

614.1 CDD (22.ed.) UFPE/CCB-2014-238

**BANCA EXAMINADORA**

---

**PRESIDENTE:** Prof. Dr. Paulo Roberto Eleutério de Souza

**Laboratório Genoma da Universidade Federal Rural de Pernambuco**

Assinatura: \_\_\_\_\_

**EXAMINADOR INTERNO:** Prof. Dr. Luiz Bezerra de Carvalho Júnior

**Departamento de Bioquímica da Universidade Federal de Pernambuco**

Assinatura: \_\_\_\_\_

**EXAMINADOR INTERNO:** Prof<sup>a</sup>. Dr<sup>a</sup>. Paula Sandrin Garcia

**Departamento de Genética da Universidade Federal de Pernambuco**

Assinatura: \_\_\_\_\_

**EXAMINADOR EXTERNO:** Prof. Dr. Lucas André Cavalcanti Brandão

**Departamento de Patologia da Universidade Federal de Pernambuco**

Assinatura: \_\_\_\_\_

**EXAMINADOR EXTERNO:** Prof. Dr. Rafael Lima Guimarães

**Departamento de Genética da Universidade Federal de Pernambuco**

Assinatura: \_\_\_\_\_

**APROVADA**

**RECIFE, 22 DE SETEMBRO DE 2014**

Dedico esta tese a meus filhos Bernardo e Ana Beatriz. Que a busca pelo conhecimento seja uma constante em suas vidas.

## **AGRADECIMENTOS**

---

Ao mestre Sergio Crovella, exemplo raro de inteligência, conhecimento, dedicação, humildade e principalmente comprometimento social.

Ao professor Paulo Eleutério, pela paciência, amizade e valiosa orientação.

A todos os professores e funcionários do LIKA, especialmente a Luiz de Carvalho e José Luiz de Lima, pela confiança e constantes incentivos. A Eliete Rodrigues um agradecimento especial pelo constante apoio.

Aos pesquisadores do LIKA, especialmente Tatiana Oliveira.

Ao pesquisador do Departamento de Genética da UFPE, Ronald Moura.

Ao meu esposo Frederico Maranhão pela paciência.

Aos meus pais Maria de Lourdes Silva e José Gomes dos Santos, pelo apoio incondicional.

Aos Secretários de Defesa Social de Pernambuco, Dr. Wilson Salles Damázio, Dr. Alessandro Carvalho Mattos e Dr. Alciomar Goersch, grandes e inesquecíveis gestores de segurança pública, fazendo acontecer mudanças significativas para o estado de Pernambuco.

Aos Peritos Criminais e Auxiliares do Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, especialmente Carlos Souza.

A todos aqueles que de alguma forma contribuíram na idealização, desenvolvimento e conclusão do presente trabalho.

*“Cada pessoa deve trabalhar para o seu aperfeiçoamento e, ao mesmo tempo, participar da responsabilidade coletiva por toda a humanidade”.*

*Marie Curie.*

## RESUMO

---

A análise de polimorfismos de DNA como ferramenta de identificação humana é de grande relevância, sendo amplamente utilizada pela Criminalística objetivando a determinação de parentesco biológico e elucidação de crimes. As principais regiões polimórficas do DNA utilizadas na rotina forense são os STRs presentes nos cromossomos autossômicos, todavia, marcadores do cromossomo Y e do cromossomo X devem receber atenção especial por serem muito eficientes quando aplicados em investigações de casos complexos de paternidade, de crimes sexuais e pela capacidade de complementar eficazmente a análise de marcadores autossômicos e do DNA mitocondrial. O presente trabalho descreve as frequências alélicas de marcadores de interesse forense presentes em 21 STRs autossômicos (*D3S1358*, *vWA*, *D16S539*, *CSF1PO*, *TPOX*, *D8S1179*, *D21S11*, *D18S51*, *D2S441*, *D19S443*, *TH01*, *FGA*, *D22S1045*, *D5S818*, *D13S317*, *D7S820*, *SE33*, *D10S1246*, *D1S1656*, *D12S391*, *D2S1338*) e as frequências alélicas de 12 X-STRs (*DXS7132*, *DXS7423*, *DXS8378*, *DXS10074*, *DXS10079*, *DXS10101*, *DXS10103*, *DXS10134*, *DXS10135*, *DXS10146*, *DXS10148* e *HPRTB*), em uma amostragem de 879 indivíduos não relacionados de oito territórios do Estado de Pernambuco, Brasil; além de Fernando de Noronha. Não foram observadas diferenças significativas para quaisquer loci analisados em comparação com outras publicações relativas a diversas regiões do Brasil tanto para as regiões autossômicas quanto para o X-STR. A maioria dos marcadores observados apresentou-se em equilíbrio de Hardy-Weinberg. Os demais parâmetros estatísticos analisados indicam que os marcadores estudados são informativos quando se trata de identificação humana com fins forenses na população pernambucana.

**Palavras-chaves:** Genética Forense; Genética Populacional; STR; cromossomo X.

## **ABSTRACT**

---

The analysis of DNA polymorphisms as a tool for human identification is of great importance, being widely used by the Criminalistics, aiming the determination of biological relationship and crimes elucidation. The main polymorphic areas of DNA are routinely used in forensic. They are STRs present in the autosomes, however, markers of the Y chromosome and the X chromosome should receive special attention since they are very efficient when applied in investigations of complex paternity cases, sex crimes and the ability to effectively complement the analysis of autosomal markers and mitochondrial DNA. This paper describes the allele frequencies of markers of forensic interest present in 21 autosomal STRs (*D3S1358*, *vWA*, *D16S539*, *CSF1PO*, *TPOX*, *D8S1179*, *D21S11*, *D18S51*, *D2S441*, *D19S443*, *TH01*, *FGA*, *D22S1*) and allele frequencies of 12 X-STRs (*DXS7132*, *DXS7423*, *DXS8378*, *DXS10074*, *DXS10079*, *DXS10101*, *DXS10103*, *DXS10134*, *DXS10135*, *DXS10146*, *DXS10148* and *HPRTB*), in samples of 879 unrelated individuals from eight territories of the State of Pernambuco, Brazil; and also beyond Fernando de Noronha. No significant differences for any loci analyzed in comparison with other publications relating to various regions of Brazil for both autosomal regions as for the X-STR were observed. Most of the markers studied were in Hardy-Weinberg equilibrium. Other statistical parameters analyzed indicated that the studied markers are informative when it comes to human identification for forensic purposes in the Pernambuco population.

**Keywords:** Forensic Genetics; Genetics Population data; STR; X chromosome.

## LISTA DE FIGURAS

---

Figura 1: Métodos de Identificação. Adaptado	20
Figura 2: Laboratórios de Genética Forense no Brasil no ano de 2004.	23
Figura 3: Laboratórios de Genética Forense no Brasil no ano de 2012. As estrelas verdes indicam os estados com laboratórios em fase de conclusão.	24
Figura 4: A estrutura do genoma e seus principais elementos.	25
Figura 5: Os cromossomos e a determinação de sexo.	25
Figura 6: Estrutura do DNA composta por introns e éxons.	26
Figura 7: Diferença entre polimorfismos de sequência e de comprimento. Adaptado	27
Figura 8: Regiões minissatélites e microssatélites do DNA. Adaptado.	28
Figura 9: Nomenclatura de um marcador STR autossômico. Adaptado.	31
Figura 10: SNPs. Adaptado.	32
Figura 11: Marcador genético do tipo indel. Adaptado.	32
Figura 12: Eletroferograma de indivíduo do sexo feminino, com 24 marcadores, sendo 21 STRs autossômicos e 03 regiões de determinação de sexo.	36
Figura 13: Eletroferograma de indivíduo do sexo masculino, com 24 marcadores, sendo 21 STRs autossômicos e 03 regiões de determinação de sexo.	37
Figura 14: Idiograma do cromossomo X.	42
Figura 15: Eletroferograma de indivíduo do sexo feminino, exibindo 12 regiões do cromossomo X.	43
Figura 16: Eletroferograma de indivíduo do sexo masculino, exibindo 12 regiões do cromossomo X.	44
Figura 17: Regiões pseudo-autossômicas dos cromossomos sexuais.	45
Figura 18: Idiograma do cromossomo Y.	47
Figura 19: Eletroferograma exibindo 17 regiões do cromossomo Y.	48
Figura 20: Organização do DNA mitocondrial humano. Adaptado.	51

Figura 21: Reação em cadeia da polimerase do DNA. Adaptado.	53
Figura 22: Analisador genético de DNA, ABI 3500.	54
Figura 23: Confronto entre perfis genéticos.	56
Figura 24: Representação esquemática do funcionamento de um banco de dados de perfis genéticos de criminosos e de vestígios coletados em locais de crime.	68
Figura 25: Representação esquemática do funcionamento de um banco de dados de perfis genéticos de cadáveres de identidade ignorada	68
Figura 26: Trabalho do perito criminal em local de crime	71
Figura 27: Coleta de Mucosa Bucal	72

---

## LISTA DE TABELAS

---

Tabela 1: Rede de Laboratórios Regionais de Genética Forense. Fonte: Secretaria Nacional de Segurança Pública, 2004	22
Tabela 2: Alguns STRs usados e recomendados pelo FBI	30
Tabela 3: <i>Loci</i> recomendados pelo FBI para definição do perfil de DNA forense	34
Tabela 4: Herança dos alelos num caso de paternidade oriunda de estupro.	58

## **LISTA DE ABREVIATURAS E SIGLAS**

---

A	Adenina
AIS	Área Integrada de Segurança
BDPF	Banco de dados de perfis genéticos
C	Citosina
CIP	Conteúdo de informação polimórfica
CODIS	Sistema Indexado Combinado de DNA
CS	Crimes sexuais
CV	Confronto de vestígios
DL	Desequilíbrio de ligação
DNA	Ácido Desoxirribonucléico
FBI	Serviço Federal de Investigação
G	Guanina
H	Heterozigosidade
HLA	Antígeno Leucocitário Humano
HV1	Região hipervariável 1
HV2	Região hipervariável 2
HWE	Equilíbrio de Hardy-Weinberg
IH	Identificação humana de cadáver de identidade ignorada
IP	Índice de paternidade típico
IPC	Índice de paternidade combinado
ISFG	Sociedade Internacional de Genética Forense
LPPGF	Laboratório de Perícia e Pesquisa em Genética Forense
mtDNA	DNA mitocondrial
PAR1	Região pseudo-autossomal 1
PAR2	Região pseudo-autossomal 2
PCR	Reação em Cadeia da Polimerase
PD	Poder de discriminação
PE	Poder de exclusão
PM	Probabilidade de coincidência
RFLP	Polimorfismo de Tamanho de Fragmento de Restrição
SENASP	Secretaria Nacional de Segurança Pública
SNPs	Polimorfismo de um único nucleotídeo
STRs	Repetições Consecutivas Curtas
SWGAM	Grupo de Trabalho Científico sobre Métodos de Análise de DNA
T	Timina
VNTRs	Repetições Consecutivas de Número Variável

## SUMÁRIO

---

RESUMO	
ABSTRACT	
LISTA DE ILUSTRAÇÕES	
LISTA DE TABELAS	
LISTA DE ABREVIATURAS E SIGLAS	
1 INTRODUÇÃO	16
2 REVISÃO DA LITERATURA	19
2.1 A história da arte da identificação humana no contexto da genética forense	19
2.2 Estrutura do genoma humano e a fundamentação molecular da identificação humana pelo DNA	23
2.3 Polimorfismos de DNA e a informação genética	27
2.4 Marcadores moleculares utilizados na rotina forense	33
2.4.1 STR autossômico	33
2.4.2 Cromossomo X	38
2.4.3 Cromossomo Y	45
2.4.4 DNA mitocondrial	50
2.5 O exame de DNA	52
2.5.1 Extração do DNA	52
2.5.2 Quantificação	53
2.5.3 Amplificação dos marcadores genéticos	53
2.5.4 Detecção dos marcadores genéticos	54
2.5.5 Interpretação dos perfis genéticos	54
2.5.6 O confronto genético	55
2.6 Princípios de Genética de Populações aplicados ao estudo da Genética Forense	59
2.7 Conceitos e fatores que influenciam diretamente a análise de DNA nos laboratórios forenses do Brasil	63
2.7.1 Identificação de cadáveres de identidade ignorada	64
2.7.2 Crimes sexuais	64
2.7.3 Confronto de vestígios	65
2.7.4 Bancos de dados de perfis genéticos	66
2.7.5 Local de crime e Genética Forense	70
2.7.6 Coleta de amostras biológicas destinadas a exames de DNA	71
3 REFERÊNCIAS	74
4 OBJETIVOS	87
<b>Artigo I</b>	88
Distribution of forensic marker allelic frequencies in Pernambuco, Northeastern Brazil.	
<b>Artigo II</b>	104
Forensic genetic data of six new generation STR loci ( <i>D1S1656</i> , <i>D2S441</i> , <i>D10S1248</i> , <i>D12S391</i> , <i>D22S1045</i> and <i>SE33</i> ): an expanded Pernambuco-Northeastern Brazil Population database.	
<b>Artigo III</b>	124
Genetic profile characterization of twenty one STR loci and twelve X-STR in sample from Fernando de Noronha, Pernambuco, Brazil	

5	GRÁFICOS COMPLEMENTARES	144
6	CONSIDERAÇÕES FINAIS	156
7	PERSPECTIVAS	159
	ANEXOS	161

# ***INTRODUÇÃO***

## 1 INTRODUÇÃO

---

A busca por metodologias confiáveis de identificação humana sempre foi uma grande necessidade social. Em diversas situações, como guerras, distúrbios sócio-políticos e desastres em massa, evidenciam-se a urgência por ações eficazes de identificação das vítimas, objetivando amenizar de alguma forma a angústia social que sempre se instala (DHANARDHONO et al., 2013).

Por outro lado, a violência urbana, responsável por inúmeros crimes contra a vida, contra o patrimônio, crimes de natureza sexual, entre outros, torna obrigatória a busca por alternativas e metodologias que permitam responsabilizar criminalmente os autores de ilícitos penais. Dessa forma, governo, autoridades, profissionais de segurança pública e cientistas devem, conjuntamente, pesquisar e aplicar toda e qualquer ferramenta que possa ser útil na luta contra o crime.

Entre todas as técnicas atualmente utilizadas para individualização de pessoas, o DNA (ácido desoxirribonucléico) se destaca especialmente em locais de crime, aonde comumente são encontrados materiais biológicos de natureza diversa; muitas vezes em quantidades pequenas, mascarados, misturados, degradados, contaminados, putrefeitos, entre outras situações (ZIĘTKIEWICZ et al., 2012). Tratando-se de local de crime, qualquer vestígio produzido pelo criminoso ou pela vítima pode ser fundamental para o esclarecimento dos fatos, independentemente do tipo, da qualidade, ou da quantidade do material. Cabe aos cientistas forenses, desenvolver e aplicar técnicas que colaborem para a interpretação de tais vestígios (HOU, 2009).

Nos locais de crime, a busca por materiais biológicos que possam fornecer material genético para estudos periciais é amplamente utilizada e apresenta alto grau de sucesso. Por este motivo, o DNA tornou-se prova material contundente, revolucionou e continua revolucionando as investigações criminais, destacando-se não somente para acusação de criminosos, mas também como prova inquestionável utilizada para inocentar pessoas injustamente acusadas (FRUMKIN, 2009).

Nos últimos anos inúmeras técnicas de tipagem do DNA foram desenvolvidas e introduzidas nas ciências forenses e a capacidade de diferenciar indivíduos utilizando marcadores genéticos tornou-se indispensável para a investigação da maioria dos crimes (GERSHAW et al., 2011).

É fato que o estudo da composição genética das populações deve ocupar lugar de destaque na Genética Forense. As principais regiões polimórficas do DNA utilizadas na rotina forense são os STRs (short tandem repeats) presentes nos cromossomos autossômicos. O tipo e a quantidade mínima de marcadores são validados e sugeridos por renomadas instituições nacionais e internacionais. Quanto maior o número de marcadores, e quanto mais estudada e conhecida a frequência alélica desses marcadores na população local, maior as chances de sucesso na identificação forense (BUTLER, 2006).

Devido a natureza das amostras biológicas analisadas em um laboratório de genética forense, e não raro, pela ausência de ascendentes e/ou descendentes diretos que possam fornecer amostras biológicas para comparação de perfis ou ainda devido a dificuldade de obtenção de perfis a partir de amostras biológicas degradadas, frequentemente utilizam-se marcadores do cromossomo Y e do cromossomo X, além do DNA mitocondrial.

Inúmeros estudos sobre frequências alélicas de marcadores de interesse forense foram realizados no Brasil, porém, continuamente são incorporados novos marcadores não contemplados em estudos anteriores (POIARES et al., 2009; AGUIAR et al., 2012). No Estado de Pernambuco, nenhum estudo direcionado à rotina forense ou com amostragem significativa e distribuição uniforme foi realizado.

A implantação do Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco no ano de 2012, demanda a necessidade de estudar a real distribuição da frequência alélica dos principais marcadores utilizados, permitindo a criação de bancos de referência que serão validados e aplicados na rotina. O presente estudo reforça os dados já existentes, fornece novas informações para a população pernambucana que contribuem quanto aos cálculos estatísticos para identificação de cadáver de identidade ignorada, na identificação de criminosos e de vítimas, influenciando fortemente as investigações criminais instauradas no estado de Pernambuco.

# ***REVISÃO***

## 2 REVISÃO DA LITERATURA

---

### 2.1 A HISTÓRIA DA ARTE DA IDENTIFICAÇÃO HUMANA NO CONTEXTO DA GENÉTICA FORENSE

A identificação humana é necessária para a caracterização do início e do fim da personalidade civil. A morte de um indivíduo precisa ser confirmada, pois com ela extinguem-se o poder familiar, o vínculo conjugal, os contratos personalíssimos, o usufruto e, sem sua confirmação não se inicia o processo de sucessão (SCORALICK et al., 2013). Além disso, a identificação humana constitui-se em um problema de importância jurídica em casos criminais, que demandam necessidade de investigação (SILVA et al., 2013a).

A identificação estabelece-se pelo uso de técnicas e métodos cientificamente comprovados (TORNAVOI & SILVA, 2010). Qualquer método de identificação deve obedecer a determinados critérios, como: **Imutabilidade**: As características não devem alterar-se com o passar do tempo. **Perenidade**: Os dados devem perdurar por toda a vida. **Variabilidade**: Distinção precisa e clara, entre o identificando e os demais. **Classificabilidade**: Deve permitir a comparação entre os dados, de forma sistemática e precisa (SILVA & PASSOS, 2006; ARAÚJO, et al., 2013).

De acordo com o Guia para Identificação de Vítimas de Desastres da Interpol, os métodos de identificação podem ser classificados como **primários**: impressões digitais, odontologia legal e DNA e **secundários**: roupas e pertences, achados médicos, tatuagens, etc. (Figura 1) (INTERPOL, 2009). A identificação humana pelo DNA é, na atualidade, considerada importantíssima para a resolução de situações que envolvam questões criminais, bem como aquelas relacionadas a paternidade e continua revolucionando as áreas jurídicas e criminais (SCHUMM et al., 2013).

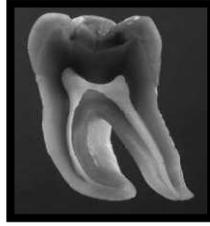
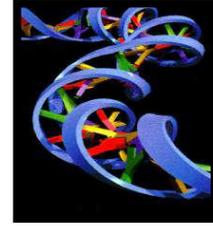
**PRIMÁRIOS****Impressões digitais****Odontologia****DNA****SECUNDÁRIOS****Pertences****Achados médicos**

Figura 1: Métodos de Identificação. Fonte: Interpol, 2009. Adaptado.

Mesmo com toda a diversidade constatada nas populações humanas e apesar das diferenças que observamos entre os indivíduos de nossa espécie, 99% da sequência do nosso DNA é idêntica. A identificação humana pelo DNA é baseada principalmente nas diferenças observadas entre os indivíduos em regiões não codificantes do genoma. Estas diferenças constituem o princípio básico da metodologia usada na Genética Forense (ROEWER, 2013).

A metodologia de identificação humana através do DNA foi desenvolvida por Alec Jeffreys, na década de 80 e a primeira aplicação prática ocorreu em 1985 objetivando a resolução de um problema de imigração na Inglaterra (JEFFREYS et al., 1985a). Uma família de origem africana (Gana), radicada na Inglaterra, teve um dos filhos, proibido de ingressar no país. As autoridades locais contestavam a identidade da criança e a veracidade de seu passaporte, acusando-a de tentar entrar ilegalmente no Reino Unido. A luta se estendeu por dois anos, até que Alec Jeffreys provou o vínculo materno existente entre mãe e filho questionado. Este fato atraiu a atenção do mundo e em pouco tempo revolucionou as ciências forenses (JEFFREYS, 2013).

Posteriormente, as descobertas de Jeffrey foram aplicadas, pela primeira vez para relacionar um único indivíduo ao estupro e assassinato de duas jovens. Lynda Mann e Dawn Ashworth, assassinadas nos anos de 1983 e 1986, respectivamente, no condado Leicester (Reino Unido). Com objetivo de solucionar os casos, todos os homens adultos da vila de Narborough e cercanias foram convocados para doar amostras de sangue. Comparações feitas entre amostras de sêmen coletadas das vítimas e o perfil genético de Colin Pitchfork estabeleceu sua relação com os crimes. Foi a primeira vez que uma Corte de Justiça aceitou o exame de DNA como evidência criminal, ensejando na prisão e condenação, decretada em 1987, do autor dos crimes (JOBBLING, 2013).

A partir de então o DNA é usado na investigação de crimes variados, sendo responsável pela descoberta e condenação de inúmeros criminosos (MORETI, 2009). Modernas metodologias de identificação humana estão associadas ao uso de diferentes técnicas para análise de caracteres, notadamente moleculares, que são compartilhados sob formas variadas na população (GUIMARÃES et al., 2009).

O potencial de variabilidade individual do DNA eleva-o ao “padrão ouro” entre as técnicas de identificação humana. A resistência ao tempo e ao calor; o fato de ser encontrado em vários tipos de materiais biológicos e o advento da amplificação de DNA *in vitro* por meio da PCR (Reação em Cadeia da Polimerase) reforçam a importância do DNA para as investigações criminais (GUIMARÃES et al., 2009).

A partir dos acontecimentos citados, os estudos com marcadores genéticos e sua utilização em exames de DNA, demonstraram grande aplicabilidade em várias situações de identificação humana e abriram um precedente, sendo usados em outros casos históricos e conseqüentemente, apresentando resultados relevantes no campo da análise do DNA para fins criminais (GERSHAW et al., 2011; COBLE, 2011).

A história da genética forense no Brasil teve início em 1995, com a inauguração do Laboratório de DNA Criminal da Polícia Civil do Distrito Federal. Este laboratório foi o responsável pelos primeiros treinamentos para peritos criminais de diversos estados da federação, incentivando a implantação dos laboratórios de genética forense da Paraíba, Rio Grande do Sul, Minas Gerais, Paraná, Mato Grosso do Sul e São Paulo (SILVA et al., 2013a).

No Brasil, o Ministério da Ciência e Tecnologia e o Ministério da Justiça lançaram no ano de 2004, o “Programa de Ciência e Tecnologia Aplicada na Segurança Pública”. Alguns dos objetivos traçados relacionavam ações diretamente voltadas para o DNA Forense. A política adotada pelo Ministério da Justiça/SENASP concentrou-se **primeiro** na formação, qualificação e capacitação de peritos criminais para a realização de exames periciais e gerenciamento de laboratórios de DNA Criminal; **segundo**, na implementação física e logística dos laboratórios de DNA Criminal; **terceiro**, na garantia da execução das atividades periciais através da continuidade no fornecimento de equipamentos e consumíveis, bem como no treinamento continuado; **quarto**, no apoio e estímulo à independência das unidades estaduais na realização das atividades periciais sem prejuízo da qualidade tecnológica (SENASP, 2004).

Com o apoio e priorização da tecnologia do DNA pelo Ministério da Justiça/SENASP/Coordenação de Políticas, Programas e Projetos, foi estruturada uma Rede de Laboratórios Regionais, que seriam distribuídos conforme o quadro abaixo e funcionariam como pólo de capacitação, pesquisa e desenvolvimento em cada região geográfica brasileira (ou conjunto de estados circunvizinhos) (Tabela 1).

Tabela 1. Rede de Laboratórios Regionais de Genética Forense

<b>PÓLO REGIONAL</b>	<b>ÁREA DE ATENDIMENTO</b>	<b>INTERCÂMBIOS AMÉRICA DO SUL</b>
RJ	RJ, SP, MG e ES	-
AL – BA	AL, BA, SE, PE, PB, RN e CE	-
AM – PA	AM, PA, AP, RR, AC e RO	Peru, Equador, Colômbia, Venezuela, Guiana, Suriname e Guiana Francesa
DF	DF, MT, GO, TO, MA e PI	Bolívia e Chile
RS	RS, PR, SC e MS	Paraguai, Argentina e Uruguai

Fonte: Secretaria Nacional de Segurança Pública, 2004.

A partir daí, observou-se crescimento relevante da Genética Forense em território brasileiro. Hoje, a maioria dos estados brasileiros dispõe de unidades laboratoriais de Genética Forense. São dezenove laboratórios com infra-estrutura e com profissionais capacitados tecnicamente na utilização do DNA para a resolução de crimes, situados nos Estados do Amapá, Amazonas, Bahia, Ceará, Distrito Federal, Espírito Santo, Goiás, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Pará, Paraíba, Paraná, Pernambuco, Polícia Federal, Rio de Janeiro, Rio Grande do Sul, Santa Catarina, São Paulo; todos integrantes de órgãos de segurança e que analisam o DNA dentro dos padrões utilizados internacionalmente. Quatro laboratórios estão em fase final de implantação: Acre, Sergipe, Maranhão e Tocantins (VELHO et al. 2013b) (Figuras 2 e 3).

É uma história de sucesso, visto que estas unidades periciais oferecem à sociedade um número considerável de resultados que contribuem significativamente para as investigações criminais num país carente de justiça.

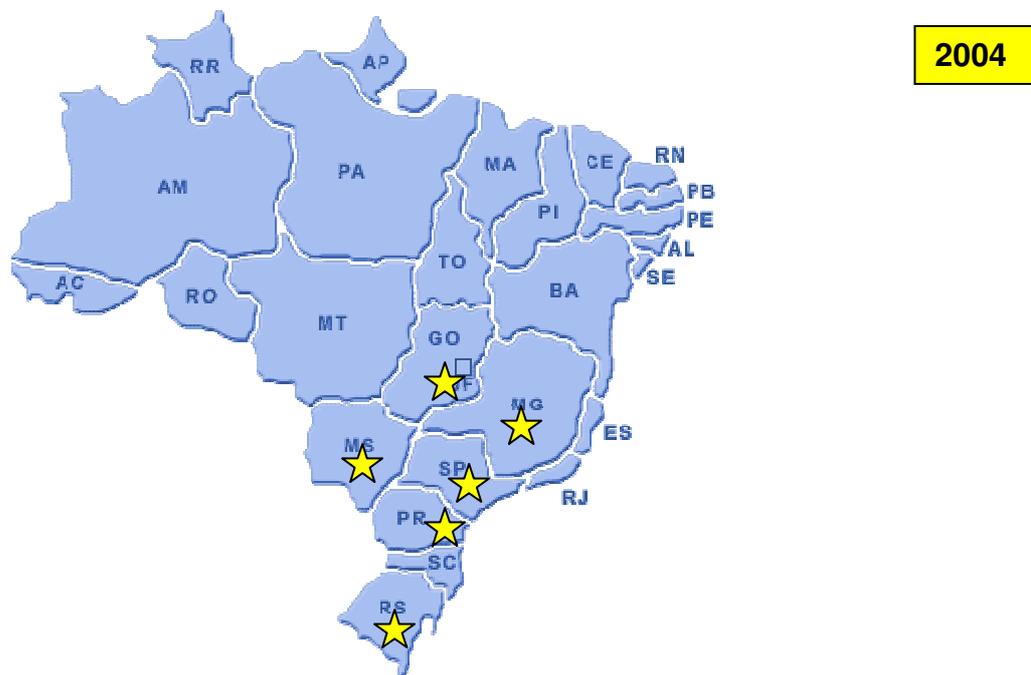


Figura 2: Laboratórios de Genética Forense no Brasil no ano de 2004. Fonte: BRASIL, 2009.

2014

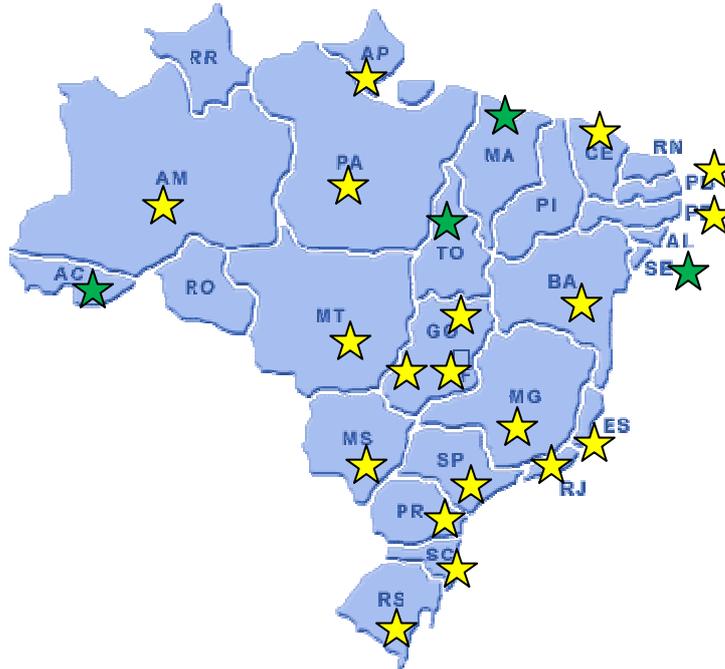


Figura 3: Laboratórios de Genética Forense no Brasil no ano de 2012. As estrelas verdes indicam os estados com laboratórios em fase de conclusão. Fonte: VELHO, 2013b.

## 2.2 ESTRUTURA DO GENOMA HUMANO E A FUNDAMENTAÇÃO MOLECULAR DA IDENTIFICAÇÃO HUMANA PELO DNA

Toda a informação genética humana está codificada no DNA e o seu conjunto é denominado genoma. A estrutura do DNA foi descoberta em 1953 pelos cientistas Francis Crick e James Watson. Trata-se de uma macromolécula composta pelo encadeamento de subunidades individuais denominadas nucleotídeos. O genoma nuclear possui 3 bilhões de pares de nucleotídeos. Cada nucleotídeo é composto de três unidades básicas: um grupamento fosfato, uma pentose e uma base nitrogenada. A pentose é a desoxirribose e a base nitrogenada pode ser classificada como purínica: adenina (A) e guanina (G) ou pirimidínica: timina (T) e a citosina (C) (BUDOWLE, 2013) (Figura 4).

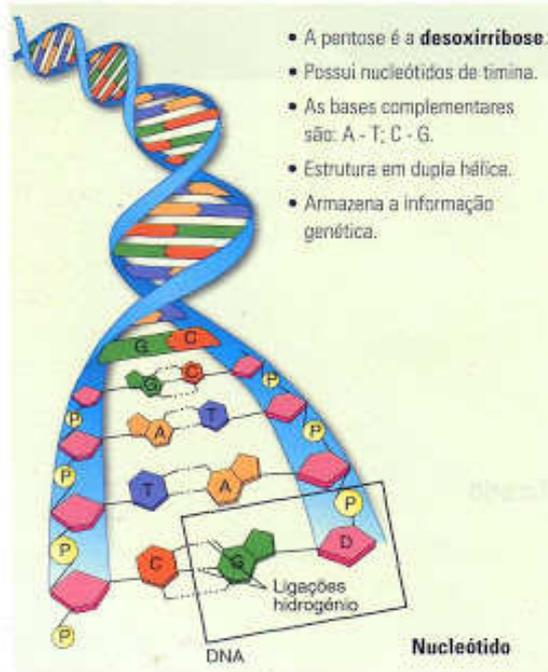


Figura 4: A estrutura do DNA e seus principais elementos. Fonte: LEWIN, 2009.

As sequências nucleotídicas estão organizadas e compactadas em cromossomos. São 22 pares de cromossomos autossômicos e 1 par de cromossomos sexuais, totalizando 46 cromossomos em cada célula diplóide, conforme esquematizado na figura 5 (VELHO et al., 2013b).

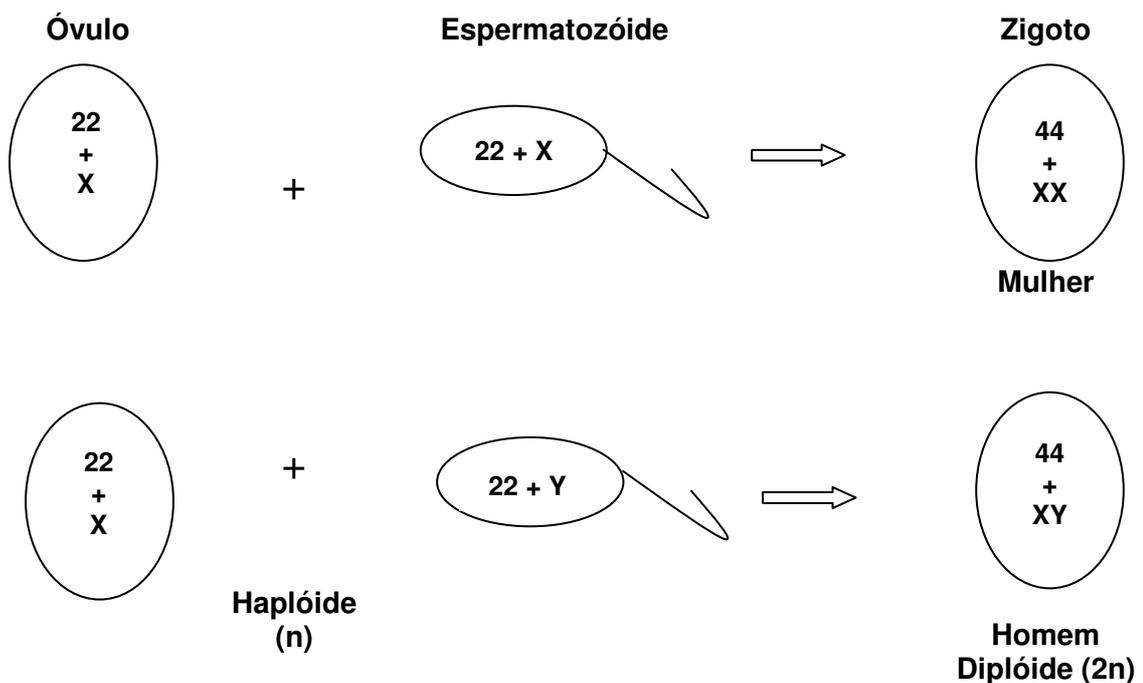


Figura 5: Os cromossomos e a determinação de sexo. Fonte: VELHO et al., 2013b.

Os pares de cromossomos são chamados de homólogos. Os cromossomos autossômicos são nomeados de acordo com o seu tamanho aproximado. O maior é o cromossomo 1, até o cromossomo 22 que é o menor deles. Os cromossomos sexuais são chamados de X e Y, e possuem em suas sequências os genes responsáveis pelas características sexuais secundárias humanas, que atribuem o dimorfismo sexual entre machos e fêmeas observado em nossa espécie. As mulheres possuem dois cromossomos X e os homens possuem um cromossomo X e outro Y, sendo representados como XX e XY, respectivamente (Figura 5) (GRIFFITHS et al., 2009).

Os cromossomos são compostos de DNA e de proteínas. O DNA carrega a informação genética e as proteínas exibem entre outras, funções estruturais e enzimáticas indispensáveis para inúmeros processos, a exemplo da replicação, transcrição, reparo e recombinação (JORDE et al. 2004).

As regiões do DNA que codificam proteínas são chamadas de éxons e as regiões que não codificam proteínas são chamadas de íntron (Figura 6). Os íntrons contêm uma grande porção de DNA repetitivo (NAKAMURA, 2009).

O que diferencia o genoma de cada indivíduo é a sequência das bases nitrogenadas no DNA. No genoma existem mais de três bilhões de nucleotídeos, destes apenas uma pequena porção, cerca de 0,1% apresentam variações na sequência de nucleotídeos, são os chamados polimorfismos do DNA. As metodologias de identificação humana se apóiam na análise dessas regiões variáveis (LI et al., 2009).

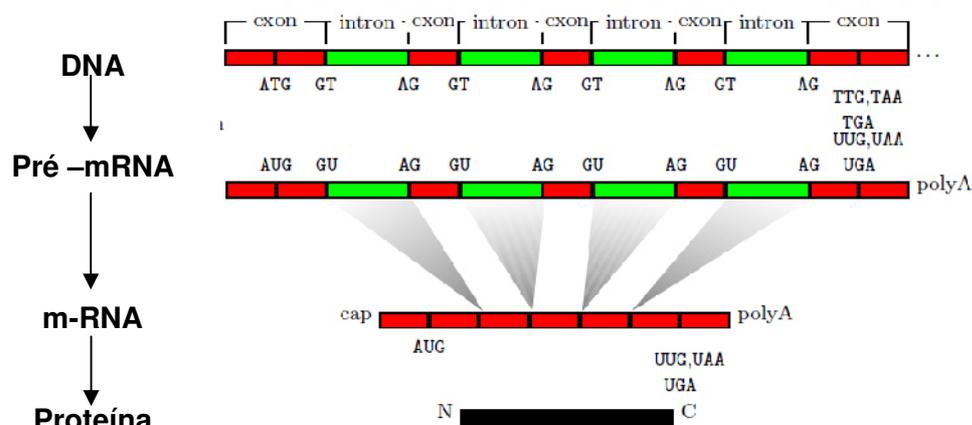


Figura 6: Estrutura do DNA composta por íntrons e éxons. Fonte: FARAH, 2007.

### 2.3 Polimorfismos do DNA e a Informação Genética

Em algumas regiões do genoma, sequências de DNA podem se repetir dezenas ou até centenas de vezes, e o mais interessante, o número de repetições varia entre os indivíduos (KOBACHUK, 2012).

Os marcadores de maior valor para a identificação humana são aqueles que apresentam alto poder de discriminação (maior que 90%), maior frequência de heterozigotos (maior que 70%), baixa taxa de mutação, menor tamanho e baixa formação de *stutters* (fragmentos de DNA com uma unidade a mais ou a menos que o alelo verdadeiro) (HARES, 2012).

Os marcadores de DNA utilizados na identificação humana exibem polimorfismos de sequência ou de tamanho. As sequências variáveis que codificam os antígenos do sistema HLA, sequências hipervariáveis do DNA mitocondrial e sequências polimórficas denominadas SNPs (single nucleotide polymorfism) são exemplos de marcadores com polimorfismos de sequência (Figura 7) (GATTÁS, 2006).

Os minissatélites e microsatélites são os polimorfismos de comprimento mais utilizados e são as regiões de maior interesse para a identificação humana (OLIVEIRA, 2012). A quantidade de repetições presentes determinará para cada indivíduo um tamanho diferente para a região do DNA estudada. O alelo vai representar cada possibilidade de tamanho (ou de número de repetições) que pode ser encontrada (LIMA, 2006).

<b>(a) Polimorfismo de Sequência</b>	-----AGACTCGACATT----- -----AGATTAGGCATT-----
<b>(b) Polimorfismo de Comprimento</b>	------(AATG)(AATG)(AATG)----- 3 repetições ------(AATG)(AATG)----- 2 repetições

Figura 7 - Diferença entre polimorfismos de sequência e de comprimento. Fonte: BUTTLER, 2005. Adaptado.

O polimorfismo de comprimento pode ser agrupado em duas classes: Os **minissatélites** ou VNTRs (*Variable Number Tandem Repeats*) e os **microssatélites** ou STRs (*Short Tandem Repeats*) (FAN & CHU 2007) (Figura 8).

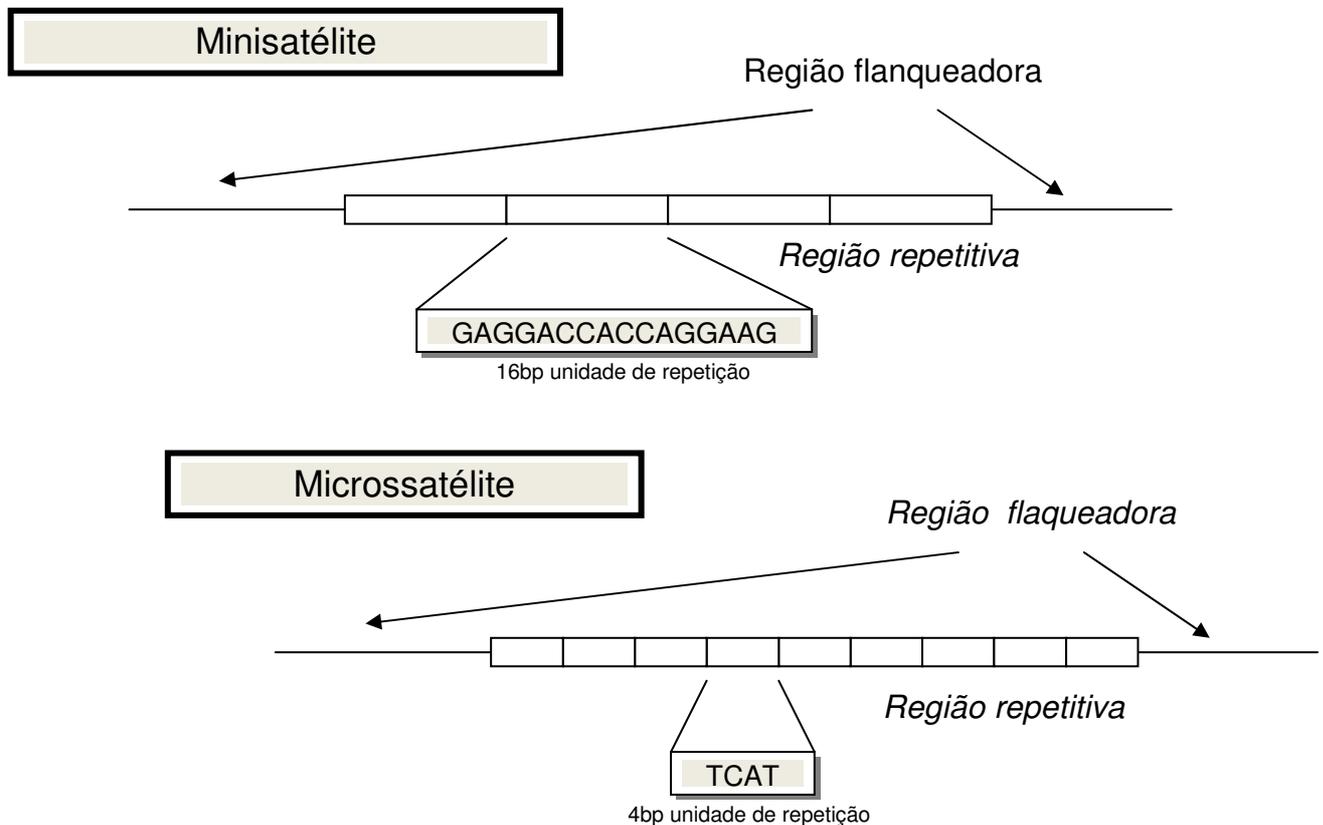


Figura 8. Regiões minissatélites e microssatélites do DNA. Fonte: BUTLER, 2005. Adaptado.

