Universidade do Estado do Rio de Janeiro
Centro Biomédico
Faculdade de Ciências Médicas

Juliana Gozi de Aquino

Desenvolvimento e validação de sistema multiplex X-STR para identificação humana por DNA

Rio de Janeiro
2011

Juliana Gozi de Aquino

Desenvolvimento e validação de sistema multiplex X-STR para identificação humana por DNA

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

Orientadora: Prof[a]. Dra. Dayse Aparecida da Silva

Coorientador: Prof. Dr. Elizeu Fagundes de Carvalho

Rio de Janeiro

2011

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

A657 Aquino, Juliana Gozi de.
Desenvolvimento e validação de sistema multiplex X-STR para identificação humana por DNA / Juliana Gozi de Aquino. - 2011.
136 f.: il.

Orientadores: Dayse Aparecida da Silva, Elizeu Fagundes de Carvalho
Dissertação (Mestrado) – Universidade do Estado do Rio de Janeiro, Faculdade de Ciências Médicas. Pós-graduação em Ciências Médicas.

1. Genética forense - Teses. 2. Antropologia forense - Teses. 3. DNA - Teses. 4. Genética da população humana – Teses. 5. Técnicas de diagnóstico molecular. I. Silva, Dayse Aparecida da. II. Carvalho, Elizeu Fagundes de. . III. Universidade do Estado do Rio de Janeiro. Faculdade de Ciências Médicas. IV. Título.

CDU 575.113.1

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

_____ _____

Assinatura Data

Juliana Gozi de Aquino

Desenvolvimento e validação de sistema multiplex X-STR para identificação humana por DNA

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

Aprovada em 24 de maio de 2011.

Orientadores:

_____
Profª. Dra. Dayse Aparecida da Silva (Orientadora)
Universidade do Estado do Rio de Janeiro - UERJ

_____
Prof. Dr. Elizeu Fagundes de Carvalho (Coorientador)
Universidade do Estado do Rio de Janeiro - UERJ

Banca Examinadora:

_____
Prof. Dr. Luís Cristóvão de Moraes Sobrino Porto
Universidade do Estado do Rio de Janeiro - UERJ

_____
Profª. Dra. Márcia Mattos Gonçalves Pimentel
Universidade do Estado do Rio de Janeiro - UERJ

_____
Prof. Dr. Celso Teixeira Mendes Junior
Universidade de São Paulo - USP

Rio de Janeiro

2011

DEDICATÓRIA

Dedico este trabalho...

Primeiramente a Deus, por me guiar em todos os momentos da vida.
Aos meus pais, Telmo e Rosangela, pelo amor incondicional, esforço e incentivo.
Aos meus irmãos, Fabiano e Fabio, pelo apoio e carinho.
Ao meu sobrinho, João Vitor, que com apenas 10 anos de vida muito nos ensina.
Ao meu grande amor, Vitor, pelo amor, carinho, incentivo, crescimento e paciência nos momentos difíceis.

AGRADECIMENTOS

A minha orientadora Dr^{a}. Dayse Aparecida da Silva, pelos conhecimentos e ensinamentos transmitidos necessários para o desenvolvimento deste trabalho.

Ao meu orientador Elizeu Fagundes de Carvalho pela orientação e grande oportunidade de fazer parte de sua equipe.

Ao professor Celso Cruz Tavares pelo incentivo e orientação desde a minha graduação.

Aos amigos Fê, Su, Cris, Paty, Esther, João e Alexandre por compartilhar conhecimentos técnicos, momentos felizes e estressantes ao longo destes anos. Muito obrigada!

A Monique, Saulo, Silvia, Rossana, Rose e Márcia pela ajuda técnica. E ao Rui Pereira do IPATIMUP, pela atenção e pelos valiosos ensinamentos passados.

A Andréa, Marcinha, Helô, Jaque, Flávio, Agda, Henrique, Rosângela, Marcos e Ednardo por toda ajuda e colaboração obtida nas mais diversas etapas deste trabalho.

*“Uma vida sem desafios não vale a pena ser vivida” (*Sócrates)

“A atenção é a mais importante de todas as faculdades para o desenvolvimento da inteligência humana” (Charles Darwin)

RESUMO

AQUINO, Juliana Gozi. D*esenvolvimento e validação de sistema multiplex X-STR para identificação humana por DNA*. 2011. 135 f . Dissertação (Mestrado em Ciências Médicas) – Faculdade de Ciências Médicas, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2011.