O primeiro marcador de DNA empregado na identificação humana foi baseado no polimorfismo de tamanho, consistindo de regiões com número variável de repetições em tandem, chamado de minissatélites ou VNTRs. São sequências hipervariáveis, caracterizados por apresentarem tamanho entre 10 e 100 pares de bases. O primeiro *locus* VNTR foi identificado por meio do rastreamento de uma biblioteca genômica de DNA humano em uma pesquisa sobre genes implicados em doenças (KOBACHUK, 2012).

A descoberta das regiões altamente polimórficas foi determinante para o desenvolvimento da metodologia de identificação humana através do DNA. Em 1985 Jeffreys desenvolveu uma técnica, por ele denominada "DNA fingerprinting", cuja técnica utilizada para a identificação das sequências hipervariáveis foi denominada RFLP (restriction fragment length polymorphisms) (JEFFREYS et al., 1985b).

Apesar do grande poder de discriminação que os VNTRs oferecem, para que a tipagem seja realizada com sucesso é necessário DNA de alto peso molecular, bem preservado e numa quantidade considerada grande para os padrões forenses. Portanto, apesar do seu alto grau de polimorfismo, essas regiões não são consideradas adequadas para uso forense, cujas amostras biológicas encontrem-se degradadas ou em quantidades reduzidas (VELHO et al., 2013a).

A partir dos anos 90, a técnica de reação em polimerase (PCR) surgiu como uma alternativa extremamente eficiente em casos cuja quantidade de DNA seja exígua ou degradada. Esta metodologia ganhou rápida aceitação da comunidade científica como ferramenta de biologia molecular para estudos de genética de populações, no diagnóstico molecular, no mapeamento do genoma humano e na identificação humana (VAN OORSCHOT et al., 2010).

Atualmente, a grande maioria das identificações humanas é realizada por meio de marcadores genéticos denominados microssatélites ou STRs. Sequências curtas repetidas em tandem no genoma. Os STRs geralmente possuem tamanho entre 2 a 10 pares de bases. Devido ao seu alto poder de discriminação, a análise de STRs se tornou rotina em testes de paternidade e genética forense. Aproximadamente 30% das regiões repetitivas do genoma são de microssatélites. A maioria delas são repetições de dois nucleotídeos, como (AC) $n$ , por exemplo. As mais utilizadas são as repetições de 4 a 5 nucleotídeos (GGAA) $n$ , que ocorrem numa região específica do genoma, num loco único (CHEN et al. 2009).

A Sociedade Internacional de Genética Forense (ISFG) recomenda que marcadores microssatélites só sejam aplicados na prática forense mediante estudos das propriedades genético-populacionais e somente se estiverem em equilíbrio de Hardy-Weinberg, apresentarem alto nível de polimorfismo e possuírem desequilíbrio de ligação conhecido (SZIBOR, 2007).

Um dos eventos mais famosos da aplicação desta técnica está relacionado com O.J. Simpson, jogador de beisebol norte-americano acusado de duplo homicídio em 1994. Um dos problemas visualizados neste tipo de metodologia aparece frequentemente quando amostras biológicas analisadas apresentam padrão de mistura (SILVA, 2004). Os loci representados na tabela 2 são exemplos de regiões utilizadas para identificação humana, especialmente na área forense (KOBILINSKY et al., 2005).

Tabela 2: Alguns STRs usados e recomendados pelo FBI (Serviço Federal de Investigação) (KOBILINSKY et al., 2005).

<b>Loco</b>	<b>Repetições</b>	<b>Número de Alelos</b>	<b>Possibilidade de Genótipos</b>
D3S1358	[TCTG] [TCTA]	10	91
FGA	CTT	19	528
CSF1PO	TAGA	15	55
D5S818	AGAT	10	55
D7S820	GATA	11	66
D8S1179	[TCTA] [TCTG]	10	78
D13S317	TATC	8	45
D16S539	GATA	8	45
D18S51	AGAA	15	231
D21S11	[TCTA] [TCTG]	69	300
TH01	TCAT	7	55
vWa	[TCTG] [TCTA]	10	105
Amelogenina	106bp/112bp	2	

O estudo de marcadores STRs depende de conhecimento anterior das regiões conservadas que flanqueiam a seqüência com repetições consecutivas, para que oligonucleotídeos iniciadores possam se parear nelas e permitir o sequenciamento do STR (SILVA, 2004).

A nomenclatura dos STRs segue uma ordem pré-estabelecida. A primeira letra do nome do marcador é “D” e significa DNA, depois vem um número que representa em que cromossomo o marcador está localizado, no caso de estar em cromossomos sexuais usa-se “X” ou “Y”, o terceiro elemento do nome é a letra “S”, do inglês, *single copy sequence* que significa seqüência de cópia única e finalmente o número do loco em que o marcador foi descrito (BUTLER, 2005).

A figura 9 exemplifica a nomenclatura de um marcador STR de um cromossomo autossômico. Além da nomenclatura clássica, alguns STRs são nomeados baseados na sequência que compõem a unidade de repetição (ex: GATA), ou ainda, quando estão localizados dentro de uma sequência gênica podem ter a nomenclatura relacionada com o gene. Um exemplo é o STR TH01 que está localizado no íntron 01 do gene da Tirosina Hidroxilase (BUTLER, 2005).

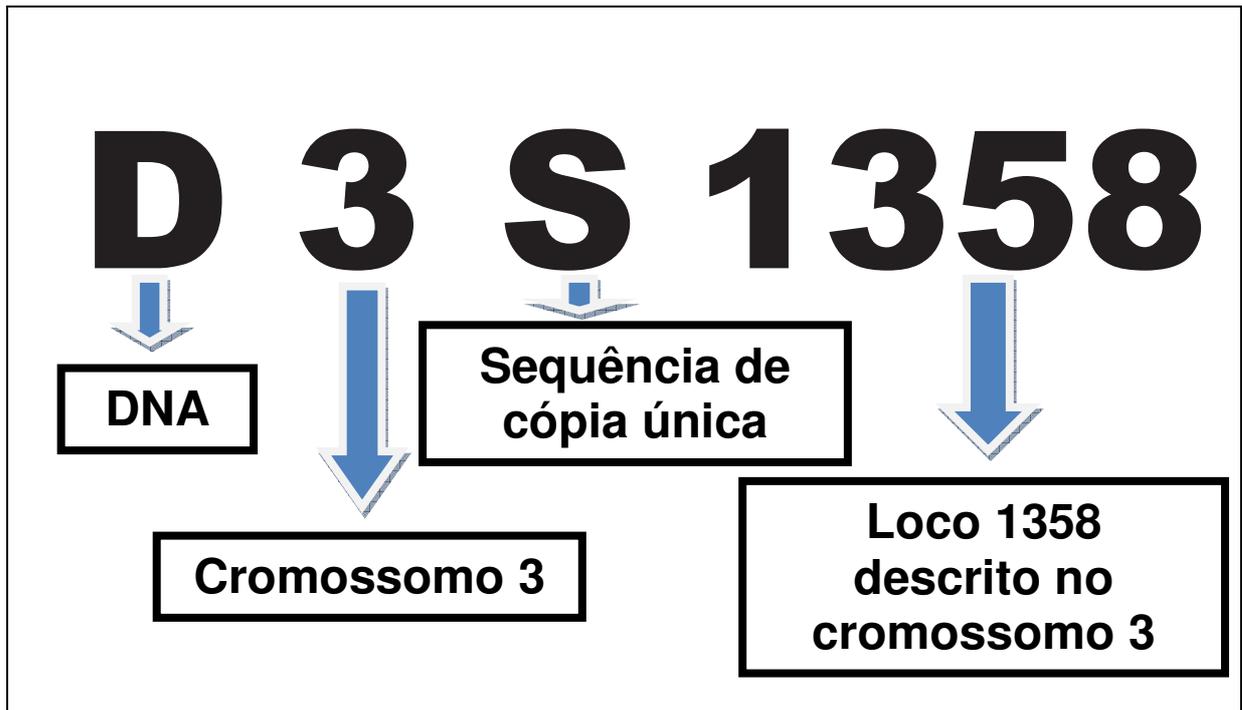


Figura 9: Nomenclatura de um marcador STR autossômico. FONTE: BUTLER, 2005. Adaptado.

A variação genética mais comum no genoma humano são os polimorfismos de uma única base ou SNPs (Single nucleotide polymorphisms). Os SNPs são variações de um único nucleotídeo em uma determinada posição da sequência do DNA quando diferentes alelos de um mesmo indivíduo são comparados, ou mesmo quando os alelos observados em uma determinada população são comparados (Figura 10) (SILVA, 2013b).

Ao comparar os cromossomos homólogos de uma mesma pessoa, estima-se a possibilidade de observarmos um SNP a cada 1000 pares de bases. Os SNPs são na maioria das vezes bialélicos. São geralmente o resultado de mutações do tipo transição, onde em determinada posição da sequência do DNA pode ser verificada um C/T ou um A/G. Por apresentarem menos variação, são menos informativos que os STRs (BUDOWLE & DAAL, 2008).

A possibilidade de análise de amostras altamente degradadas, que exibem DNA muito fragmentado atribui uma grande utilidade a esta ferramenta. Além disso, muitos SNPs, localizados em regiões codificantes de determinados genes podem ser usados para a determinação de características fenotípicas dos indivíduos, como cor dos olhos, cabelo e pele (DECORTE, 2010)

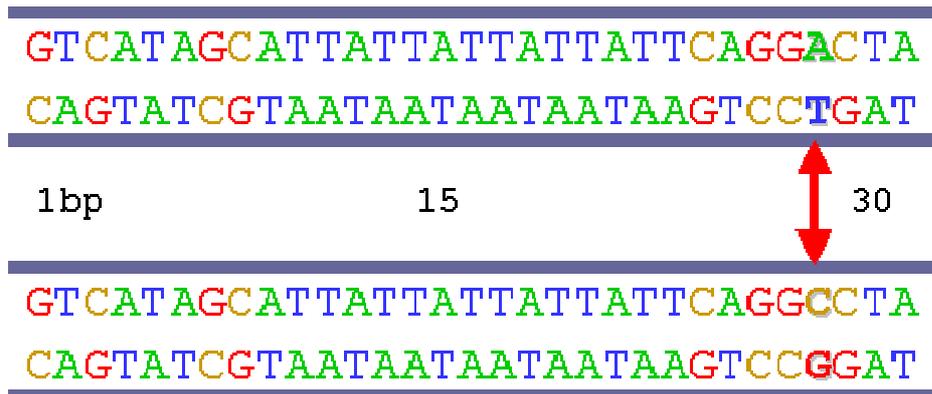


Figura 10: SNPs. Fonte: SILVA, 2013b. Adaptado.

Outra forma de polimorfismo observado no genoma humano são os marcadores do tipo Indels. Tratam-se de inserções ou deleções da sequência de DNA em determinados locus que podem variar de um nucleotídeo a centenas deles (Figura 11) (LI et al., 2013). Geralmente são bialélicos, apresentando dois alelos para cada loco analisado. Ainda não são amplamente utilizados na identificação humana, todavia, observa-se grande potencial futuro (SILVA, 2013b).



Figura 11: Marcador genético do tipo indel. Fonte: SILVA, 2013b. Adaptado.

## 2.4 MARCADORES MOLECULARES UTILIZADOS NA ROTINA FORENSE

### 2.4.1 STR AUTOSSÔMICO

Os STRs são considerados uma poderosa e confiável ferramenta disponível para as análises de DNA que envolvem a identificação humana (TAN et al., 2013).

O FBI (Federal Bureau of Investigation) adota e recomenda uma série padrão de *loci* STRs que são seguidos pelos laboratórios de genética forense em vários países. Durante muitos anos o FBI preconizou que a análise genética de no mínimo 13 regiões autossômicas, além de uma sonda adicional para o gene da amelogenina, usado para identificação do sexo, era o suficiente para resolver casos forenses (BUTLER & HILL 2012).

Recentemente o FBI publicou novas recomendações que sugerem a utilização de um número maior de marcadores, sob a alegação de que o conjunto de 13 *loci* propostos inicialmente apresenta eficácia comprovada em situações que envolvem comparações de perfis de fonte única. No entanto, para amostras forenses desafiadoras e que envolvam misturas, a utilização de um número maior de marcadores representa um melhor desempenho. Diante do exposto, o FBI passou a recomendar o uso de no mínimo 20 marcadores, conforme a tabela 3 (HARES, 2012; KATSANIS & WAGNER, 2013).

Os critérios utilizados para manter ou retirar determinado *loci* foi o grau de informatividade, o poder elevado de discriminação e baixa taxa de mutação, todavia o principal fator é o melhor desempenho em casos forenses. Os *loci* recomendados foram divididos em 02 seções. A seção A, denominada “necessária” na qual constam 20 *loci*. Nesta seção observa-se a exclusão do *TPOX* e a inclusão dos *loci* *D2S1338*, *D19S433*, *PENTA E* e *D10S1248* já costumeiramente utilizados na rotina dos laboratórios forenses brasileiros e os *loci* *D1S1656*, *D12S391*, *D10S1248*, usados na rotina européia e o *DYS391*. O *TPOX*, assim como o *D22S1045*, *SE33* e o *PENTA D* foram colocados na seção B e podem ser utilizados se ainda houver necessidade e por ordem de preferência (HARES, 2012).

Tabela 3: Loci recomendados pelo FBI para definição do perfil de DNA forense (HARES, 2012).

<b>Ordem</b>	<b>Loci inicialmente recomendados</b>	<b>Loci atualmente recomendados</b>
<b>SEÇÃO A (necessário)</b>		
	AMELOGENINA	AMELOGENINA
01	D18S51	D18S51
02	FGA	FGA
03	D21S11	D21S11
04	D8S1179	D8S1179
05	VWA	VWA
06	D13S317	D13S317
07	D16S539	D16S539
08	D7S820	D7S820
09	TH01	TH01
10	D3S1358	D3S1358
11	D5S818	D5S818
12	CSF1PO	CSF1PO
13	TPOX	D2S1338
14		D19S433
15		D1S1656
16		D12S391
17		D2S441
18		D10S1248
19		PENTA E
20		DYS391
<b>SEÇÃO B</b>		
<b>Ordem de Preferência</b>		
21		TPOX
22		D22S1045
23		SE33
24		PENTA D

A análise de STR's foi bastante facilitada, com a introdução de sistemas de amplificação simultânea (*multiplex*) que permitem obter informação de vários *loci* numa única reação, diminuindo o tempo de estudo, assim como da quantidade de DNA e de reagentes. Diversos *kits* comerciais validados, fornecidos por várias empresas, que permitem a amplificação de múltiplos *loci*, constitui um considerável avanço na resolução de perícias do âmbito da Genética e Biologia Forense, mesmo nas situações mais complexas (JOBIM et al, 2005).

Atualmente os laboratórios de genética forense brasileiros estão adaptando-se ao uso de kits multiplex que amplificam 24 regiões. Estes, geralmente exibem 21 regiões autossômicas, além de regiões de determinação de sexo (Figuras 12 e 13). Esta é uma maneira de se adequar às orientações do FBI relacionadas ao tipo e a quantidade de marcadores, padronização que favorece a utilização do CODIS (Combined DNA Index System), Banco de Dados de DNA, adotado em âmbito nacional (SILVA, 2013a; HOFFMAN & FENGER, 2010).

É importante salientar que os STRs localizados nos cromossomos autossômicos continuam sendo os marcadores eleitos nos exames de investigação de paternidade. Todavia, em situações complexas, independentemente da quantidade de polimorfismos estudados, os marcadores autossômicos podem expressar um número de informações que impedem a efetiva identificação de um indivíduo. Diante desta realidade, a genética forense também utiliza outros polimorfismos (BECKER et al., 2008; LIU et al., 2008; NADEEM et al., 2008; RIBEIRO et al., 2008; GUSMÃO et al., 2009).

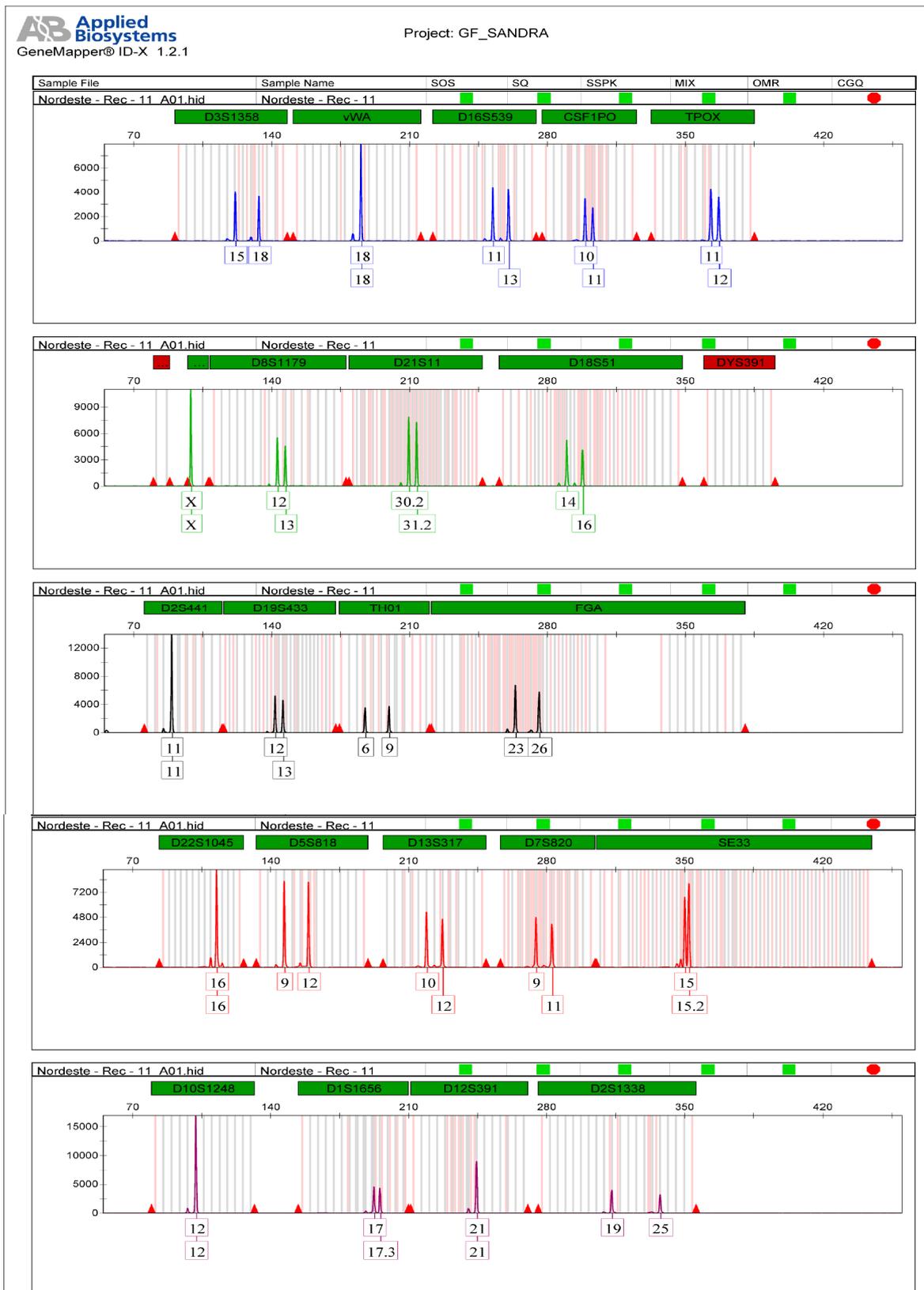


Figura 12. Eletroferograma de indivíduo do sexo feminino, com 24 marcadores, sendo 21 STR autossômicos e 03 regiões de determinação de sexo. Fonte: Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, 2014.

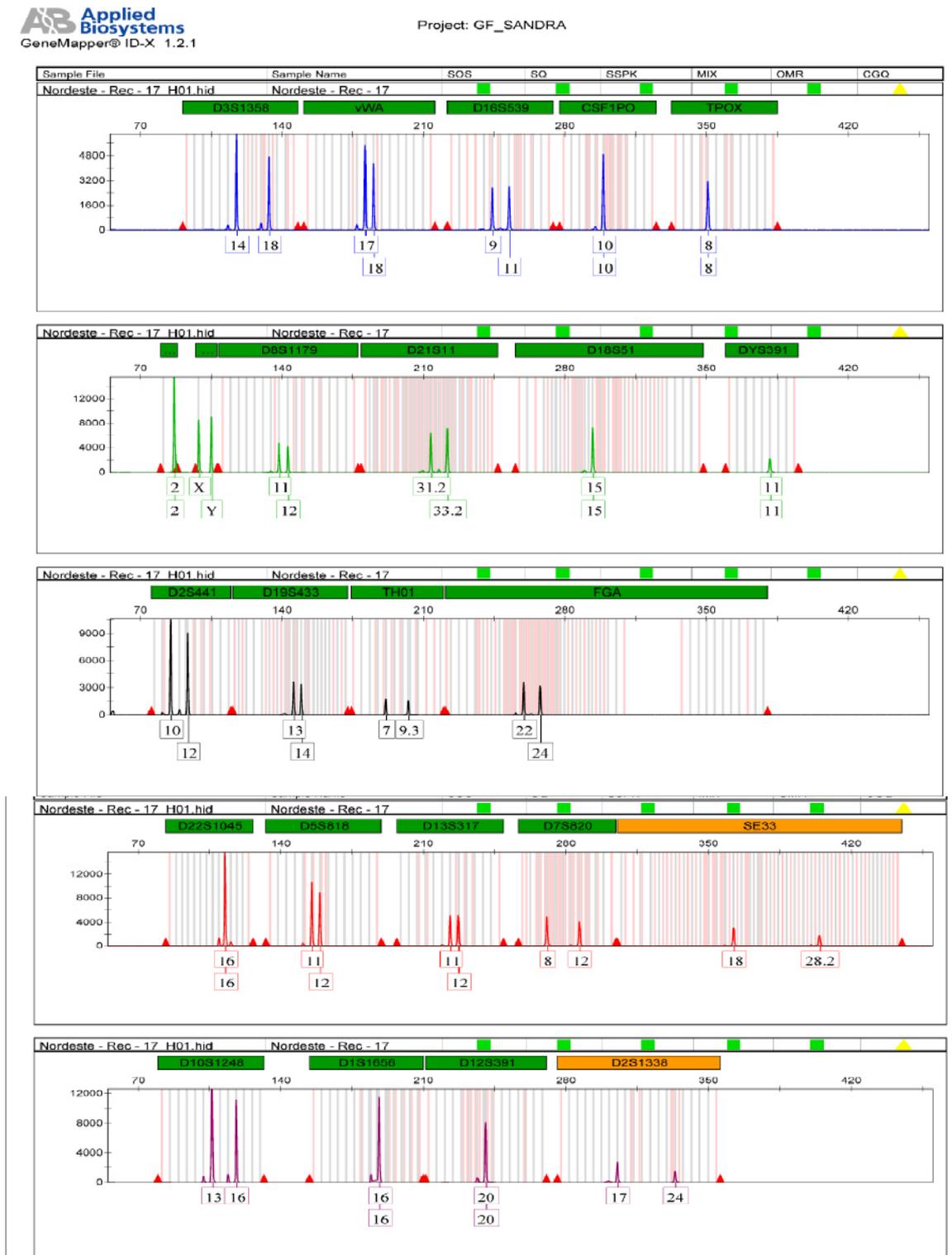


Figura 13. Eletroferograma de indivíduo do sexo masculino, exibindo 24 marcadores, sendo 21 STR autossômicos e 03 regiões de determinação de sexo. Fonte: Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, 2014.

### 2.4.2 CROMOSSOMO X

O cromossomo sexual X apresenta características peculiares no genoma humano: as mulheres herdaram um cromossomo X do pai e um da mãe, já os homens recebem apenas o X materno. Este cromossomo é composto por aproximadamente 155 mega pares de bases e representa cerca de 5% do genoma das células femininas e 2,5% do das células masculinas humanas. Possui uma baixa densidade gênica, com 1.098 genes identificados, representando 4% do total de genes humanos. A quantidade de regiões repetitivas é elevada, correspondendo a 56% do DNA deste cromossomo. A média do genoma é de 45% (ROSS et al. 2005).

Pessoas do sexo feminino têm o dobro de cópias dos genes ligados ao X em relação ao sexo masculino. Para tornar proporcional o produto gênico do cromossomo X entre homens e mulheres, um dos cromossomos X da mulher é inativado, aleatoriamente, no início do desenvolvimento embrionário e este estado inativo é transmitido para todas as células (CHOW et al., 2005).

Na linhagem germinativa o segundo cromossomo X é reativado durante a ovogênese para recombinar-se com o outro cromossomo X feminino. No homem, o cromossomo X não se recombina ao longo de todo o seu comprimento, sendo tal processo limitado a algumas regiões pseudoautosômicas localizadas nas extremidades dos cromossomos X e Y. Tais regiões são indispensáveis para a normal segregação dos cromossomos sexuais durante o processo de meiose (GRIFFITHS et al., 2009).

Os marcadores STRs do cromossomo X, utilizados na identificação humana, estão localizados em regiões não recombinantes entre os cromossomos X e Y. Todavia, ocorre recombinação entre os cromossomos X na mulher, garantido a variabilidade desses marcadores. Nos homens isso é inviável devido a hemizigose do cromossomo X nestas regiões. Assim, da mesma forma que os autossomos, o cromossomo X se recombina (mulheres) e, similarmente ao cromossomo Y e DNA mitocondrial, tem um modelo que permite determinar diretamente o haplótipo (homens) (SZIBOR et al., 2006).

O cromossomo X exibe características genéticas distintas em relação aos autossomos, especialmente: (1) menor tamanho populacional efetivo e (2) taxa de mutação menor. Todos os cromossomos masculinos possuem maior taxa de mutação que os femininos devido ao elevado número de mitoses nas células masculinas. Como dois cromossomos X estão presentes nas mulheres e apenas um nos homens a taxa de mutação global deste cromossomo é inferior a dos autossomos e (3) maior desequilíbrio de ligação, uma vez que o cromossomo X recombina apenas na meiose feminina (ZARRABEITIA et al., 2009).

Estudos sobre polimorfismos no cromossomo X estão atraindo a atenção de pesquisadores, principalmente devido à aplicabilidade destes em investigação de casos complexos de paternidade. Embora vários X-STRs estejam sendo validados para uso forense, pouco se sabe sobre a variação destes polimorfismos em diferentes populações do mundo. Diante deste fato, estudos para avaliar a distribuição das frequências alélicas em diferentes populações, taxas de mutação e desequilíbrio de ligação, viabilizando a produção de bases de dados úteis para aplicações forenses e estudos antropológicos, tornam-se indispensáveis (GOMES et al., 2007; RIBEIRO et al., 2008).

Os X-STRs são capazes de complementar a análise dos marcadores autossomos e do cromossomo Y de forma muito eficaz (NADEEM et al., 2008; RIBEIRO et al., 2008). Estes marcadores se caracterizam por apresentarem, em determinadas situações, maiores valores de poder de exclusão (PE) que os autossomos. Os STRs do cromossomo X possuem maior capacidade de excluir de uma situação de investigação de paternidade, indivíduos não relacionados à criança testada e, em razão disso, suficiente poder estatístico é obtido mesmo quando poucos X-STRs são analisados (GOMES et al., 2007; LIM et al., 2009).

A maior capacidade de exclusão dos X-STRs é devido a diferença de alelos existente entre os marcadores autossomos e os do cromossomo X para os indivíduos do sexo masculino. Estes apresentam sempre dois alelos para os marcadores autossomos e, apenas um, para os do cromossomo X. Dessa forma, no mecanismo de herança genética, o pai transfere 100% do seu perfil haplotípico do cromossomo X para a filha (exceto em casos de mutação) diferentemente do que ocorre com os autossomos, nos quais apenas 50% da informação genética paterna é herdada (MARTINS, 2008).

Considerando essa propriedade do cromossomo X, a maior vantagem da sua aplicação está nos testes de exclusão de paternidade, cujo material biológico do suposto pai não está disponível e o DNA de seus parentes é analisado para a reconstrução do seu perfil genético. Analisando apenas os STRs autossômicos, em muitas situações estes casos permanecem sem solução, pois para se ter adequada reconstrução do perfil do suposto pai ausente, depende da participação de inúmeros parentes genotipados, que na maioria das vezes não estão disponíveis. Isso ocorre com frequência em situações de guerras e migração, quando há necessidade de aproximar familiares, bem como para identificar cadáveres, vítimas de guerras e de desastres em massa (BECKER et al., 2008; LIU et al., 2008; TARIQ et al., 2008).

Testes de paternidade que envolve trios: mãe, filho e suposto pai, são na maioria das vezes solucionados apenas com STRs autossomos e não necessitam de marcadores adicionais. Se o teste de paternidade abrange uma dupla pai e filho, os X-STRs também não são utilizados, já que não há herança paterna do cromossomo X. Entretanto, se a dupla for pai e filha, o uso de X-STRs é importante, principalmente quando se trata de amostra degradada, oriunda de exumação. Para duplas mãe e filha os STRs do cromossomo X não têm nenhuma vantagem em relação aos STRs dos autossomos. Entretanto, em duos mãe e filho os X-STRs são mais eficientes que os marcadores autossomos, neste caso a chance de exclusão é igual à obtida com X-STRs na dupla pai e filha. A opção por utilizar microssatélites do cromossomo X deve ser considerada principalmente em casos de cadáveres putrefeitos ou em outras amostras de difícil análise. Nestas situações obtém-se um significativo poder estatístico com a análise de poucos X-STRs (SZIBOR, 2007).

Estudos de irmandade ou meia-irmandade, em pessoas do sexo feminino, exibem excelentes resultados com X-STRs. Quando dois indivíduos testados possuem o mesmo pai, o cromossomo X de origem paterna será comum entre ambos. Dessa forma, em uma investigação em que duas irmãs ou meio-irmãs são tipadas, a presença de quatro diferentes alelos em um marcador do cromossomo X permite a exclusão da paternidade comum entre elas, mesmo se nenhum dos pais biológicos ou seus parentes forem analisados (KOBACHUK, 2012).

O cromossomo X também é muito útil em casos de paternidade em que os supostos pais são parentes. Imaginemos que os supostos pais são pai e filho, analisando os STRs autossomos, estes obrigatoriamente compartilham 50% dos alelos e o exame com um deles pode aparecer com apenas uma inconsistência genética, o que dificulta a conclusão dos exames, pois esta inconsistência pode ser atribuída a uma mutação paterna. No caso dos X-STRs, nenhum alelo entre pai e filho é idêntico por descendência, uma vez que estes herdam o cromossomo X de suas mães e ambas são diferentes, esta informação facilita a resolução desses casos. Já no caso dos supostos pais serem dois irmãos, ambos compartilham o cromossomo X da mesma mãe e, portanto há uma chance de 50% destes possuírem alelos idênticos por descendência. Nestas situações, os X-STRs e os STRs autossomos podem ser considerados sistemas equivalentes (KOBACHUK, 2012).

Considerando o grande potencial de contribuição dos cromossomos X em análises forenses, alguns marcadores estão sendo propostos para esta finalidade. Todavia, para interpretar os resultados dos exames periciais de DNA e utilizar os dados obtidos como ferramenta forense é preciso criar bancos de frequências alélicas de referência (ZARRABEITIA et al., 2007).

Grande parte dos estudos de marcadores STRs do cromossomo X foram realizados em populações da Ásia (GU & LI, 2006; KANG & LI, 2006; LIU et al.,2008; CHEN et al.,2009; LIM et al.,2009) e da Europa (ROBINO et al.,2006; ZARRABEITIA et al.,2007; ALER et al.,2007; TURRINA et al.,2007; BECKER et al.,2008; ZALÁN et al.,2008; FRACASSO et al.,2008).

Para as populações africanas (POESTSC et al.,2009; GOMES et al.,2009) e latino americanas (RIBEIRO et al.,2008; TAVARES et al.,2008; GUSMÃO et al.,2009; MARTINS et al.,2009; MARTINS et al.,2010) as publicações são menos frequentes. Vários STRs do cromossomo X foram validados para uso forense (figura 14), todavia, mais estudos de frequências alélicas são necessários, especialmente em nível regional, para avaliação dos polimorfismos em diferentes populações e no estabelecimento de um banco de dados útil para aplicações forenses, bem como para estudos antropológicos (MARTINS et al.,2010).

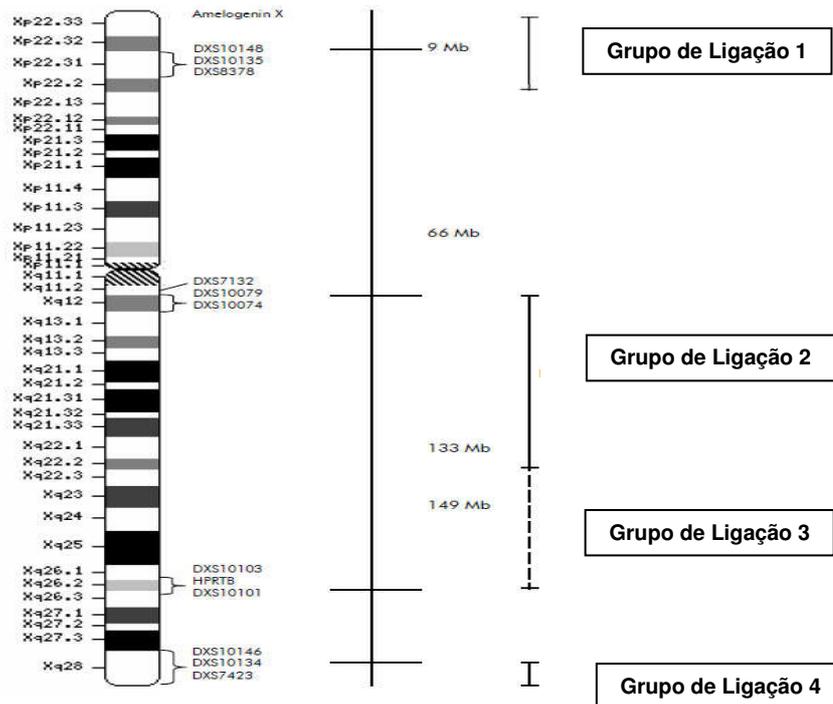


Figura 14: Idiograma do cromossomo X.. Fonte: Investigator Argus X-12 Handbook Qiagen 2013.

O mercado já disponibiliza *kits* comerciais validados, que permitem a amplificação de múltiplos *loci* do cromossomo X, facilitando assim a amplificação e genotipagem (Figura 15 e 16).

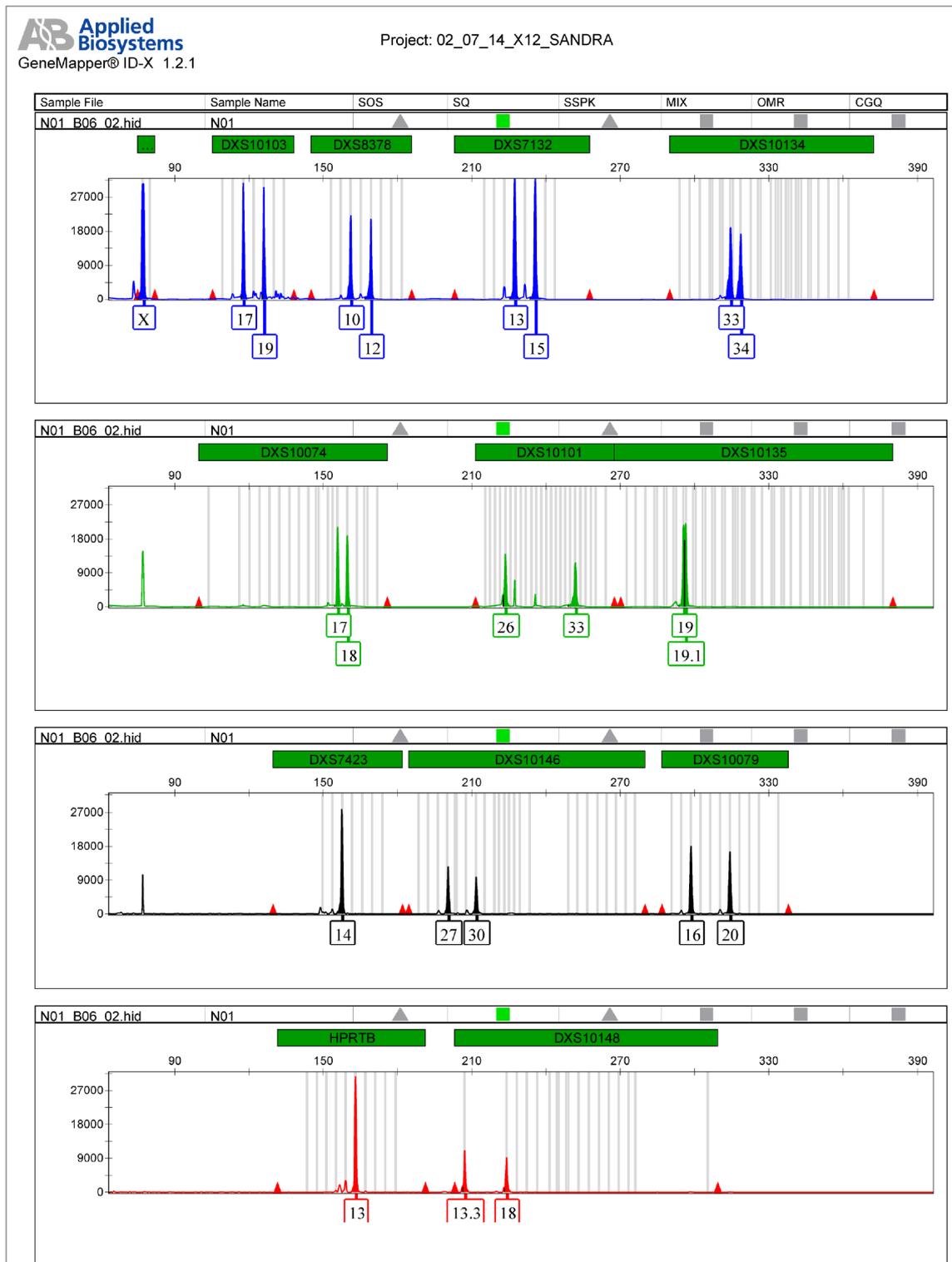


Figura 15. Eletroferograma de indivíduo do sexo feminino, exibindo 12 regiões do cromossomo X.

Fonte: Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, 2014.

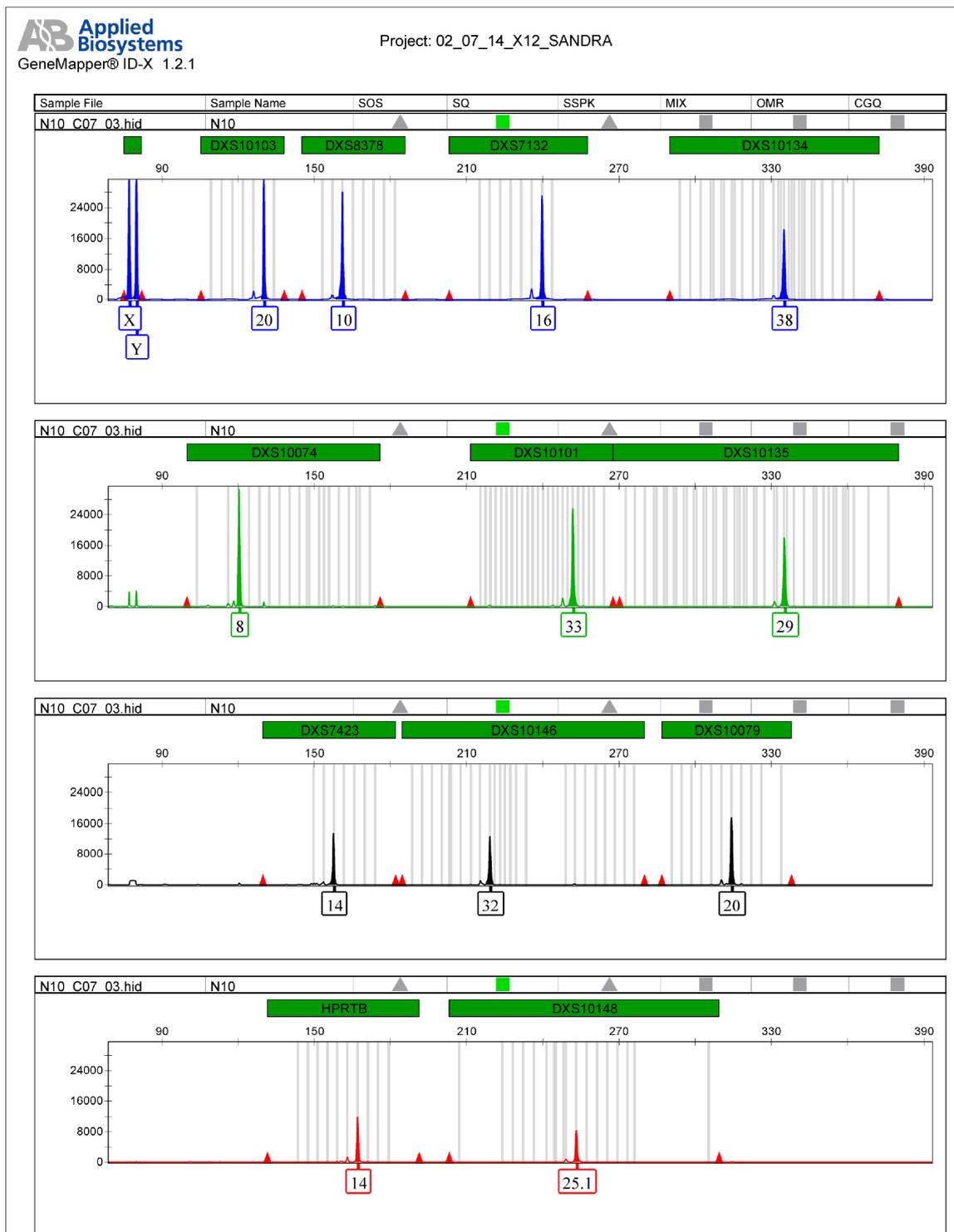


Figura 16. Eletroferograma de indivíduo do sexo masculino, exibindo 12 regiões do cromossomo X.  
 Fonte: Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, 2014.

### 2.4.3 Cromossomo Y

Entre os cromossomos humanos, o cromossomo Y é único em que a maior parte dele, excluindo a região denominada pseudo-autossomal não participa da recombinação meiótica sendo herdado inalterado de pai para filho. Por esta razão, análises do cromossomo Y humano têm sido muito úteis em estudos evolutivos, através da linhagem masculina, bem como para estudos forenses, incluindo casos de violência sexual e os casos de paternidade, onde o suposto pai ou outros parentes estão ausentes (GRIFFITHS et al., 2009).

Devido a ausência de um elemento homólogo a maior parte do cromossomo Y não recombina durante a meiose. Todas as sequências localizadas na região não-recombinante do cromossomo Y, que compreende aproximadamente 95% do cromossomo e corresponde a porção exclusivamente masculina do genoma humano, são herdadas como um bloco. Dessa forma, a informação passa de pai para filho sem que haja variação ou alterações, exceto aquelas originadas por mutações durante a formação dos gametas (Figura 17) (KRENKE, et al., 2005).

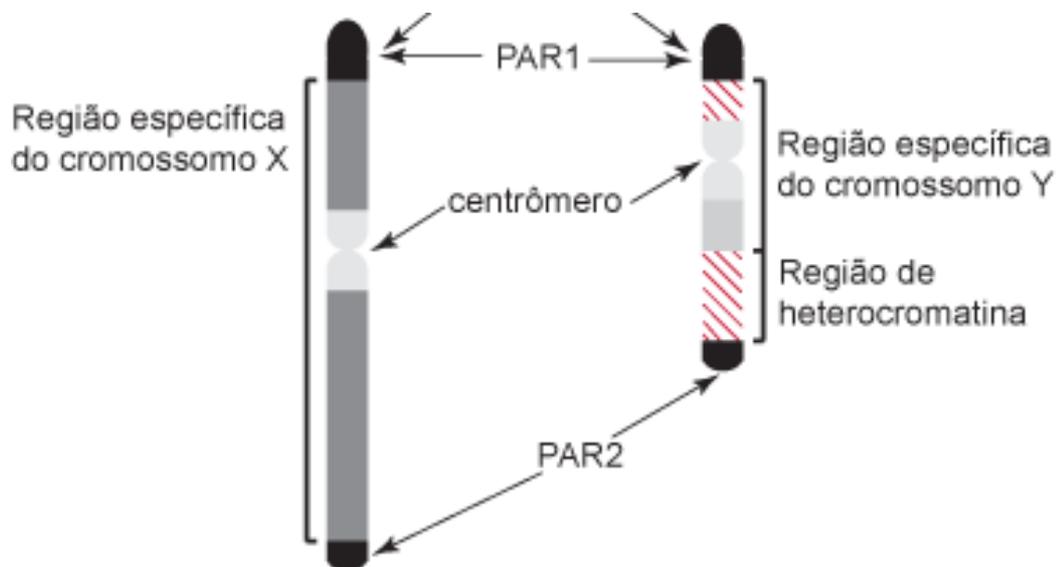


Figura 17: Regiões pseudo-autossômicas dos cromossomos sexuais (JORGE et al., 2008).

Tal como acontece com o mtDNA (DNA mitocondrial), os marcadores do cromossoma Y são conhecidos como “marcadores de linhagem”, tendo em vista que são herdados exclusivamente de um dos progenitores. São transmitidos de geração em geração sem alterações, definindo linhagens, neste caso linhagens paternas (LJUBKOVIC et al., 2008)

Determinadas características apresentadas pelo cromossomo Y o tornam particularmente úteis para estudos de vínculo de parentesco, bem como para análise de vestígios forenses. No entanto, o modo de transmissão implica limitações no contexto da criminalística biológica, uma vez que a mesma informação genética é partilhada por todos os indivíduos do sexo masculino pertencentes a mesma linhagem paterna (LJUBKOVIC et al., 2008).

Após o lançamento a divulgação da sequência do genoma humano, o conhecimento sobre os *loci* STR do cromossomo Y sofreu significativa evolução. Na década de 90 poucos Y-STR estavam caracterizados e disponíveis para utilização em testes de identificação humana pelo DNA. Em 2002, eram somente 30 os Y-STR disponíveis (NIST, 2003). Em 2006, já havia citações de 417 *loci* Y-STR disponíveis para potencial uso na prática forense (HANSON & BALLANTYNE, 2006).

Segundo a comunidade forense europeia, os marcadores DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 e DYS385a/b, constituem o chamado haplótipo mínimo dos Y-STR (TURI et al., 2009).

Em Janeiro de 2003 o U.S. Scientific Working Group on DNA Analysis Methods (SWGDM) recomendou o acréscimo de dois novos marcadores: DYS438 e DYS439 (BUTLER, 2003) que haviam sido descritos três anos antes, acarretando assim, a expansão do haplótipo mínimo. Posteriormente diversos *loci* adicionais foram descritos, que combinados com os do haplótipo mínimo, potencializam o poder de discriminação (ROEWER, 2009).

O primeiro kit comercial de Y-STR foi disponibilizado comercialmente em Janeiro de 2001 e incluía seis marcadores: DYS19, DYS389II, DYS390, DYS391, DYS393 e DYS385a/b. Hoje, diversos Y-STR estão sendo usados como marcadores (Figura 18). O mercado disponibiliza diferentes kits multiplex, devidamente validados para a casuística forense (GUSMÃO et al., 2009) (Figura 19).

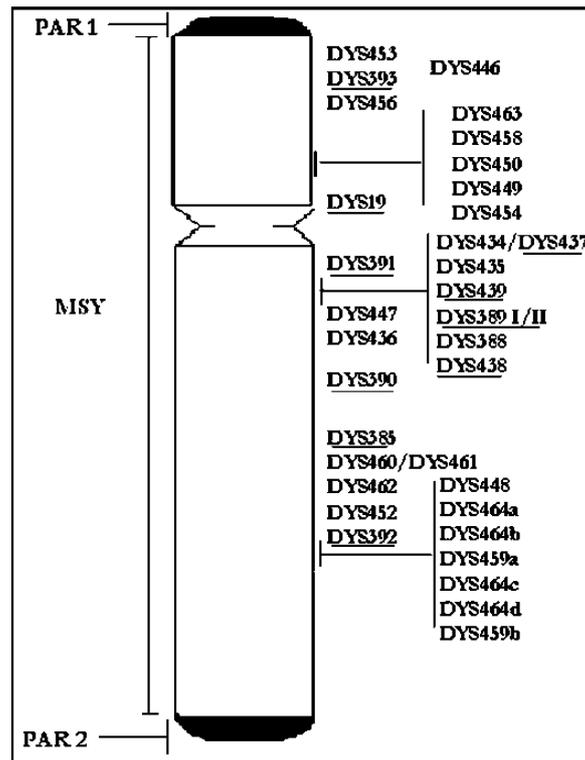


Figura 18: Idiograma do cromossomo Y. Fonte: LEITE, 2006b.

A importância do cromossomo Y nas análises forenses está diretamente ligada ao fato deste se encontrar somente nos indivíduos do sexo masculino. Considerando que determinados crimes, particularmente os estupros, envolve como perpetradores indivíduos do sexo masculino, testes de DNA desenvolvidos especificamente para examinar a porção masculina da amostra biológica tornaram-se extremamente valiosos. Assim, STRs do cromossomo Y têm sido extensivamente utilizados na Genética Forense para a identificação de pessoas do sexo masculino, como prova material de violência sexual (RODIG et al., 2007).

Em ocasiões em que as informações genéticas obtidas a partir dos marcadores autossômicos se revelam insuficientes, muitas vezes devido a quantidade e/ou qualidade da amostra biológica, situação bastante comum em amostras biológicas originadas de agressões sexuais, quando existe uma pequena quantidade de DNA masculino misturado com grandes quantidades de DNA feminino, a análise do cromossomo Y poderá decidir o rumo das investigações (TURI et al., 2009).

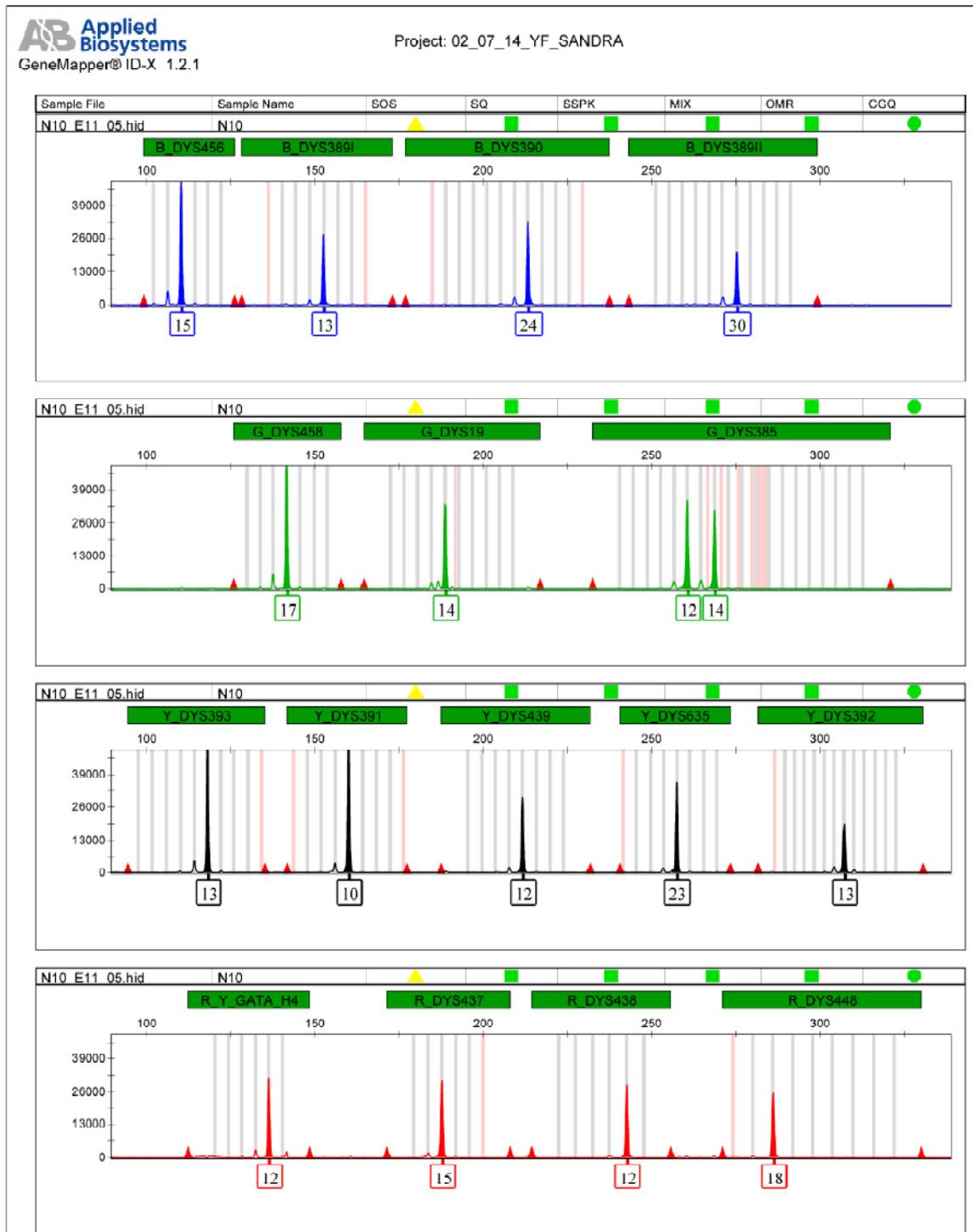


Figura 19. Eletroferograma exibindo 17 regiões do cromossomo Y. Fonte: Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, 2014.

Quando uma amostra contém o DNA de mais de um contribuinte, a interpretação do seu perfil genético torna-se complicada. A incidência, a complexidade e a importância dos perfis mistos estão aumentando devido à sensibilidade da reação em cadeia da polimerase. Amostras de DNA de pelo menos dois participantes podem ser originadas na cena do crime ou no âmbito da manipulação da amostra, durante a coleta, armazenamento e/ou exames laboratoriais (ROEWER, 2009).

A análise de áreas polimórficas do cromossomo Y contribui para aumentar o painel de *loci* STR autossômicos utilizados para a identificação humana. Estes testes podem ser usados tanto para estudos genealógicos, quanto para crimes sexuais que envolvam mistura de fluidos biológicos masculino e feminino. Nestas situações, havendo mais material procedente da vítima do que do violador, tornar-se difícil a tipagem do DNA masculino pelos testes convencionais de STR. Os testes Y-STR ocupam-se apenas do DNA dos homens, facilitando a detecção e a tipagem do componente masculino. As desvantagens são o baixo poder discriminatório, quando comparados aos testes convencionais de STR, e a não distinção de homens da mesma linhagem paterna (GUSMÃO et al., 2009).

Nos últimos anos diversos estudos demonstrando a aplicabilidade e eficácia dos Y-STR foram publicados. A possibilidade de se determinarem vínculos de filiação na ausência do progenitor, a investigação da identidade e do sexo das vítimas de acidentes e/ou desastres de massa e a identificação de criminosos, especialmente os agressores sexuais são importantes aplicações dos marcadores do cromossomo Y (DELFIN et al., 2005).

A literatura e a rotina dos laboratórios de DNA forense demonstram que a tipagem de Y-STR é especialmente útil na análise de amostras de mistura envolvendo: (1) um ou mais indivíduos do sexo masculino (2) homens azoospérmicos ou vasectomizados e (3) a presença de mistura de outros fluidos corporais (ex.: mucosa bucal, pele, etc.) de vítimas e suspeitos de sexo distinto (ROEWER, 2009).

A capacidade de detectar e discriminar DNA masculino torna os marcadores STR do cromossomo Y um vantajoso acréscimo aos já bem estabelecidos STR autossômicos. Além disso, o cromossomo Y pode também ser usado para outros fins, tais como: traçar padrões das migrações humanas e para responder a questões antropológicas, históricas e genealógicas (TURI et al., 2009).

#### 2.4.4 DNA MITOCONDRIAL

O genoma mitocondrial (mtDNA) não possui marcadores genéticos do tipo STRs. A sequência de nucleotídeos do genoma mitocondrial foi estabelecida a mais de 30 anos. As variações observadas entre os diferentes mtDNAs são principalmente de polimorfismos de um único nucleotídeo. Mesmo sendo observados em todo mtDNA, os polimorfismos são muito mais frequentes na “região controle”, local da abertura das fitas do DNA, que promovem a replicação do genoma. Nessa região existem três regiões hipervariáveis. Os laboratórios de Genética Forense normalmente analisam duas dessas regiões: HV1 e HV2 (Figura 20) (BUDOWLE et al., 2003).

As mitocôndrias seguem o padrão de transmissão materno, pois são repassadas à prole através do óvulo. Apesar de todas as pessoas, independentemente do sexo, possuírem mitocôndrias, elas são transmitidas exclusivamente pelas mães (BRATIC & LARSSON, 2013).

Da mesma forma que o cromossomo Y, a análise do mtDNA não permite a individualização, visto que todos os indivíduos de uma mesma linhagem possuem o mtDNA exatamente igual. Cada célula possui dezenas e até centenas de mitocôndrias, dependendo do tecido no qual estão localizadas. Além disso, cada mitocôndria possui várias cópias do seu genoma. Estas características permitem a ocorrência de inúmeras cópias de mtDNA em cada célula, superando e muito a quantidade do DNA nuclear (WANG et al., 2011).

Diante do exposto, a análise do DNA mitocondrial é muito utilizada quando o DNA autossômico encontra-se muito degradado e já não oferece condições de análise. Além disso, como todos os integrantes de uma mesma linhagem materna possuem o mesmo mtDNA, o estudo desse marcador é muito utilizada quando se deseja estabelecer vínculos maternos distante (BUDOWLE et al., 2003).

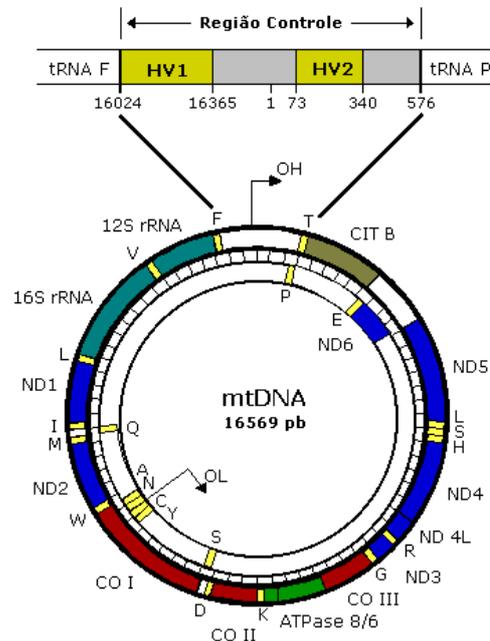


Figura 20. Organização do DNA mitocondrial humano. Fonte: BUTLER, 2005. Adaptado.

## 2.5 O EXAME DE DNA

Passaram-se quase 30 anos desde que o primeiro exame de DNA foi utilizado para identificar um indivíduo. Desde as primeiras aplicações da Genética com o objetivo de solucionar crimes ficou claro que amostras forenses podem fornecer DNA de alta qualidade e em número suficiente para individualizar pessoas, partindo de quantidades ínfimas de diversos materiais biológicos, tais como: sangue, cabelos, sêmen, saliva, pele ou qualquer outro fluido biológico, possibilitando a associação de um suspeito com um determinado crime (KAYSER et al. 2013).

Para a realização de um exame de DNA, existem etapas básicas que são executadas independentemente da finalidade do exame, são elas: extração e purificação do DNA, quantificação, amplificação, separação e detecção dos marcadores genéticos por eletroforese capilar, interpretação dos perfis e confronto genético (FOSTER & LAURIN, 2012).

### **2.5.1 EXTRAÇÃO DO DNA**

A extração do DNA é a etapa na qual o DNA total da célula é separado (nuclear e mitocondrial) dos demais componentes celulares, como proteínas, lipídeos e polissacarídeos (BUTLER, 2012).

Inúmeras metodologias de extração foram desenvolvidas, objetivando a separação das proteínas e outros materiais celulares das moléculas de DNA. Nesta fase, também é feita a purificação, visando livrar o material biológico de possíveis contaminantes. Considerando que grande parte das amostras biológicas que chegam a um laboratório de genética forense tem origem em locais de crimes diversos, diferentes substratos podem fornecer componentes capazes de interferir nas outras análises, são os inibidores. O objetivo final desta etapa é obter DNA de boa qualidade e em quantidade suficiente que possibilitem a realização das etapas posteriores. Esta etapa é composta basicamente de lise celular, separação das moléculas de DNA de outros materiais celulares e isolamento do DNA (STRAY et al., 2010).

### **2.5.2 QUANTIFICAÇÃO**

Após a extração do DNA, é recomendável determinar a concentração de DNA humano presente na amostra. Esta etapa pode ser feita de várias maneiras. Os laboratórios forenses comumente utilizam a PCR em tempo real. Além da concentração do DNA, pode-se determinar a presença de inibidores e ainda verificar se há sinais de degradação. Outra importante informação que pode ser obtida nesta fase é se o DNA presente na amostra pertence a um indivíduo do sexo masculino, feminino ou ambos (SILVA, et al., 2013a).

### **2.5.3 AMPLIFICAÇÃO DOS MARCADORES GENÉTICOS**

Após a extração e a quantificação inicia-se a amplificação dos marcadores genéticos. A amplificação é feita por meio da PCR, técnica que revolucionou a Biologia Molecular, desde a sua invenção (FOSTER & LAURIN, 2012).

Trata-se de uma reação bioquímica que possibilitou a amplificação de regiões específicas do DNA a partir de quantidades muito pequenas. O impacto para a Genética Forense foi e continua sendo espetacular, uma vez que permite a obtenção de perfis genéticos a partir de amostras ínfimas. A PCR é uma amplificação *in vitro* de uma sequência específica do DNA que acontece toda vez que uma célula entra no processo de divisão celular (Figura 21) (SILVA, et al. 2013a).

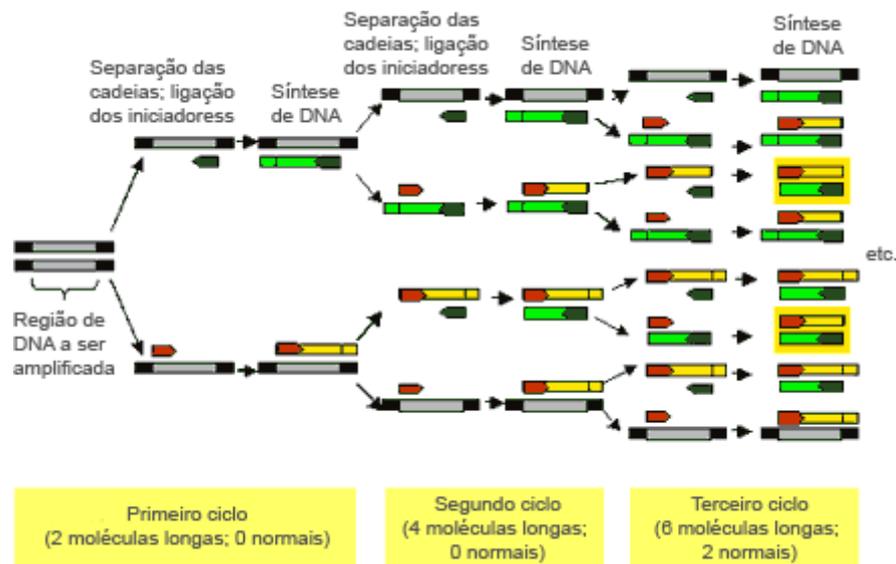


Figura 21: Reação em cadeia da polimerase do DNA. Fonte: SILVA, 2013a. Adaptado.

#### 2.5.4 DETECÇÃO DOS MARCADORES GENÉTICOS

Após a amplificação é chegada a hora de separar e detectar os produtos de amplificação. Nos laboratórios forenses essa etapa é feita com a utilização de analisadores genéticos ou seqüenciadores de DNA (Figura 22). No seqüenciador ocorre a separação por tamanho dos produtos de amplificação e posteriormente a detecção dos produtos por meio de fluorescência (Figuras 12, 13, 15, 16 e 19) (BARBOSA, 2010; BERGLUND et al. 2011).



Figura 22: Analisador genético de DNA, ABI 3500. Fonte: Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, 2014.

### 2.5.5 INTERPRETAÇÃO DOS PERFIS GENÉTICOS

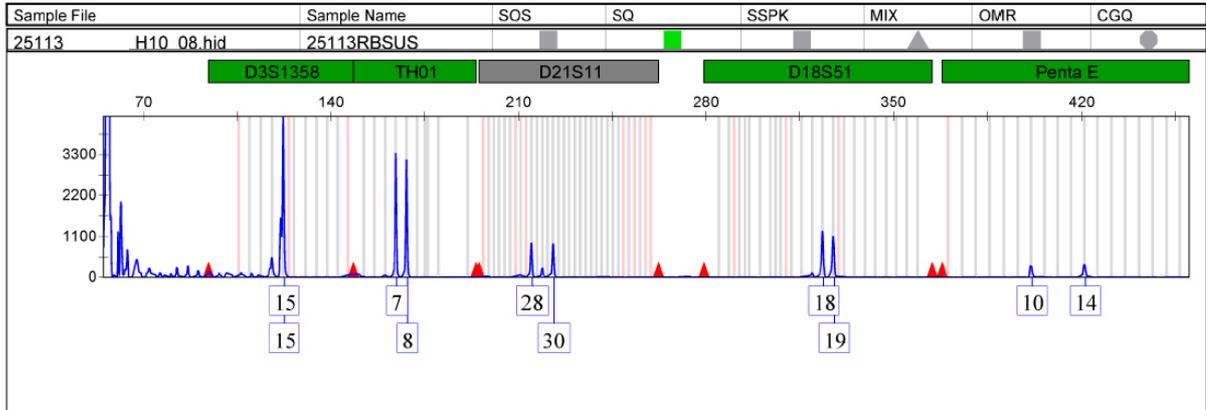
Ao final da amplificação e detecção dos marcadores genéticos, segue-se com a interpretação dos perfis genéticos, utilizando-se programas específicos. O eletroferograma (Figuras 12, 13, 15, 16 e 19) exibe em cada pico um produto de amplificação, que corresponde a um alelo. O número de repetições de STRs de cada alelo amplificado é estabelecido e o perfil genético da amostra é obtido (VELHO, et al. 2013b).

### 2.5.6 O CONFRONTO GENÉTICO

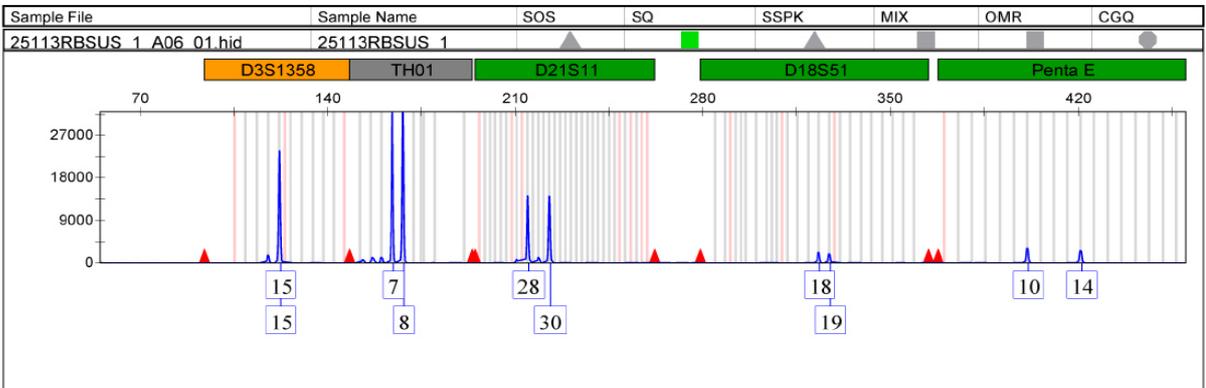
Nesta etapa, o perfil genético obtido a partir de um vestígio é comparado com os perfis dos suspeitos ou da vítima. No caso de um crime, por exemplo, o perfil genético obtido a partir de um vestígio é comparado com os perfis dos suspeitos. Quando o perfil genético do suspeito e do vestígio não são idênticos se estabelece uma “exclusão”. No caso de exclusão do suspeito não é feita nenhuma análise estatística. Por outro lado, quando o perfil do suspeito é idêntico ao do vestígio, calcula-se a probabilidade ou a frequência da ocorrência do genótipo observado na população da qual o indivíduo faz parte (Figura 23). Neste momento, são utilizadas as frequências alélicas para cada um dos marcadores analisados obtidas de indivíduos escolhidos ao acaso na população em questão, objeto do presente estudo (VELHO et al., 2013b).

Dessa forma, fundamentos da Genética de Populações são utilizados para a realização dos cálculos da frequência esperada do genótipo na população em que pertence o suspeito. Quando muitos marcadores genéticos são analisados, a frequência esperada de um determinado genótipo pode ser muito baixa, como de um indivíduo em alguns milhões e até bilhões, trilhões de indivíduos. Se a frequência esperada do genótipo observado na população for de 1/1.000.000.000, podemos afirmar que a probabilidade de observarmos o mesmo perfil genético em outro indivíduo, não tipado geneticamente e não aparentado ao suspeito é de 1 em 1 bilhão. Quanto menor for a frequência do genótipo obtido, maior será o peso da evidência genética (BUCKLETON et al., 2004).

### Suspeito 1 (INCLUÍDO)



### VESTÍGIO



### Suspeito 2 (EXCLUÍDO)

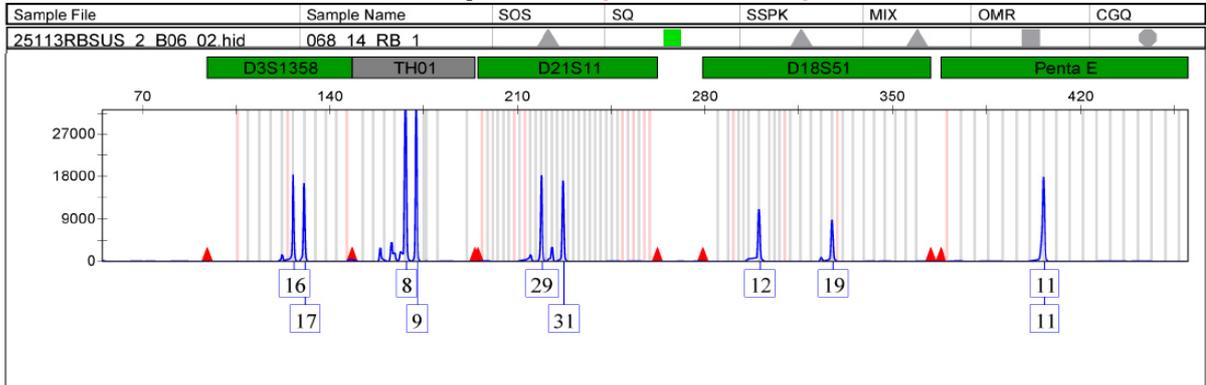


Figura 23: Confronto entre perfis genéticos. Fonte: Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, 2014.

Outra forma de analisarmos o peso da evidência genética é por meio da Razão de Verossimilhança, que é uma razão de duas hipóteses mutuamente excludentes. Duas hipóteses são consideradas quando o perfil genético do suspeito é idêntico ao obtido a partir do vestígio. A primeira hipótese ( $H_1$ ), também conhecida como a hipótese da acusação, supõe que o material é oriundo do suspeito. A segunda hipótese ( $H_2$ ), conhecida como a hipótese de defesa, supõe que o material biológico é oriundo de outro indivíduo qualquer da população que possui o mesmo perfil genético do suspeito (SILVA et al., 2013a).

Os casos para determinação de vínculo genético, utilizados para paternidade oriunda de estupro, por exemplo, são mais complexos, e o confronto depende do tipo de vínculo biológico que está sendo determinado. Se o vínculo genético de paternidade está sendo testado, a mulher é considerada a mãe biológica do filho (a), uma vez que não há dúvidas quanto esse vínculo biológico. Desta forma, é possível estabelecer os alelos paternos obrigatórios, ou seja, os alelos do filho(a) que foram obrigatoriamente herdados do pai (Tabela 4). Quando o suposto pai não possui todos os alelos paternos obrigatórios, é possível afirmar que houve uma “exclusão” de paternidade. Quando o suposto pai não pode ser excluído, é calculado o Índice de Paternidade (IP). O IP é também uma razão de probabilidades ( $IP=X/Y$ ) (BUTLER et al., 2011).

Tabela 4: Herança dos alelos num caso de paternidade oriundo de estupro. Fonte: Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, 2014.

<b>Marcadores</b>	<b>Suposto Pai</b>	<b>Filho</b>	<b>Mãe</b>
D3S1358	13-15	<b>13-15</b>	15-15
vWA	16-16	<b>16-16</b>	16-17
D16S539	9-13	<b>9-11</b>	11-13
CSF1PO	10-12	<b>10-10</b>	10-10
TPOX	6-8	<b>6-8</b>	6-8
Indel	2	<b>2</b>	—
AMELOGENINA	XY	<b>XY</b>	XX
D8S1179	13-14	<b>13-13</b>	13-15
D21S11	30-32.2	<b>30-30</b>	30-31.2
D18S51	16-19	<b>12-19</b>	12-19
DYS391	11-11	<b>11-11</b>	—
D2S441	11-14	<b>10-11</b>	10-11
D19S433	12.2-14.2	<b>12.2-14</b>	12.2-14
TH01	7-8	<b>7-8</b>	7-8
FGA	20-22	<b>20-22</b>	20-20
D22S1045	16-17	<b>11-16</b>	11-17
D5S818	8-12	<b>8-12</b>	9-12
D13S317	12-12	<b>12-12</b>	12-12
D7S820	10-12	<b>10-12</b>	10-11
SE33	18-19	<b>18-19</b>	14-18
D10S1248	13-14	<b>14-14</b>	14-15
D1S1656	15-16	<b>15-16.3</b>	16-16.3
D12S391	21-24	<b>24-24</b>	18-24
D2S1338	22-25	<b>22-23</b>	23-25

Nas situações e que os marcadores uniparentais são utilizados, como os STRs do cromossomo Y ou DNA mitocondrial, a análise é diferente, uma vez que não ocorre recombinação nesses marcadores e eles são herdados somente de um dos progenitores (BUCKLETON et al., 2004).

## 2.6 PRINCÍPIOS DE GENÉTICA DE POPULAÇÕES APLICADOS AO ESTUDO DA GENÉTICA FORENSE

Os estudos genéticos de populações humanas são úteis especialmente quando se busca a compreensão da história, demografia e evolução da espécie humana, além de estudos de ancestralidade e de mapeamento de genes de relevância clínica. Outra aplicação prática dá-se na investigação forense, incluindo testes de paternidade, bem como exames de identificação humana destinada a investigação criminal (LEITE, 2006a).

A genética de populações tem como um dos objetivos estudar as consequências das aplicações dos princípios mendelianos em determinada população, analisando os efeitos das mutações, seleção natural, migração e deriva genética. É de grande interesse para as ciências forenses, quando se faz necessário conhecer alguns parâmetros estatísticos populacionais de referência que permitam aos cientistas forenses concluir de forma satisfatória suas análises periciais de DNA (BEIGUELMAN, 2008).

A distribuição da frequência alélica dos marcadores de DNA difere entre populações geograficamente separadas ou que apresentam origem distinta. Assim, quando se deseja identificar um indivíduo oriundo de determinada população, é indispensável o estudo prévio de diferentes marcadores naquela população para saber quais são os alelos presentes e em que frequência, a fim de se definir quais são os melhores marcadores a serem utilizados (FRANCEZ et al., 2012).

Os microssatélites são úteis para identificação individual e estabelecimento de bancos de dados populacionais, aplicáveis na área forense. A maioria dos estudos publicados objetiva o estabelecimento de bancos de dados regionais em diversos países, enfatizando a aplicação destes marcadores em estudos de identificação humana (KOHLRAUSCH et al. 2005).

Para a verificação da aplicabilidade de um banco de dados de frequências alélicas deve-se primeiro averiguar se o quantitativo de indivíduos da amostra consegue representar seguramente a população em questão. Estudos concluem que de 100 a 150 indivíduos por população consegue prover uma amostra adequada para todos os *locus* genéticos utilizados em Genética Forense. Amostras acima de 200 indivíduos não acrescentam genótipos significativamente representativos na população (KOBACHUK, 2012).

A seguir algumas definições básicas a respeito dos parâmetros estatísticos, de interesse forense, apresentados neste estudo:

### **Frequências alélicas**

Normalmente as frequências alélicas de todos os marcadores analisados são obtidas pelo método de contagem direta, ou seja, verifica-se o número de vezes em que o alelo é observado, dividido pelo número total de alelos averiguados para o marcador. Para se fazer este tipo de cálculo é comum o uso de softwares que auxiliam a contagem, a exemplo do Arlequin v.3.1 (EXCOFFIER, 2010).

### **Heterozigose (ou heterozigosidade)**

É o estudo da proporção de indivíduos com dois alelos diferentes em determinado *locus*, ou seja, corresponde a quantidade de indivíduos heterozigóticos. Quanto maior a heterozigose, maior é a variabilidade genética do *locus* na população em questão. Dessa forma, a probabilidade de que dois indivíduos tenham o mesmo perfil torna-se menor. Considerando tal parâmetro, deve-se utilizar marcadores genéticos que apresentem heterozigose superior a 70%, pois são mais informativos (BUTLER, 2005).

A heterozigose pode ser esperada ou observada. A heterozigose observada estabelece a proporção real de heterozigotos para um determinado *locus* na amostra analisada. É calculada diretamente dos genótipos encontrados na população. Trata-se número de heterozigotos em cada *locus* dividido pelo número total de indivíduos pesquisados. A heterozigose esperada é a fração estimada de heterozigotos para um *locus*, assumindo a condição de equilíbrio de Hardy-Weinberg na população. Difere-se da observada por ser calculada a partir das frequências alélicas segundo a fórmula (McMANUS et al., 2011):

$$\begin{array}{l} \text{Heterozigotos} + \text{Homozigotos} = 1 \\ \text{Heterozigotos} = 1 - \text{Homozigotos} \end{array}$$

Homozigotos = soma das frequências alélicas ao quadrado (Freq. Alelo 1<sup>2</sup> + Freq. Alelo 2<sup>2</sup> + Freq. Alelo 3<sup>2</sup> + Freq. Alelo 4<sup>2</sup> + .... + Freq. Alelo N<sup>2</sup>)

### Equilíbrio de Hardy-Weinberg

A partir das frequências alélicas observadas são obtidas as frequências genotípicas esperadas considerando que as frequências alélicas não devem sofrer modificações ao longo das gerações se o *locus* é estável. Quando a frequência genotípica esperada é semelhante à frequência genotípica observada pressupõe-se que o *locus* na população analisada está em equilíbrio de Hardy-Weinberg (BUTLER, 2005). O importante é definir se os alelos encontrados dentro do *locus* são independentes, caso contrário, a população está em desequilíbrio.

A população pode estar em desequilíbrio quando se analisa sub populações ou se tem excesso de homozigose devido a amplificação preferencial de um alelo em detrimento de outro. Quando necessário, os valores do equilíbrio de Hardy-Weinberg podem ser corrigidos pelo cálculo de Bonferroni para neutralizar a análise de comparações múltiplas (BUTLER, 2005).

### Desequilíbrio de ligação (DL)

Os STRs localizados no mesmo cromossomo são denominados sintênicos e quando utilizados em cálculos estatísticos objetivando a identificação humana, o ideal é que estejam a uma distância física de pelo menos 50 mega bases entre si. Esta regra é utilizada para garantir a independência entre os *loci*. Para STRs localizados à uma distância menor é necessário calcular o desequilíbrio de ligação, ou seja, a independência entre eles (O'CONNOR & TILLMAR, 2012).

**Conteúdo de informação polimórfica (CIP)**

Este parâmetro é utilizado para aferir o quanto um marcador genético é informativo. Medida capaz de estimar o quão polimórfico é o STR. O grau de polimorfismo é determinado pelo número de alelos e suas frequências. Recomenda-se utilizar marcadores com conteúdo polimórfico maior que 0,5 (50%), pois possuem maior eficácia em diferenciar indivíduos (BUTLER, 2005).

**Probabilidade de coincidência (“*matching*”) (PM)**

É a probabilidade de que dois indivíduos selecionados ao acaso tenham genótipo idêntico para um determinado *locus*. É calculado pela soma dos quadrados das frequências genotípicas observadas. Pois, a chance de encontrar um determinado genótipo = sua frequência observada; a chance de encontrar um determinado genótipo e, em seguida, o mesmo genótipo = sua frequência observada ao quadrado. Dessa forma, a soma das frequências observadas ao quadrado é igual a probabilidade de coincidência para o *locus* (TILLMAR , 2010).

**Poder de discriminação (PD)**

É a probabilidade de encontrar dois indivíduos ao acaso na população e eles possuírem perfis genéticos diferentes. Refere-se da probabilidade contrária de se encontrar dois perfis iguais (probabilidade de coincidência). É calculado pela fórmula  $PD = 1 - PM$  (BUTLER, 2005).

**Poder de exclusão (PE)**

Corresponde à probabilidade de se excluir qualquer indivíduo que tenha sido falsamente acusado em teste de paternidade. Equivalente a frequência de todos os homens, em uma determinada população, que não tem os alelos concordantes com o alelo paterno obrigatório (alelo herdado do pai biológico) da criança em um teste de paternidade. O valor é estabelecido pela seguinte fórmula:  $PE = h^2 (1 - 2hH^2)$ , onde  $h$  significa a frequência observada de heterozigotose  $H$  é a frequência observada de homozigotos (BUTLER, 2005).

### Índice de paternidade típico

Estabelece a probabilidade de que o indivíduo que está sendo testado seja realmente o pai biológico da criança, do que um homem selecionado aleatoriamente. Os índices de paternidade (IP) de cada *locus* representam a razão entre a probabilidade do suposto pai e a probabilidade de qualquer outro indivíduo da população, transmitir o alelo em questão para o filho. A probabilidade de qualquer outro indivíduo transmitir o alelo paterno obrigatório corresponde a frequência do alelo na população avaliada (BUTLER, 2005).

O **índice de paternidade combinado (IPC)** é o resultado dos índices de paternidade individuais para cada *locus* investigado e determina a força da evidência genética do suposto pai ser o pai biológico. Dessa forma, quanto maior o índice maior é a evidência genética (BUTLER, 2005).

### Distância genética

A finalidade desta análise é investigar a semelhança genética entre as populações analisadas (BUTLER, 2005).

## 2.7 CONCEITOS E FATORES QUE INFLUENCIAM DIRETAMENTE A ANÁLISE DE DNA NOS LABORATÓRIOS FORENSES DO BRASIL

Dentre as atividades de um laboratório forense de genética deve-se destacar a busca de perfis genéticos presentes em vestígios relacionados à violência sexual, homicídios, crimes contra o patrimônio, entre outras situações. Estes laboratórios também são os responsáveis pela identificação de vítimas de acidentes em massa, além do gerenciamento do banco de dados de perfis genéticos (SILVA, et al. 2013a).

No âmbito pericial, as amostras biológicas são comumente classificadas como: **amostra de referência** (aquelas de origem conhecida, oriundas de vítimas, suspeitos, acusados, réus ou de seus parentes diretos (pai, mãe irmãos, filhos, etc) e **amostra questionada** (vestígios derivados do local de crime, de objetos relacionados a ocorrências criminais ou de quaisquer outros pontos ou pessoas, cujas origens não sejam conhecidas) (PERNAMBUCO, 2012).

Considerando as peculiaridades e o grau de dificuldade exibidos nas amostras biológicas destinadas à análise pericial do DNA, é possível classificar estas amostras de acordo com sua origem. Dessa forma, destacam-se: a **identificação humana de cadáver de identidade ignorada** (IH), os **crimes sexuais** (CS), o **confronto de vestígios** (CV) e o **banco de dados de perfis genéticos** (BDPF).

### 2.7.1 Identificação de Cadáveres de Identidade Ignorada

Procedimentos modernos para a identificação humana estão diretamente associados ao uso de técnicas variadas para análise de caracteres, notadamente moleculares, que são compartilhados de diversas formas na população (BROOKFIELD, 1995).

A identificação de vítimas de grandes desastres ou vítimas de qualquer situação forense (homicídio, ocorrência de trânsito, incêndio, etc) que resultem em cadáveres com grave fragmentação, decomposição e/ou mistura de restos mortais, incluindo diversos indivíduos, fazem parte da rotina de um laboratório de genética forense (ZIETKIEWICZ et al. 2012.)

Quando o corpo está bem preservado a identificação pode ser feita através da impressão digital, desde que existam dados *ante mortem* destas características (MALAVER & YUNIS, 2003). Quando o corpo está em estado de decomposição muito avançado, carbonizado, esqueletizado ou outras situações que impossibilitem a utilização da impressão digital ou odontologia legal, faz-se necessária a análise do DNA (PRIMORAC & SCHANFIELD, 2000).

### 2.7.2 Crimes Sexuais

O estupro é um crime violento com implicações médicas, psicológicas, legais e sociais. Mais do que um ato sexual, trata-se de um crime agressivo e covarde, cujas principais vítimas são as crianças e adolescentes. É capaz de desencadear reações emocionais complexas na vítima, frequentemente mais significativas do que o dano físico (CERQUEIRA & CRUZ, 2014).

A violência sexual é um fenômeno de caráter universal, que atinge indistintamente todas as classes sociais, etnias, religiões e culturas. Ocorre em populações de diferentes níveis de desenvolvimento social e econômico, em espaços públicos ou dentro dos lares. Embora possa atingir tanto mulheres quanto homens, em qualquer etapa de suas vidas, os dados acerca deste tipo de violência, mostram que as principais vítimas são mulheres jovens e crianças. Pelos danos que acarreta e pelos altos índices de reincidência que a caracterizam, representa um sério problema de segurança e saúde pública em todo o mundo (CERQUEIRA & CRUZ, 2014).

As amostras biológicas relacionadas à violência sexual devem ser coletadas no local do crime, por perito criminal, podendo ser produzidas por vítimas e/ou criminosos, a exemplo de sangue, esperma e saliva encontrados em diversos suportes, como: vestimentas, copos, garrafas, preservativos, etc. O perito médico legista é o profissional responsável pela coleta de amostras biológicas na vítima, durante a perícia sexológica correspondente. Nestes casos, secreções e/ou pelos sobre a vítima, secreções obtidas da cavidade vaginal e/ou anal, restos de pele sob as unhas de vítimas e/ou suspeitos podem oferecer a prova material necessária a identificação do agressor (VUICHARD et al., 2011).