Novas metodologias de análise molecular voltadas para estudos populacionais, clínicos, evolutivos, da biodiversidade e identificação forense foram desenvolvidas com base em marcadores microssátelites ou STR – "Short Tandem Repeats". Os marcadores STR, que estão amplamente espalhados nos genomas e se caracterizam por apresentar alto grau de polimorfismo, podem ser analisados a partir da amplificação por PCR (Reação em Cadeia da polimerase). A análise foi facilitada a partir do desenvolvimento de sistemas de amplificação simultânea de múltiplos STR (multiplex STR) e com a detecção automatizada dos produtos de amplificação marcados por fluorescência. Recentemente, o uso de marcadores STR do cromossomo X (X-STR) tornou-se significativo na prática forense. Devido ao seu modo de transmissão, os X-STR são úteis em situações particulares de investigação de relações de parentesco, apresentando vantagens sobre o uso de STR autossômicos. Este estudo teve como principal objetivo o desenvolvimento e validação de sistema multiplex, denominado LDD (X-STR) Decaplex, capaz de amplificar dez *loci* X-STR (DXS7133, DXS7424, DXS8378, DXS6807, DXS7132, DXS10074, DXS7423, DXS8377, GATA172D05 e DXS10101) para aplicação em genética populacional, identificação e análises forenses. Utilizando o LDD (X-STR) Decaplex 170 indivíduos autodenominados afrodescendentes, não aparentados geneticamente, foram genotipados. As freqüências alélicas e genotípicas não apresentaram desvio do equilíbrio de Hardy-Weinberg e estão em concordância com aquelas observadas em outros estudos. Os haplótipos observados foram únicos em indivíduos de amostra masculina. A análise de desequilíbrio de ligação não revelou associação entre os marcadores X-STR. A diversidade genética foi elevada, variando entre 0,6218 para o *locus* DXS7133 a 0,9327 para o *locus* DXS8377. Os parâmetros de Probabilidade de Vinculação (PV), Índice de Vinculação (IV), Poder de Exclusão (PE), Poder de Discriminação e Razão de Verossimilhança foram também elevados, demonstraram que os dez X-STRs são altamente polimórficos e discriminativos na população estudada. A concentração mínima de DNA para a amplificação dos *loci* do LDD (X-STR) Decaplex é de 0,5 ng e verificamos que amplificação por PCR pode ser afetada quando são adicionados mais de 5 ng de DNA nas reações. Os percentuais de bandas "*stutter*" foram elevados para os *loci* DXS7132 e DXS8377. No teste de reprodutibilidade observamos consistência entre as tipagem de diferentes amostras biológicas, incluindo as de restos mortais. No teste de mistura a proporção limite em que observamos a coexistência de duas espécies biológicas foi de 2,5:1ng (feminino-masculino). Os resultados evidenciaram que os *loci* do LDD (X-STR) Decaplex são altamente informativos, consistindo, em conjunto, uma ferramenta importante em estudos de identificação humana e de relações de parentesco.

Palavras-chave: X-STR. Multiplex. PCR. Identificação humana.