A luta contra a impunidade de crimes sexuais é complexa e requer esforços contínuos. O caminho mais curto e eficiente para a aplicação da justiça é a produção de uma prova material irrefutável. Nestas situações, a coleta e o tratamento adequado de vestígios encontrados, a manutenção de cadeia de custódia e o estabelecimento de um banco de dados são indispensáveis (MICHELIN et al., 2013)

### **2.7.3 Confronto de Vestígios**

O confronto de vestígio está relacionado às amostras biológicas coletadas em locais de crimes diversos e apresentam grau de dificuldade significativo, especialmente no que diz respeito à diversidade, a quantidade e a qualidade das amostras disponíveis (VELHO et al. 2013a).

Para um perito criminal, independentemente da sua especialidade, a preservação do local de crime é tema de grande importância e necessita ser difundido para facilitar o entendimento das questões relacionadas a esses espaços.

O sucesso das análises biológicas de materiais coletados em locais de crime depende visceralmente do isolamento e preservação do mesmo. Amostras biológicas podem ser localizadas em vários suportes, entre eles: ferramentas, impressões digitais, pelos e fibras, manchas de sangue, saliva, esperma, objetos ou suportes tocados por suspeitos, indumentos, armas, etc. (NORÉN et al., 2013).

A existência do vestígio pressupõe a existência de um agente provocador (que o causou ou contribuiu para tanto) e de um suporte adequado (local em que o vestígio se materializou) e este agente provocador na maioria das vezes deixa vestígios biológicos potencialmente analisáveis através do DNA (JOBIM et al., 2006).

#### **2.7.4 Banco de Dados de Perfis Genéticos**

A baixa taxa de elucidação de homicídios no Brasil é preocupante. A fim de reverter este quadro e de avançar a legislação brasileira na direção das mais modernas do mundo, o Governo brasileiro deu um passo histórico para a implantação do banco de dados de perfis genéticos no país. No dia 28 de maio de 2012, foi publicada a Lei nº 12.654, que *“Altera as Leis nos 12.037, de 1º de outubro de 2009, e 7.210, de 11 de julho de 1984 - Lei de Execução Penal, para prever a coleta de perfil genético como forma de identificação criminal, e dá outras providências”* (BRASIL, 2012).

Um dos principais fatores para o sucesso de um banco de DNA é a legislação que o regulamenta. Na maioria dos países a legislação começou de forma mais restritiva, armazenando-se apenas os perfis de condenados por crimes hediondos. Este é o modelo ora adotado no Brasil. O tempo mostrou que quanto mais abrangente for a inserção de perfis de criminosos, mais eficiente será o banco de dados. A Lei acima foi regulamentada no dia 12 de março de 2013, através do Decreto nº 7.950, que *“Institui o Banco Nacional de Perfis Genéticos e a Rede Integrada de Bancos de Perfis Genéticos”* (BRASIL, 2013a).

. Estudos de laboratórios criminais norte-americanos mostram que em mais de 40% dos vestígios encontrados em um local de ocorrência de homicídio ou violência sexual é possível encontrar vestígios biológicos passíveis de serem examinados por análise de DNA. Da mesma maneira, estudos do serviço britânico de ciências forenses indicam que 50% dos delitos de crime contra o patrimônio possuem vestígios biológicos passíveis de serem examinados geneticamente. Contudo, em menos de 1% dos casos, um suspeito é apresentado para a comparação (BRASIL, 2009).

A solução encontrada por países como Inglaterra e Estados Unidos para este problema foi a criação de um banco de dados de DNA. A partir deste ponto, mesmo os crimes sem suspeitos são investigados com exames de DNA, com uma alta taxa de sucesso, uma vez que os bancos de dados daqueles países já possuem armazenados, milhões de perfis genéticos de pessoas que já foram tipadas em outras oportunidades. Os Estados Unidos foram os pioneiros. Hoje diversos países utilizam esta poderosa ferramenta de investigação, entre eles: Alemanha, Áustria, Canadá, Croácia, Bélgica, Chile, Colômbia, Espanha. Hungria, etc (SILVA, et al. 2013b).

Os Bancos de Dados de DNA podem acelerar as investigações e permitir troca de informações nacional e internacionalmente. O estudo de diferentes regiões polimórficas do DNA humano, obtidas de vestígios coletados em locais de crimes, organizadas em Bancos de Dados, expande consideravelmente as chances de êxito em investigações criminais (BRASIL, 2009).

Os Bancos de Dados de Perfis Genéticos tornaram-se fundamentais para as investigações criminais. Uma das principais utilizações é armazenar perfis de DNA a partir de amostras biológicas coletadas em locais de crime e fazer comparações de perfis posteriormente. Os Bancos de Dados de DNA denotam grande utilidade para identificar criminosos em série (HARES, 2012).

O funcionamento de um banco de dados é simples. Na realidade existem dois bancos de dados: um de perfis genéticos obtidos de amostras coletadas em locais de crime ou de cadáveres de identidade ignorada, etc e outro banco com os perfis genéticos de referência (LIMA, 2007) (Figura 24 e 25).

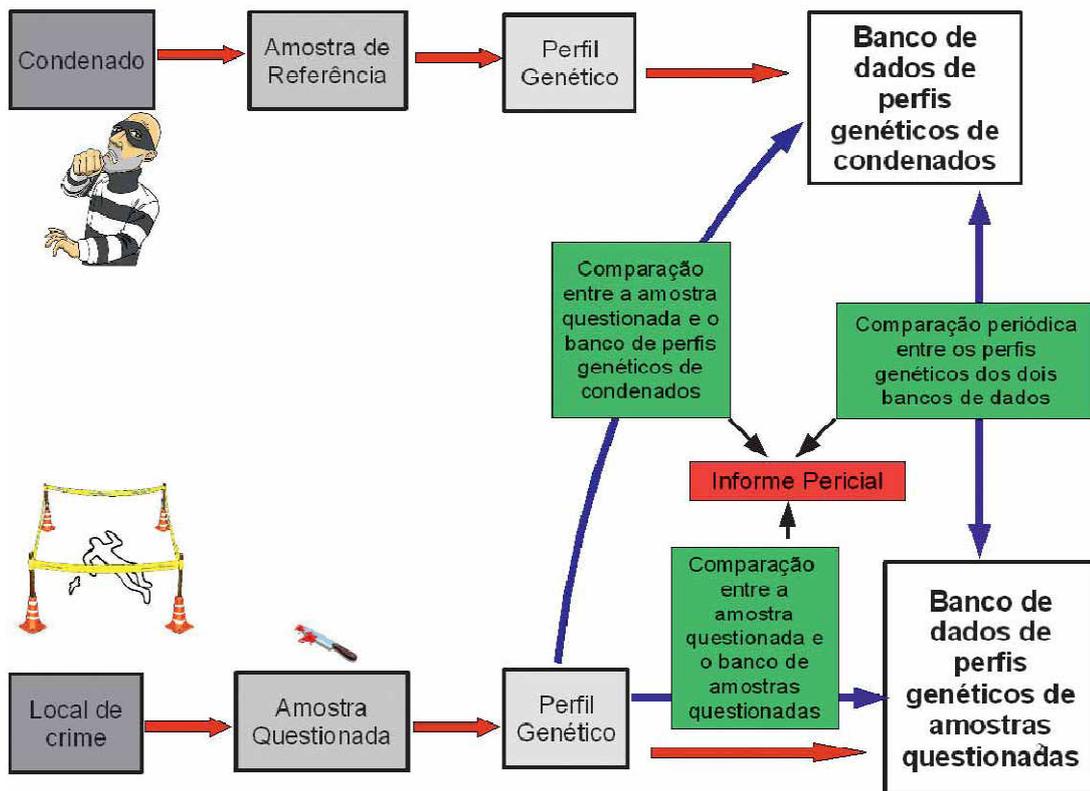


Figura 24: Representação esquemática do funcionamento de um banco de dados de perfis genéticos de criminosos e de vestígios coletados em locais de crime. Fonte: LIMA, 2007.



Figura 25: Representação esquemática do funcionamento de um banco de dados de perfis genéticos de cadáveres de identidade ignorada. Fonte: UFAL (2008).

Com a implantação de um banco de dados de DNA, os perfis genéticos obtidos de vestígios das vítimas e dos locais de crime, podem ser confrontados a qualquer tempo com perfis de criminosos condenados, mesmo que por outro tipo de crime e em qualquer estado da federação, aumentando bastante a possibilidade de identificação do agressor. Também é possível fazer o cruzamento de dados com perfis de vestígios obtidos de outras vítimas e de outros locais de crime, estabelecendo a relação entre diversos crimes cometidos pelo mesmo indivíduo (JACQUES & MINERVINO, 2007).

Um banco de dados de perfis genéticos permite a resolução de crimes para os quais não há nenhum suspeito, e que de outra forma permaneceriam insolúveis. Também temos que considerar o número de crimes que um banco de dados pode evitar, pois cria a oportunidade de deter o criminoso antes que faça novas vítimas. A quantidade de crimes que podem ser evitados e o benefício social advindo da prisão de um criminoso sexual antes que faça novas vítimas são os aspectos mais importantes de um banco de dados de perfis genéticos, e justificam os investimentos em sua criação e manutenção (JACQUES & MINERVINO, 2007).

No Brasil, a Rede Integrada de Bancos de Perfis Genéticos coordena a implantação de bancos de dados estaduais conectados a um banco de dados nacional por meio do programa CODIS. Trata-se de uma rede efetiva em seu propósito de auxílio à Perícia, aos órgãos policiais e à Justiça, por servir como instrumento de prova na investigação criminal, desde que asseguradas as condições técnicas e científicas para a produção da informação e estabelecidos os parâmetros legais que regulamentem a inclusão, acesso, utilização e exclusão das informações armazenadas nesses bancos (BRASIL, 2013b).

O CODIS (*Combined DNA Index System*) é um programa de gerenciamento de perfis genéticos desenvolvido na década de 1990 pelo *Federal Bureau of Investigation* (FBI) estadunidense. O FBI disponibiliza gratuitamente este programa aos governos dos países interessados em utilizá-lo na área de segurança pública. Além dos aproximadamente 180 laboratórios de DNA a utilizarem o CODIS nos EUA, mais de 40 laboratórios também utilizam em outros 27 países (GE et al., 2012). O CODIS permite a organização de perfis genéticos em diferentes categorias, permitindo que os mesmos sejam tratados de forma diferenciada (BRASIL, 2009).

Os perfis podem ser agrupados pelos seguintes índices: “Condenados”, “Forense”, “Detidos”, “Pessoas Desaparecidas”, “Restos Mortais não-Identificados” e “Parentes Biológicos de Pessoas Desaparecidas”. Assim, familiares de pessoas desaparecidas, por exemplo, podem doar voluntariamente seu DNA para que seja confrontado apenas com restos mortais e não com amostras de locais de crime (BRASIL, 2009).

### **2.7.5 Local de Crime e Genética Forense**

Local de Crime é toda área onde tenha ocorrido um fato que necessite de providências da polícia. Este conceito abrange os crimes de qualquer espécie, assim como, todo fato que, não constituindo crime, deva chegar ao conhecimento da polícia, a fim de ser convenientemente esclarecido (SANTOS, 2012).

Um dos conceitos mais difundidos e respeitados sobre local de crime é que se trata de uma porção do espaço compreendido num raio que, tendo por origem o ponto no qual é constatado o fato, se entenda de modo a abranger todos os lugares em que, aparente, necessária ou presumidamente, hajam sido praticados, pelo criminoso, ou criminosos, os atos materiais, preliminares ou posteriores à consumação do delito, e com este diretamente relacionado (UNODOC, 2010).

O tema local de crime está estreitamente relacionado a necessidade do isolamento e preservação, assumindo grande importância para a investigação criminal, pois além de concretizar a materialidade do crime, pode fornecer elementos relevantes para se chegar à autoria do crime (Figura 26) (BOTELHO, 2011).

Isolar um local significa delimitar a área onde ocorreu o fato. Tal procedimento deve ser realizado de forma efetiva para que o menor número de pessoas tenha acesso ao local, evitando-se que evidências sejam modificadas de suas posições e até destruídas antes mesmo de seu reconhecimento. Preservar um local de crime significa não alterar o estado das coisas, garantindo a sua integridade e propiciando a colheita de vestígios que fornecerão os elementos preliminares de uma investigação (MELO, 2009).



Figura 26. Trabalho do Perito Criminal em local de crime. Fonte: RAMOS & FIGUEIREDO, 2012.

O sucesso das análises biológicas de materiais coletados em locais de crime depende visceralmente do isolamento e preservação do mesmo. Amostras biológicas podem ser localizadas em vários suportes, entre eles: ferramentas, impressões digitais, pelos e fibras, manchas de sangue, saliva, esperma, objetos ou suportes tocados por suspeitos, indumentes, armas, etc. (VELHO et al. 2013a).

Toda e qualquer amostra biológica coletada em local de crime poderá ser submetida às técnicas de identificação baseadas nas regiões polimórficas do DNA, processo seguro, com alto poder de discriminação e alta confiabilidade, sendo aceitos como prova legal em casos judiciais (MICHELIN et al., 2013).

### 2.7.6 Coleta de Amostras Biológicas Destinadas a Exames de DNA

Em qualquer que seja a situação deve-se atentar e valorizar ao máximo o momento da coleta de uma amostra biológica destinada a análise de DNA. O uso adequado de EPI (equipamento de proteção individual); bem como a utilização de material descartável para acondicionar e transportar as amostras coletadas são indispensáveis para garantia da integridade da amostra e da cadeia de custódia (Figura 27) (HEDMAN et. al, 2011).

Qualquer tipo de material biológico pode ser usado como fonte de material genético. Contudo, a qualidade do material genético a ser analisado é fortemente influenciada pelo tempo decorrido entre o evento e a coleta da amostra, pelas condições ambientais, pelo manuseio e armazenamento (MICHELIN et al., 2013).

Ao utilizar amostras degradadas e em quantidades limitadas as técnicas utilizadas no laboratório exigem otimização que possibilitem a obtenção máxima do DNA e garantam o sucesso nos resultados laboratoriais, todavia, fatores como a alta temperatura, a contaminação e o processo de putrefação devido a ação de microorganismos podem prejudicar irreversivelmente a qualidade do DNA (VIRKLER & LEDNEY, 2009)



Figura 27: Coleta de Mucosa Bucal. Fonte: LPPGF, 2013.

# ***REFERÊNCIAS***

## REFERÊNCIAS BIBLIOGRÁFICAS

---

1. ALER M, SÁNCHEZ-DIZ P, GOMES I, GISBERT M, CARRACEDO A, AMORIM A, GUSMÃO L (2007) Genetic data of 10 X-STRs in a Spanish population sample. *Forensic Sci Int.* 173:193-6.
2. ARAÚJO LG, BIANCALANA RC, TERADA ASSD, PARANHOS LR, MACHADO CEP, SILVA RHA (2013) Human identification of victims of mass disasters: the importance and role of Forensic Dentistry Human identification of victims of mass disasters: the importance and role of Forensic Dentistry. *RFO* 18:224-229.
3. BARBOSA ACOF (2010) Avaliação de critérios de compatibilidade entre pares de primers para otimização de sistemas multiplex de genotipagem. Dissertação. Universidade Federal de Goiás.
4. BECKER D, RODIG H, AUGUSTIN C, EDELMANN J, GÖTZ F, HERING S, SZIBOR R, BRABETZ W (2008) Population genetic evaluation of eight X-chromosomal short tandem repeat loci using Mentype Argus X-8 PCR amplification kit. *Forensic Sci Int Genet.* 2:69-74.
5. BEIGUELMAN B (2008) Dinâmica dos genes nas famílias e nas populações. Ribeirão Preto. SBG 48-49.
6. BERGLUND EC, KIIALAINEN A, SYVÄNEN A-C (2011) Next-generation sequencing technologies and applications for human genetic history and forensics. *Investigative Genetics* 2:23.
7. BOTELHO, J. A (2013) Necessidade de se Preservar o Local do Crime à Luz da Moderna Investigação e seus Reflexos no CPP. [www.novacriminologia.com.br](http://www.novacriminologia.com.br). Acesso em: 07-01-2013.
8. BRASIL (2009). Projeto Rede Integrada de Bancos de Perfis Genéticos. A Implantação do CODIS.
9. BRASIL (2012). Lei nº 12.654, de 28 de maio de 2012. Casa Civil. Subchefia para assuntos jurídicos. Publicada no DOU de 29.5.2012. <http://www4.planalto.gov.br/legislacao/legislacao-1/leis-ordinarias>. Acesso 03/02/2014.

10. BRASIL (2013a). Decreto 7.950, de 12 de março de 2013. Casa Civil. Subchefia para assuntos jurídicos. Publicado no DOU de 13.3.2013. <http://www4.planalto.gov.br/legislacao/legislacao-1/decretos>. Acesso em 10/02/2014.
11. BRASIL (2013b). Portaria 2774, de 19 de agosto de 2013. Ministério da Justiça. Publicado no DOU de 21.8.2013. <http://www4.planalto.gov.br/legislacao/legislacao>. Accessed 10/02/2014.
12. BRATIC A, LARSSON N-G (2013) The role of mitochondria in aging. *J Clin Invest.* 123:951–957.
13. BROOKFIELD JFY (1995) Statistical issue in DNA evidence. *Electrophoresis.* 16:1665-1669.
14. BUCKLETON JS, TRIGGS CM, WALSH SJ (2004) *Forensic DNA Evidence Interpretation.* CRC Pres.
15. BUDOWLE B, ALLARD MW, WILSON MR, CHAKRABORTY R (2003) Forensic and mitochondrial DNA: applications, debates and foundations. *Genomis Human Genetics* 4:119-141.
16. BUDOWLE B, DAAL AV (2008) Forensically relevant SNP classes. *Biotechniques* 44:603-610.
17. BUDOWLE B (2013) ENCODE and its first impractical application *Investigative Genetics* 2013, 4:4.
18. BUTLER JM (2003) Recent developments in Y-shirt tandem repeat an Y-single nucleotide polymorphism analysis. *Forensic Sci Rev* 15:91-111.
19. BUTLER JM (2005) *Forensic DNA Typing: Biology, Technology, and Genetics of STR markers.* Press EA.
20. BUTLER JM (2006) Genetics and Genomics of Core Short Tandem Repeat Loci Used in Human Identity Testing. *J Forensic Sci* 51:2.
21. BUTLER JM (2007) Short tandem repeat typing technologies used in human identity testing. *BioTechniques* 43.
22. BUTLER JM, HILL CR (2012) Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic Sci. Rev.* 21:15–26.
23. BUTLER J (2011) *Advanced topics in forensic DNA typing: methodology* academic press.
24. CERQUEIRA D, CRUZ DSC (2014) Estupro no Brasil: uma radiografia segundo os dados da Saúde. IPEA Instituto de Pesquisa Econômica Aplicada.

25. CHEN DP, TSENG CP, TSAI SH, WANG MC, LU SC, WU TL, CHANG PY, SUN CF (2009) Use of X-linked short tandem repeats loci to confirm mutations in parentage caseworks. *Clin Chim Acta* 408:29-33.
26. CHEN MY, PU CE (2004) Population data on the X chromosome short tandem repeat loci DXS10011, DXS101, DXS6789, DXS7132, DXS8377, and DXS9895 in Taiwan. *Forensic Sci Int.* 146:65-67.
27. CHOW JC, YEN Z, ZIESCHE SM, BROWN CJ (2005) Silencing of the mammalian X chromosome. *Review Genomics Human Genetics* 6:69–92.
28. CHOW ST, TAN WF, YAP KH, NG TL (1993) The development of DNA profiling database in an HAE III based RFLP system for Chinese, Malays and Indians in Singapore. *J Forensic Sci* 38: 874-884.
29. COBLE MD (2011) The identification of the Romanovs: Can we (finally) put the controversies to rest? *Coble Investigative Genetics* 2011, 2:20.
30. DECORTE R (2010) Genetic identification in the 21<sup>st</sup> century – current status and future developments. *Forensic Sci Int.* 201:160-164.
31. DELFIN FC, MADRID BJ, TAN MP, DE UNGRIA CA (2005) Y-STR analysis for detection and objective confirmation of child sexual abuse. *Int J Legal Med* 119:158-163.
32. DHANARDHONO T, WULANDARI N, BHIMA SKL, AHMADB HJ, WIDODO PT (2013) DNA profiling of disaster victim identification in Trenggalek shipwreck case. *Forensic Science International: Genetics Supplement Series* 4:5-6.
33. EXCOFFIER L (2010) Arlequin ver. 3.5.1.2. <http://cmpq.unibe.ch/software/arlequin3>. Accessed 12/11/2013.
34. FAN H, CHU JY (2007) A brief review of short tandem repeat mutation. *Geno Prot Bioinfo* 5:7-14.
35. FARAH SB (2007) *DNA: Segredos e Mistérios*. Sarvier.
36. FOSTER A, LAURIN N (2012) Development of a fast PCR protocol enabling rapid generation of AmpFtSTR® Identifiler® profiles for genotyping of human DNA. *Investigative Genetics* 3:6.
37. FRACASSO T, SCHÜRENKAMP M, BRINKMANN B, HOHOFF C (2008) An X-STR meiosis study in Kurds and Germans: allele frequencies and mutation rates. *Int J Legal Med.*122:353-356.

38. FRANCEZ PAC, RODRIGUES EMR, FRAZÃO GF, BORGES NDR, SANTOS SEB (2011) Allelic frequencies and statistical data obtained from 12 codis STR loci in an admixed population of the Brazilian Amazon. *Genetics and Molecular Biology* 34:35-39.
39. FRUMKIN D, WASSERSTROM A, DAVIDSON A, GRAFIT A (2009) Authentication of forensic DNA samples. *Forensic Science International: Genetics*.
40. GATTÁS GJF (2006) Polimorfismos de DNA e suas aplicações em ciências forenses da saúde. Tese. Universidade de São Paulo.
41. GE J, EISENBERG A, BUDOWLE B (2012). Developing criteria and data to determine best options for expanding the core CODIS loci. *Invest Genet* 3: 1.
42. GERSHAW CJ, SCHWEIGHARDT AJ, ROURKE LC, WALLACE MM (2011) Forensic utilization of familial searches in DNA databases. *Forensic Science International Genetics* 16–20.
43. GOMES I, PEREIRA V, GOMES V, PRATA MJ, PINTO N, CARRACEDO A, AMORIM A, GUSMÃO L (2009) The Karimojong from Uganda: genetic characterization using an X-STR decaplex system. *Forensic Sci Int Genet.* 3:127-128.
44. GOMES I, PRINZ M, PEREIRA R, MEYERS C, MIKULASOVICH RS, AMORIM A, CARRACEDO A, GUSMÃO L (2007) Genetic analysis of three US population groups using na X-chromosomal STR decaplex. *International Journal of Legal Medicine* 121:198 – 203.
45. GRIFFITHS AJF, WESSLER SR, LEWONTIN RC, CARROLL SB (2009) *Introdução à Genética*. 9. Rio de Janeiro, Guanabara Koogan.
46. GU S, LI S (2006) X-chromosome STRs analysis of Ewenke ethnic population. *Forensic Science International* 158:72–75.
47. GUIMARÃES MA, SOARES-VIEIRA JA, SILVA RHA, Evison MP (2009) A standard procedure for accommodating forensic anthropological and genetic analysis of decomposing human remains from tropical climates. *Forensic Science International: Genetics Supplement Series* 2:165–166

48. GUSMÃO L, SÁNCHEZ-DIZ P, ALVES C, GOMES I, ZARRABEITIA MT, ABOVICH M, ATMETLLA I, BOBILLO C, BRAVO L, BUILES J, CAINÉ L, CALVO R, CARVALHO E, CARVALHO M, CICARELLI R, CATELLI L, CORACH D, ESPINOZA M, GARCÍA O, MALAGHINI M, MARTINS J, PINHEIRO F, JOÃO PORTO M, RAIMONDI E, RIANCHO JA, RODRÍGUEZ A, RODRÍGUEZ A, RODRÍGUEZ CARDOZO B, SCHNEIDER V, SILVA S, TAVARES C, TOSCANINI U, VULLO C, WHITTLE M, YURREBASO I, CARRACEDO A, AMORIM A (2009) A GEP-ISFG collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations. *Int J Legal Med.* 1233:227-34.
49. HANSON EK, BALLANTYNE J (2006) Comprehensive annotated STR physical map of the human Y chromosome: Forensic implications. *Legal Med* 8:110-120.
50. HANSON EK, BALLANTYNE J, (2004) A Highly Discriminating 21 Locus Y-STR “Megaplex” System Designed to Augment the Minimal Haplotype Loci for Forensic Casework. *J Forensic Sci* 49:40-51.
51. HARES DR (2012) Expanding the codis core loci in the United States. *Forensic Science International: Genetics* 6:52-54.
52. HEDMAN J, DALIN E, RASMUSSEN B, ANSELL R (2011) Evaluation of amylase testing as a tool for saliva screening of crime scene trace swabs. *Forensic Science International Genetics* 5:194-199.
53. HOFFMAN NH, FENGER T (2010) Validation of Half-Reaction Amplification Using Promega PowerPlex\_16. *Journal Forensic Sci* 55:1044-1049.
54. HOU YP (2009) Forensic DNA typing in China. *Legal Medicine* 11:S103–S105.
55. INTERPOL. International Criminal Police Organization. Interpol DVI Guide. Lyon: OIPC – Interpol; 2009. Disponível em: <http://www.interpol.int/INTERPOL-expertise/Forensics/DVI-Pages/DVI-guide>. Acesso em 05/02/2014.
56. JACQUES GS, MINERVINO AC (2007). Aspectos éticos e legais dos bancos de dados de perfis genéticos. *Perícia Federal, Brasília*, 26: 17-20. [www.apcf.org.br](http://www.apcf.org.br). Acesso: 04/02/2014.

57. JEFFREYS AJ, BROOKFIELD JFY, SEMEONOFF R (1985a) Positive Identification of an Immigration Test-Case Using Human DNA Fingerprints. *Nature* 317:818 – 819.
58. JEFFREYS AJ, WILSON V, THEIN SL (1985b) Hipervariable “minisatelites” Regions in Human DNA. *Nature* 314:73.
59. JEFFREYS AJ (2013) The man behind the DNA fingerprints: an interview with Professor Sir Alec Jeffreys. *Investigative Genetics* 4:21.
60. JOBLING MA (2013) Curiosity in the genes: the DNA fingerprinting story *Investigative Genetics* 2013, 4:20.
61. JOBIM LF, JOBIM MR, BRENNER C (1999) Identificação humana pelo DNA: investigação de paternidade e análise de casos forense. Porto Alegre. Sagra Luzzato.
62. JOBIM LF, COSTA LRS, SILVA M (2005) Identificação Humana. Campinas: Millennium.
63. JORDE LB, CAREY JC, BAMSHAD MJ, WHITE RL (2004) *Genética Médica*. Rio de Janeiro. Elsevier.
64. JORGE AAL, NISHI MY, FUNARI MFA, SOUZA SC, ARNHOLD IJP, MENDONÇA BB (2008) Short stature caused by SHOX gene haploinsufficiency: from diagnosis to treatment. *Arquivos Brasileiros de Endocrinologia & Metabologia*.
65. KATSANIS SH, WAGNER JK (2013) Characterization of the Standard and Recommended CODIS Markers. *J Forensic Sci* 58.
66. KANG L, LI S (2006) X-chromosome STR polymorphism of Luoba Ethnic Group living in Tibet (SW China). *Forensic Science International* 156: 88–90
67. KAYSER M, KITTLER R, ERLER A, HEDMAN M, LEE AC, MOHYUDDIN A, MEHDI SQ, ROSSER Z, STONEKING M, JOBLING MA, SAJANTILA A, TYLER-SMITH C (2004) A Comprehensive Survey of Human Y-Chromosomal Microsatellites. *Am J Hum Genet* 74:1183-1197.
68. KAYSER M, SAJANTILA A, BUDOWLE B (2013) A tribute to DNA fingerprinting. *Investigative Genetics* 4:19.
69. KOBACHUK LDG (2012) Estudo de Frequências Alélicas de Dez Locus STR s do Cromossomo X na População do Estado do Paraná e sua Contribuição na Identificação Humana. Dissertação de Mestrado. Universidade Federal do Paraná.

70. KOBILINSKY L, LIOTTI TF, OESER-SWEAT J (2005) DNA: Forensic and Legal Applications. John Wiley & Sons, Inc. Publication, New Jersey.
71. KOHLRAUSCH F, CALLEGARI-JACQUES SM, TSUNETO LT, PETZLERLER ML, HILL K, HURTADO M, SALZANO FM, HUTZ MR (2005) Geography influences microsatellite polymorphism diversity in Amerindians. *Am J Phys Anthropol* 126:468-470.
72. KRENKE BE, VICULIS L, RICHARD ML, PRINZ M, MILNE SC, LADD C, GROSS AM, GORNALL T, FRAPPIER JR, EISENBERG AJ, BARNA C, ARANDA XG, ADAMOWICZ MS, BUDOWLE B (2005) Validation of male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex. *Forensic Sci Int*. 151:111-124.
73. LEITE FP, SANTOS SE, RODRÍGUEZ EM, CALLEGARI-JACQUES SM, DEMARCHI DA, TSUNETO LT, PETZLERLER ML, SALZANO FM, HUTZ MH (2009) Linkage disequilibrium patterns and genetic structure of Amerindian and non-Amerindian Brazilian populations revealed by long-range X-STR markers. *Am J Phys Anthropol*. 139:404-412.
74. LEITE FPN (2006) Análise da estrutura genética da população do Rio Grande do Sul através de microssatélites autossômicos e de cromossomos sexuais. Tese de Doutorado em Genética e Biologia Molecular. Instituto de Biociências. Universidade Federal do Rio Grande do Sul.
75. LEWIN B (2009) Genes IX. Porto Alegre: Artmed.
76. LI H, TANG H, ZHANG Q, BAI J, SANG S (2009) A multiplex PCR for 4 X chromosome STR markers and population data from Beijing Han ethnic group. *Legal Medicine* 11:248-250.
77. LI S, LI R, LI H, LU J, LI Y, BOLUND L, SCHIERUP MH, WANG J (2013) SOAPindel: Efficient identification of indels from short paired reads. *Genome Res*. 23: 195-200.
78. LIM EJ, LEE HY, SIM JE, YANG WI, SHIN KJ (2003) Genetic polymorphism and haplotype analysis of 4 tightly linked X-STR duos in Koreans. *Croat Med J*. 50:305-312.
79. LIMA HB (2007). DNA e Criminalidade. *Perícia Federal*, Brasília, 26: 17-20. [www.apcf.org.br](http://www.apcf.org.br). Acesso: 04/02/2014.
80. LIMA LO (2006) Direito médico – Utilização de polimorfismo em Análises forenses.

81. LIU Q-L, DE-JIAN LV, XIANG-LIN WU, HONG-YU SUN, XIN-YAO WU, HUI-LING LU (2008) Development of a five ChX STRs loci typing system. *Int J Legal Med* 122:261-265.
82. LJUBKOVIC J, STIPISIC A, SUTIOVIC D, DEFINIS-GOJANOVIC M, BUCAN K, ANDELINOVIC S (2008) Y-Chromosomal Short Tandem Repeat Haplotypes in Southern Croatian Male Population Defined by 17 loci. *Croatian Medical Journal* 49:201-6.
83. MALAVER PC, YUNIS JJ (2003) Different Dental Tissues as Source of DNA for Human Identification in Forensic Cases. *Croatian Medical Journal. Forensic Sciences* 44:306-309.
84. MARTINS JA (2008) Estudo de frequências alélicas de STRs do cromossomo X na população brasileira de Araraquara, SP. 122f. Dissertação (Mestrado) – Universidade Estadual Paulista Julio de Mesquita Filho, Araraquara.
85. MARTINS JA, COSTA JC, PANETO GG, FIGUEIREDO RF, GUSMÃO L, (2010) Genetic profile characterization of 10 X-STRs in four populations of the southeastern region of Brazil. *Int J Leg Med* 124:427–432.
86. MARTINS JA, COSTA JC, PANETO GG, GUSMÃO L, SÁNCHEZ-DIZ P, CARRACEDO A, CICARELLI RMB (2009) Genetic data of 10 X-chromosomal loci in Vitória population (Espírito Santo State, Brazil). *Forensic Science International: Genetics* 2:394–395.
87. McMANUS C, PAIVA S, CORRÊA PS, SEIXAS L, MELO CB (2011) Estatísticas para descrever genética de populações. INCT: Informação genético-sanitária da pecuária brasileira.
88. MELO GJ (2009) Atuação dos policiais militares do Pará em isolamento e preservação de local de crime. Centro Universitário do Pará.
89. MICHELIN K, FREITAS JM, KORTMANN GL (2013) Locais de Crime dos Vestígios a Dinâmica Criminosa. Campinas. Milenium.
90. MICHELIN K, PACHECO AC, AULER-BITTENCOURT EA, LIMA MJM. (2008). Banco de Perfis Genéticos no Combate aos Crimes Sexuais. *Perícia Federal, Brasília*. 26: 13-16. [www.apcf.org.br](http://www.apcf.org.br). Acesso 02/02/2014.
91. MORETI T (2009) Identificação humana: Uma proposta metodológica para obtenção de DNA de ossos e implementação de banco de dados de frequências alélicas de STRs autossômicos na população de Santa Catarina. 145f. Dissertação (Mestrado)- Universidade Federal de Santa Catarina.

92. NADEEM A, BABAR ME, HUSSAIN M, TAHIR MA (2008) Development of pentaplex PCR and genetic analysis of X chromosomal STRs in Punjabi population of Pakistan. *Mol Biol Rep.* 367:1671-1675.
93. NAKAMURA Y (2009) DNA variations in human and medical genetics: 25 years of my experience. *Journal of Human Genetics* 54:1-8.
94. NIST (2003) Recent developments in y-short tandem repeat and y-single nucleotide polymorphism analysis. *Forensic Sci Rev.* 15:91-111.
95. NORÉN L, HEDELL R, ANSELL R, HEDMAN J (2013) Purification of crime scene DNA extracts using centrifugal filter devices. *Investigative Genetics* 4:8.
96. O'CONNOR KL, TILLMAR AO (2012) Effect of linkage between vWA and D12S391 in kinship analysis. *Forensic Sci Int Genet.* 6:840-844.
97. OLIVEIRA TMM (2012) Análise da frequência alélica de 15 loci STR na população do Rio Grande do Norte. Dissertação. Universidade do Rio Grande do Norte.
98. PERNAMBUCO (2012) Secretaria de Defesa Social de Pernambuco. Portaria GAB/SDS Nº 1440/2012.
99. POIARES LA, OSORIO OS, SPANHOL FA, COLTRE SC, RODENBUSCH R, GUSMÃO L, A LARGURA, SANDRINI F, SILVA CMD (2009). Allele frequencies of 15 STRs in a representative sample of the Brazilian population . *Forensic Science International: Genetics* 473.
100. RAMOS D, FIGUEIREDO T (2012) Local de Crime. Perícia Federal, Brasília. 29: 26-30. [www.apcf.org.br](http://www.apcf.org.br). Acesso 02/02/2014.
101. RIBEIRO EM, LEITE FPN, HUTZ MH, PALHA TJBF, SANTOS AKCR, SANTOS SEB (2008) A multiplex PCR for 11 X chromosome STR markers and population data from a Brazilian Amazon Region. *Forensic Science International: Genetics* 2:154-158.
102. ROBINO C, GIOLITTI A, GINO S, TORRE C (2006) Development of two multiplex PCR systems for the analysis of 12 X-chromosomal STR loci in a northwestern Italian population sample. *Int. J. Legal Med.* 120:315-318.
103. RODIG H, GRUM M, HANS-DIETER GRIMMECKE HD (2007) Population study and evaluation of 20 Y-chromosome. STR loci in Germans. *Int J Legal Med* 121:24–27.
104. ROEWER L (2009) Y Chromosome STR typing in crime casework. *Forensic Sci Med Pathol* 5:77-84.

105. ROEWER L (2013) DNA fingerprinting in forensics: past, present, future. *Investigative Genetics* 4:22.
106. ROSS MT, GRAFHAN DV, COFFEY AJ, SCHERER S, MCLAY K, MUZNY D (2005) The DNA sequence of the human X chromosome. *Nature* 434:325-327.
107. SANTOS SM (2012) Isolamento e preservação do local do crime. Governo de Pernambuco. Secretaria de Defesa Social. Gerência Geral de Polícia Científica.
108. SCORALICK RA, BARBIERI AA, MORAES ZM, FRANCESQUINI JÚNIOR L, DARUGE JÚNIOR E; NARESSI SCM (2013) Human identification through dental radiographs study: case report. *Rev. odontol. UNESP* 42:1.
109. SCHUMM JW, GUTIERREZ-MATEO C, TAN E, SELDEN R (2013) A 27-Locus STR Assay to Meet All United States and European Law Enforcement Agency Standards. *J Forensic Sci* 58:6
110. SENASP (2004) Programa de Ciência e Tecnologia Aplicada na Segurança Pública – Secretaria Nacional de Segurança Pública. Brasil.
111. SILVA DA (2004). Metodologia para análise forense por DNA: estudo de validação para 14 loci STR e construção de banco de dados. Tese. Doutorado em Biologia. Universidade do Rio de Janeiro.
112. SILVA EFA, JACQUES GS, CHEMALE G, FRANCEZ PA (2013a) *Genética Forense*. Millenium.
113. SILVA EC, SOUZA AS, BARRETO ES, AARESTRUP JR (2013b) Strategies for human identification: From the general to the genome *REBES* 3:46-52.
114. SILVA LAF, PASSOS NS (2006) *DNA Forense: Coleta de Amostra Biológica em Locais de Crime para Estudo do DNA*. Maceió: UFAL.
115. STRAY JE, LIU JY, BREVNOV MG, SHEWALE JG (2010) Extraction of DNA from forensic biological samples for genotyping. *Forensic Sci Rev.* 22:159-175.
116. SZIBOR R (2007) X-Chromosomal markers: Past; present and future. *Forensic Science International Genetics* 1:93-99.
117. SZIBOR R, HERING S, EDELMANN J (2006) A new Web site compiling forensic chromosome X research is now online. *International Journal of Legal Medicine* 120:252-254.

118. TAN E, TURINGAN RS, HOGAN C, VASANTGADKAR S, PALOMBO L, SCHUMM JW, SELDEN RF (2013) Fully integrated, fully automated generation of short tandem repeat profiles *Investigative Genetics* 2013, 4:16.
119. TARIQ MA, ULLAH O, RIAZUDDIN SA, RIAZUDDIN S (2008) Allele frequency distribution of 13 X-chromosomal STR loci in Pakistani population. *Int J Legal Med.* 122:525-528.
120. TAVARES CC, GUSMÃO L, DOMINGUES C, DOMINGUES P, SILVA DA, AQUINO J, PEIXE C, AMORIM A, CARVALHO EF (2008) Population data for six X-chromosome STR loci in a Rio de Janeiro (Brazil) sample: Usefulness in forensic casework. *Forensic Science International: Genetics* 1:164–166.
121. TILLMAR A (2010) *Populations and Statistics in Forensic Genetics*. Linköping University, Faculty of Health Sciences.
122. TORNAVOI DC, SILVA RHA (2010) Rugoscopia palatina e a aplicabilidade na identificação humana em odontologia legal: revisão de literatura. *Saúde, Ética & Justiça* 15:1.
123. TURI E, KING, JOBLING M (2009) What's in a name? Y chromosomes, surnames and the genetic genealogy revolution. *Trends Genetic* 25: 351-360.
124. TURRINA S, ATZEI R, D (2007) Polymorphism of four X-chromosomal STRs: DXS7423, DXS7424, DXS8378 and DXS6809 in a North Italian population sample and their use in kinship testing. *Forensic Science International* 168:241-243.
125. UFAL - UNIVERSIDADE FEDERAL DE ALAGOAS. LABORATÓRIO DE DNA FORENSE. Disponível em <http://www.labdnaforense.org>. Acesso em 04/01/2014.
126. UNODOC (2010) Conscientização sobre o local de crime e as evidências materiais em especial para pessoal não forense. Escritório das Nações Unidas.
127. VAN OORSCHOT RAH, BALLANTYNE KN, MITCHELL RJ (2010) Forensic trace DNA: a review *Investigative Genetics* 1:14.
128. VELHO JA, COSTA KA, DAMASCENO CTM (2013a) *Locais de Crime dos Vestígios à Dinâmica Criminosa*. Campina. Millenium.

129. VELHO JA, GEISER GC, ESPÍNDULA A (2013b) *Ciências Forenses. Uma Introdução às Principais Áreas da Criminalística Moderna*. Campina. Millenium.
130. VIRKLER K, LEDNEV IK (2009) Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene. *Forensic Science International* 188:1-17.
131. VUICHARD S, BORER U, BOTTINELLI M, COSSU C, MALIK N, MEIER V, GEHRIG C, SULZER A, MOREROD M-L, CASTELLA V (2011) Differential DNA extraction of challenging simulated sexual-assault samples: a Swiss collaborative study. *Investigative Genetics* 2:11.
132. ZALÁN A, VÖLGYI A, BRABETZ W, SCHLEINITZ D, PAMJAV H (2008) Hungarian population data of eight X-linked markers in four linkage groups. *Forensic Science International* 175: 73–78.
133. ZARRABEITIA MT, MIJARES V, RIANCHO JÁ (2007) Forensic efficiency of microsatellites and single nucleotide polymorphisms on the X chromosome. *International Journal of Legal Medicine* 121:433-437.
134. ZARRABEITIA MT, PINHEIRO F, PANCORBO MM, CAINÉ L, CARDOSO S, GUSMÃO L, RIANCHO JÁ (2009) Analysis of 10 X-linked tetranucleotide markers in mixed and isolated populations. *Forensic Science International: Genetics* 3:63-66.
135. ZIĘTKIEWICZ E, WITT M, DACA P, ŻEBRACKA-GALA J, GONIEWICZ M, JARZĄB B, WITT M (2012). Current genetic methodologies in the identification of disaster victims and in forensic analysis. *J Appl Genetics* 53:41–60.
136. WANG T, PRADHAN K, YE K, WONG LJ, ROHAN TE (2011) Estimating allele frequency from next-generation sequencing of pooled mitochondrial DNA samples. *Front Genet.* 2:51.

# ***OBJETIVOS***

---

## 4 OBJETIVOS

---

### 4.1 Geral

Estudar e descrever a frequência alélica de marcadores de interesse forense no Estado de Pernambuco.

### 4.2 Específicos

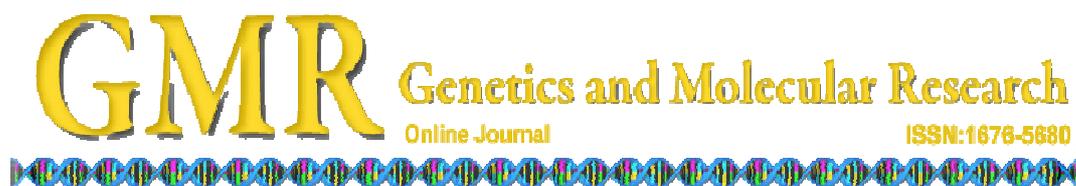
1. Determinar a frequência alélica de 21 STRs autossômicos (*D19S433*, *D2S1338*, *D21S11*, *TH01*, *D3S1358*, *FGA*, *TPOX*, *D8S1179*, *D16S539*, *D7S820*, *D13S317*, *D5S818*, *vWA*, *D18S51*, *CSF1PO*, *D2S441*, *D22S1045*, *SE33*, *D10S1248*, *D1S1656* e *D12S391*) em indivíduos da população brasileira, nascidos e residentes no Estado de Pernambuco;
2. Verificar, na população de Fernando de Noronha, Pernambuco, a frequência alélica de 12 marcadores X-STRs (*DXS7132*, *DXS7423*, *DXS8378*, *DXS10074*, *DXS10079*, *DXS10101*, *DXS10103*, *DXS10134*, *DXS10135*, *DXS10146*, *DXS10148* and *HPRTB*);
3. Comparar os dados obtidos com os de outras populações do mundo quanto ao grau de variabilidade genética;
4. Fornecer dados para a conclusão de casos forenses realizados no Laboratório de Perícia e Pesquisa em Genética Forense de Pernambuco, Brasil;
5. Fornecer dados que permitam a ampliação de Bancos de Dados de frequências alélicas da população brasileira.

# ***ARTIGOS***

**ARTIGO I**

**Distribution of forensic marker allelic frequencies in Pernambuco, Northeastern Brazil.**

**ACEITO**



**Qualis: B2**

## **Distribution of forensic marker allelic frequencies in Pernambuco, Northeastern Brazil**

Forensic markers in Pernambuco

S.M. Santos<sup>1,2\*</sup>, C.A. Souza<sup>2</sup>, K.C.N Rabelo<sup>1,2</sup>, P.R.E. Souza<sup>3</sup>, R.R. Moura<sup>1,4</sup>, T.C. Oliveira<sup>1,2</sup> and S. Crovella<sup>1,4</sup>

<sup>1</sup>Laboratório de Imunopatologia Keizo Asami (LIKA) da Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil

<sup>2</sup>Laboratório de Perícia e Pesquisa em Genética Forense (LPPGF) da Gerência Geral de Polícia Científica da Secretaria de Defesa Social de Pernambuco (SDS), Recife (PE), Brasil

<sup>3</sup>Laboratório Genoma da Universidade Federal Rural de Pernambuco (UFRPE), Recife (PE), Brasil

<sup>4</sup>Departamento de Genética, Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil

\*Corresponding author: SANTOS Sandra M,

Tel.: 55 81 3183 5389

Fax: 55 81 31835388

E-mail: sandra.santos@sds.pe.gov.br

**ABSTRACT.** Pernambuco is one of the 27 federal units of Brazil, ranking seventh in the number of inhabitants. We examined the allele frequencies of 13 short tandem repeat loci (*CFS1PO*, *D3S1358*, *D5S818*, *D7S820*, *D8S1179*, *D13S317*, *D16S539*, *D18S51*, *D21S11*, *FGA*, *TH01*, *vWA*, and *TPOX*), the minimum recommended by the Federal Bureau of Investigation and commonly used in forensic genetics laboratories in Brazil, in a sample of 609 unrelated individuals from all geographic regions of Pernambuco. The allele frequencies ranged from 5 to 47.2%. No significant differences for any loci analyzed were observed compared with other publications in other various regions of Brazil. Most of the markers observed were in Hardy-Weinberg equilibrium. The occurrence of the allele 47.2 (locus *FGA*) and alleles 35.1 and 39 (locus *D21S11*), also described in a single study of the Brazilian population, was observed. The other forensic parameters analyzed (matching probability, power of discrimination, polymorphic information content, paternity exclusion, complement factor I, observed heterozygosity, expected heterozygosity) indicated that the studied markers are very informative for human forensic identification purposes in the Pernambuco population.

**Key words:** Autosomal STRs; Human identification; Population data

## **INTRODUCTION**

Studies estimating the genetic admixture in established populations indicate that the most striking migratory movement in recent centuries has been the colonization of the Americas (Bonilla et al, 2004; Bedoya et al, 2006; Martinez et al., 2007; Tang et al., 2007). One method of increasing the understanding of the genetic makeup of a population is to estimate the contribution of parental populations based on the distribution of allele frequencies of genetic markers (Salzano, 2004).

The most frequently used genetic markers used in forensic laboratories and paternity tests include short tandem repeats (STRs), the Y chromosome, and mitochondrial DNA (Martins et al., 2010; Ziętkiewicz et al., 2012). Those with high discriminatory power, higher frequency of heterozygotes, low mutation rate, high degree of informativeness, smaller size, and lower training stutters, are considered to be the most important for human identification (Hares, 2012).

The distribution of allele frequencies of DNA markers differs between biogeographically separated populations and those from a different origin (Gusmão et al., 2001). The Brazilian populations, including Pernambuco, are considered to be the most heterogeneous in the world (Aguiar et al., 2012). Several studies have been conducted to examine the real contribution of Native Americans, Europeans, and Africans, but the only consensus is that the dynamics of crossovers in Brazil are both unique and complex (Pena et al., 2009).

Material evidence based on DNA analysis is valuable in the process of criminal investigation and identification (Gershaw et al., 2011). Because of the importance and applicability of STR typing (Choi and Seo, 2009) targeting human identification, in November 1997, the Federal Bureau of Investigation (FBI) established 13 loci that would be used as the Combined DNA Index System (CODIS), which were tested and validated for forensic applications (Budowle et al., 1998; Butler, 2005; Michelin et al., 2008).

These markers have been widely used worldwide; even after more than a decade, the markers continue to play an important role in the field of forensics, particularly in the construction of criminal databases (FBI, 1997; Krzyżanińska et al., 2006; Li et al., 2008). When a biological sample is derived from a single source, these 13 markers satisfactorily meet the needs of any forensic laboratory (Ge et al, 2012).

A database of genetic profiles, for forensic purposes or for locating missing persons, has been adopted in many countries, including Brazil; this justified the importance of STRs in the forensic community (Corte-Real, 2004; Walsh, 2004). DNA databases have become central to criminal investigations and for understanding local allele frequencies (Gill et al., 2006; Minervino and Jacques, 2007; Martins et al., 2010).

In 2012 and 2013, Brazil took an important step in combating crime by publishing Law No. 12.654/12, which establishes DNA as a form of criminal identification (Brazil, 2012), and Decree No. 7.950/2013, which establishes the National Bank of Genetic Profiles and Integrated Bank Network of Genetic Profiles (Brazil, 2013). In this study, we examined the allele frequencies of 13 STR loci. The results will serve to support the actions of forensic genetics in Northeastern Brazil.

## **MATERIAL and METHODS**

### **DNA samples**

Because of the forensic application of our research, sampling was determined in proportion to the number of inhabitants residing in each of the 8 areas of Pernambuco, with territorial division proposed by the Pact for Life Government Program aimed at preventing and reducing violence. There are 8 territories and 26 integrated security areas, linking overt police actions with police actions in the state (Pernambuco, 2007). Samples were collected mainly in the host cities of each region, resulting in very diverse samples. Buccal mucosa samples derived from 609 adult subjects, unrelated, all born and residing in the State of Pernambuco, were analyzed. The samples were collected using a mouth swab after obtaining informed consent from each subject. The sample was composed of individuals living throughout the territory, respecting the proportionality of the number of inhabitants, according to the census of the Brazilian Institute of Geography and Statistics (IBGE, 2010): territory 1 - capital (N = 115); territory 2 - metropolitan region (N = 150); territory 3 - south forest (N = 37); territory 4 - north forest (N = 60); territory 5 - wild 1 (N = 65); territory 6 - wild 2 (N = 56); territory 7 - backcountry 1 (N = 61); territory 8 - backcountry 2 (N = 65).

This study was approved by the Ethics Committee on Human Research of the Centro de Ciências da Saúde, Universidade Federal de Pernambuco, CEP/CCS/UFPE. The Certificate Presentation to Ethical Consideration (CAAE) was 4961813.2.0000.5208. The territorial division according to the model adopted by the Government of the State of Pernambuco: "Pact for Life" program is displayed in Figure 1.



**Figure A:** Map of the State of Pernambuco, divided into 08 territories. Adpted from: <http://www.sds.pe.gov.br/>

### DNA extraction

Biological samples were extracted using the DNA IQ™ System kit (Promega; Madison, WI, USA) as recommended by the manufacturer and adopted at the “Laboratório de Perícia e Pesquisa em Genética Forense” (Recife).

### STR genotyping

To amplify the STR 13 regions (loci) (*D21S11*, *TH01*, *D3S1358*, *FGA*, *TPOX*, *D8S1179*, *D16S539*, *D7S820*, *D13S317*, *D5S818*, *VWA*, *D18S51*, and *CSF1PO*), as well as amelogenin to identify gender, we used the AMPF!STR Identifiler® Plus® PCR Kit (Life Technologies; Carlsbad, CA, USA) and PowerPlex®16HS (Promega) commercial kits following the manufacturer protocols. Amplicon separation and detection were performed using the ABI3500 automated sequencer (Applied Biosystems; Foster City, CA, USA) with the GS-600 LIZ standard and POP 4 polymer (Applied Biosystems) size. Genotyping was performed using GeneMapper IDX software (Life Technologies).

### Statistical analysis

Allele frequencies, Hardy-Weinberg equilibrium, expected heterozygosity ( $H_e$ ), and observed heterozygosity ( $H_o$ ) were calculated using Arlequin software version 3.5.1.2 (Excoffier, 2010). Other statistical parameters of forensic and paternity importance included matching probability (MP), power of discrimination (PD), polymorphism information content (PIC), power of exclusion (PE), and typical and paternity index (TPI), and were obtained using PowerStats, version 12 (Tereba, 1999).

### RESULTS and DISCUSSION

Deviations from Hardy-Weinberg equilibrium were observed for the markers *FGA*, *D5S818*, *vWA*, and *D18S51*, but these deviations disappeared after Bonferroni correction (level of significance: 0.0038). The presence of rare alleles or the characteristic substructure of Brazilian populations may account for this finding (Soloum et al., 2013). For all markers,  $H_e$  values were very similar to  $H_o$  values.

Of the 13 loci studied, *D18S51* showed a PIC value of 0.87, suggesting high informational power. The *TPOX* locus showed a peak of 0.65, indicating that this marker was the least informative. The *D18S51* and *FGA* loci showed higher discrimination powers of 0.973 and 0.976, respectively. Our results for the Brazilian population (Grattapaglia et al, 2001; Fridman et al, 2007; Aguiar, et al., 2012) agreed with those for the populations of Pará – North Region (Ribeiro-Rodrigues et al., 2008), the Mato Grosso do Sul – Midwest Region (Silva et al., 2004), Minas Gerais - Southeast Region (del Castillo et al, 2009), and Rio Grande South - South Region (Leite et al., 2003) for most markers and indicators assessed.

In a previous study examining STR allele frequencies in the population of Pernambuco, the *D5S818* and *D3S1338* loci were not evaluated (Dellalibera et al., 2004). The present study confirmed that both are important in forensic laboratories, with higher  $H_e$ , high informative power based on the PIC values, and high PD observed (Table 1).

In this study, the occurrence of rare alleles, described in a single study for the Brazilian population (Aguiar et al. 2012), was also observed. These included allele 47.2 (locus *FGA*) and alleles 35.1 and 39 (locus *D21S11*). Allele frequencies and other statistical parameters of forensic interest for the 13 loci suggested by CODIS are shown in Tables 1 and 2, respectively.

Our study examining the allele frequencies of the 13 STR loci recommended by CODIS was more robust than previous reports for the state of Pernambuco. The results presented here will be applied for routine statistical calculations of the Laboratório de Perícia e Pesquisa em Genética Forense and will serve as basis for criminal studies performed in Pernambuco.

**Table 1.** Allele frequency distribution of 13 STR loci observed in the present study.

Allele	D21S11	TH01	D3S1358	FGA	TPOX	D8S1179	D16S539	D7S820	D13S317	D5S818	vWA	D18S51	CSF1PO
5		0.000821											
6		0.206897			0.032841				0.000821				
7		0.241379			0.006568			0.007389	0.000821	0.018062		0.000000	0.013957
8		0.166667			0.477833	0.006568	0.014778	0.171593	0.097701	0.022989	0.001642		0.021346
9		0.17734			0.106732	0.006568	0.176515	0.128079	0.061576	0.046798		0.000000	0.017241
9.3		0.202791											
10		0.003284			0.073071	0.064039	0.091133	0.273399	0.055829	0.055829		0.000011	0.295567
10.2												0.000000	
11		0.000821			0.252874	0.070608	0.302956	0.239737	0.310345	0.311987	0.005747	0.000218	0.280788
12	0.000821		0.007389		0.044335	0.119869	0.2578	0.147783	0.325944	0.353859		0.998305	0.300493
13	0.000821		0.004105		0.004926	0.272578	0.128079	0.027094	0.10509	0.178161	0.004105		0.05665
13.2												0.000011	
14			0.092775		0.000821	0.258621	0.028736	0.004926	0.04023	0.00821	0.082102	0.021598	0.00027
15			0.29064	0.000821		0.157636			0.000001	0.000821	0.132184	0.030295	0.002463
16			0.29803	0.000821		0.039409				0.000821	0.285714	0.021118	
17			0.172414	0.003284		0.003284				0.000821	0.261905	0.016827	
17.2										0.000821			
18			0.119869	0.013957							0.155993	0.006212	
18.2				0.002463									
19	0.000821		0.013957	0.051724		0.000821					0.050082	0.002427	
20			0.000821	0.114122							0.016420	0.001189	
21				0.140394							0.003284	0.000357	
21.2				0.002463									
22				0.162562								0.000006	
22.2				0.004926									
23				0.145320								0.000006	



39 0.000821

cont.

Allele	<i>D21S11</i>	<i>TH01</i>	<i>D3S1358</i>	<i>FGA</i>	<i>TPOX</i>	<i>D8S1179</i>	<i>D16S539</i>	<i>D7S820</i>	<i>D13S317</i>	<i>D5S818</i>	<i>vWA</i>	<i>D18S51</i>	<i>CSF1PO</i>
43.2				0.000821									
47.2				0.000821									

**Table 2.** Forensic parameters in Pernambuco, Brazil.

STRs	MP	PD	PIC	PE	TPI	$H_o$	$H_e$	P(HWE)
<i>D21S11</i>	0.073	0.927	0.78	0.614	2.6	0.85878	0.85878	0.70609
<i>TH01</i>	0.112	0.888	0.7	0.497	1.94	0.79146	0.79922	0.57657
<i>D3S1358</i>	0.091	0.909	0.74	0.617	2.62	0.79146	0.77437	0.19426
<i>FGA</i>	0.033	0.967	0.86	0.719	3.63	0.867	0.87552	0.00567
<i>TPOX</i>	0.138	0.862	0.65	0.376	1.49	0.66502	0.68845	0.07229
<i>D8S1179</i>	0.061	0.939	0.78	0.595	2.47	0.79967	0.80953	0.42405
<i>D16S539</i>	0.08	0.92	0.75	0.54	2.15	0.76683	0.78549	0.09766
<i>D7S820</i>	0.074	0.926	0.77	0.614	2.61	0.81117	0.79993	0.16464
<i>D13S317</i>	0.087	0.913	0.74	0.52	2.05	0.75698	0.76896	0.15552
<i>D5S818</i>	0.114	0.886	0.7	0.521	2.06	0.76026	0.74008	0.03337
<i>vWA</i>	0.071	0.929	0.77	0.546	2.18	0.77011	0.79904	0.00002
<i>D18S51</i>	0.027	0.973	0.87	0.699	3.38	0.85386	0.883	0.00001
<i>CSF1PO</i>	0.116	0.884	0.69	0.503	1.97	0.74056	0.73981	0.14177

MP: matching probability; PD: power of discrimination; PIC: polymorphism information content; PE: power of exclusion; TPI: typical paternity index;  $H_e$ : expected heterozygosity;  $H_o$ : observed heterozygosity; P (HWE): uncorrected P values from Hardy–Weinberg equilibrium test.

---

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the Secretaria de Defesa Social de Pernambuco (SDS/PE/ Brasil). R. Moura is recipient of a post-graduate scholarship from Fundação de Amparo à Pesquisa do Estado de Pernambuco (FACEPE, IBPG 0926-2.02/11).

## REFERENCES

- Aguiar VRC, Wolfgramm Ede V, Malta FS, Bosque AG, et al. (2012). Updated Brazilian STR allele frequency data using over 100,000 individuals: an analysis of CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, Penta D, Penta E, TH01, TPOX and vWA loci. *Forensic Sci. Int. Genet.* 6: 504-509.
- Bedoya G, Montoya P, García J, Sato I, et al. (2006). Admixture dynamics in Hispanics: A shift in the nuclear genetic ancestry of South American Population isolate. *Proc. Natl. Acad. Sci. USA* 103: 7234-7239.
- Bonilla C, Parra EJ, Plaff CL, Dias S, et al. (2004). Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Ann. Hum. Genet.* 68: 139-153.
- Butler JM (2005). *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*. Elsevier Academic Press, New York.
- Budowle B, Moretti TR, Niezgoda SJ and Brown BL (1998). CODIS and PCR-based short tandem repeat loci: law enforcement tools. In: *Proceedings of the second European symposium on human identification*, Promega Corporation, Madison, WI, pp. 73-88.
- Brazil (2012). Lei nº 12.654, de 28 de maio de 2012. Casa Civil. Subchefia para assuntos jurídicos. Publicada no DOU de 29.5.2012. <http://www4.planalto.gov.br/legislacao/legislacao-1/leis-ordinarias>. Accessed 03/02/2014.

- Brazil (2013). Decreto 7.950, de 12 de março de 2013. Casa Civil. Subchefia para assuntos jurídicos. Publicado no DOU de 13.3.2013. <http://www4.planalto.gov.br/legislacao/legislacao-1/decretos>. Accessed 10/02/2014.
- Choi JY and Seo TS (2009). An integrated microdevice for high-performance short tandem repeat genotyping. *Biotechnol. J.* 4: 1530-1541.
- Corte-Real F (2004). Forensic DNA databases. *Forensic Sci. Int.* 146 Suppl: S143-S144.
- del Castillo DM, Perone C, de Queiroz AR, Mourão PH, et al. (2009). Populational genetic data of 15 STR markers in the Brazilian population of Minas Gerais. *Legal Med.* 11: 45-47.
- Dellalibera E, Havro MLB, Souza M, Kajihara K, et al. (2004). Genetic analysis of 13 STR loci in the population from the State of Pernambuco, northeast Brazil. *ForensicSci. Int.* 146: 57-59.
- Excoffier L (2010) Arlequin ver. 3.5.1.2. <http://cmpq.unibe.ch/software/arlequin3>. Accessed 12/11/2013.
- FBI/Codis (1997). <http://www.fbi.gov/about-us/lab/codis/ndis-statistics>. Accessed 16/01/2014.
- Fridman C, dos Santos PC, Kohler P, Garcia CF, et al. (2007). Brazilian populations profile of 15 STR markers. *Forensic Sci Int. Genet.* 2: 1-4.
- Ge J, Eisenberg A and Budowle B (2012). Developing criteria and data to determine best options for expanding the core CODIS loci. *Invest. Genet.* 3: 1.
- Gershaw CJ, Schweighardt AJ, Rourke LC and Wallace MM (2011). Forensic utilization of familial searches in DNA databases. *Forensic Sci. Int. Genet.* 5: 16-20.
- Gill P, Fereday L, Morling N and Schneider PM (2006). The evolution of DNA databases: recommendations for new European STR loci. *Forensic Sci. Int.* 156: 242-244.
- Grattapaglia D, Schmidt AB, Costa e Silva C, Stringher C, et al. (2001). Brazilian population database for the 13 STR loci of the AmpFISTR® Profile Plus and Cofiler multiplex kits. *Forensic Sci. Int.* 118: 91-94.

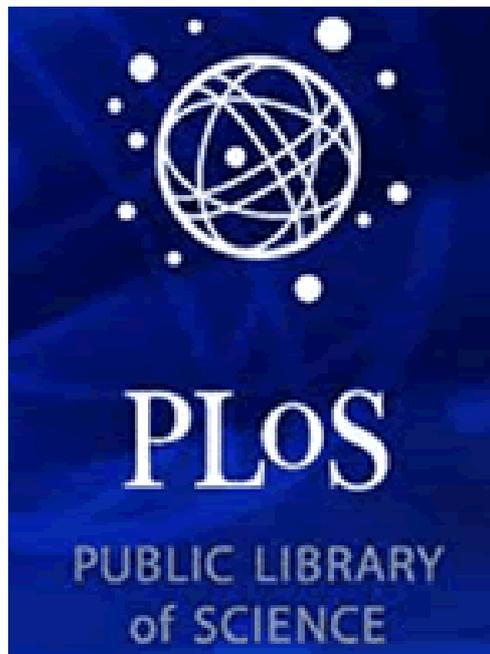
- Gusmão L, Alves C and Amorim A (2001). Molecular characterization of four human Y-specific microsatellites (DYS434, DYS437, DYS438, DYS439) for population and forensic studies. *Ann. Hum. Genet.* 65: 285-291.
- Hares DR (2012). Expanding the CODIS core loci in the United States. *Forensic Sci. Int. Genet.* 6: e52-e54.
- Brazilian Institute of Geography and Statistics (IBGE) (2010). Censo Demográfico 2010. <http://censo2010.ibge.gov.br/resultados>. Accessed: 23/01/2014.
- Jacques GS and Minervino AC (2007). Aspectos éticos e legais dos bancos de dados de perfis genéticos. *Perícia Federal, Brasília*, 26: 17-20. [www.apcf.org.br](http://www.apcf.org.br). Accessed: 04/02/2014.
- Krzyzannska A, Kowalczyk E, Markowska J and Dobosz T (2006). STR loci D2S1338 and D19S433 in a population sample from the Lower Silesia region. *Arch. Med. Sadowej. Kryminol.* 56: 236-238.
- Leite FP, Menegassi FJ, Schwengber SP, Raimann PE, et al. (2003). STR data for 09 autosomal STR markers from Rio Grande do Sul (southern Brazil). *Forensic Sci. Int.* 132: 223-224.
- Li CT, Guo H, Zhao ZM and Li L (2008). Analysis of genetics and genomics of short tandem repeat loci commonly used in Kinship Testing. *Fa Yi Xue Za Zhi* 24: 214-220.
- Martinez-Marignac VL, Valladares A, Cameron, A, Chan, A, et al. (2007). Admixture in Mexico City: implications for admixture mapping of type 2 diabetes genetic risk factors. *Hum. Genet.* 120: 807-819.
- Martins JA, Costa JC, Paneto GG, Figueiredo RF, et al. (2010). Genetic profile characterization of 10 X-STRs in four populations of the southeastern region of Brazil. *Int. J. Leg. Med.* 124: 427-432.
- Michelin K, Pacheco AC, Auler-Bittencourt EA, Lima MJM, et al. (2008). Banco de Perfis Genéticos no Combate aos Crimes Sexuais. *Perícia Federal, Brasília*. 26: 13-16. [www.apcf.org.br](http://www.apcf.org.br). Accessed 02/02/2014.
- Pena S D J, Bastos-Rodrigues L, Pimenta JR and Bydlowski SP (2009). DNA Tests Probe the Genomic Ancestry of Brazilians. *Braz. J. Med. Biol. Res.* 42: 870-876.
- Pernambuco (2007). Pacto pela Vida. <http://www.pactopelavida.pe.gov.br>. Accessed: 29/01/2014.

- Ribeiro-Rodrigues EM, Leite FPN, Hutz MH, Palha Tde J, et al. (2008). A multiplex PCR for 11 X chromosome STR markers and population data from a Brazilian Amazon Region. *Forensic Sci. Int. Genet.* 2: 154-158.
- Salzano MF (2004). Interethnic variability and admixture in Latin América--social implications. *Rev. Biol. Trop.* 52: 405-415.
- Silva DA, Crouse CA, Chakraborty R, Góes ACS, et al. (2004). 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 Sci. Int.* 139: 173-176.
- Soloum de Neves Manta F, Pereira R, Vianna R, Rodolfo Beuttenmüller de Araújo A, et al. (2013). Revisiting the genetic ancestry of Brazilians using autossomal AIM-INDELS. *PLoS One* 8: e75145.
- Tang H, Choudhry S, Mei R, Morgan M, et al. (2007). Recent genetic selection in the ancestral admixture of Puerto Ricans. *Am. J. Hum. Genet.* 81: 626-633.
- Tereba A (1999). Tools for analysis of population statistics, *Profile DNA* 2:14–16. <http://www.promega.com/geneticidtools/powerstats/>. Accessed 22/01/2014.
- Zietkiewicz E, Witt M, Daca P, Zebracka-Gala J, et al. (2012). Current genetic methodologies in the identification of disaster victims and in forensic analysis. *J. Appl. Genet.* 53: 41-60.
- Walsh SJ (2004). Recent advances in forensic Ggenetics. *Expert Rev Mol. Diagon.*4: 34-40.

**ARTIGO II**

**Forensic genetic data of six new generation STR loci (*D1S1656*, *D2S441*, *D10S1248*, *D12S391*, *D22S1045* and *SE33*): an expanded Pernambuco-Northeastern Brazil Population database.**

***Submetido***



***Qualis: A2***

**Forensic genetic data of six new generation STR loci (*D1S1656*, *D2S441*, *D10S1248*, *D12S391*, *D22S1045* and *SE33*): an expanded Pernambuco-Northeastern Brazil population database.**

New generation forensic markers in Pernambuco.

Sandra M Santos<sup>1,2\*</sup>, Amanda F Flor<sup>1</sup>, Carlos A Souza<sup>2</sup>, Kaynara CN Rabelo<sup>1</sup>, Lucas A C Brandão<sup>6</sup>, Paulo E Raimann<sup>5</sup>, Paulo RE Souza<sup>3</sup>, Ronald R Moura<sup>4</sup>, Tatiana C Oliveira<sup>1</sup>, Wesley M Tsutsumida<sup>5</sup>, Sergio Crovella<sup>1,4</sup>

1. Laboratório de Imunopatologia Keizo Asami (LIKA) da Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil.

2. Laboratório de Perícia e Pesquisa em Genética Forense (LPPGF) da Gerência Geral de Polícia Científica da Secretaria de Defesa Social de Pernambuco (SDS), Recife (PE), Brasil.

3. Laboratório Genoma da Universidade Federal Rural de Pernambuco (UFRPE), Recife (PE), Brasil.

4. Departamento de Genética, Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil.

5. Applied Biosystems by Thermo Fisher

6. Departamento de Patologia, Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil.

\*Corresponding author: SANTOS Sandra M,

Tel.: 55 81 31835389

Fax: 55 81 31835388

e-mail:sandra.santos@sds.pe.gov.br

**ABSTRACT.** Pernambuco is located in Northeastern Brazil, it ranks seventh in the number of inhabitants and stands out in the national scene due to the reducing violence policy adopted in the last seven years called "Pact for Life". It is sought, among the actions, studies and research especially in the area of Forensic Genetics the strengthening of the human profiling methodologies through DNA aiming the identification of criminals. The present study describes the allele frequencies of six STR new generation STR loci (D2S441, D22S1045, D10S1248, D1S1656, D12S391 and SE33) and other genetic parameters in the population of Pernambuco, compared with other populations. These markers were added to those initially recommended by the FBI in criminal investigations, as well as to feed databases of genetic profiles. 250 samples of unrelated individuals from all geographic regions of Pernambuco were analyzed. All loci analyzed showed a Hardy-Weinberg equilibrium in the population studied ( $P > 0.05$ ). The heterozygosity numbers observed ranged from 0.75600 (D2S441) to 0.94000 (SE33), except the marker D22S1045, which showed  $H_{obs} \leq 0.75$ . The other statistical parameters analyzed indicate that the studied markers are very informative when it comes to human identification in forensic purposes in the population of Pernambuco. The allele frequencies for each locus in the State of Pernambuco (Northeastern Brazil) were compared with ten different populations spread out in Brazil, Europe and Africa, with results that indicated that the most divergent population for the loci studied was the Brazilian from Pernambuco and the Lebanese (average  $DN = 0.1215$ ), whereas the two populations of Germany were the least divergent from each other (average  $DN = 0.0096$ ) and the nearest to Pernambuco. The results show that the studied markers are sufficiently informative and should be incorporated into the routine of human identification with forensic purposes in the northeastern population of Brazil.

Key words: Autosomal genetic markers, CODIS STRs, DNA typing, Human identification, population data, forensic DNA analysis.

## **AUTHOR SUMMARY**

At Present time, fighting violence is a major concern of any Brazilian citizen. Recent studies have noted that some cities, notably located in the northeastern region, are among the most violent in the world. Across the country, numerous actions are taken, in order to reduce crimes, among these, it is highlighted the investments given to the production of physical evidence, that means, in criminal forensics. This study describes the allelic frequency of 6 markers of forensic interest, commonly used in genetic-based human identification, aiming to explain crimes of sexual nature, crimes against life, crimes against the property, among others. Currently these markers are recommended by renowned international institutions and are already part of the forensic genetics laboratories routine in Brazil. Considering that to identify an individual from a given population it is necessary a previous study of the allele frequency of the region, the results here described represent an expansion of the database of the state of Pernambuco and it will be directly applied in the DNA forensic analysis of the Laboratory of Expertise and Research in Forensic Genetics, contributing to the control of violence at the state level and thus favoring the growth of criminal forensics in the Northeast of Brazil.

## **INTRODUCTION**

The most used DNA markers in human identification and genetic linkage studies are the *short tandem repeats* (STRs) [1, 2, 3]. Short tandem repeats are regions of tandemly repeated DNA segments found throughout the human genome that vary in length (through insertion, deletion, or mutation) with a core repeated DNA sequence [4].

The STRs gained even more importance and prominence due to its applicability in human identification [45] when the FBI in November 1997 instituted the 13 *loci* that would be the basis of the Combined DNA Index System (CODIS) [5, 6, 7]. The purpose of a genetic profiles database is to store DNA profiles, from biological samples collected from crime scenes and those that allow to locate missing persons. Continuous comparison of profiles also allows the detection of several crimes committed by a single individual, being fundamentally important to help solving future crimes [8].

The success of the DNA databases, the rate of crime solving supported by this technology, linking criminals to crime scenes led many countries to sanction specific legislation establishing their own databases [9]. In Brazil, in the year 2012 the Law No. 12.654/12 was sanctioned establishing the DNA as a form of criminal identification [10]. In 2013 the Decree No. 7.950/2013 established the National Database of Genetic Profiles and the Integrated Network of Genetic Profiles Database [11].

These 13 markers are still been widely used throughout the world, and even after more than a decade it has an important role in the field of forensic analysis, and when dealing with the biological sample of a single source they meet satisfactorily the needs of any forensic lab [12, 13 , 14, 15].

In May 2010, the FBI established a working group to review the 13 loci initially determined and suggest new loci with the purpose of expanding the base of CODIS. It is important to point out that the genetic markers used in human identification must present significant performance in forensic cases, high discrimination power, higher frequency of heterozygotes, low mutation rate, high degree of information, smaller size and lower formation of stutters [16].

The main reasons for expanding the base of the CODIS was the need to reduce the probability of accidental error, as the CODIS database progressively increases every year; to increase the compatibility of markers between different countries and to increase the power of discrimination aiming at locating missing people [8].

Numerous debates about which types of genetic markers should be implemented as a new basis for the databases of genetic profiles were made over the last few years [17]. It is known, however, that regardless of the database used when you want to identify an individual from a given population, the previous study of the allele frequency in this population is necessary [18], given that the distribution of the allelic frequency of DNA markers differs between geographically separated populations or with distinct origin [19].

From the moment that a set of STRs markers was standardized aiming human identification for forensic purposes numerous populations started to be investigated in order to establish the allele frequencies of the markers. The main results of these studies were used to generate baseline data on the allele frequencies of the populations studied, and to establish which markers were significant enough to specific populations [20].

In Brazil, there have been many studies on allele frequencies, whether statewide or covering the entire national territory. The vast majority of studies report allelic frequencies of the 13 *loci* recommended by the CODIS, database of genetic profiles used by Brazil, plus two additional markers related to the multiplex kit used [21,22,23,24,25,26,27, 28,29,30,31,32,33,34,35,36, 37].

After the new recommendations of the CODIS, as well as according to the new Brazilian reality, of which the deployment and implementation of the Database of Genetic Profiles continues in progress and currently covers the vast majority of the states, the need to study allele frequencies of the new markers suggested by CODIS arises.

In this context, the present study reports the allele frequency of six STR *loci* present in the new recommendations of CODIS, not yet described to the Pernambuco population, comparing them with other populations and thus expanding the information about allele frequencies in Pernambuco.

## RESULTS

### Intrapopulation analysis

Tables 1 and 2 show the descriptive statistics data for the 06 STR markers evaluated in the state of Pernambuco population, such as allele frequency, observed and expected heterozygosity, the exact test number for Hardy-Weinberg equilibrium (HWE) and other statistical parameters of forensic interest.

All *loci* analyzed presented themselves in Hardy-Weinberg equilibrium, considering the significance level of 5%. The estimated content of polymorphism for the 06 *loci* ranged from 0.70 to 0.94. The marker that showed lower informative polymorphism content was the D22S1045, with 11 alleles detected. The marker with the highest polymorphism information content was the SE33 with 27 alleles detected.

The numbers of heterozygosity observed ranged from 0.75600 (D2S441) to 0.94000 (SE33), except the marker D22S1045, which showed a Hobs  $\leq$  0.75. The loci SE33 and D1S1656 indicated a higher discrimination power, respectively, 0.990 and 0.978.

Table 1. Allele frequency distribution of 6 STR loci observed in the present study.

<i>Allele</i>	<i>D2S441</i>	<i>D22S1045</i>	<i>D12S391</i>	<i>D10S1248</i>	<i>D1S1656</i>	<i>SE33</i>
<i>2n</i>	<i>500</i>	<i>500</i>	<i>500</i>	<i>500</i>	<i>500</i>	<i>500</i>
8		0.002000				
9	0.004000			0.002000		
10	0.228000	0.004000		0.004000	0.004000	
11	0.314000	0.136000		0.040000	0.062000	0.004000
11.3	0.050000					
12	0.064000	0.018000		0.058000	0.086000	0.006000
12.3	0.002000					
13	0.034000	0.004000		0.278000	0.090000	0.010000
13.3	0.002000					
14	0.272000	0.044000		0.298000	0.122000	0.032000
14.3					0.010000	
15	0.022000	0.346000	0.056000	0.194000	0.162000	0.058000
15.3					0.060000	
16	0.008000	0.330000	0.042000	0.098000	0.134000	0.096000
16.3					0.056000	
17		0.098000	0.114000	0.022000	0.034000	0.068000
17.3					0.128000	
18		0.012000	0.194000	0.006000	0.002000	0.108000
18.3					0.038000	
19		0.006000	0.140000			0.086000
19.3			0.010000		0.012000	
20			0.158000			0.050000
20.2						0.030000
21			0.102000			0.008000
21.2						0.008000
22			0.104000			0.028000
22.2						0.014000
23			0.060000			
23.2						0.022000
24			0.016000			
24.2						0.002000
25			0.004000			0.032000
25.2						0.060000
26.2						0.074000
27.2						0.080000
28.2						0.040000
29.2						0.040000
30.2						0.002000
31						0.016000
31.2						0.020000
32.2						0.006000

Table 2. Forensic parameters in Pernambuco, Brazil.

<i>Allele</i>	<i>D2S441</i>	<i>D22S1045</i>	<i>D12S391</i>	<i>D10S1248</i>	<i>D1S1656</i>	<i>SE33</i>
<i>2n</i>	500	500	500	500	500	500
<b><i>MP</i></b>	0.090	0.106	0.031	0.089	0.022	0.010
<b><i>PD</i></b>	0.910	0.894	0.969	0.911	0.978	0.990
<b><i>PIC</i></b>	0.73	0.70	0.86	0.75	0.89	0.94
<b><i>PE</i></b>	0.513	0.447	0.606	0.637	0.683	0.894
<b><i>TPI</i></b>	2.02	1.74	2.55	2.78	3.21	9.62
<b><i>Ho</i></b>	0.75600	0.71200	0.80800	0.82000	0.88800	0.94000
<b><i>He</i></b>	0.76865	0.74229	0.87646	0.78273	0.89739	0.93905
<b><i>P<sub>(HWE)</sub></i></b>	0.48347	0.13330	0.05868	0.39755	0.91781	0.84822

MP: matching probability; PD: power of discrimination; PIC: polymorphism information content; PE: power of exclusion; TPI: typical paternity index; *He*: expected heterozygosity; *Ho*: observed heterozygosity; P (HWE): uncorrected P values from Hardy–Weinberg equilibrium test.

### Interpopulational analysis

The allele frequencies for each locus in the State of Pernambuco (Northeastern Brazil) were compared with ten populations, distributed as follows: **Brazil**: Rio Grande do Sul [38]. **Europe**: Germany [populations Nordrhein-Westfalen (NRW) and Schleswig-Holstein (SH)] [39], Portugal [40], Eastern Croatia [41] Macedonia [42], Poland [43] and UK [44]. **Africa**: Maghreb [45] and Lebanon [46]. Based on gene frequencies of the six STRs, to which data is available [38, 39, 40, 41, 42, 43, 44, 45, 46], pairwise genetic distances were calculated between populations using the Nei's genetic distance ( $D_N$ ) [47] implemented in Microsoft Excel<sup>®</sup>. The  $D_N$  numbers for each *locus* are listed in Table 3.

Table 3. Nei's Genetic distances between the studied sample and other populations.

	Brazil PE	Brazil_R S	Croatia East	Maghreb	GermanyN RW	Germany_ SH	Lebanon	Macedó nia	Polonia	Portugal
Brazil_RS	0.0662									
Croatia_East	0.1060	0.0247								
Maghreb	0.0939	0.0447	0.0437							
Germany_NRW	0.0515	0.0769	0.1000	0.0986						
Germany_SH	0.0411	0.0776	0.0976	0.0906	0.0096					
Lebanon	0.1215	0.0468	0.0304	0.0286	0.1063	0.0978				
Macedônia	0.1117	0.0496	0.0175	0.0371	0.1010	0.0946	0.0231			
Polonia	0.1101	0.0321	0.0139	0.0603	0.1018	0.1058	0.0450	0.0348		
Portugal	0.0676	0.0160	0.0174	0.0328	0.0729	0.0667	0.0264	0.0248	0.0361	
UK	0.1076	0.0302	0.0160	0.0357	0.1024	0.0941	0.0352	0.0252	0.0327	0.0197

Based on the NJ tree, built from the average distance between the populations (figure 1), we are able to notice that the most divergent populations for the loci studied were the Brazilian from Pernambuco and the Lebanese (average  $D_N = 0.1215$ ), whereas the two populations from Germany were the less divergent between each other (average  $D_N = 0.0096$ ). The Brazilian population of Pernambuco had more similarities with the German European population than with the Brazilian population of Rio Grande do Sul.

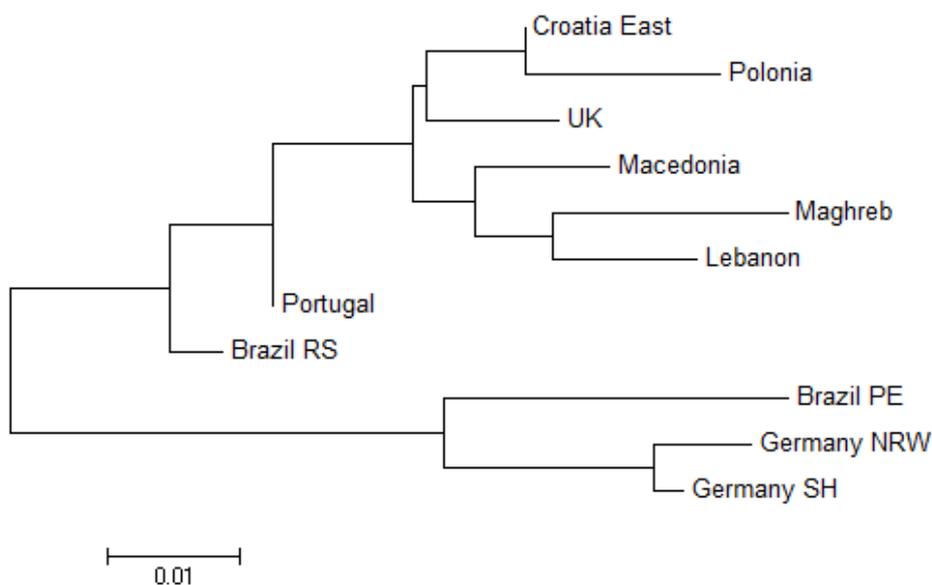


Figure 1. Neighbour-joining tree based on pairwise Nei's genetic distances calculated between the seven populations. References for the population data used: Northeastern Brazil [this study]; Southern Brazil [38]; Germany [39]; South Portuguese [40]; Eastern Croatia [41]; Republic of Macedonia [42]; Poland [43], UK [44], Maghreb [45] and Lebanon [46].

## DISCUSSION

The Brazilian population is one of the most heterogeneous populations in the world, consisting mainly of a mixture of three groups, they are: amerindians, europeans and africans [24]. This genetic variability of STR markers in populations located in different geographical areas of our country can be observed in several studies [21, 24, 48]. The ethnic and cultural mix of these three elements has also been the pillar for the configuration of the Northeastern population and thus the Pernambuco population. Pernambuco was one of the most important social centers of Brazil from the Northeast at the time. The most remarkable colonization was the Portuguese, then the Spanish, the Dutch, the Germans and the English, besides the strong African influence and of course, the indigenous [49].

These historical facts support the results obtained in this paper and demonstrate consistency with regard to the formation of the Pernambuco population.

The results of the statistical analysis of the 06 STR markers in the state of Pernambuco population showed high genetic diversity by the analysis of three statistical parameters. First, the markers showed heterozygosity between 0.74229 to 0.93905. In the sampling it was possible to detect the presence of rare alleles, that is, with frequency numbers below 0.5% for all markers analyzed. The high heterozygosity suggests that this population has a high variability, probably maintained by the random mating between individuals. Another parameter analysed, informative polymorphism content reflects the diversity of alleles for each marker and the frequency of these alleles in the population. Finally, we observed that the discrimination power ranged between 0.894 and 0.990. According to Butler [6], STRs markers useful in human identification must have higher than 90% discriminatory power and above 70% Heterozygosity.

The results obtained with the analysis of Hardy-Weinberg equilibrium suggest that the population of the state of Pernambuco is structured, considering only the 06 STRs in common. According to the International Society of Forensic Genetics, for an STR marker to be used in the forensic practice it is necessary that it is in Hardy-Weinberg equilibrium in the population that will be used [50].

A population that is in Hardy-Weinberg equilibrium holds the unchanged gene frequencies and the genotype proportions in a stable balance, presenting the same constant relation to each other over time. Thus, it is possible to determine the proportion of the present different genotypes in the population from the allelic frequencies found [51].

In Brazil, there are few studies on allelic frequency of the 06 markers analyzed here. Considering the data available, it was sought to compare with Rio Grande do Sul-Brazil and nine other populations located in Europe and Africa. The results indicate that the population of Pernambuco is genetically closer to the German population, a result that may be referred to the German presence during the colonization of Brazil, according to historical descriptions.