ABSTRACT

New molecular analysis technologies used to evolving clinic and population studies of biodiversity and forensic identification have been developed based on microsatellite markers or STR – “Short Tandem Repeats”. These STR markers, which are widely spread on genomes and characterized by their high degree of polymorphism, can be analyzed by PCR (Polymerase Chain Reaction) amplification technology. This analysis was facilitated from the development of simultaneous amplification system of multiples STR (multiplex STR) and automatic detection of amplification products by fluorescence. Lately, the use of STR markers of chromosome X (X-STR) has become significant in forensic practice. Due to their transmission ways, X-STR are useful at especial kinship investigation, showing advantages in relation to the use of STR autosomal. This study main focus was the development and validation of the multiplex system called LDD (X-STR) Decaplex, which was able to amplify ten X-STR loci (DXS7133, DXS7424, DXS8378, DXS6807, DXS7132, DXS10074, DXS7423, DXS8377, GATA172D05 and DXS10101) in order to be used in population genetics application, identification and forensic analysis. By using LDD (X-STR) Decaplex 170, individuals, who were self-appointed Africandescendants and genetically unrelated, were genotyped. Allele and genotype frequencies have not shown Hardy-Weinberg equilibrium deviation and are in agreement with the ones observed in other studies. The observed haplotypes were unique in male individual samples. The linkage disequilibrium analysis has not shown any association among the X-STR markers. The genetic diversity was high, ranging from 0.6218 (DXS7133 locus) to 0.9327 (DXS8377 locus). The parameters of Probability of Relatedness (PR), Avuncular Index (AI), Power of Exclusion (PE), Power of Discrimination and Likelihood Ratios were also high, showing that these ten X-STRs are highly polymorphous and discriminating on the studied population. The minimum DNA concentration for the amplification of the LDD (X-STR) Decaplex loci is 0.5 ng and it was also verified that the PCR amplification can be affected when more than 5 ng of DNA are added to the reactions. The "stutter" bands percentages were high to DXS7132 and DXS8377 loci. At the reproducing test, consistency among typing of different biological samples was observed, including the ones from remains. At the mixing test, the limit proportion in which two biological species coexisted was 2.5:1 ng (female-male). The results point out that LDD (X-STR) Decaplex loci are able to provide many important piece of information which, as a whole, are a very important tool at human identification and of kinship studies.

Keywords: X-STR. Multiplex. PCR. Human identification

LISTA DE FIGURAS

Figura 1 – Modelo esquemático do genoma nuclear e mitocondrial humano. Adaptado de Butler, 2005..... 24

Figura 2 – Distribuição de marcadores STR ao longo do cromossomo X. Adaptado de www.chrx-str.org..... 30

Figura 3 – Alelos X-STR a partir de DNA controle 9947 utilizando o LDD (X-STR) Decaplex..... 54

Figura 4 – Alelos X-STR amplificados a partir de DNA de um indivíduo afrodescendente, do sexo feminino, do Rio de Janeiro utilizando o LDD (X-STR) Decaplex..... 56

Figura 5 – Alelos X-STR amplificados a partir de DNA de um indivíduo afrodescendente, do sexo masculino, do Rio de Janeiro utilizando o LDD (X-STR) Decaplex..... 56

Figura 6 – Resultado da reação de sequenciamento do fragmento 24.2 do *locus* DXS10101 (orientação 3'-5)..... 62

Figura 7 – Resultado da reação de sequenciamento do fragmento 26 do *locus* DXS10101(orientação 3´-5')..... 63

Figura 8- Perfil alélico gerado pela amplificação de 10ng de DNA 9947A com o sistema LDD (X-STR) Decaplex..... 71

Figura 9 – Perfil alélico gerado pela amplificação de 5ng de DNA 9947A com o sistema LDD (X-STR) Decaplex..... 71

Figura 10– Perfil alélico gerado pela amplificação de 2ng de DNA 9947A com o sistema LDD (X-STR) Decaplex..... 72

Figura 11– Perfil alélico gerado pela amplificação de 1ng de DNA 9947A com o sistema LDD (X-STR) Decaplex..... 72

Figura 12- Perfil alélico gerado pela amplificação de 0,5ng de DNA 9947A com o sistema LDD (X-STR) Decaplex..... 73

Figura 13– Perfil alélico gerado pela amplificação de 0,25ng de DNA 9947A com o sistema LDD (X-STR) Decaplex..... 73

Figura 14– Perfil alélico gerado pela amplificação de 0,125ng de DNA 9947A com o sistema LDD (X-STR) Decaplex..... 74

Figura 15– Perfil alélico gerado pela amplificação de 0,0625ng de DNA 9947A com o sistema LDD (X-STR) Decaplex..... 74

Figura 16- Perfil alélico gerado pela amplificação de 5ng de DNA de dente com o sistema *Identifiler*..... 76

Figura 17– Perfil alélico gerado pela amplificação de 5ng de DNA de dente com o sistema LDD (X-STR) Decaplex..... 76

Figura 18– Perfil alélico gerado pela amplificação de 5ng de DNA de saliva (mucosa oral) com o sistema *Identifiler*..... 77

Figura 19- Perfil alélico gerado pela amplificação de 5ng de DNA de saliva (mucosa oral) com o sistema LDD (X-STR) Decaplex..... 77

Figura 20– Perfil alélico gerado pela amplificação de 5ng de DNA de cabelo (bulbo) com o sistema *Identifiler*..... 78

Figura 21– Perfil alélico gerado pela amplificação de 5ng de DNA de cabelo (bulbo) com o sistema LDD (X-STR) Decaplex..... 78

Figura 22– Perfil alélico gerado pela amplificação de 5ng de DNA de urina com o sistema *Identifiler*..... 79

Figura 23- Perfil alélico gerado pela amplificação de 5ng de DNA de urina com o sistema LDD (X-STR) Decaplex..... 79

Figura 24- Perfil alélico gerado pela amplificação de 5ng de DNA de sangue (FTA) com o sistema *Identifiler*..... 80

Figura 25– Perfil alélico gerado pela amplificação de 5ng de DNA de sangue (FTA) com o sistema LDD (X-STR) Decaplex..... 80

Figura 26– Perfil alélico gerado pela amplificação de 5ng DNA de sêmen com o sistema *Identifiler*..... 81

Figura 27– Perfil alélico gerado pela amplificação de 5ng de DNA de sêmen com o sistema LDD (X-STR) Decaplex 81

Figura 28- Relação entre as intensidades de bandas *stutter* e alelos verdadeiros verificados nos *loci* do sistema LDD (X-STR) Decaplex. As barras verticais representam o desvio padrão referente a cada marcador..... 83

Figura 29A- Padrão de amplificação no *locus* DXS7133..... 85

Figura 29B- Padrão de amplificação no *locus* DXS7424..... 85

Figura 29C– Padrão de amplificação no *locus* DXS8378..... 85

Figura 29D- Padrão de amplificação no *locus* DXS6807..... 85

Figura 29E- Padrão de amplificação no *locus* DXS7132..... 86

Figura 29F- Padrão de amplificação no *locus* DXS7423..... 86

Figura 29G- Padrão de amplificação no *locus* DXS8377..... 86

Figura 29H- Padrão de amplificação no *locus* GATA172D05..... 86

Figura 29I- Padrão de amplificação no *locus* DXS10074..... 87

Figura 29J- Padrão de amplificação no *locus* DXS10101..... 87

Figura 29L- Padrão de amplificação no *locus* D2S1338..... 87

Figura 29M- Padrão de amplificação para o *locus* autossômico FGA..... 87

Figura 29N- Padrão de amplificação no *locus* D7S820..... 88

Figura 29O- Padrão de amplificação no *locus* D16S539..... 88

Figura 30- Heredogramas representativos de três meio-irmãs paternas..... 89

Figura 31- Heredograma representativo de vínculo genético paterno entre neta e avó paterna..... 90

Figura 32- Eletroferograma mostrando bandas *stutter* (setas) em relação ao alelo verdadeiro..... 106

LISTA DE TABELAS

Tabela 1– *Loci* utilizados na construção do LDD (X-STR) Decaplex com respectivas unidades de repetição, fluoróforos, comprimentos de onda de excitação e emissão dos fluoróforos e referências bibliográficas..... 42

Tabela 2– *Loci* utilizados na construção do LDD (X-STR) Decaplex e respectivas seqüências de *primers*..... 43

Tabela 3– Condições termocíclicas utilizadas para amplificação do LDD (X-STR) Decaplex..... 45

Tabela 4– Distribuição das freqüências dos alelos verificados nos dez *loci* X-STR estudados na fração masculina da amostra populacional constituída por indivíduos do Rio de Janeiro autodenominados afrodescendentes..... 58

Tabela 5– Distribuição das freqüências dos alelos verificados nos dez *loci* X-STR estudados na fração feminina da amostra populacional constituída por indivíduos do Rio de Janeiro autodenominada afrodescendentes..... 59

Tabela 6– Distribuição das freqüências dos alelos verificados nos dez *loci* X-STR estudados na fração total da amostra populacional constituída por indivíduos do Rio de Janeiro autodenominados afrodescendentes..... 60

Tabela 7– Parâmetros estatísticos para os *loci* do sistema LDD (X-STR) Decaplex estudados na amostra populacional do Rio de Janeiro autodenominada afrodescendente..... 66

Tabela 8– Análise do desequilíbrio de ligação para os pares de marcadores na população masculina do Rio de janeiro autodenominada afrodescendente..... 67

Tabela 9– Parâmetros estatísticos de avaliação da eficiência forense para os *loci* do sistema LDD (X-STR) Decaplex na amostra populacional do Rio de Janeiro autodenominada afrodescendente..... 69