## **MATERIAL and METHODS**

### **DNA samples**

On the application of forensic study, the sampling was determined in proportion to the number of inhabitants residing in each of the eight territories of the State of Pernambuco, a territorial division implemented by the "Pact for Life" Government Program which aims to prevent and reduce violence. There are 08 territories and 26 integrated areas of security, associating actions of an overt police with actions of a judiciary police in the state [52]. The territorial division according to the model adopted by the Government of the State of Pernambuco: the "Pact for Life" program is displayed in Figure 2.



Figure 2. Map of the state of Pernambuco, divided into 8 territories. Adapted from: <http://www.sds.pe.gov.br/>.

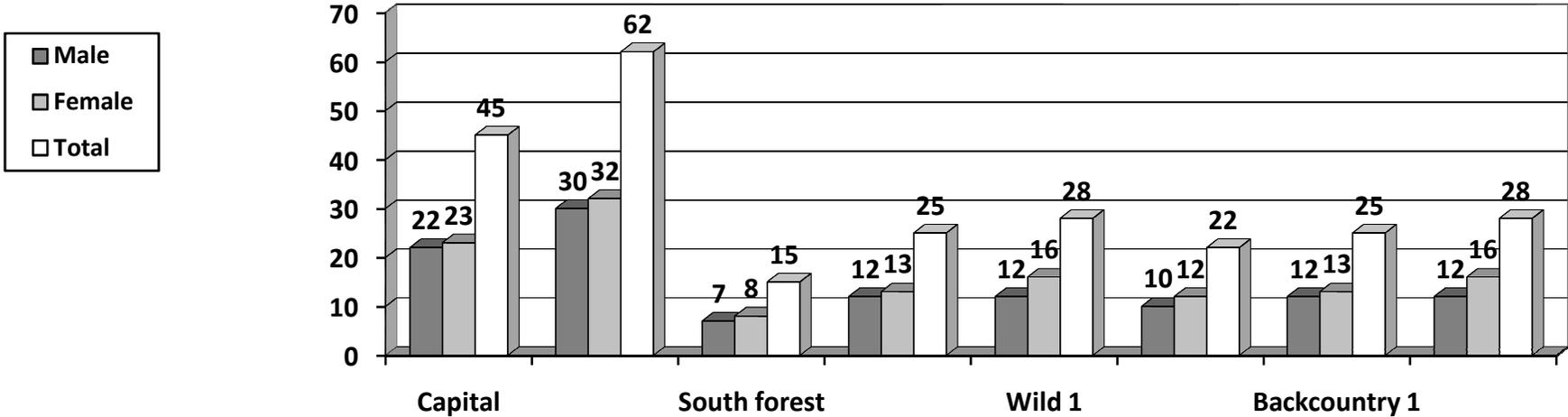
The collections were made in the host cities of each integrated area of security (AIS), summing up 22 cities, resulting in a very diversified sampling. Buccal mucosa samples were analyzed, collected from 250 adult individuals, unrelated, all born and living in the State of Pernambuco. All participants were aged between 18 and 70 years.

The samples were obtained using a mouth swab, by signing a clear and informed consent. The sample was composed of individuals spread over territory (capital, metropolitan region, matasul, matanorte, agreste 1 and 2, sertão1 and 2), respecting the proportionality of the number of inhabitants, according to the census of the Brazilian Institute of Geography and Statistics (IBGE) [53]. The distribution of the sampling by territory and sex can be viewed in Figure 3.

### **Ethics Statement**

This study was approved by the Ethics Committee on Human Research of the Centro de Ciências da Saúde, Universidade Federal de Pernambuco, CEP/CCS/UFPE. The Certificate Presentation to Ethical Consideration (CAAE) was 4961813.2.0000.5208. Our laboratory successfully participated in proficiency testing provided by the GITAD - 2013.

Figure 3. sampling territory and sex.



### **DNA extraction**

Biological samples were extracted using the *PrepFiler Express™ Forensic DNA Extraction* (Applied Biosystems by Thermo Fisher), through *Automate Express™*, as recommended by the manufacturer and adopted at the “Laboratório de Perícia e Pesquisa em Genética Forense” (Recife-Pernambuco).

### **STR's genotyping**

To amplify the 21 regions (loci) (*D21S11*, *TH01*, *D3S1358*, *FGA*, *TPOX*, *D8S1179*, *D16S539*, *D7S820*, *D13S317*, *D5S818*, *vWA*, *D18S51* and *CSF1PO*) as well as 3 to identify gender, we used the *GlobalFiler™ Express Kit*, (Applied Biosystems by Thermo Fisher) commercial kits following the manufacturer protocols. Amplicon separation and detection were performed using the ABI3500 automated sequencer (Applied Biosystems by Thermo Fisher) with the GS-600 LIZ standard and POP 4 polymer (Applied Biosystems by Thermo Fisher) size. Genotyping was performed using GeneMapper IDX software (Applied Biosystems by Thermo Fisher).

### **Statistical analysis**

Allele frequencies, Hardy-Weinberg equilibrium, expected heterozygosity ( $H_e$ ), and observed heterozygosity ( $H_o$ ) were calculated using the Arlequin software version 3.5.1.2 [55]. Other statistical parameters of forensic and paternity importance included matching probability (MP), power of discrimination (PD), polymorphism information content (PIC), power of exclusion (PE), and typical and paternity index (TPI), and were obtained using PowerStats, version 12 [55]. A tree using the Neighbor Joining method was built from Nei matrix of genetic distance [47].

### **ACKNOWLEDGEMENTS**

The authors would like to acknowledge the Eduardo Campos (*in memoriam*), the Secretaria de Defesa Social de Pernambuco (SDS/PE/ Brasil). R. Moura is recipient of a post-graduate from Fundação de Amparo à Pesquisa do Estado de Pernambuco (FACEPE, IBPG 0926-2.02/11) and Applied Biosystems by Thermo Fisher.

## AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: SMS PRES SC. Performed the experiments: SMS WMT PER. Analyzed the data: SMS CAS RRM TCO SC. Contributed reagents/materials/analysis tools: SMS AFF CAS KCNR LACB PER PRES RRM WMT TCO SC. Wrote the paper: SMS RRM. Performed sample collection: SMS AFF CAS KCNR TCO. Read and provided comments on the manuscript: SMS AFF CAS KCNR LACB PER PRES RRM WMT TCO SC.

## REFERENCES

1. Kobachuk LAG, Malaghini M, Silva JS, Gusmão L, Bicalho MG (2012) Genetic profile characterization of ten X-STRs in a sample from Paraná, Brazil. *International Journal of Legal Medicine* 126:975-976.
2. Martins JA, Costa JC, Paneto GG, Figueiredo RF, Gusmão L, et al. (2010) Genetic profile characterization of 10 X-STRs in four populations of the southeastern region of Brazil. *Int J. Leg Med.* 124:427–432.
3. Ziętkiewicz E, Witt M, Daca P, Zebracka-Gala J, Goniewicz M, et al. (2012) Current genetic methodologies in the identification of disaster victims and in forensic analysis. *J. Appl. Genet.* 53:41-60.
4. Butler JM, Hill CR (2012). Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic Sci. Rev* 24:15.
5. Choi JY and Seo TS (2009). An integrated microdevice for high-performance short tandem repeat genotyping. *Biotechnol. J.* 4 :1530-1541.
6. Butler JM (2005). *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*. Elsevier Academic Press, New York.
7. Budowle B, Moretti TR, Niezgoda SJ and Brown BL (1998). CODIS and PCR-based short tandem repeat loci: law enforcement tools, in: *Proceedings of the Second European Symposium on Human Identification*, Promega Corporation, Madison, WI, pp. 73–88.
8. Ge J, Eisenberg A and Budowle B (2012). Developing criteria and data to determine best options for expanding the core CODIS loci. *Invest Genet.* 3: 1.
9. Martin PD (2004). National DNA databases—practice and practicability. A forum for discussion. *International Congress Series* 1261:1 – 8.

10. Brasil (2012). Lei nº 12.654, de 28 de maio de 2012. Casa Civil. Subchefia para assuntos jurídicos. Publicada no DOU de 29.5.2012. <http://www4.planalto.gov.br/legislacao/legislacao-1/leis-ordinarias>. Accessed: 03/02/2014.
11. Brasil (2013). Decreto 7.950, de 12 de março de 2013. Casa Civil. Subchefia para assuntos jurídicos. Publicado no DOU de 13.3.2013. <http://www4.planalto.gov.br/legislacao/legislacao-1/decretos>. Accessed: 10/02/2014.
12. Michelin K, Pacheco AC, Auler-Bittencourt EA, Lima MJM, Komers AT (2008). Banco de Perfis Genéticos no Combate aos Crimes Sexuais. Perícia Federal, Brasília, 26:13-16. [www.apcf.org.br](http://www.apcf.org.br). Accessed:02/02/2014.
13. FBI/Codis (1997). <http://www.fbi.gov/about-us/lab/codis/ndis-statistics>. Accessed: 16/01/2014.
14. Li CT, Guo H, Zhao ZM, Li L(2008). Analysis of genetics and genomics of short tandem repeat loci commonly used in Kinship Testing. *Fa Yi Xue Za Zhi* 24:214-20.
15. Krzyżanińska A, Kowalczyk E, Markowska J, Dobosz T (2006). STR loci D2S1338 and D19S433 in a population sample from the Lower Silesia region. *Arch Med Sadowej Kryminol* 56:236-238.
16. Hares DR (2012). Expanding the CODIS core loci in the United States. *Forensic Sci Int Genet.* 6: e52-e54.
17. Dixon LA, Dobbins AE, Pulker HK, Butler JM, Vallone PM et al. (2006). Analysis of artificially degraded DNA using STRs and SNPs—results of a collaborative European (EDNAP) exercise. *Forensic Sci. Int.* 164:3–44.
18. Chakraborty R, Stiver DN, Su B, Zhong Y, Budowle B (1999). The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. *Electrophoresis* 20:1682-1696.
19. Gusmão L, Alves C and Amorim A (2001). Molecular characterization of four human Y-specific microsatellites (DYS434, DYS437, DYS438, DYS439) for population and forensic studies. *Ann. Hum. Genet.* 65:285–291.

20. Babiker HMA, Schlebusch CM, Hassan HY, Jakobsson M (2011). Genetic variation and population structure of Sudanese populations as indicated by 15 Identifiler sequence-tagged repeat (STR) loci. *Investigative Genetics* 2:12.
21. Grattapaglia D, Schimidt AB, Costa e Silva C, Stringuer C, Fernandes AP, et al. (2001). Brazilian population database for the 13 STR loci of the AmpfISTR® Profiler Plus™ and Cofiler™ multiplex Kits. *Forensic Sci. Int.* 118:91-94.
22. Fridman C, Santos PC, Kohler P, Garcia CF, Lopez LF, et al. (2008). Brazilian population profile of 15 STR markers. *Forensic Sci Int Genet* 2:1-4.
23. Poiaras LA, Osorio PS, Spanhol FA, Coltre SC, Rodenbusch R, et al. (2010). Allele frequencies of 15 STRs in a representative sample of the Brazilian population. *Forensic Sci Int Genet.* 4:61-63.
24. Aguiar VRC, Wolfgramm Ede V, Malta FS, Bosque AGA, Mafia AC et al. (2012) Updated Brazilian STR allele frequency data using over 100,000 individuals: an analysis of CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, Penta D, Penta E, TH01, TPOX and vWA loci. *Forensic Sci Int Genet.* 6:504–509.
25. Leite FP, Menegassi FJ, Schwengber SP, Raimann PE, Albuquerque TK (2003). STR data for 09 autosomal STR markers from Rio Grande do Sul (southern Brazil). *Forensic Sci. Int.* 132:223-224.
26. Bydlowski SP, Moura-Neto RS, Soares RPS, Silva R, Debes-Bravo AA, et al. (2003). Genetic data on 12 STRs (F13A01, F13B, FESFPS, LPL, CSF1PO, TPOX, TH01, vWA, D16S539, D7S820, D13S317, D5S818) from four ethnic groups of São Paulo, Brazil. *Forensic Sci. Int.* 135:67–71.
27. Caine LM, Corte-Real F, Anjos MJ, Carvalho M, Serra A, et al. (2003) Allele frequencies of 13 loci in the Santa Catarina population of Southern Brazil. *J Forensic Sci.* 48:901-902.
28. Santos MV, Anjos MJ, Andrade L, Vide MC, Corte-Real F, et al. (2004). Population genetic data for the STR loci using the AmpFISTR Identifiler Kit in Bahia, Brazil. *International Congress Series* 1261:219-222.
29. Dellalibera E, Havro MLB, Souza M, Kajihara K, Silva LM, et al. (2004). Genetic analysis of 13 STR loci in the population from the State of Pernambuco, northeast Brazil. *Forensic Sci Int* 146: 57-59.

30. Silva DA, Crouse CA, Chakraborty R, Góes ACS, Carvalho EF (2004). 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 Sci. Int.* 139:173-176.
31. Góes AC, Silva DA, Fonseca GEH, Silva MT, Pereira RW, et al. (2004). Allele frequencies data and statistic parameters for 16 STR loci - D19S433, D2S1338, CSF1PO, D16S539, D7S820, D21S11, D18S51, D13S317, D5S818, FGA, PENTA E, TH01, vWA, D8S1179, TPOX, D3S1358 - in the Rio de Janeiro population, Brazil. *Forensic Sci. Int.* 140:131-132.
32. Ferreira FL, Leal-Mesquita ER, Santos SEB, Ribeiro-dos-Santos AKC (2005). Genetic characterization of the population of São Luís, MA, Brazil. *Genetics and Molecular Biology* 28:1.
33. Gomes AV, Silva LM, Raposo G, Vieira JRC, Silva RS (2007). 13 STR loci frequencies in the population from Paraíba, Northeast Brazil. *Forensic Sci. Int.* 173:231-234.
34. Rodrigues EMR, Palha TJBFde and Santos SEB (2007). Allele frequencies data and statistic parameters for 13 STR loci in population of the Brazilian Amazon Region. *Forensic Sci. Int.* 168:244-247.
35. del Castillo DM, Perone C, de Queiroz AR, Mourão PH, Vasconcellos LS, et al. (2009). Population genetic data for 15 STR markers in the Brazilian population of Minas Gerais. *Legal Medicine* 11:45-47.
36. Wolggramm FV, Silva BC, Aguiar VRC, Malta FSV, Castro AM, et al. (2010). Genetic analysis of 15 autosomal and 12 Y-STR loci in the Espírito Santo State population, Brazil. *Forensic Sci. Int.* 5:41-43.
37. Francez PAC, Rodrigues EMR, Frazão GF, Borges NDR, Santos SEB (2011). Allelic frequencies and statistical data obtained from 12 codis STR loci in an admixed population of the Brazilian Amazon. *Genetics and Molecular Biology* 34:35-39.
38. Rodenbusch R, Schumacher S, Mardini AC, Chula FGL, Michelon CT (2012). Allele frequencies of the five new generation forensic STR (D1S1656, D2S441, D10S1248, D12S391 and D22S1045) in the population from Rio Grande do Sul, Southern Brazil. *Forensic Sci. Int.* 6:55-57.

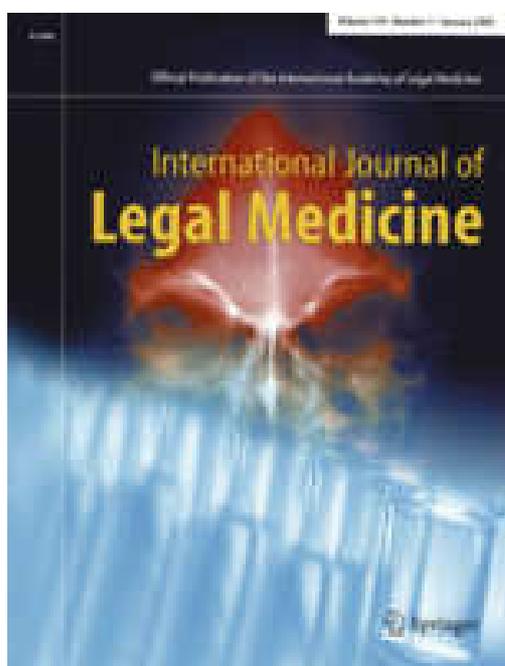
39. Poetsch M, Bayer K, Ergin Z, Milbrath M, Schwark T, et al. (2011). First experiences using the new Powerplex® ESX17 and ESI17 kits in casework analysis and allele frequencies for two different regions in Germany. *International Journal Legal Medicine*, 125:733–739.
40. Ribeiro T, Dario P, Sanches S, Espinheira R, Geada H et al. (2013). Population data of the AmpFISTR1 NGMTM loci in South Portuguese population. *Forensic Sci Int Genet.*7:37–39.
41. Ćurić G, Gašić V, Plužarić Vand Smiljčić D (2012). Genetic parameters of five new European Standard Set STR loci (D10S1248, D22S1045, D2S441, D1S1656, D12S391) in the population of eastern Croatia. *Croat Med J.* 53:409-15.
42. Jakovski Z, Nikolova K, Jankova-Ajanovska R, Janeska B, Pojskic N, Marjanovicab D (2012). Allele frequencies of the new European Standard Set (ESS) loci plus SE33 locus in a population from the Republic of Macedonia. *Forensic Sci Int Genet.*6:90–92.
43. Parys-Proszek A, Kupiec T, Ska-Nowak PW, Branicki W (2010). Genetic variation of 15 autosomal STR loci in a population sample from Poland. *Legal Medicine* 12:246–248.
44. Tucker VC, Baumgartner C, Stead GR and Hopwood AJ (2012). UK population data generated with the PowerPlex1 ESI 16 System. *Forensic Sci Int Genet.* 6:112–118.
45. Cortellini V, Cerri N and Verzeletti A (2011). Genetic variation at 5 new autosomal short tandem repeat markers (D10S1248, D22S1045, D2S441, D1S1656, D12S391) in a populationbased sample from Maghreb region. *Croat Med J.* 52: 368-71.
46. Andari AEI, Othman H, Taroni F and Mansour I (2013). Population genetic data for 23 STR markers from Lebanon. *Forensic Sci Int Genet.*7:108–113.
47. Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York.
48. Whittle MR, Romano NL, and Negreiros VAC (2004). Updated Brazilian genetic data, together with mutation rates, on 19 STR loci, including D10S1237. *Forensic Sci. Int.*139: 207–210.

49. Pernambuco (2014). Governo de Pernambuco: <http://www.pe.gov.br/conheca/populacao>. Accessed: 29/03/2014.
50. Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E at al. (2003). Use of X-linked markers for forensic purposes. *International Journal of Legal Medicine*. 117: 67-74.
51. Beiguelman B (1995). *Dinâmica dos genes nas famílias e nas populações*. 2ª ed. Ribeirão Preto: Editora SBG.
52. Pernambuco (2007). Pacto pela Vida. <http://www.pactopelavida.pe.gov.br>. Accessed: 27/01/2014.
53. IBGE (2010). Censo Demográfico 2010. <http://censo2010.ibge.gov.br/resultados>. Accessed: 30/04/2014.
54. Excoffier L (2010) Arlequin ver 3.5.1.2. <http://cmpq.unibe.ch/software/arlequin3>. Accessed: 27/05/2014.
55. PowerStats version 12. Promega Corporation. Available on: <http://www.promega.com/geneticidtools/powerstats>.

**ARTIGO III**

**Genetic profile characterization of twenty one STR loci and twelve X-STR in sample from Fernando de Noronha, Pernambuco, Brazil.**

***A ser submetido***



***Qualis: B1***

**Genetic profile characterization of twenty one STR loci and twelve X-STR in sample from Fernando de Noronha, Pernambuco, Brazil.**

Forensic markers in Fernando de Noronha.

Sandra Maria dos Santos<sup>1,2\*</sup>, Amanda Ferreira Flor<sup>2</sup>, Carlos Antônio de Souza<sup>2</sup>, Kaynara Cecília Nery Rabelo<sup>1</sup>, Lucas André Cavalcanti Brandão<sup>5</sup>, Luciellen d`Avila Giacomel Kobachuk<sup>6</sup>, Paulo Roberto Eleutério de Souza<sup>3</sup>, Ronald Rodrigues Moura<sup>4</sup>, Tatiana C Oliveira<sup>1</sup>, Sergio Crovella<sup>1,4</sup>

**Affiliations:**

1. Laboratório de Imunopatologia Keizo Asami (LIKA) da Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil.
2. Laboratório de Perícia e Pesquisa em Genética Forense (LPPGF) da Gerência Geral de Polícia Científica da Secretaria de Defesa Social de Pernambuco (SDS), Recife (PE), Brasil.
3. Laboratório Genoma da Universidade Federal Rural de Pernambuco (UFRPE), Recife (PE), Brasil.
4. Departamento de Genética, Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil.
5. Departamento de Patologia, Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil.
6. Laboratório de Genética Molecular Forense, Instituto de Criminalística do Paraná, Curitiba (PR), Brasil.

\*Corresponding author: SANTOS Sandra M,

Tel.: 55 81 3183 5389

Fax: 55 81 31835388

E-mail: sandra.santos@sds.pe.gov.br

## Abstract

Fernando de Noronha is an isolated archipelago in the Equatorial Atlantic, belonging to the Brazilian state of Pernambuco, composed of 21 islands of volcanic origin, taking (maybe, taking up) an area of 26 square kilometers, to the northeast of the capital of Pernambuco, Recife. The main island is the only Brazilian oceanic island inhabited constantly for over 500 years. It has a population of 2,630 inhabitants and it is heavily visited by tourists, a factor which adds further complexity to an already multiethnic population just like the Brazilian population. The present study describes the allele frequencies of markers of forensic interest in the population of Fernando de Noronha, including 21 autosomal STRs and 12 X chromosome STRs, and other genetic parameters comparing them with other populations. 29 non-related families were analyzed, representing all the native people. All the autosomal markers were observed in Hardy-Weinberg equilibrium. The estimated content of polymorphism among the 21 loci ranged from 0.60 to 0.93. The observed heterozygosity values ranged from 0.72414 to 0.96552. The loci SE33 showed higher discrimination power. Hardy-Weinberg equilibrium was tested in female samples, and significant deviations were found. The linkage disequilibrium analysis did not reveal consistent evidence of association between the X-STRs used. The average diversity per locus varied between 69% and 93% for *DXS8378*, and *DXS10135*, with the average diversity for all markers being 81%. The observed values in this study indicate that these markers constitute a powerful tool for human identification purposes in the forensic population of Fernando de Noronha. However, since the significant differences between the population and the local people from the island and the continent were observed, these numbers will be added to the data base of genetic profiles of the population of Pernambuco.

## Keywords

Autosomal genetic markers, X chromosome, DNA typing, Human identification, genetic population data, forensic DNA analysis.

## Introduction

The occupation of Fernando de Noronha is almost as old as the continent. Due to its geographical position, the archipelago was one of the first lands found in the New World. The island was forgotten for over two centuries and located on the route of the great journeys it was approached by many different people, among them the Dutch, French, German and English [1].

Only in 1737, Portugal decided to occupy and populate the island, through the captaincy of Pernambuco. Thus, a defense system was deployed with several fortifications, and the construction of two urban cores for the operation of a correctional colony for ordinary prisoners coming from Pernambuco. In many national political phases there were also sheltered political prisoners: the gypsies of Brazil (in 1739), the farroupilhas (in 1844) and the capoeiristas (in 1890). In 1938 the island was requested by the Union to become officially a Political Prison [2].

Today Fernando de Noronha is a State District, managed by Pernambuco, with a remaining population from the several experienced periods, and added to those who went there for different reasons. Descendants of common criminals or politicians, guards, military or people who have been deployed there to provide services, or simply to follow fellow islanders sightseeing, make up this population, which gets over to 2,630 people, mostly of whom live off of tourism [3]. The Brazilian population shows high levels of genetic variability, a result of the contribution of three ethnic groups: amerindians, europeans, and African [4]. The major contribution of Amerindians was observed in the north, while the larger African and European contributions were found in the northeast and south regions, respectively [5]. The markers most commonly used in forensic genetics are autosomal short tandem repeats (STRs), followed by Y chromosome STRs and mitochondrial DNA [6].

In Brazil, there have been many studies on allele frequencies, whether statewide or covering the entire national territory. The vast majority of studies report allelic frequencies of the 13 *loci* recommended by the CODIS, database of genetic profiles used by Brazil, plus two additional markers related to the multiplex kit used [7 - 23].

The analysis of X-chromosome STRs (X-STRs) has recently become of importance in some complex cases of biological relationships, when the alleged father is not available, in paternity cases involving close blood relatives, in maternity cases involving a son, and in some other identification cases. In these situations, X-STRs may efficiently complement the autosomal analysis as they present higher mean exclusion chances [24, 25, 26 and 27]. In recent years, several studies had validated panels of markers of the X chromosome to be used in forensic genetics. [28, 29, 30, 31 and 32].

In this context, the present study reports the allele frequency of 21 STR loci present in the new recommendations of CODIS loci and 12X-STRs, never described in the population of Fernando de Noronha, comparing them with other populations and thus expanding the information about allele frequencies in Pernambuco.

## Materials and methods

### DNA samples

The samples were collected on the island of Fernando de Noronha, after careful study of kinship among the natives, resulting in a sample of only 29 families which were born and lived on the island for generations. Thus we ensure that donors of biological sample are not related

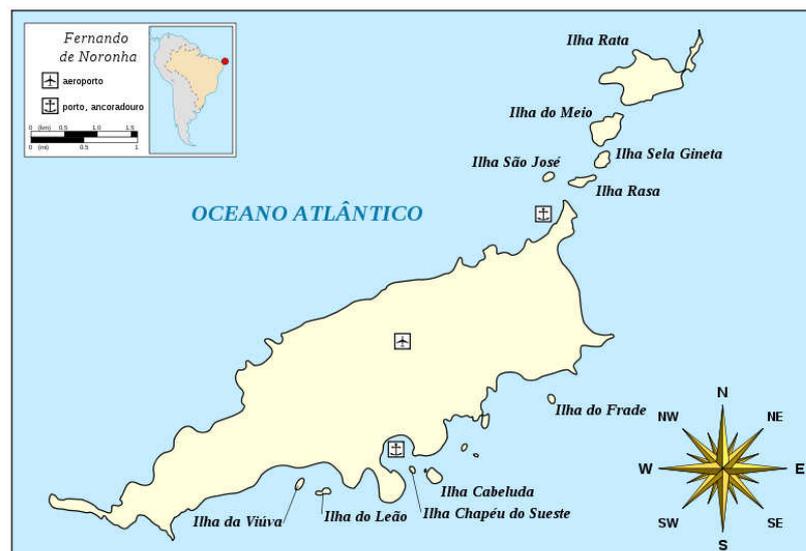


Figure 1: Map of Fernando de Noronha. <http://www.ladatco.com/FEN-Map.HTM> [2].

The samples were obtained using a mouth swab, by signing a clear and informed consent. According to the census of the Brazilian Institute of Geography and Statistics [21] Fernando de Noronha has 2,630 inhabitants.

This study was approved by the Ethics Committee on Human Research of the Centro de Ciências da Saúde, Universidade Federal de Pernambuco, CEP/CCS/UFPE. Certificate Presentation to Ethical Consideration (CAAE): 14961813.2.0000.5208. Our laboratory successfully participated in proficiency testing provided by the GITAD - 2013.

### **DNA extraction**

Biological samples were extracted using the *Resin Chelex® 100 (Bio-Rad Laboratories)*, as recommended by the manufacturer [34, 35].

### **STR's genotyping**

To amplify the STR 21 regions (loci) (*D19S433, D2S1338, D21S11, TH01, D3S1358, FGA, TPOX, D8S1179, D16S539, D7S820, D13S317, D5S818, vWA, D18S51, CSF1PO, D2S441, D22S1045, SE33, D10S1248, D1S1656 e D12S391*), as well as three to identify gender, we used the *GlobalFiler™ Express Kit, (Life Technologies by Termofisher)*.

Amplification of the twelve X-STR loci (*DXS7132, DXS7423, DXS8378, DXS10074, DXS10079, DXS10101, DXS10103, DXS10134, DXS10135, DXS10146, DXS10148* and *HPRTB*) was performed in a single PCR multiplex reaction using the Investigator Argus X-12 Kit (Qiagen) contains primers for Amelogenin for gender determination.

Separation and detection were performed in an ABI3500 automated sequencer (Applied Biosystems) using the GS-600 LIZ size standard and POP 4 polymer (Applied Biosystems). The Genotyping was performed used GeneMapper IDX software (Applied Biosystems).

### Statistical analysis

Allele frequencies, Hardy-Weinberg equilibrium (HWE), the expected heterozygosity ( $H_e$ ) and observed heterozygosity ( $H_o$ ) were calculated using the Arlequin software version 3.5.1.2 [36]. The polymorphism information content, power of discrimination, power of exclusion, matching probability, typical paternity index, expected heterozygosity, and observed heterozygosity were calculated using PowerStats v.1.2 [37].

The allele frequencies, gene diversities, exact test of the Hardy-Weinbergequilibrium for female samples, pairwise exact test of linkage disequilibrium (LD) formale samples, were calculated using ARLEQUIN software version 3.1 [36]. Statistics for forensic efficiency evaluation of each locus, namely, the expected probability of exclusion in trios involving daughters (PET) and in father/daughter duos (PED) and the power of discrimination in females (PDF) and in males (PDM), were computed according to Desmarais et al. [38].

## RESULTS and DISCUSSION

### Autosomal STR

Tables 1 and 2 present data descriptive statistics for the 21 STR markers evaluated in the population of Fernando de Noronha, as allele frequencies, observed and expected heterozygosity, the value of the exact test for Hardy-Weinberg equilibrium (HWE) and other statistical parameters of forensic interest.

All loci analyzed showed up in Hardy-Weinberg, considering the significance level of 5%. The estimated content of polymorphism for the 21 loci ranged from 0.60 to 0.93. The *D3S1358*, *TPOX*, *TH01* and *D22S1045* markers showed the lowest polymorphism information content, all with only 05 alleles detected. The marker with the highest polymorphism information content was the *SE33* with 22 alleles detected. The observed heterozygosity values ranged from 0.72414 to 0.96552. The loci *SE33* showed higher discrimination power.

Comparing the data obtained from the population of Fernando de Noronha with available data for the same markers (*CFS1PO*, *D3S1358*, *D5S818*, *D7S820*, *D8S1179*, *D13S317*, *D16S539*, *D18S51*, *D21S11*, *FGA*, *TH01*, *VWA* and *TPOX*) population of Pernambuco, the continent [39] as well as other regions of Brazil [23,40,17,41], no significant differences were observed. Fernando de Noronha, even for a relatively small sample compared with other studies, the allele 24.3 (*D21S11*) first appears at a frequency study for the state of Pernambuco [39,15]. For the remainder of the markers analyzed in this study (*D19S433*, *D2S1338*, *D2S441*, *D22S1045*, *D10S1248*, *D1S1656* and *D12S391*) comparisons are limited to the few published studies [41, 42, 43], and also followed the pattern observed in Fernando de Noronha. No publications on studies of allelic frequency of the locus SE33 in the Brazilian population were found in Pernambuco. The results presented here can complement satisfactorily allele frequencies already published, as also described markers never be for studied in Pernambuco.

### **X-STR**

Tables 3 and 4 present the descriptive statistics data for the 12 X-STR markers for the population of Fernando de Noronha and Forensic statistical parameters. Gene diversity (the same as power of discrimination in males;  $PD_M$  [44]) per locus varied between 69% for *DXS8378* and 93% for *DXS10135*, with the average diversity for all markers being 81%. For a significance level of 0,005 (after the Bonferroni correction), no deviations from the Hardy-Weinberg equilibrium were observed for any of the analyzed loci. In our study, *DXS10135* was the most informative marker, and *DXS8378* was the least informative. Overall, the power of discrimination was 0,9999999999 for females and 0,9999999995 for males. The combined probability of exclusion was 0,999999998 for trios and 0,99999995 for duos. The linkage disequilibrium analysis did not reveal consistent evidence of association between the X-STRs used. We had great difficulty comparing our data with those of other populations from Brazil because only a small subset of previously studied markers was coincident with our 12 X-STRs. Further, the studied markers should be adopted in routine forensic of Pernambuco, because the multiplex system used is available in the market.

Most markers were more discriminative in Fernando de Noronha than in other Brazilian [5, 24, 45, 46,47]. In general, the X-STRs used in this work proved to be highly discriminating and therefore useful for forensic purposes, should be incorporated in to routine forensic of Pernambuco, complementing previous studies. In conclusion, in the Fernando de Noronha population, the 12 X-STRs and 21 autosomal short tandem repeats (STRs) are in Hardy-Weinberg equilibrium, could be used in forensic casework and kinship analysis. This study will also compose the data base allelic profiles for the people of Pernambuco, Brazil.

Table 1. STR allele frequencies for each of 21STR loci found in the Fernando de Noronha- Pernambuco - Brazil sample.

Allele	D3S1358	vWA	D16S539	CSF1PO	TPOX	D8S1179	D21S11	D18S51	D2S441	D19S443	TH01	FGA	D22S1045	D5S818	D13S317	D7S820	SE33	D10S1246	D1S1656	D12S391	D2S133
2n																					8
6											0,1724					0,0172					
7											0,3276			0,0517		0,1896					
8			0,0172	0,0345	0,4827						0,0517				0,0345	0,1552					
9			0,2069	0,0345	0,1034	0,0172					0,1552			0,0690	0,0690	0,2757					
9.3											0,2931										
10			0,1724	0,2414	0,0349	0,0345		0,0172	0,1207					0,0517	0,1207	0,2069		0,0172	0,0862		
11			0,2241	0,3793	0,3103				0,3448	0,0517			0,1207	0,2241	0,1724	0,1552					
11.3						0,0690			9,0517						0,4655				0,0690		
12			0,2241	0,2759	0,06897	0,1379		0,0172	0,1207	0,1552				0,3793				0,0345			
12.2										0,0172											
12.3															0,0517						
13			0,1379	0,0345		0,2241		0,0690	0,0345	0,2759				0,2241				0,2241	0,1552		
13.2										0,0172											
13.3															0,0862						
14	0,1207	0,1034	0,0172			0,2931		0,1896	0,3103	0,1724			0,0690				0,0172	0,3448	0,1034		
14.2								0,0172		0,0517											
14.3																			0,0345		
15	0,2241	0,1207				0,1724		0,1552	0,0172	0,1552			0,3965				0,0517	0,1379	0,0862	0,5172	
15.2										0,0517											
15.3																			0,0345		
16	0,3448	0,3103				0,0345		0,1724		0,05172			0,3276				0,0862	0,1896	0,1379	0,2831	0,0690
16.3																			0,0345		
17	0,2241	0,2069				0,0172		0,1896					0,0862				0,0517	0,0517	0,0345	0,1034	0,2414
17.3																			0,1552	0,0345	
18	0,0862	0,1896						0,0862									0,1034		0,0172	0,1379	0,0517
18.3																			0,0172	0,0345	
19		0,0345						0,0345			0,1034						0,1207			0,1207	0,1207
19.3																			0,0172	0,0172	
20		0,0345						0,0517				0,1207					0,0517			0,0690	0,1552
20.3																			0,0172		
21												0,1552					0,0172			0,2069	0,0517
21.2																	0,0345				
22												0,1207					0,0172			0,0690	0,0690
23												0,0862					0,0172			0,0690	0,1034
24												0,1552									0,0517
24.2																	0,0172				
24.3						0,0172															
25												0,1379					0,0172			0,0172	0,0690

25.2				0,0517		
26			0,0345			
26.2				0,0295	0,0172	0,0172
27	0,0862		0,0172			
27.2				0,0690		
28	0,1207					
28.2				0,0862		
29	0,2069					
29.2				0,0517		
29.9				0,0172		
30	0,2414					
31	0,1207					
31.2	0,0172			0,0345		
32.2	0,1379			0,0179		
33				0,0172		
33.2	0,0517					

---

Table 2. Forensic Parameters in Fernando de Noronha, Pernambuco, Brazil.

STRs	MP	PD	PIC	PE	TPI	H <sub>o</sub>	H <sub>e</sub>	P (HWE)
<i>D3S1358</i>	0,099	0,901	0,72	0,525	2,07	0,75862	0,77193	0,98854
<i>vWA</i>	0,084	0,916	0,77	0,651	2,90	0,82759	0,81125	0,90239
<i>D16S539</i>	0,094	0,906	0,78	0,467	1,81	0,72414	0,82154	0,27713
<i>CSF1PO</i>	0,180	0,820	0,67	0,651	2,90	0,82759	0,73079	0,13141
<i>TPOX</i>	0,210	0,790	0,60	0,467	1,81	0,72414	0,66546	0,70077
<i>D8S1179</i>	0,089	0,911	0,78	0,719	3,63	0,86207	0,82154	0,82048
<i>D21S11</i>	0,073	0,927	0,82	0,586	2,42	0,79310	0,85481	0,22460
<i>D18S51</i>	0,075	0,925	0,84	0,859	7,25	0,96552	0,87235	0,20707
<i>D2S441</i>	0,120	0,880	0,72	0,719	3,63	0,86207	0,76467	0,93701
<i>D19S443</i>	0,084	0,916	0,82	0,788	4,83	0,89655	0,84936	0,11943
<i>TH01</i>	0,146	0,854	0,71	0,651	2,90	0,82759	0,76346	0,12776
<i>FGA</i>	0,049	0,951	0,87	0,859	7,25	0,93103	0,89776	0,91487
<i>D22S1045</i>	0,141	0,859	0,66	0,467	1,81	0,72414	0,72111	0,66018
<i>D5S818</i>	0,137	0,863	0,71	0,651	2,90	0,82759	0,75862	0,49415
<i>D13S317</i>	0,122	0,878	0,69	0,467	1,82	0,72414	0,73563	0,55153
<i>D7S820</i>	0,089	0,911	0,77	0,525	2,07	0,75862	0,81065	0,57036
<i>SE33</i>	0,034	0,966	0,93	0,930	14,50	0,96552	0,95039	0,96949
<i>D10S1246</i>	0,108	0,892	0,74	0,467	1,81	0,72414	0,78524	0,10095
<i>D1S1656</i>	0,046	0,954	0,89	4,83	0,044	0,89655	0,91228	0,46328
<i>D12S391</i>	0,89	0,956	0,88	0,859	7,25	0,93103	0,90563	0,92687
<i>D2S1338</i>	0,058	0,942	0,86	0,651	2,90	0,82759	0,88506	0,26234

MP: matching probability; PD: power of discrimination; PIC: polymorphism information content; PE: power of exclusion; TPI: typical paternity index; H<sub>o</sub>: observed heterozygosity; H<sub>e</sub>: expected heterozygosity; P (HWE): uncorrected P values from Hardy–Weinberg equilibrium test.

**Table 3.** Allele frequency distribution and gene diversity X-STRs in Fernando de Noronha, Brazil.

Alleles	DXS10103	DXS8378	DXS7132	DXS10134	DXS10074	DXS10101	DXS10135	DXS7423	DXS10146	DXS10079	HRPTB	DXS10148
7					0,0232							
8					0,0930							
9											0,0232	
10		0,3256									0,0232	
11		0,2791									0,1395	
12		0,3488	0,1163		0,0465			0,0232			0,2093	
13		0,0465	0,1860		0,0698			0,0698			0,2558	
13.3												0,0465
14			0,1860		0,0465			0,3488		0,0232	0,2325	
15	0,0232		0,3721		0,1628			0,3953			0,0930	
16	0,0930		0,0930		0,2093			0,1395		0,0465	0,0232	
16.1							0,0232					
17	0,0930		0,0465		0,1395		0,0698	0,0232		0,0465		
17.1							0,0232					
18	0,2093				0,1860					0,1163		0,2325
18.1							0,0232					
19	0,4651				0,0232		0,0698			0,1628		0,0232
19.1							0,0232					
20	0,0930						0,0465			0,2790		0,0232
21	0,0232						0,0465			0,2325		
21.1							0,0232					
22							0,0930			0,0465		
23							0,1395			0,0465		0,0930
24							0,0698					
24.1							0,0232					0,1395
24.2						0,0698						
25							0,0232					
25.1												0,2791
26						0,0232	0,0465		0,1163			
26.1												0,0930
26.2						0,0232						
27							0,0232		0,1860			
27.1												0,0698
27.2						0,0465						

28				0,0232	0,0465	0,0930
29				0,0465	0,1163	0,0930
29.2				0,1163		
30		0,0232		0,0465	0,0232	0,1163

Cont.

Alleles	DXS10103	DXS8378	DXS7132	DXS10134	DXS10074	DXS10101	DXS10135	DXS7423	DXS10146	DXS10079	HRPTB	DXS10148
30.2						0,0698						
31				0,0232		0,0930			0,0930			
31.2						0,0465						
32				0,0698		0,1395	0,0232					
32.2									0,0232			
33				0,0465		0,1163			0,0698			
34				0,0465		0,0232						
35				0,1395								
35.2									0,0465			
36				0,1163								
37				0,2093								
37.2							0,0232					
38				0,1628								
38.3				0,0465								
39				0,0465								
39.2									0,0698			
40.3				0,0465								
43				0,0232								
43.2									0,0698			
46.2									0,0232			

---

**Table 4.** Forensic parameters in Fernando de Noronha, Pernambuco, Brazil.

X- STRs	HE	HO	$p$ - HWE	PE <sub>T</sub>	PE <sub>D</sub>	PD <sub>F</sub>	PD <sub>M</sub>
<i>DXS10103</i>	0,857	0,714	0,940	0,679	0,538	0,884	0,713
<i>DXS8378</i>	0,857	0,690	0,306	0,630	0,483	0,843	0,692
<i>DXS7132</i>	0,857	0,799	0,276	0,736	0,603	0,914	0,768
<i>DXS10134</i>	0,857	0,886	0,493	0,869	0,777	0,974	0,880
<i>DXS10074</i>	1,000	0,873	0,103	0,837	0,738	0,959	0,857
<i>DXS10101</i>	0,857	0,944	0,004	0,904	0,831	0,985	0,910
<i>DXS10135</i>	1,000	0,955	0,626	0,925	0,865	0,991	0,930
<i>DXS7423</i>	0,500	0,746	0,065	0,644	0,501	0,855	0,697
<i>DXS10146</i>	0,785	0,913	0,182	0,885	0,802	0,980	0,895
<i>DXS10079</i>	0,857	0,876	0,8677	0,796	0,678	0,944	0,819
<i>HRPTB</i>	0,571	0,831	0,008	0,779	0,656	0,935	0,807
<i>DXS10148</i>	0,571	0,857	0,021	0,801	0,686	0,947	0,823

*HE*: expected heterozygosity, *HO*: observed heterozygosity, *p-HWE*:  $p$  value for Hardy-Weinberg equilibrium test, *PET*: probability of exclusion in trios involving daughters, *PED*: probability of exclusion in father/daughter duos, *PDF*: power of discrimination in females, *PDM*: power of discrimination in males.

### Acknowledgements

The authors would like to acknowledge the Secretaria de Defesa Social de Pernambuco (SDS/PE/ Brasil). R. Moura is recipient of a post-graduate from Fundação de Amparo à Pesquisa do Estado de Pernambuco (FACEPE, IBPG 0926-2.02/11).

### Conflict of interest

The authors declare that they have no conflicts of interest.