Tabela10- Parâmetros estatísticos em investigação de vínculo genético paterno com a participação de uma suposta filha, sua genitora e da mãe do suposto pai utilizando o sistema LDD (X-STR)..... 91

Tabela11– Parâmetros estatísticos na análise forense calculado com base na freqüência alélica dos *loci* STR do sistema LDD (X-STR) Decaplex, conforme mostrado na tabela 6..... 92

Tabela12– Parâmetros estatísticos de vínculo genético e de interesse forense calculado com base nos sistemas de tipagem LDD (X-STR)

Decaplex e *Identifiler*..... 93

Tabela13– *Locus* X-STR estudado em diferentes populações mundiais..... 95

LISTA DE ABREVIATURAS E SIGLAS

A-	Adenina
ARA -	*Androgen Receptor* – Receptor Andrógeno
C -	Citosina
cM -	centimorgan
DNA	*DeoxyriboNucleic Acid* –- Ácido desoxirribonucléico
dNTP –	desoxirribonucleotídeo 5’ trifosfato
dp –	desvio padrão
DXS –	DNA CromossomoX Seqüências de cópia única
EDTA -	ácido etilenodiaminotetracético
EHW-	Equilíbrio de Hardy-Weinberg
F -	Foward
FTA –	Cartão destinado armazenagem de sangue, células bucais, saliva, secreções, tecidos, etc.
g -	gravidade
G -	Guanina
G6PD-	Glicose-6-fosfato desidrogenase
GATA –	Repetição Guanina Adenina Timina Adenina
GHEP –	Grupo de Língua Espanhola e Portuguesa
He -	Heterosigozidade esperada
Ho -	Heterosigosidade observada
HPRTB-	*Human Phosphoribosyl Transferase* – Fosforribosil Transferase Humana
ICV –	Índice Cumulativo de Vinculação
IV -	Índice de Vinculação
ISFG –	*International Society for Forensic Genetics* – Sociedade Internacional de Genética Forense
$KHCO_3$-	Bicarbonato de potássio
M -	molar
Mbp -	milhões de pares de bases
MEC-	*Mean Exclusion Chance* – chance Média de Exclusão
Mg -	miligrama

min - minuto

mL - mililitro

mm^2 - milímetro quadrado

NA9947A- linhagem celular feminina

Nacl – Cloreto de sódio

ng - nanograma

NH_4CL – Cloreto de amônio

NLB - Tampão de lise de membrana nuclear

pb- pares de bases

PBS - Tampão Salina Fosfato

PC - Probabilidade de Coincidência

PCR – *Polymerase Chain Reaction* - Reação em cadeia da polimerase Poder

PD - de Discriminação

PE - Poder de Exclusão

pH – potencial de hidrogênio

PIC - Polymorphism Information Content - Conteúdo de Informação Polimórfica

POP – *Performance Optimizer Polymer*

PV - Probabilidade de Vinculação

R - *Reverse*

RFU – *Relative Fluorescent Unit* - unidade relativa de fluorescência Razão de

RV – Verossimilhança

SDS - Dodecil Sulfato de Sódio

SEB - Tampão de extração

seg - segundo

STR – *Short Tandem Repeat* - Seqüências curtas repetidas em tandem

T - Timina

Taq – Thermus aquaticus

TE - Tampão Tris-EDTA

TLH - Tampão de lise de hemácias

Tris – Tris-Hidroxametil-Aminometano

U - unidade de enzima

VNTR – Número variado de seqüências repetidas em tandem

X-STR - Seqüências curtas repetidas em tandem do cromossomo X

LISTA DE SIMBOLOS

λ –	comprimento de onda
%	porcentagem
µL -	microlitro
µM -	micromolar
^{0}C -	grau Celsius
∞ -	infinito
6-FAM –	6-carboxi-fluoresceína

SUMÁRIO

	INTRODUÇÂO.....	22
1	**OBJETIVOS**.....	36
1.1	**Objetivo Geral**.....	36
1.2	**Objetivos Específicos**	36
2	**MATERIAL E MÉTODOS**	37
2.1	**Amostras biológicas**.....	37
2.1.1	Detalhamento das amostras.....	37
2.1.2	Detalhamento das amostras biológicas utilizadas no estudo de validação do sistema multiplex de tipagem.....	38
2.2	**Extração de DNA**.....	38
2.2.1	Extração por Salting-out.....	38
2.2.2	Extração por Chelex.....	39
2.2.3	3.2.3 Extração por método orgânico/ fenol:clorofórmio:álcool isoamílico / filtro microcon®.....	40
2.3	**Determinação da concentração de DNA**.....	41
2.4	**Seleção dos *loci***.....	41
2.5	**Padronização da concentração dos *primers* no LDD (X-STR) Decaplex**.....	42 44
2.6	**Amplificação dos STR no sistema LDD (X-STR) Decaplex**.....	44
2.7	**Eletroforese Capilar**.....	45
2.7.1	Preparação das amostras (produtos da PCR)	45
2.7.2	Eletroforese capilar e detecção.....	46
2.7.3	Análise dos produtos da PCR.....	46
2.7.4	Construção dos bins.....	46
2.7.5	Sequenciamento.....	47
2.8	**Estratégias de validação**.....	48
2.8.1	Determinação do grau de sensibilidade em amostra de DNA.....	48
2.8.2	Estudo de reprodutibilidade das tipagens utilizando o LDD (X-STR) Decaplex.....	48
2.8.3	Análise das bandas *stutter*.....	49
2.8.4	Análise do padrão de misturas de DNA.....	49
2.9	**Análise estatística**.....	49
2.10	**Lista de soluções e tampões utilizados**.....	51

2.10.1 Extração de DNA Genômico 51
2.10.2 Amplificação da Polimerase em cadeia 52
3 **RESULTADOS** 53
3.1 **Desenvolvimento de sistema multiplex** 53
3.2 **Construção do banco de freqüências alélicas** 54
3.2.1 Estimativas das freqüências alélicas 57
3.2.2 Confirmação da ocorrência de alelo microvariante 61
3.2.3 Verificação da freqüência haplotípica 64
3.2.4 Análise de parâmetros estatísticos 64
3.2.5 Análise do desequilíbrio de ligação 66
3.2.6 Parâmetros forenses 68
3.3 **Estratégias de validação** 70
3.3.1 Determinação do grau de sensibilidade 70
3.3.2 Análise de reprodutibilidade da tipagem de alelos X-STR a partir de diferentes amostras biológicas 75
3.3.3 Análises de bandas *stutter* 82
3.3.4 Análises de padrão de misturas 83
3.4 **Estudo de casos** 88
3.4.1 Reconstrução de perfil genético do investigado a partir de suas filhas. 89
3.4.2 Pesquisa de vínculo genético a partir da suposta avó paterna. 89
3.4.3 Parâmetros estatísticos de análise forense 92
4 **DISCUSSÃO** 94
4.1 **Construção do sistema LDD (X-STR) Decaplex** 94
4.2 **Construção do banco de freqüência de alelos X-STR** 96
4.2.1 Estimativas das freqüências alélicas 96
4.2.2 Estimativas das freqüências haplotípicas 99
4.2.3 Análise de parâmetros estatísticos 99
4.2.4 Análise do desequilíbrio de ligação 101
4.2.5 Parâmetros forenses 102
4.3 **Estratégias de validação** 104
4.3.1 Efeito da quantidade de DNA na eficiência do sistema multiplex. 104
4.3.2 Estudo de reprodutibilidade do LDD (X-STR) Decaplex 105
4.3.3 Análises de bandas *stutter* 106

4.3.4 Padrão de mistura de DNA..... 107
4.4 **Estudo de casos**..... 108
5 **CONCLUSÃO**..... 111
REFERÊNCIAS..... 112
APÊNDICE A - Marcadores X-STR descritos e sua localização no cromossomo. A localização dos marcadores foi obtida a partir das bases de dados NCBI "National Center for Biotechnology Information" e Rutgers Map. Adaptado de www.chrx-str.org. 120
APÊNDICE B - Perfis genéticos constituídos pelos alelos tipados após amplificação, por PCR, com o LDD (X-STR) Decaplex, de alíquotas dos DNA de 102 mulheres afrodescendentes do Rio de Janeiro, conforme autodenominação..... 122
APÊNDICE C – Haplótipos formados pelos alelos tipados após amplificação, por PCR, com o LDD (X-STR) Decaplex, de alíquotas dos DNA de 68 homens afrodescendentes do Rio de Janeiro, conforme autodenominação..... 129
APÊNDICE D – Population data for six X-chromosome STR loci in a Rio de Janeiro (Brazil) sample: Usefulness in forensic casework……. 132
APÊNDICE E – A X-chromosome STR hexaplex as a powerful tool in deficiency paternity cases…….. 134