### References

1. Roukema, E. (1963) Brazil in the Cantino Map", *Imago Mundi*, Vol. 17, p. 7-26.
2. Pernambuco (2014) Fernando de Noronha. [http:// www.noronha.pe.gov.br](http://www.noronha.pe.gov.br). Accessed 08/06/2014. IBGE (2014) <http://cod.ibge.gov.br/13S8>. Accessed 08/06/2014.
3. Santos SEB, Guerreiro JF (1995) The indigenous contribution to the formation of the population of the Brazilian Amazon Region. *Braz J Genet* 18:311–315.
4. Ribeiro-Rodrigues EM, Palha TJBF, Bittencourt EA, Ribeiro-dos-Santos A, Santos S. (2011) Extensive survey of 12 X-STRs reveals genetic heterogeneity among Brazilian populations. *Int J Legal Med* 125:445–452.
5. Aquino J, Peixe C, Silva D, Tavares C, Carvalho EF (2009) A Xchromosome STR hexaplex as a powerful tool in deficiency paternity cases. *Forensic Sci Int Genet Suppl Ser* 2:45–46.
6. Grattapaglia D, Schimidt AB, Costa e Silva C, Stringuer C, Fernandes AP, et al. (2001) Brazilian population database for the 13 STR loci of the AmpflSTR® Profiler Plus™ and Cofiler™ multiplex Kits. *Forensic Sci. Int.* 118:91-94.
7. Fridman C, Santos PC, Kohler P, Garcia CF, Lopez LF, et al. (2008) Brazilian population profile of 15 STR markers. *Forensic Sci Int Genet* 2:1-4.
8. Poiaras LA, Osorio PS, Spanhol FA, Coltre SC, Rodenbusch R, et al. (2010) Allele frequencies of 15 STRs in a representative sample of the Brazilian population. *Forensic Sci Int Genet.* 4:61-63.

9. Aguiar VRC, Wolfgramm Ede V, Malta FS, Bosque AGA, Mafia AC et al. (2012) Updated Brazilian STR allele frequency data using over 100,000 individuals: an analysis of CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, Penta D, Penta E, TH01, TPOX and vWA loci. *Forensic Sci Int Genet.*6:504–509.
10. Leite FP, Menegassi FJ, Schwengber SP, Raimann PE, Albuquerque TK (2003) STR data for 09 autosomal STR markers from Rio Grande do Sul (southern Brazil). *Forensic Sci. Int.* 132:223-224.
11. Bydlowski SP, Moura-Neto RS, Soares RPS, Silva R, Debes-Bravo AA, et al. (2003) Genetic data on 12 STRs (F13A01, F13B, FESFPS, LPL, CSF1PO, TPOX, TH01, vWA, D16S539, D7S820, D13S317, D5S818) from four ethnic groups of São Paulo, Brazil. *Forensic Sci. Int.* 135:67–71.
12. Caine LM, Corte-Real F, Anjos MJ, Carvalho M, Serra A, et al. (2003) Allele frequencies of 13 loci in the Santa Catarina population of Southern Brazil. *J Forensic Sci.* 48:901-902.
13. Santos MV, Anjos MJ, Andrade L, Vide MC, Corte-Real F, et al. (2004) Population genetic data for the STR loci using the AmpFISTR Identifiler *Kit* in Bahia, Brazil. *International Congress Series* 1261:219-222.
14. Dellalibera E, Havro MLB, Souza M, Kajihara K, Silva LM, et al. (2004) Genetic analysis of 13 STR loci in the population from the State of Pernambuco, northeast Brazil. *Forensic Sci Int* 146: 57-59.
15. Silva DA, Crouse CA, Chakraborty R, Góes ACS, Carvalho EF (2004) 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 Sci. Int.* 139:173-176.
16. Góes AC, Silva DA, Fonseca GEH, Silva MT, Pereira RW, et al. (2004) Allele frequencies data and statistic parameters for 16 STR loci - D19S433, D2S1338, CSF1PO, D16S539, D7S820, D21S11, D18S51, D13S317, D5S818, FGA, PENTA E, TH01, vWA, D8S1179, TPOX, D3S1358 - in the Rio de Janeiro population, Brazil. *Forensic Sci. Int.* 140:131-132.
17. Ferreira FL, Leal-Mesquita ER, Santos SEB, Ribeiro-dos-Santos AKC (2005) Genetic characterization of the population of São Luís, MA, Brazil. *Genetics and Molecular Biology* 28:1.

18. Gomes AV, Silva LM, Raposo G, Vieira JRC, Silva RS (2007) 13 STR loci frequencies in the population from Paraíba, Northeast Brazil. *Forensic Sci. Int.* 173:231-234.
19. Rodrigues EMR, Palha TJBFe and Santos SEB (2007) Allele frequencies data and statistic parameters for 13 STR loci in population of the Brazilian Amazon Region. *Forensic Sci. Int.* 168:244-247.
20. del Castillo DM, Perone C, de Queiroz AR, Mourão PH, Vasconcellos LS, et al. (2009) Population genetic data for 15 STR markers in the Brazilian population of Minas Gerais. *Legal Medicine* 11:45-47.
21. Wolggramm FV, Silva BC, Aguiar VRC, Malta FSV, Castro AM, et al. (2010) Genetic analysis of 15 autosomal and 12 Y-STR loci in the Espírito Santo State population, Brazil. *Forensic Sci. Int.* 5:41-43.
22. Francez PAC, Rodrigues EMR, Frazão GF, Borges NDR, Santos SEB (2011) Allelic frequencies and statistical data obtained from 12 codis STR loci in an admixed population of the Brazilian Amazon. *Genetics and Molecular Biology* 34:35-39.
23. Martins JÁ, Costa JC, Paneto GG, Figueiredo RF, Gusmão L, Sánchez-Diz P, Carracedo A, Cicarelli RMB (2009) Genetic profile characterization of 10 X-STRs in four populations of the southeastern region of Brazil. *Int J Leg Med* 124(5):427–432.
24. Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D (2003) Use of X-linked markers for forensic purposes. *Int J Leg Med* 117:67–74.
25. Builes JJ, Manrique A, Aguirre D, Puerto Y, Bravo ML, Gusmão L (2009) Utility of Y- and X-STRs in the research of complex biological relationship. *Forensic Sci Int Genet Suppl* 2:236–237.
26. Silva DA, Manta FSN, Desidério M, Tavares C, Carvalho EF (2009) Paternity testing involving human remains identification and putative half sister: usefulness of an X-hexaplex STR markers. *Forensic Sci Int Genet Suppl* 2:230–231.
27. Kobachuk LAG, Malaghini M, Silva JS, Gusmão L, Bicalho MG (2012) Genetic profile characterization of ten X-STRs in a sample from Paraná, Brazil. *International Journal of Legal Medicine* 126:975-976.

28. Ribeiro-Rodrigues EM, Palha TJBF, Santos AKCR, Santos SEB (2010) Genetic data of twelve X-STRs in a Japanese immigrant population resident in Brazil. *Forensic Sci Int Genet* 4:e57–e58.
29. Bekada A, Benhamamouch S, Boudjema A, Fodil M, Menegon S, Torre C, Robino C (2009) Analysis of 12 X-chromosomal STRs in an Algerian population sample. *Forensic Sci Int Genet Suppl Ser* 2:400–401.
30. Tamura A, Tsutsumi H, Hara M, Takada A, Saito K, Suzuki K, Komuro T (2009) Genetic studies of eight X-STRs in a Japanese population. *Leg Med* 11:S451–S452.
31. Leite FPN, Santos SEB, Rodríguez EMR, Callegari-Jacques SM, Demarchi DA, Tsuneto LT, Petzi-Erler ML, Salzano FM, Hutz MH (2009) Linkage disequilibrium patterns and genetic structure of Amerindian and non-Amerindian Brazilian populations revealed by long-range X-STR markers. *Am J Phys Anthropol* 139:404–412.
32. IBGE (2010) Censo Demográfico 2010. <http://censo2010.ibge.gov.br/resultados>. Accessed 09/07/2014.
33. Saul D (2006) Automated PCR of buccal-swab FTA cards in a 96-well format using forensic GEM™ and the Liquid-Handling Workstation epMotion® 5075 MC. Eppendorf.
34. Walsh PS, Metzger DA, Higuchi R (1991) Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10:506–13.
35. Excoffier L (2010) Arlequin ver 3.5.1.2. <http://cmpq.unibe.ch/software/arlequin3>. Accessed: 27/06/2014.
36. PowerStats version 12. Promega Corporation. <http://www.promega.com/geneticidtools/powerstats>. Accessed: 27/06/2014.
37. Desmarais D, Zhong Y, Chakraborty R, Perreault C, Busque L (1998) Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA). *J Forensic Sci* 43:1046–1049.

38. Santos SM, Souza CA, Rabelo KCN, Souza PRE, Moura RR, Oliveira TC, Crovella S (2014) Distribution of forensic marker allelic frequencies in Pernambuco, Northeastern Brazil. accepted for publication in *Genetics and Molecular Research* in 08/08/2014.
39. Leite FP, Menegassi FJ, Schwengber SP, Raimann PE, Albuquerque TK (2003). STR data for 09 autosomal STR markers from Rio Grande do Sul (southern Brazil). *Forensic Sci. Int.* 132:223-224.
40. Barra GB, Santa Rita TH, Chianca CF, Velasco LFR, Souza CF (2013) Allele frequencies for 15 autosomal STRs in a population sample from the Federal District (Brazil)—a territory that arose from nothing. *Int J Legal Med* 128:741-742.
41. Raimann PE, Hirschfeld-Campolongo G, Schumacher S, Almeida SEM, Silva CMD et al. (2012) Population genetic analyses of the AmpFISTR® NGM™ in Brazil. *Int J Legal Med* 126:337–341.
42. Rodenbusch R, Schumacher S, Mardini AC, Chula FGL, Michelon CT (2012). Allele frequencies of the five new generation forensic STR (D1S1656, D2S441, D10S1248, D12S391 and D22S1045) in the population from Rio Grande do Sul, Southern Brazil. *Forensic Sci. Int.* 6:55-57.
43. Pereira R, Gomes I, Amorim A, Gusmão L (2007) Genetic diversity of 10 X-chromosome STRs in northern Portugal. *Int J Leg Med* 121:192–197
44. Cainé LM, Zarrabeitia MT, Riancho JA, Pinheiro MF (2010) Genetic data of a Brazilian population sample (Santa Catarina) using an X-STR decaplex. *J Forensic Leg Med* 17:272-274
45. Kobachuk LAG, Malaghini M, Silva JS, Gusmão L, Bicalho MG (2012) Genetic profile characterization of ten X-STRs in a sample from Paraná, Brazil. *International Journal of Legal Medicine* 126:975-976.
46. Silva IHF, Barbosa AG, Azevedo DA, Sánchez-Diz P, Gusmão L, Tavares CC, Carvalho EF, Silva LAF (2010) An X-chromosome pentaplex in two linkage groups: Haplotype data in Alagoas and Rio de Janeiro populations from Brazil. *Forensic Science International: Genetics* 4:95-100.

## 5 GRÁFICOS COMPLEMENTARES

---

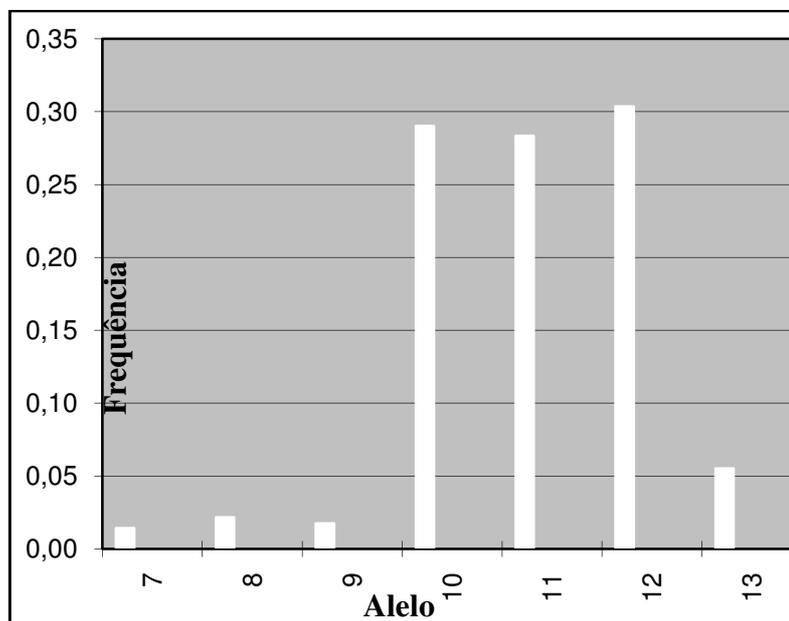


Gráfico 1: Frequência alélica do *loci* CFS1PO.

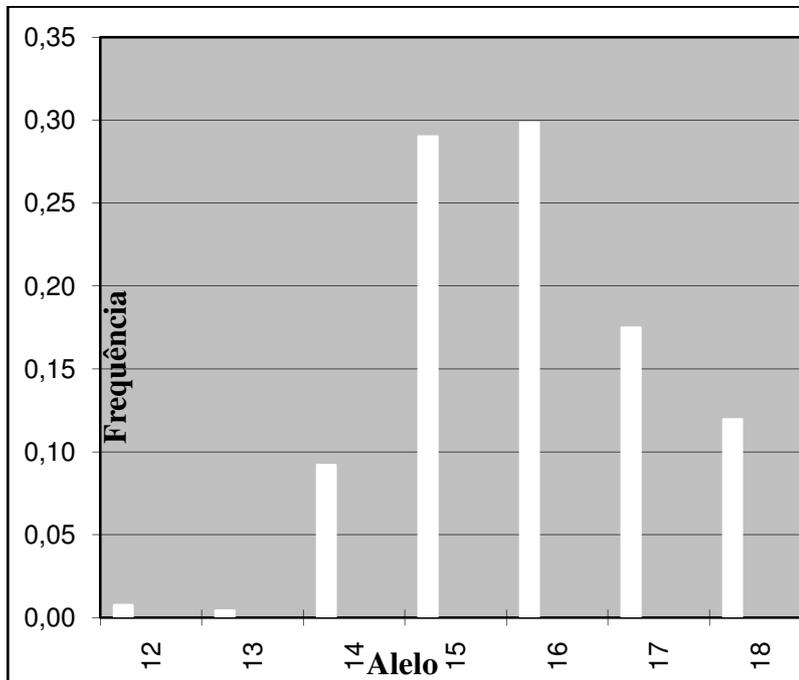


Gráfico 2: Frequência alélica do *loci* D3S1358.

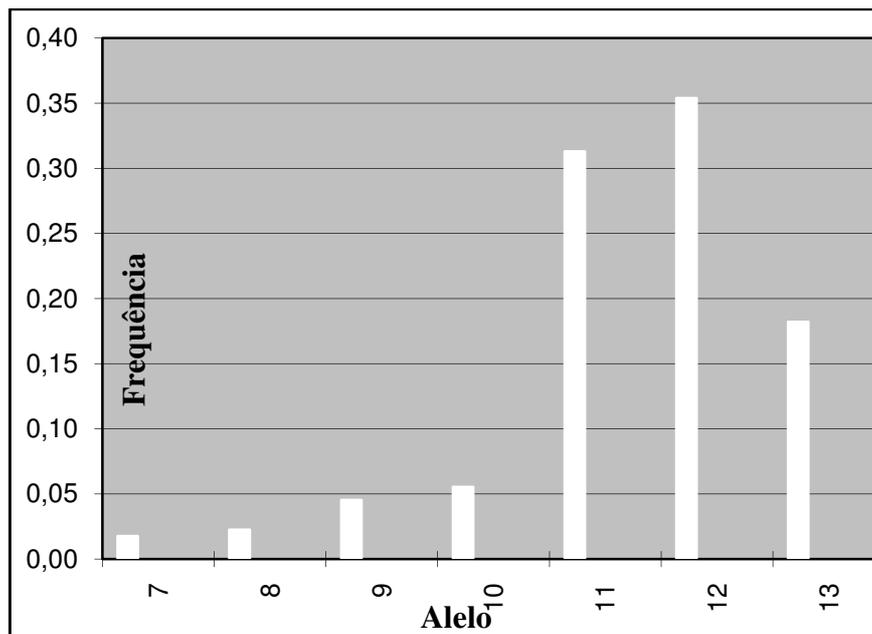


Gráfico 3: Frequência alélica do *loci* D5S818.

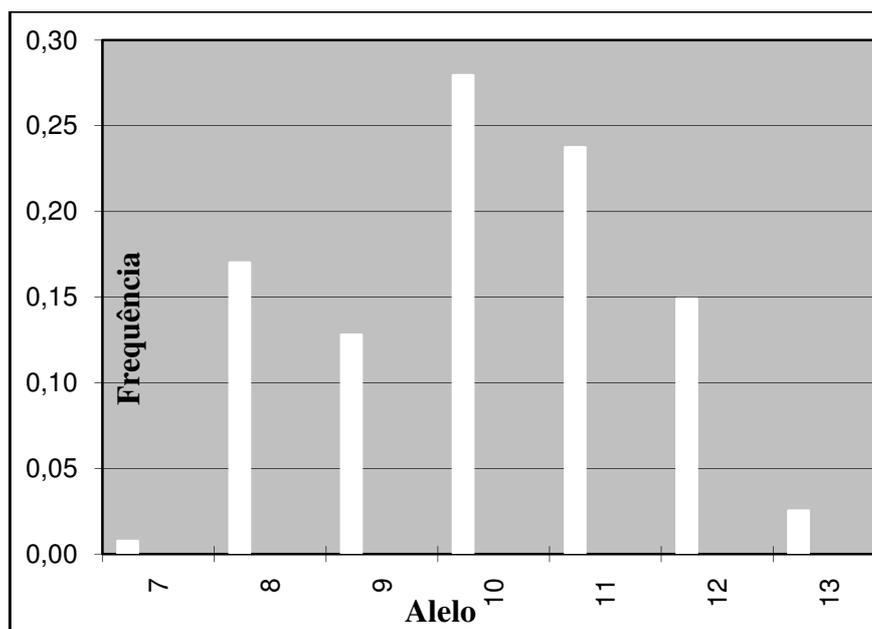


Gráfico 4: Frequência alélica do *loci* D7S820.

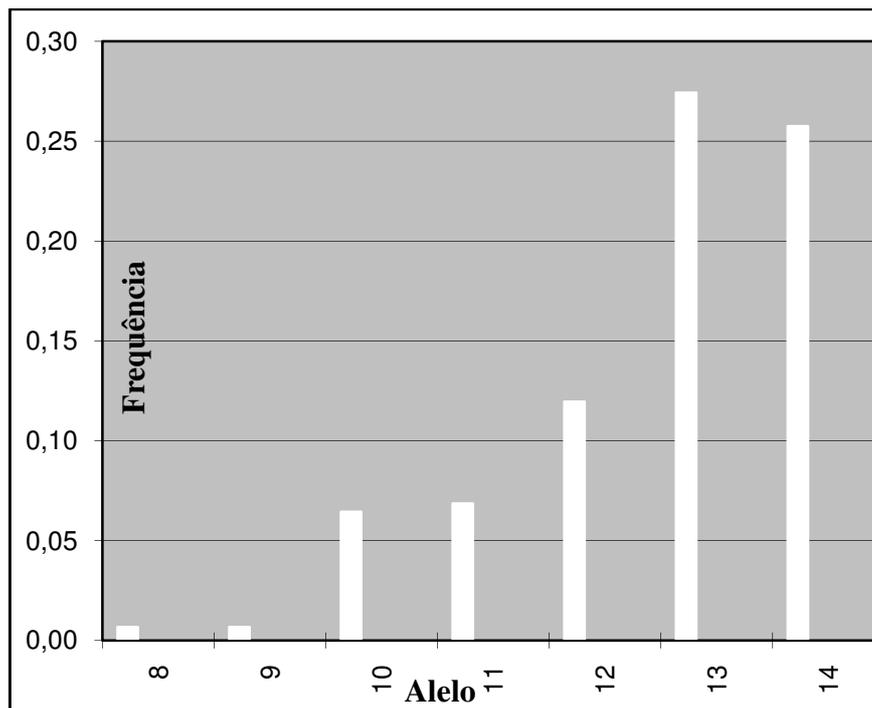


Gráfico 5: Frequência alélica do *loci* D8S1179.

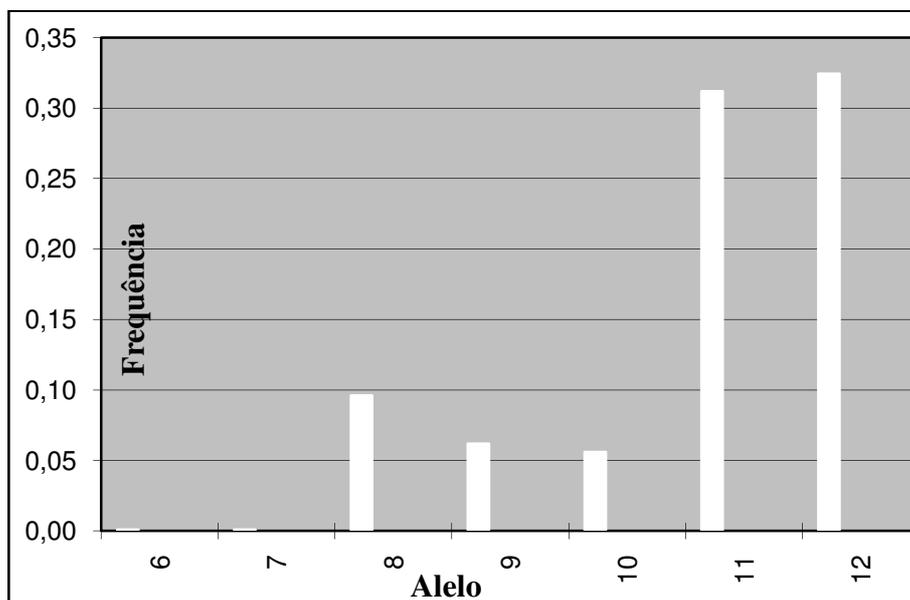


Gráfico 6: Frequência alélica do *loci* D13S317.

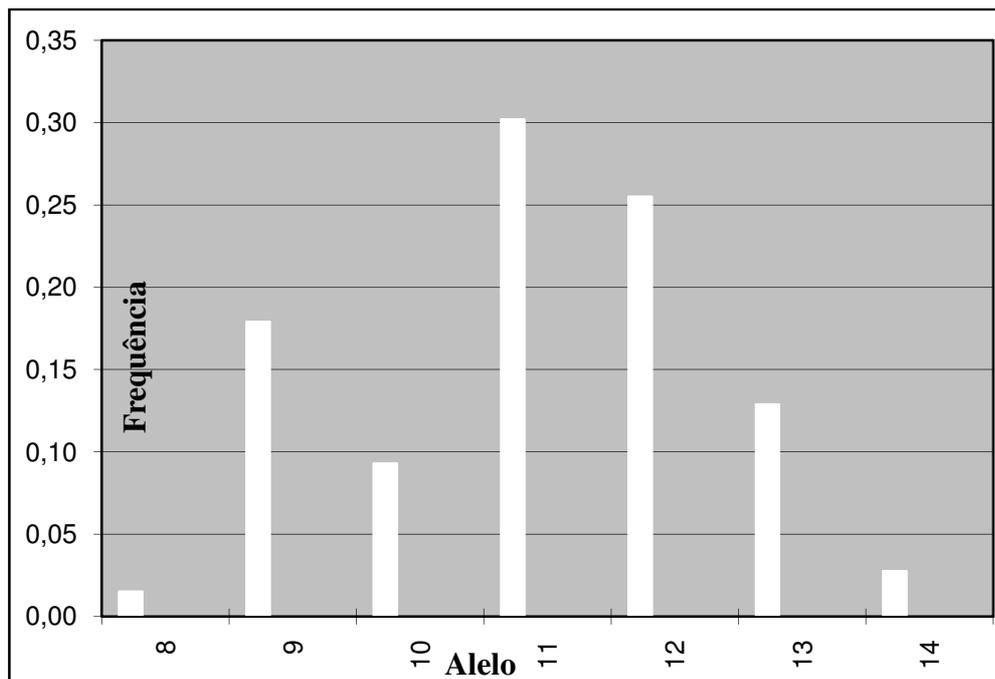


Gráfico 7: Frequência alélica do *loci* D16S539.

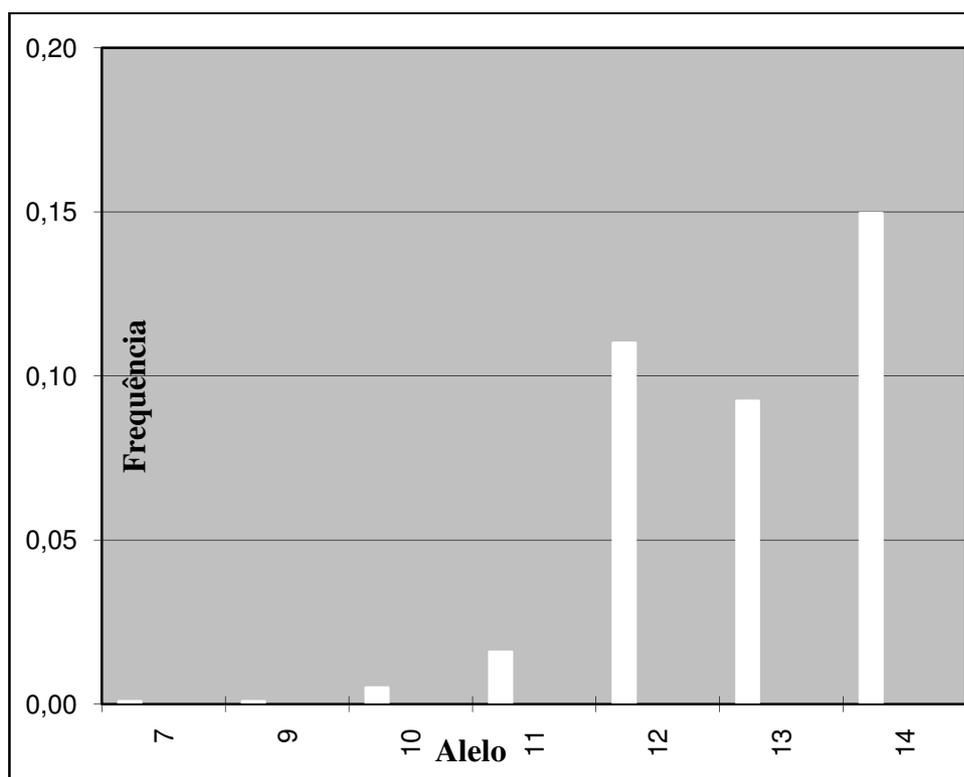


Gráfico 8: Frequência alélica do *loci* D18S51.

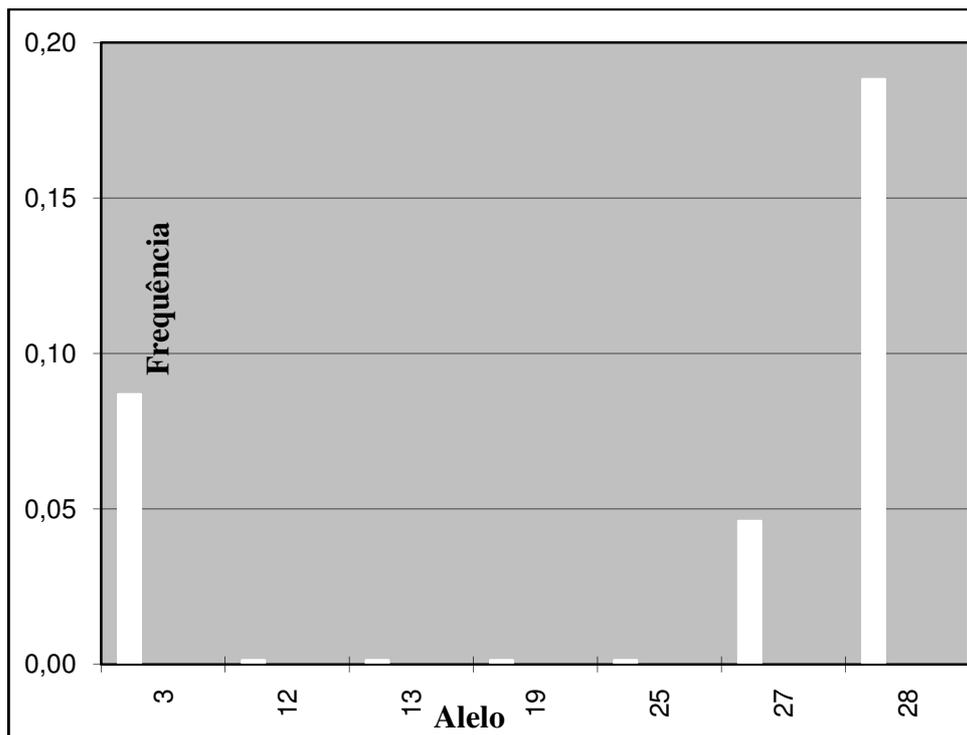


Gráfico 9: Frequência alélica do loci D21S11.

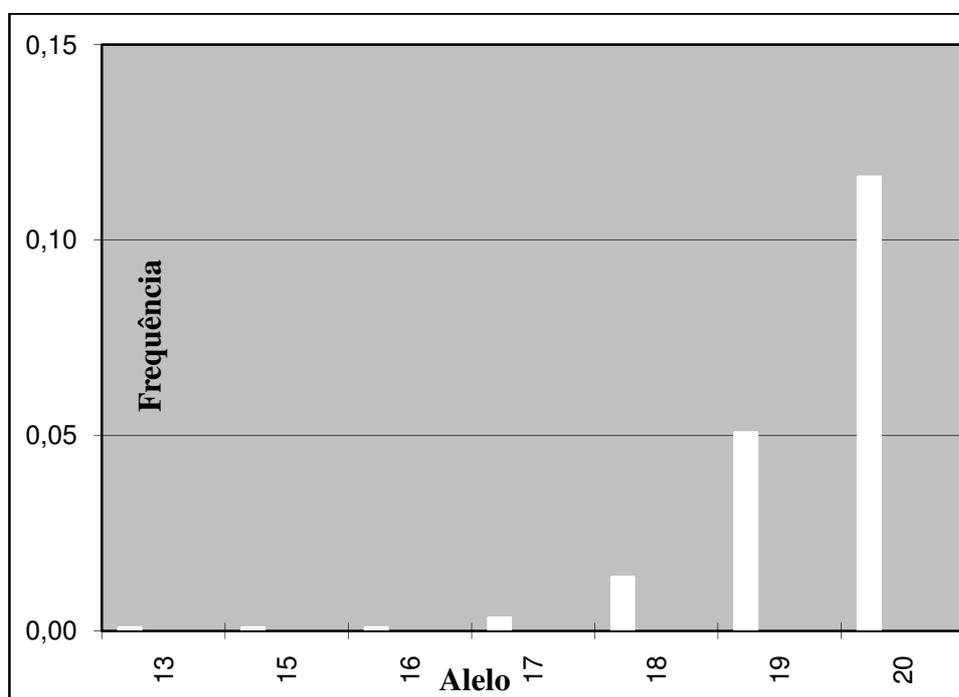


Gráfico 10: Frequência alélica do loci FGA.

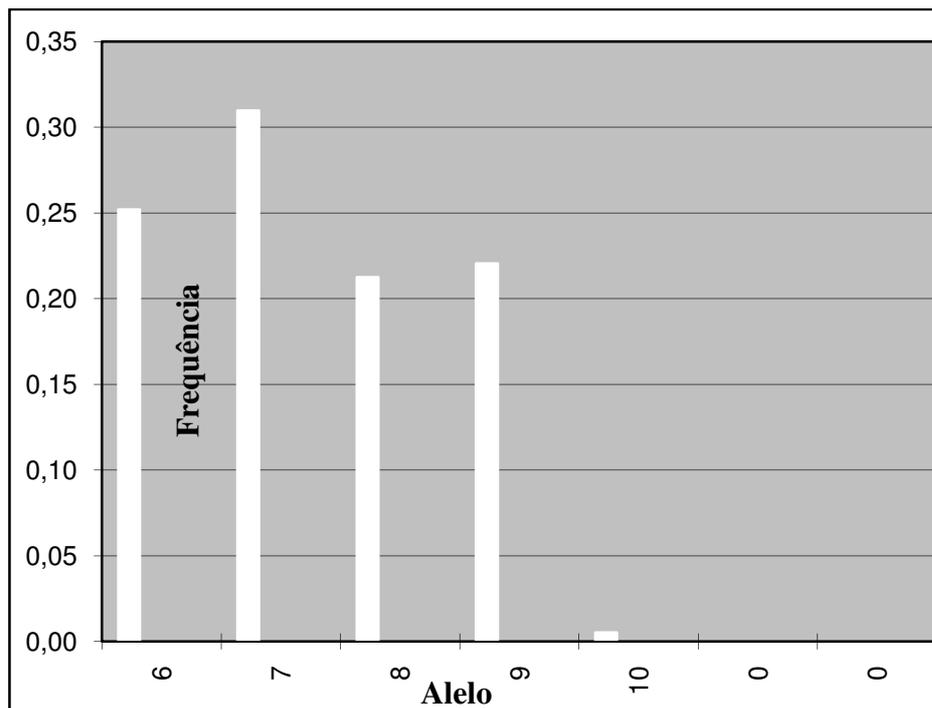


Gráfico 11: Frequência alélica do loci TH01.

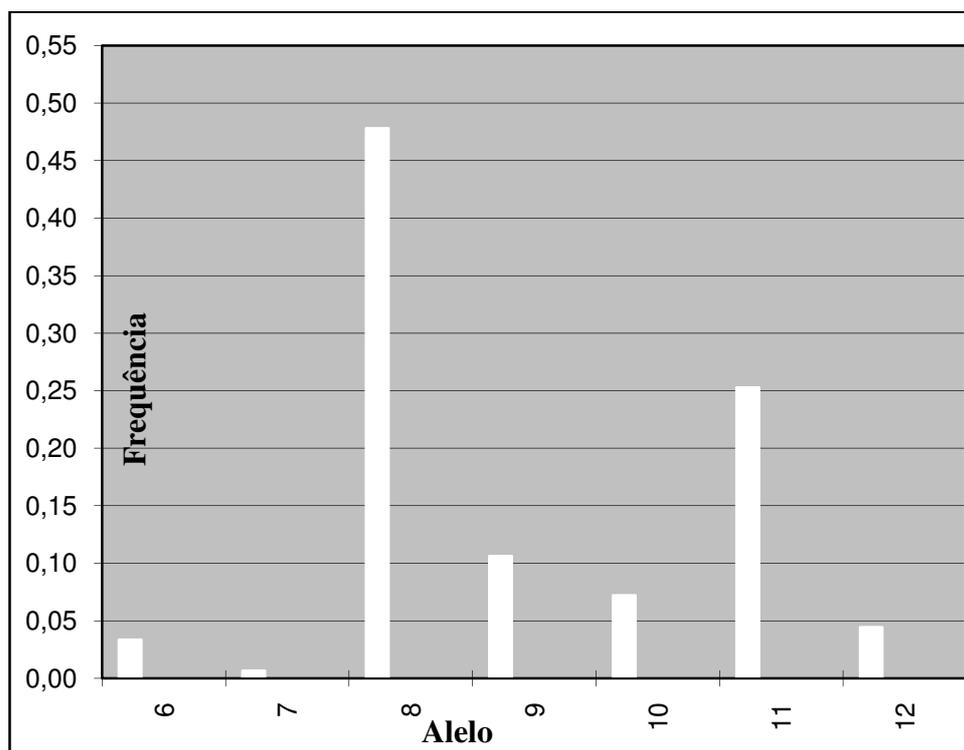


Gráfico 12: Frequência alélica do loci TPOX.

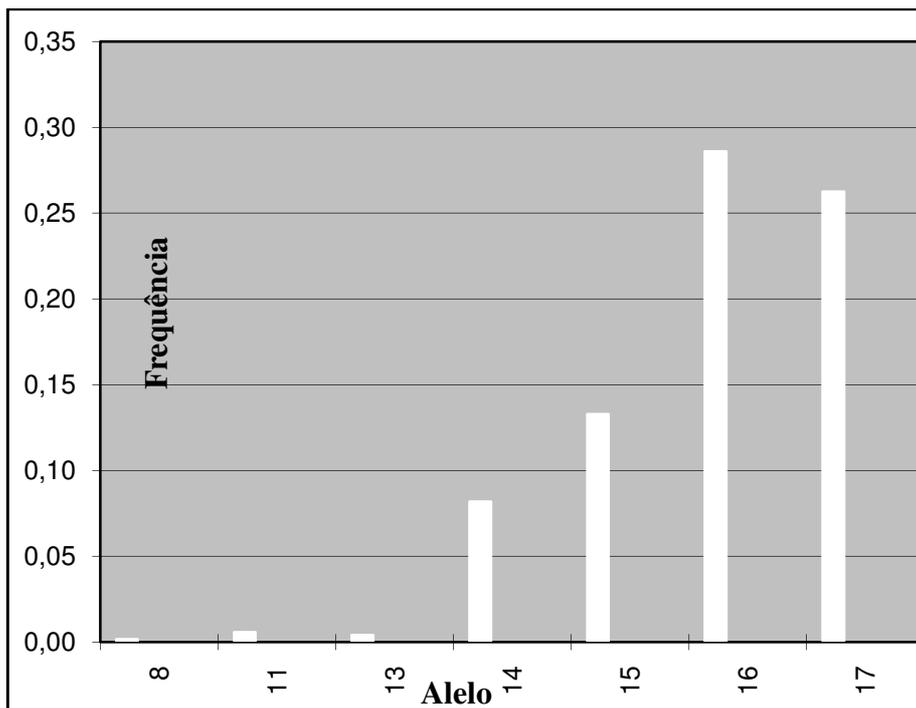


Gráfico 13: Frequência alélica do *loci vWA*.

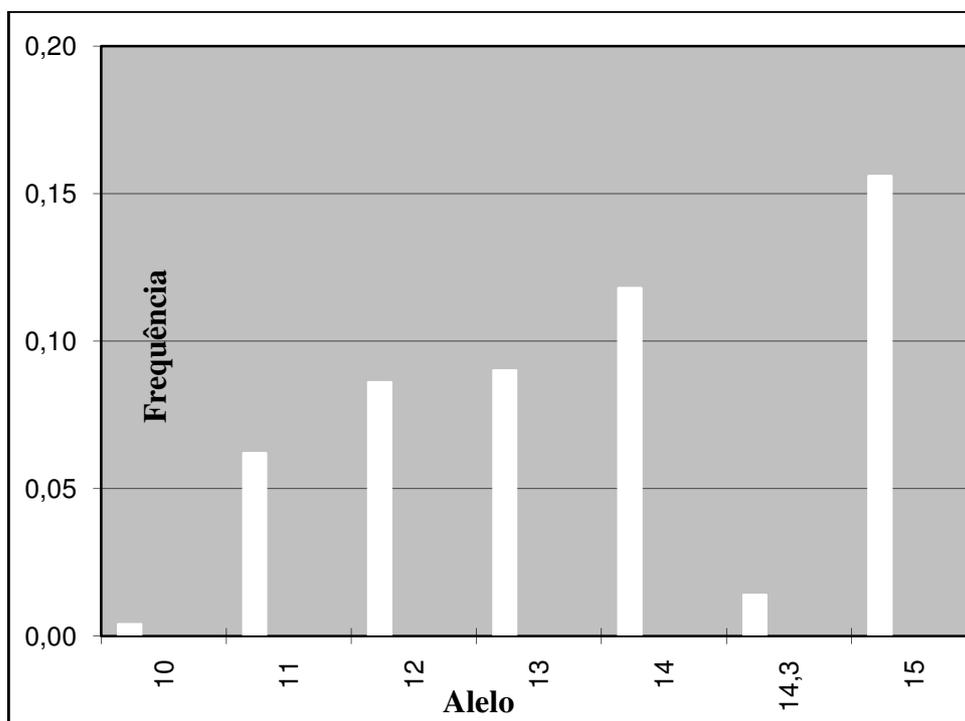


Gráfico 14: Frequência alélica do *loci D1S1656*.

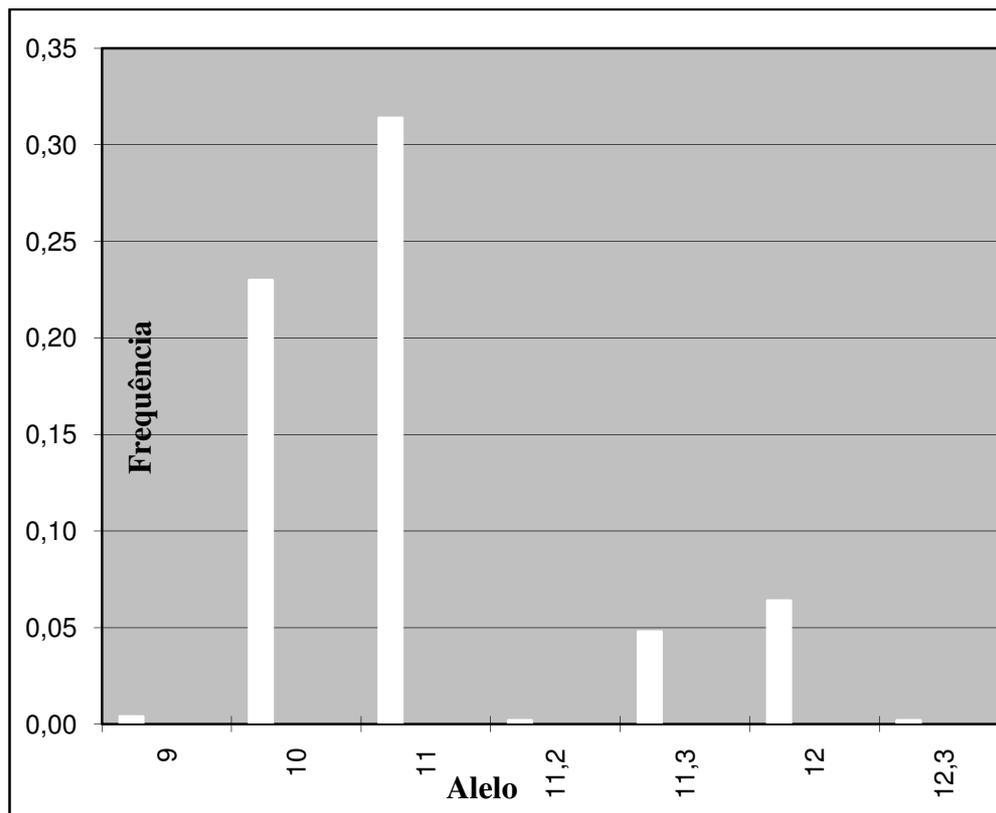
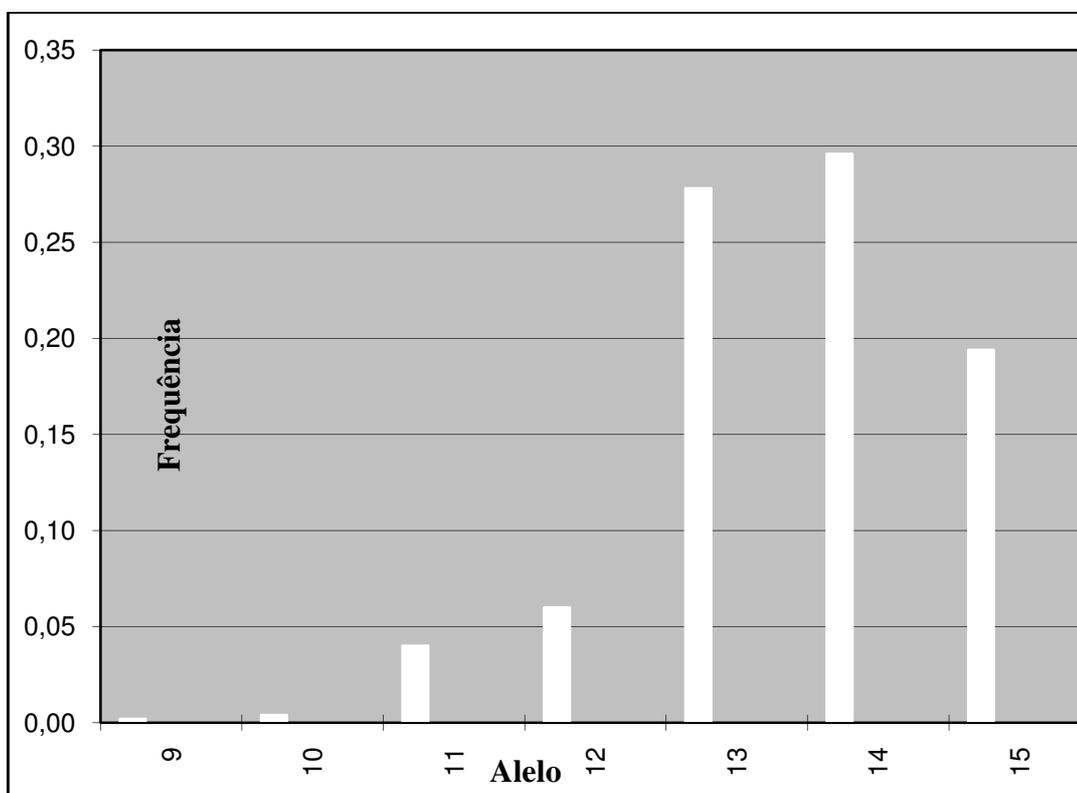


Gráfico 15: Frequência alélica do *loci* D2S441.