INTRODUÇÃO

DNA

A molécula de ácido desoxirribonucleico (DNA) ocupa, nos dias atuais, um lugar de destaque na comunidade científica. Isso se deve ao extenso conhecimento adquirido sobre a estrutura e função desta macromolélula. A aplicação do DNA aliada ao avanço do conhecimento das técnicas de biologia molecular possibilitou o aprimoramento e o desenvolvimento de diferentes áreas das ciências biológicas.

Neste contexto, destacam-se áreas como a genética clínica trazendo grandes contribuições para o diagnóstico e prevenção de distúrbios genéticos (Pasternak, 2007), a biodiversidade, através da análise de variações genéticas específicas para identificação e diferenciação de espécies (Monaghan *et al*., 2005), além da farmacogenética, com a produção e desenvolvimento de fármacos associados às variações genéticas dos indivíduos (Lesko *et al*., 2003).

A molécula de DNA foi pela primeira vez caracterizada como responsável pela informação genética em 1944 por Avery e colaboradores. Através de experimentos realizados com cepas de bactérias, eles propuseram que o DNA seria o componente responsável pela transmissão da informação genética (Avery *et al.,* 1944). Em 1949, estudos realizados por Erwin Chargaff deram uma contribuição importante para a elucidação da estrutura do DNA através da quantificação de cada um dos tipos de base nitrogenada presente na molécula (adenina, timina, citosina e guanina) de várias espécies.

Posteriormente, estudos sobre a estrutura do DNA, utilizando difração de raios-X, foram realizados por Rosalind Franklin objetivando a compreensão da estrutura da molécula do DNA. Este trabalho permitiu aos cientistas James Dewey Watson e Francis Crick, em 1953, confirmar a dupla estrutura helicoidal da molécula do DNA. Neste modelo, a estrutura do DNA apresenta duas cadeias (ou filamentos) de açúcar-fosfato em hélice ligado a um grupo de bases nitrogenadas. O açúcar é a desoxirribose e as duas fitas de DNA são unidas por de pontes de hidrogênio que se formam entre as bases nitrogenadas especificas, sendo duas pontes formadas entre as bases A e T, e três outras formadas entre as bases G e C. (Watson & Crick,

1953). As bases nitrogenadas podem se combinar de diferentes maneiras, e essa variedade de combinação é a característica responsável pela diversidade biológica entre os seres humanos.

O genoma humano

O conteúdo genético dos seres humanos está organizado em um genoma nuclear complexo, que possui a quase totalidade da informação genética (aproximadamente 26.000 genes), e em um genoma mitocondrial, mais simples, com poucos genes codificadores (37 genes), (Strachan & Read, 2010).

O genoma nuclear diplóide está distribuído em unidades denominadas cromossomos, constituído por 22 pares autossômicos e 1 par sexual. Nas mulheres, o par sexual é representado por dois distintos cromossomos X, enquanto que nos homens este ocorre uma única vez, juntamente com o cromossomo Y.

O genoma mitocondrial é representado por um DNA pequeno e circular de linhagem exclusivamente materna, localizado na matriz das mitocôndrias, em número variável dependendo do tipo celular. As células que requerem uma grande demanda energética como, por exemplo, as hepáticas, podem conter de 1.000 a 1.600 desta organela. O DNA mitocondrial totaliza pouco mais de 16.000 pares de bases (pb), ao contrário do DNA nuclear, que possui um genoma haploide com aproximadamente 3 bilhões de pares de bases (figura 1), (Butler, 2009).