16: Frequência alélica do *loci* D10S1246.

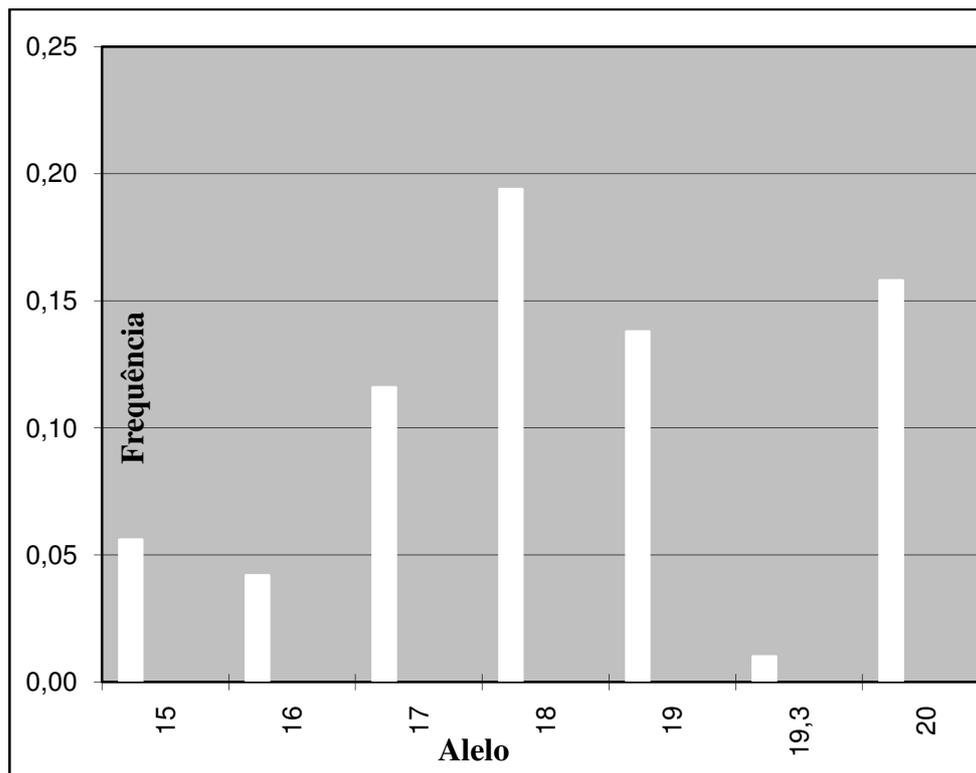


Gráfico 17: Frequência alélica do loci D12S391.

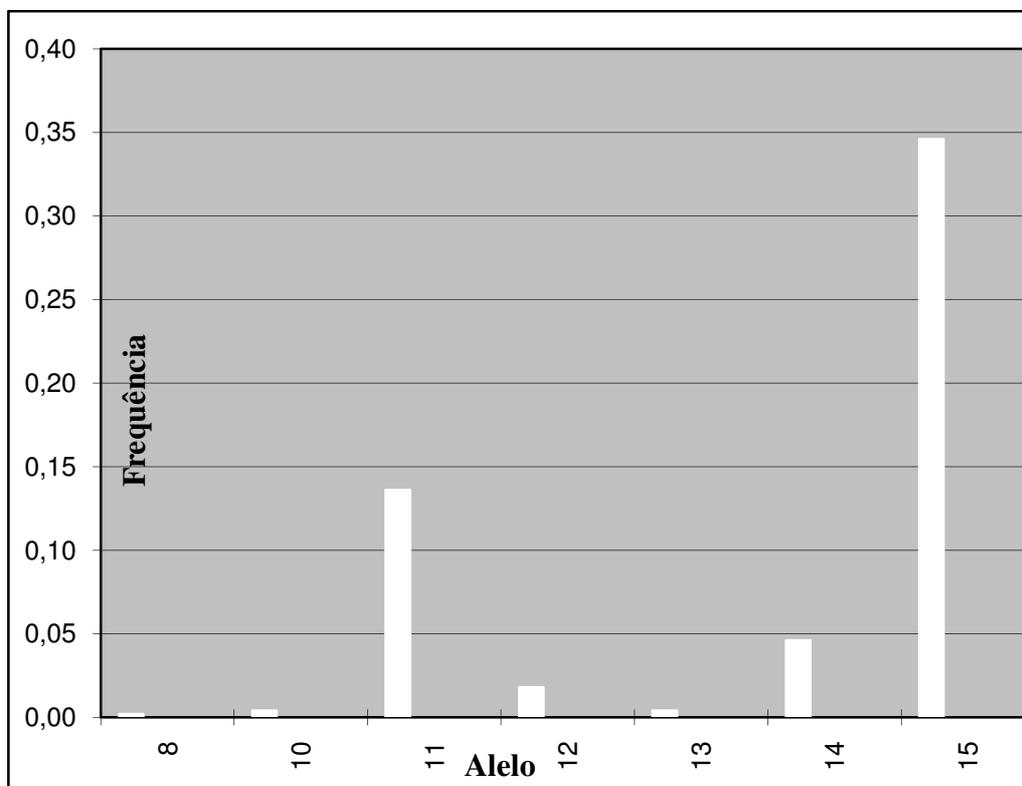


Gráfico 18: Frequência alélica do loci D22S1045.

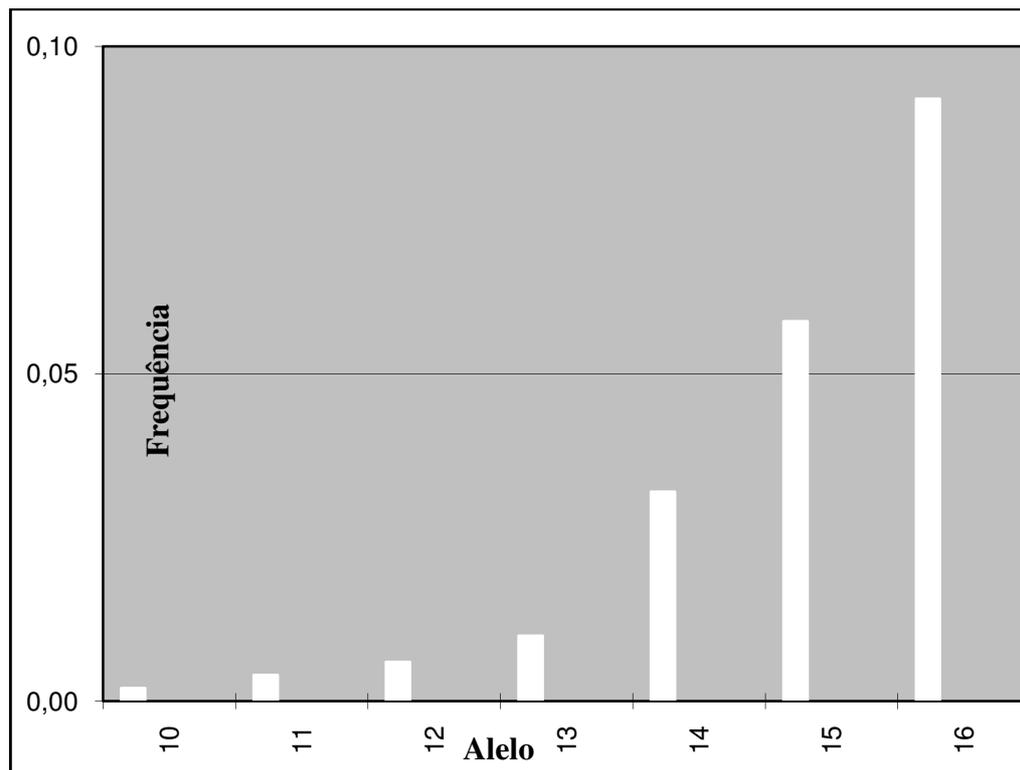


Gráfico 19: Frequência alélica do *loci* SE33.

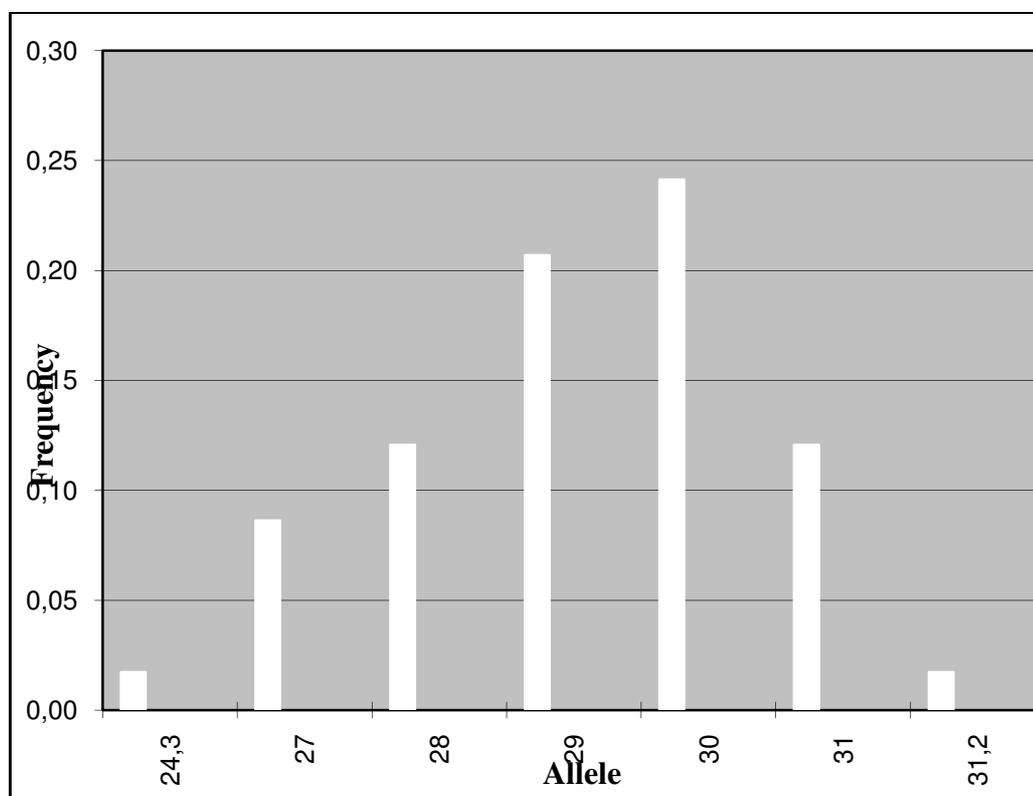


Gráfico 20: Frequência alélica do *loci* D21S11.

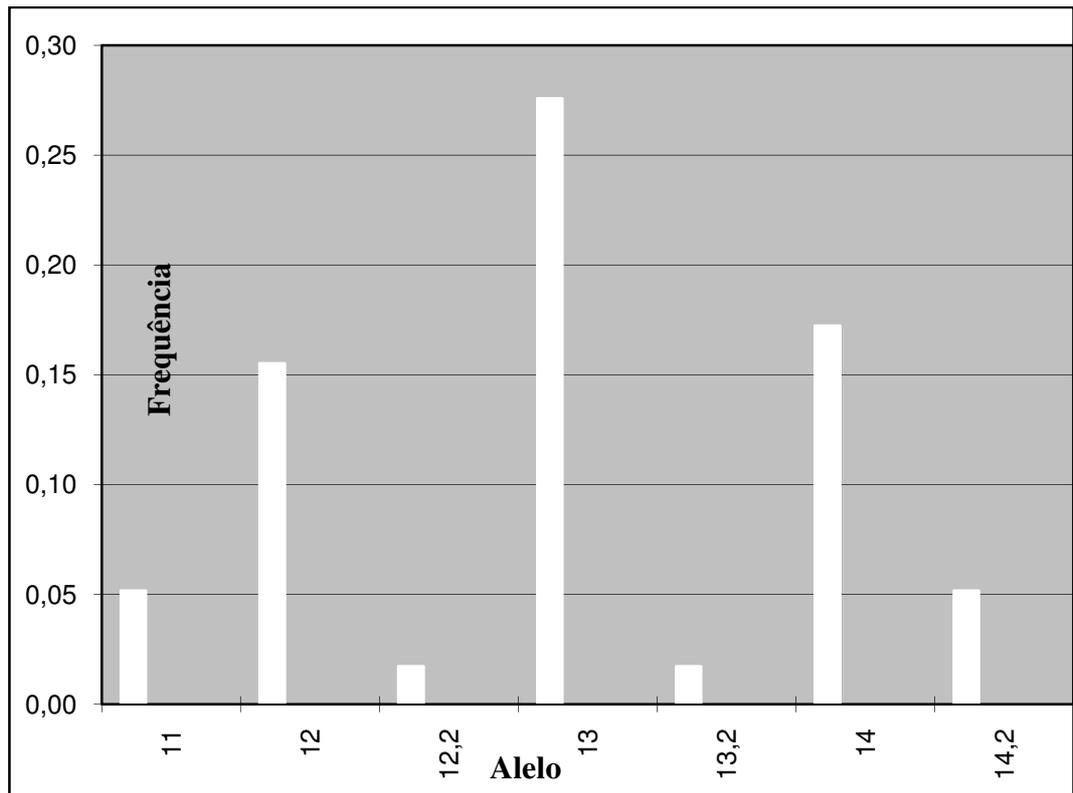


Gráfico 20: Frequência alélica do *loci* D19S443.

# ***CONSIDERAÇÕES***

## ***FINAIS***

---

## 6 CONSIDERAÇÕES FINAIS

---

Os resultados obtidos neste trabalho indicam a importância do estudo contínuo de frequências alélicas de marcadores de interesse forense, adaptando e atualizando frequentemente informações que reforcem a prova pericial do DNA, amplamente utilizada em investigações criminais, objetivando a identificação humana.

O presente estudo disponibiliza informações inéditas para a população de Pernambuco relacionada à frequência alélica dos STRs autossômicos e dos STRs do Cromossomo X, incluindo dados da população de Fernando de Noronha, nunca antes analisada. É importante salientar que devido à aplicação forense da pesquisa, a divisão territorial adotada baseou-se no Programa de Governo Pacto pela Vida.

Os vinte e um STRs autossômicos (*D19S433, D2S1338, D21S11, TH01, D3S1358, FGA, TPOX, D8S1179, D16S539, D7S820, D13S317, D5S818, vWA, D18S51, CSF1PO, D2S441, D22S1045, SE33, D10S1248, D1S1656 e D12S391*) estudados estão em Equilíbrio de Hardy-Weinberg, apresentam elevado Poder de Discriminação e Exclusão, estando suas frequências alélicas disponíveis e aptas para utilização na rotina forense local. Pela primeira vez os marcadores STR's *D1S1656, D2S441, D10S1248, D12S391, D22S1045* foram analisados a partir de uma amostra significativa para o estado de Pernambuco, além do marcador *SE33* cujos dados disponibilizados para a população pernambucana são inéditos.

As informações obtidas a partir do estudo de doze X-STRs (*DXS7132, DXS7423, DXS8378, DXS10074, DXS10079, DXS10101, DXS10103, DXS10134, DXS10135, DXS10146, DXS10148 and HPR1B*) também estabelecem que estes marcadores podem ser aplicados na rotina forense, visto que estão em Equilíbrio de Hardy-Weinberg, há estudos de desequilíbrio de ligação e apresentam elevado Poder de Discriminação e Exclusão. Para a maioria dos marcadores X-STRs aqui apresentados, não há estudos anteriores para a população de Pernambuco.

Para o Brasil, a necessidade de adequar-se a atuais recomendações quanto ao tipo e ao número de marcadores ideais para identificar pessoas, bem como para alimentar bancos de dados de perfis genéticos, tornou-se imprescindível. Portanto, os dados aqui apresentados destacam a importância deste trabalho, cuja aplicabilidade coloca o Laboratório de Perícia e Pesquisa em Genética Forense do Estado de Pernambuco entre os mais modernos do mundo.

# ***PERSPECTIVAS***

---

## 7 PERSPECTIVAS

---

Novos estudos de frequências alélicas que sirvam para implementar os bancos de dados já existentes devem ser uma constante, visto que outros marcadores continuarão a ser referendados como importantes para a prática forense.

Encontra-se em andamento projeto de determinação do perfil genético, através de marcadores autossômicos e de linhagem, por meio de diferentes técnicas e desenvolvimento de estratégias metodológicas de inferências biogeográficas e fisionômicas a partir do DNA de populações do Nordeste para utilização em Genética Forense, que viabilizará novos estudos e formação de recursos humanos, especialmente aplicáveis às investigações criminais.

Estudos que objetivam o desenvolvimento da Genética Forense são significativamente importantes para a humanidade. No Brasil, cujos índices de conclusão de inquéritos com determinação de autoria, são extremamente baixos, toda e qualquer contribuição assume papel de destaque perante uma sociedade carente de justiça.

O potencial é imenso e cabe aos pesquisadores e peritos criminais sugerir inovações, estimulando e aguçando o detetive que existe em cada um de nós.

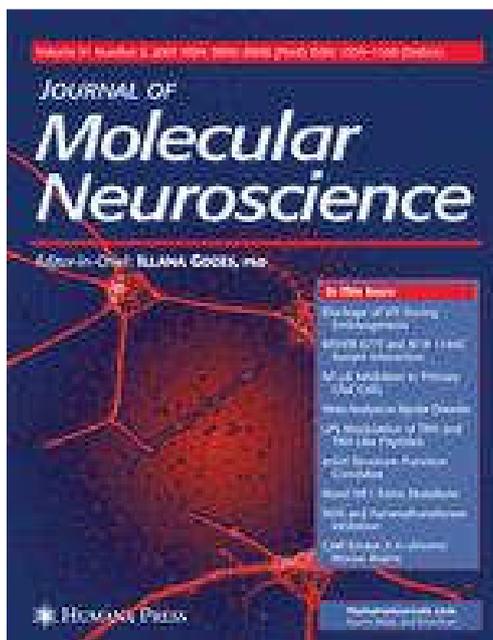
# ***ANEXOS***

## ANEXO I –ARTIGO V

---

First report of a de novo mutation at *slc20a2* in a patient with brain calcification.

Publicado no periódico científico *Journal Molecular Neuroscience* em  
27/06/2014.



J Mol Neurosci  
DOI 10.1007/s12031-014-0357-9

## First Report of a De Novo Mutation at *SLC20A2* in a Patient with Brain Calcification

J. B. Ferreira · L. Pimentel · M. P. Keasey · R. R. Lemos · L. M. Santos ·  
M. F. Oliveira · S. Santos · N. Jensen · K. Teixeira · L. Pedersen ·  
C. R. Rocha · M. R. Dias da Silva · J. R. M. Oliveira

Received: 3 June 2014 / Accepted: 12 June 2014  
© Springer Science+Business Media New York 2014

**Abstract** Primary familial brain calcification (PFBC) is identified by mineralization of the basal ganglia and other brain regions in the absence of known causes. The condition is often inherited in an autosomal dominant pattern and can manifest itself clinically with neuropsychiatric symptoms such as

Parkinsonism, headaches, psychosis, and mood swings. Mutations in the *SLC20A2* gene account for ~40 % of inherited cases, and this gene encodes an inorganic phosphate transporter (PiT-2), a transmembrane protein associated with Pi homeostasis. The p.Y386X mutation in *SLC20A2* was identified in a patient who presented migraines, brain calcification, and mild but chronic hypovitaminosis D. *SLC20A2* c.1158C>G single-nucleotide heterozygous mutation results in a premature stop codon and a putative truncated protein of 385 amino acids. Proband parents do not present the mutation, which is also not present in major public SNP databases, suggesting a de novo sporadic trait. This study describes for the first time a de novo *SLC20A2* mutation in a PFBC patient with migraine and mild hypovitaminosis D. This data further reinforces the pathogenic role of *SLC20A2* mutations as causal factors in PFBC physiopathology.

J. B. Ferreira · L. Pimentel · J. R. M. Oliveira (✉)  
Biological Sciences Graduate Program, Universidade Federal de Pernambuco (UFPE), Recife, Brazil  
e-mail: joao.ricardo@ufpe.br

J. B. Ferreira · L. Pimentel · M. P. Keasey · R. R. Lemos ·  
K. Teixeira · C. R. Rocha · J. R. M. Oliveira  
Keizo Asami Laboratory (LIKA), Universidade Federal de Pernambuco (UFPE), Recife, Brazil

L. M. Santos · M. R. Dias da Silva  
Laboratory of Molecular and Translational Endocrinology,  
Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil

M. F. Oliveira  
Neurosurgery Residency Program, Hospital do Servidor Público  
Estadual de São Paulo, São Paulo, SP, Brazil

S. Santos  
Forensics Lab-Social Defense Secretary, Governo do Estado de  
Pernambuco, Recife, Brazil

J. R. M. Oliveira  
Neuropsychiatry Department, Universidade Federal de Pernambuco,  
Av. Prof. Moraes Rego, 1235 - Cidade Universitária, Recife, PE CEP  
50670-901, Brazil

N. Jensen · L. Pedersen  
Department of Molecular Biology and Genetics, Aarhus University,  
Aarhus, Denmark

N. Jensen · L. Pedersen  
Department of Hematology, Aarhus University, Aarhus, Denmark

N. Jensen · L. Pedersen  
Department of Clinical Medicine, Aarhus University, Aarhus,  
Denmark

**Keywords** *SLC20A2* · Brain calcification · Vitamin D ·  
Inorganic phosphate · De novo mutation

### Introduction

Calcification of the brain can occur secondarily to other conditions such as infections (e.g., neurocysticercosis) and hormonal and biochemical imbalances. However, the terms “idiopathic basal ganglia calcification” or “Fahr’s disease” were often used as synonyms to refer to cases without an apparent etiology. More recently, a new nomenclature suggests “primary familial brain calcification” (PFBC) as more appropriate for diagnostic purposes (Sobrido et al. 2013).

Common clinical manifestations include Parkinsonism, psychosis, dementia, and headaches, with an autosomal dominant pattern of inheritance for calcifications in the basal ganglia, thalamus, and dentate nuclei. In 2012, the first mutations associated with PFBC were reported in *SLC20A2*

(chromosome 8) encoding the ubiquitously expressed inorganic phosphate transporter PiT-2. Currently, almost 40 *SLC20A2* mutations have been associated with PFBC (Böttger and Pedersen 2002; Wang et al. 2012; Nicolas et al. 2014b).

More recently, mutations in platelet-derived growth factor (*PDGFB*) and its receptor *PDGFRB* genes have been found in multiple PFBC patients. This data suggests that several pathways might lead to a similar phenotype (Keller et al. 2013; Nicolas et al. 2013).

Here, we report the discovery of the first de novo *SLC20A2* mutation in a Brazilian patient with brain calcification found during the investigation of a chronic migraine.

## Methods

### Subjects and Samples

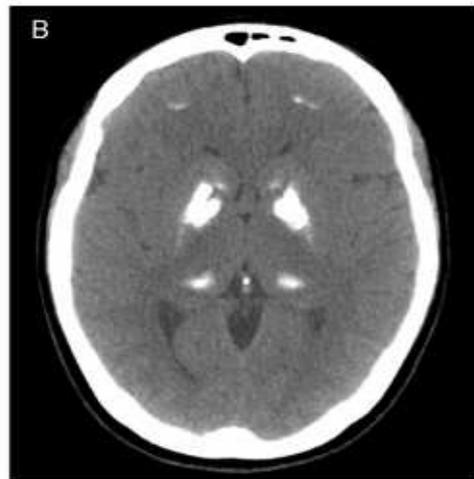
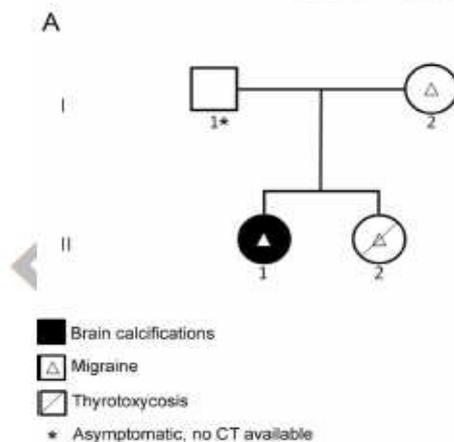
An index case presenting chronic headaches was initially investigated with full blood and urine screening as well as neuroimaging analysis. Screening was also performed in three of the patient's relatives. CT scans were available only for subjects I-2, II-1, and II-2. Peripheral blood samples were collected from four subjects: index case and three relatives (Fig. 1a). A skin biopsy was also performed on the proband for DNA and RNA analysis. This project was approved by the UFPE Ethics Committee with signed informed consent forms from each subject.

### DNA Sequencing

Genomic DNA from blood and skin was isolated by salting-out protocol. PCR amplicons for *SLC20A2* exons were generated from blood DNA and sequenced by the Sanger method, using primers and cycle conditions described previously (Wang et al. 2012). Exon variants were analyzed using CLC Main Workbench software (Qiagen, USA). The identified mutation was verified by PCR and sequencing with five replicates in the proband and two for each relative. To confirm the biological relationship between the parents and proband, DNA was amplified and short tandem repeat (STR) analysis was performed with "Identifiler® Plus" (Life Technology, USA) and "PowerPlex® 16 HS" systems (Promega, USA) using 18 loci.

### RNA Analysis

Quantitative real-time PCR (qRT-PCR) was performed to quantify *SLC20A2* messenger RNA (mRNA) expression in samples using the  $\Delta\Delta CT$  method. Total RNA was isolated from blood using PAXgene Blood RNA System (Qiagen, #762174, Switzerland). Reverse transcription was performed for each sample using the RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific Bio, #K1632, USA) in a 20- $\mu$ L reaction containing 500 ng of total RNA. Real-time reactions were as follows: 1  $\mu$ L complementary DNA (cDNA), 3.5  $\mu$ L H<sub>2</sub>O, and 5  $\mu$ L TaqMan® Gene Expression Master Mix (Applied Biosystems, #4369510, Foster City, CA) with 0.5  $\mu$ L of 20 $\times$  TaqMan Gene Expression Assay (Hs\_00198840\_m1; Applied



**Fig. 1** Pedigree analysis and positive CT scan of index case with migraine and hypovitaminosis D. **a** Pedigree of the index case indicating the subjects with brain calcifications (only II-1) and those with migraine (I-2,

II-1, and II-2). II-2 also presents thyrotoxicosis. **b** Horizontal CT scan slice showing spots of calcification in the basal ganglia for the proband

Biosystems) or 20× TaqMan® Endogenous Control (Hs\_02758991\_g1; Applied Biosystems) in a 10-μL reaction. Real-time PCR reactions were run in triplicate and run using a 7500 Fast Real-Time PCR system (Applied Biosystems).

PCR-amplified cDNA fragments were then sequenced with automatic Sanger sequencing using the primers as follows: forward (5'-TCT GCT GCA CAA AAT CCA CA-3') and reverse (5'-ACG GGT GTA GCT GCT TCT TG-3').

## Results

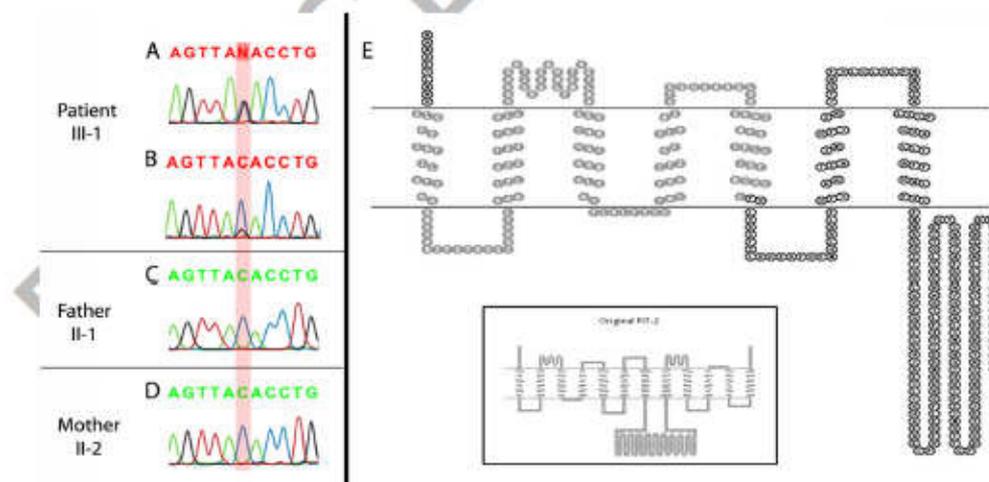
### Clinical Features of Patient with *SLC20A2* De Novo Mutation

A 33-year-old female (mother of two children, employed as a real estate manager) was under investigation for chronic migraine, with mild functional impairment that had progressed recently. The chronic migraine had characteristically a pulsating pattern, was sensitive to light, and sometimes during over 24 h. This initially took her to an ER, and a neurological follow-up detected the calcifications. Various medications were used to alleviate this intermittent symptom. The patient also referred to regular menses not linked to a regular menstrual cycle. No neurological deficits were observed on physical examination. A CT scan revealed bilateral calcifications in the *globus pallidus* and pulvinar region of the thalamus and less intensely in the dentate nucleus. Sparse calcifications also presented in the white matter of the cerebellum and frontal subcortical regions (Fig. 1b). Hemoglobin, ceruloplasmin, and serological testing for toxoplasmosis, CMV, HIV, and hepatorenal biochemical markers were normal. Adenohypophysial hormones such as prolactin, GH,

TSH, FSH, and LH, together with cortisol, glucose, and lipids, were all within normal levels. Biochemical and endocrinological tests revealed only mild hypovitaminosis D (16 nmol/L of 25 hydroxyvitamin D [25(OH)D]). A supplementation was prescribed, but no change was detected at serum level follow-up (15 nmol/L of 25(OH)D). The proband reported a familial history of dementia (mother and maternal aunt) and thyrotoxicosis (Graves' disease) in a younger sister. Neuroimaging analysis of both mother and sister showed no calcification (data not shown).

### Screening and Validation of a De Novo *SLC20A2* Mutation

We identified the new mutation p.Y386X in exon 10 (previously exon 8) of the *SLC20A2* gene (NM\_006749.4:c.1158C>G; NP\_006740.1), which results in a premature stop codon and a putative truncated protein with 385 amino acids instead of 652 (Fig. 2). This mutation was not observed in the online databases *EVS* and *1000 Genomes* or in 55 controls from the *Inserm/Ruen exome* databases. Three relatives of the patient (two generations) were screened for the mutation, all of which showed negative (Figs. 1a and 2c, d). The biological relationship between the proband and parents was confirmed by multiplex STR analyses (data not shown). Since somatic mosaicism is common in subjects with rare disorders and de novo mutations, we screened another tissue from a different embryonic origin, which also contained the mutation. We also analyzed mRNA expression in RNA extracted from peripheral blood of the proband by RT-qPCR. *SLC20A2* expression was decreased by ~10 % (data not shown) in the proband. cDNA sequence reads show a reduced peak



**Fig. 2** De novo mutation in *SLC20A2* gene. **a, b** Electropherograms showing a *SLC20A2* de novo mutation in the patient (NM\_006749.4:c.1158C>G), both in genomic DNA (**a**) and cDNA (**b**). This variant is absent in parents (**c, d**). Schematic diagram of wild-type

and mutant *SLC20A2* proteins highlighting a shorter putative PIT-2 protein due to a stop codon at the 386 amino acid residue (NP\_006740.1: p.Y386X) (**e**). Both models were generated by TOPO2 online software (<http://www.secs.ucsf.edu/TOPO2/>)

correspondent to the mutant allele (Fig. 2b), comparatively with the genomic DNA (Fig. 2a).

### Discussion

This is the first report of a *de novo* mutation in the *SLC20A2* gene, reinforcing the causative role of pathogenic variations of this gene in PFBC.

Curiously, this patient also presents mild but chronic hypovitaminosis D. In 2012, the patient was diagnosed with hypovitaminosis D, and despite prescription of supplements, there was no change (parathyroid hormone was within the normal range in both instances). Chronic hypovitaminosis D is influenced by several factors, i.e., age, skin tone, latitude, sun exposure, and season of the year. In addition, hypovitaminosis D is a common problem in the city of São Paulo (Maeda et al. 2013). Therefore, without further evidence, it was not possible to confirm if the hypovitaminosis D was a comorbidity or a new phenotype.

Analysis of *SLC20A2* mRNA expression demonstrated only a small decrease in *SLC20A2* expression (~10 % decrease relative to controls) despite the proband carrying a premature transcription termination signal (stop codon). Zhang et al. (2013) performed similar experiments and described a decrease by ~30 % in *SLC20A2* expression in patients carrying the mutation. This might be due to compensation by the wild-type allele triggered by allelic loss. In our study, a stable expression of mutant *SLC20A2* mRNA or upregulation of the wild-type allele could account, on one hand, for the small differences we observed in our RT-qPCR. On the other hand, cDNA sequence reads suggest lower expression of the mutant allele (Fig. 2b), comparatively with the genomic DNA (Fig. 2a), suggesting an RNA-mediated decay, typically found in stop codon carriers.

*SLC20A2* and other genes associated with PFBC should also be screened for mutations in patients with non-idiopathic brain calcifications. The presence of hypovitaminosis D in the proband likely might exist as a comorbidity due to the absence of this condition in previous PFBC patients with *SLC20A2* mutations. However, it is necessary to document this finding and to be aware of occasional future reports.

Nicolas et al. (2014a) recently reported a *de novo* mutation in *PDGFB*, the second most frequent mutated gene in PFBC, in a patient with brain calcification and laryngeal dystonia. Together with our report, such findings reinforce the need to

screen for mutations linked to PFBC even in cases with sporadic and atypical phenotype.

**Acknowledgments** We thank Ithgu Lucena and Dr. Juliana Machado for technical assistance with sample processing. FACEPE (APQ 1831-4.01/12) and CNPq (457556/2013-7;307909/2012-3;400540/2013-4) for the financial support; and Gaël Nicolas (Ruen University), Paula Sandrin, and Sérgio Crovella. J.B.F. and L.M.S. hold FACEPE and CAPES fellowships, respectively.

**Conflict of interest** S.S., L.P., C.R., M.R.D.S., L.S., and J.R.M.O. work in public institution and did not receive any incentive from private companies. J.R.M.O. is the author of a book about familial brain calcification, edited by Nova Publishing (USA). J.F., L.P., M. K., R.R.L., M.F.O., N.J., and K.T. receive fellowships from graduate, undergraduate and post-graduate programs. The authors report no other disclosure and no conflict of interest directly linked to this report.

### References

- Böttger P, Pedersen L (2002) Two highly conserved glutamate residues critical for type III sodium-dependent phosphate transport revealed by uncoupling transport function from retroviral receptor function. *J Biol Chem* 277:42741–42747
- Keller A, Westenberger A, Sobrido MJ et al (2013) Mutations in the gene encoding PDGF-B cause brain calcifications in humans and mice. *Nat Genet* 45:1077–1082
- Maeda SS, Saraiva GL, Kuni IS et al (2013) Factors affecting vitamin D status in different populations in the city of São Paulo, Brazil: the São Paulo vitamin D Evaluation Study (SPADES). *BMC Endocr Disord* 13:14
- Nicolas G, Pottier C, Maltête D et al (2013) Mutation of the PDGFRB gene as a cause of idiopathic basal ganglia calcification. *Neurology* 80:181–187
- Nicolas G, Rovelet-Lecrux A, Pottier C et al (2014a) PDGFB partial deletion: a new, rare mechanism causing brain calcification with leukoencephalopathy. *J Mol Neurosci* 53:171–175
- Nicolas G, Richard AC, Pottier C, Vermy C, Durif F, Roze E, Favrole P, Rudolf G, Anheim M, Tranchant C, Frebourg T, Campion D, Hanquinn D (2014b) Overall mutational spectrum of *SLC20A2*, *PDGFB* and *PDGFRB* in idiopathic basal ganglia calcification. *Neurogenetics*. doi:10.1007/s10048-014-0404-2
- Sobrido MJ, Copolla G, Oliveira JRM, Hopfer S, Geschwind DH (2013) Familial primary brain calcification. In P. RA, B. TD, & D. CR (Eds.), *Gene reviews™*, (2014 ed.). University of Washington, Seattle
- Wang C, Li Y, Shi L et al (2012) Mutations in *SLC20A2* link familial idiopathic basal ganglia calcification with phosphate homeostasis. *Nat Genet* 44:254–256
- Zhang Y, Guo X, Wu A (2013) Association between a novel mutation in *SLC20A2* and familial idiopathic basal ganglia calcification. *PLoS One* 8:e57060

**ANEXO II - CONVITE/INFORMAÇÕES AO DOADOR/TCLE**

	UNIVERSIDADE FEDERAL DE PERNAMBUCO LABORATÓRIO DE IMUNOPATOLOGIA KEIZO ASAMI PÓS-GRADUAÇÃO EM BIOLOGIA APLICADA À SAÚDE	
---	---	---

**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO****IDENTIFICAÇÃO HUMANA COMO FERRAMENTA DE INVESTIGAÇÕES CRIMINAIS: ESTUDO DE FREQUÊNCIAS ALÉLICAS DE MARCADORES DE INTERESSE FORENSE NO ESTADO DE PERNAMBUCO**

Convidamos o(a) Sr.(a) para participar, como voluntário(a), da pesquisa **“Identificação Humana como Ferramenta de Investigações Criminais: Estudo de Frequências Alélicas de Marcadores de Interesse Forense no Estado de Pernambuco**. Após ser esclarecido(a) sobre as informações a seguir, no caso de aceitar fazer parte do estudo, assine ao final deste documento, que está em duas vias. Uma delas é sua e a outra é do pesquisador responsável. Em caso de recusa você não será penalizado(a) de forma alguma. Em caso de dúvida você pode procurar o Comitê de Ética em Pesquisa Envolvendo Seres Humanos da UFPE no endereço: **(Avenida da Engenharia s/n – 1º Andar, Sla 4 - Cidade Universitária, Recife-PE, CEP: 50740- 600, Tel.: 2126.8588 – e-mail: cepccs@ufpe.br)**.

O Pesquisador Responsável é a doutoranda Sandra Maria dos Santos. **Endereço:** Laboratório de Perícia e Pesquisa em Genética Forense, localizado na Estrada da Batalha, S/N - Prazeres – Jaboatão dos Guararapes – Pernambuco – Telefone: **(81) 3183-5389 – Celular: (9081) 9488-7555. E-mail: [sandra.santos@sds.pe.gov.br](mailto:sandra.santos@sds.pe.gov.br)** e os **Pesquisadores participantes são o Professor da UFPE Dr. Sérgio Crovella (fone: 81 2101-2655), o Professor da UFRPE Dr. Paulo Roberto Eleutério de Souza (81 3320-6313).**

Objetivo da pesquisa é determinar a frequência alélica de 10 STRs do cromossomo X em indivíduos da população brasileira, residentes no Estado de Pernambuco. Este estudo será de grande importância para as investigações criminais que envolvem identificação humana pelo DNA. Para isto pedimos sua **colaboração e permissão** para coletar pequena quantidade de sua mucosa bucal e a partir disto extrair o **DNA**. As regiões do DNA analisadas nesta pesquisa são não codificantes, não permitindo identificações de estado de saúde ou de comportamento humano. Serão observadas, exclusivamente, “frequências alélicas”, não sendo permitida a relação entre o perfil genético e o doador, ou seja, não constarão dados que o identifique como doador da amostra. A amostra coletada nesta ocasião poderá ser utilizada em projetos futuros que envolvam testes genéticos, aprovados pelo sistema CEP/CONEP, desde que receba novamente sua autorização, após um novo contato.

Para sua segurança, garantimos que o material biológico, o DNA extraído e todas as informações coletadas são sigilosas e estarão sob responsabilidade da coordenadora do projeto Sandra Maria dos Santos.

Ao participar do projeto o(a) senhor(a) não corre riscos relacionados à coleta, pois o procedimento é simples, não causa dor ou lesões e todo o material usado (suabe, luvas, toucas e máscaras) são descartáveis. Todas as informações coletadas, bem como os resultados das análises genéticas serão mantidas em sigilo e serão utilizadas somente pela equipe da pesquisa. Também não há nenhum risco quanto a sua privacidade e muito menos que o relacione a investigações policiais, pois toda e qualquer informação utilizada a partir do estudo das frequências alélicas não terá nenhuma vinculação com a sua pessoa.

O(a) senhor(a) não pagará para fazer parte deste estudo e poderá cancelar a sua participação a qualquer momento, sem precisar justificar-se.

Esclarecemos que o(a) senhor(a) não terá nenhum benefício direto ao participar desta pesquisa, todavia, saiba que os resultados beneficiarão os cidadãos pernambucanos envolvidos em processos civis ou criminais em busca da garantia de seus direitos perante a lei.

### **Pesquisador Responsável**

---

**Sandra Maria dos Santos**

### **CONSENTIMENTO DA PARTICIPAÇÃO DA PESSOA COMO SUJEITO**

Eu, \_\_\_\_\_, RG/  
CPF/\_\_\_\_\_, abaixo assinado, concordo em participar do estudo  
**“Identificação Humana como Ferramenta de Investigações Criminais: Estudo de Frequências Alélicas de Marcadores de Interesse Forense no Estado de Pernambuco**, como sujeito. Fui devidamente informado(a) e esclarecido(a) pelo(a) pesquisador(a) Sandra Maria dos Santos sobre a pesquisa, os procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes de minha participação. Foi-me garantido que posso retirar meu consentimento a qualquer momento, sem que isto leve a qualquer penalidade.

\_\_\_\_\_, \_\_\_\_\_ de \_\_\_\_\_ de 201\_\_\_\_.

---

Doador

**Presenciamos a solicitação de consentimento, esclarecimentos sobre a pesquisa e aceite do sujeito em participar.**

Nome: \_\_\_\_\_

Assinatura: \_\_\_\_\_

Nome: \_\_\_\_\_

Assinatura: \_\_\_\_\_