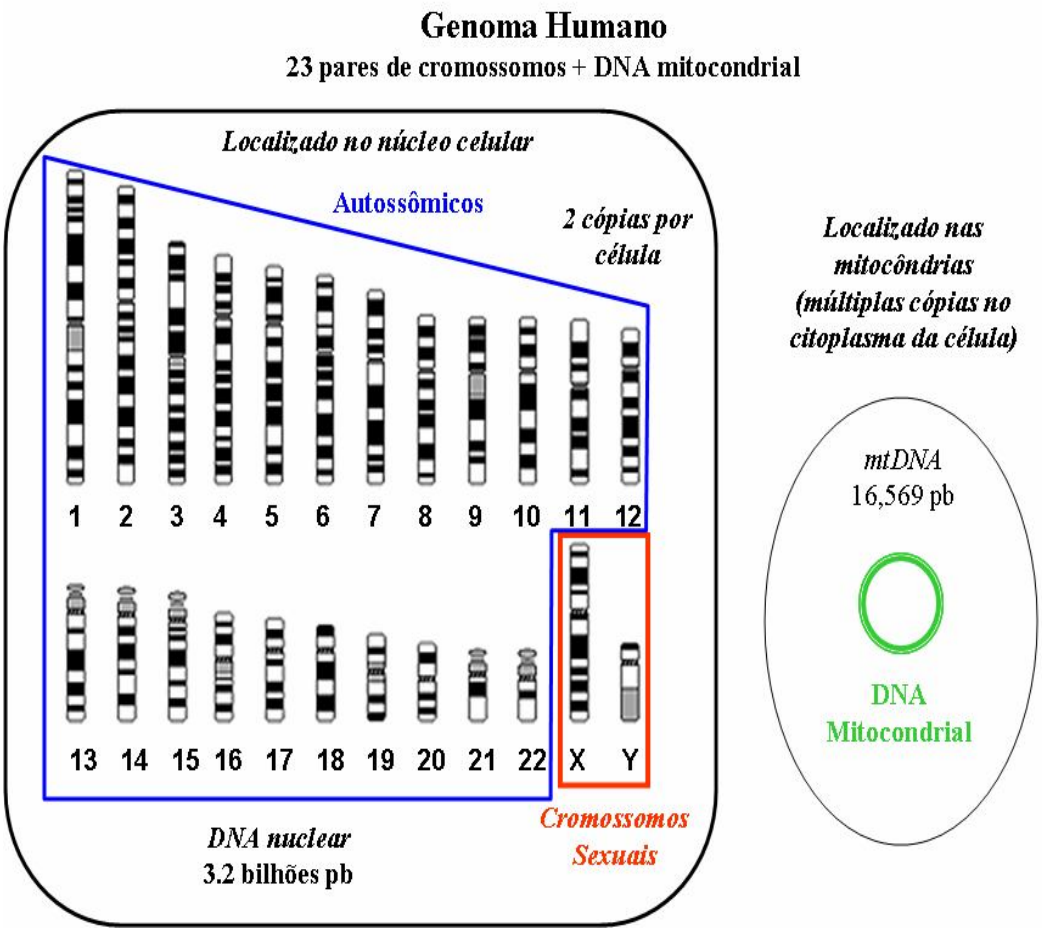


Figura 1- Modelo esquemático do genoma nuclear e mitocondrial humano. Adaptado de Butler, 2005.

A identificação humana por DNA

A base da identificação genética entre os indivíduos está presente no DNA. As diferentes frações do DNA presente em cada ser nos tornam indivíduos únicos, à exceção dos gêmeos monozigóticos, que compartilham o mesmo genoma.

O DNA humano pode ser dividido em regiões funcionais ou codificantes (genes e seqüências relacionadas), representando cerca de 20% do genoma, e em regiões extra-gênicas ou não codificantes para nenhuma característica individual. Essas regiões extra-gênicas por sua vez são divididas em seqüências únicas ou em

baixo número de cópias, representando cerca de 80% da fração não funcional do genoma, e em seqüências repetitivas, caracterizando 20% desta fração. No DNA, qualquer região que apresente um determinado polimorfismo pode ser utilizada como marcador para fins de identificação humana, porém as de maior interesse são as seqüências repetitivas, por apresentarem um elevado grau de polimorfismo (Fowler *et al.*, 1988).

Essas regiões do genoma humano contêm seqüências de DNA altamente repetitivas organizadas em blocos, podendo ser caracterizada em dois tipos principais, de acordo com a sua organização estrutural: o DNA repetitivo disperso e o DNA repetido em tandem.

O DNA repetitivo disperso é caracterizado por conter unidades de seqüências repetitivas espalhadas por todo o genoma. Os elementos de repetição mais estudados são os pertencentes à família Alu, com segmentos de aproximadamente 300 pares de bases, e os pertencentes à família Line1 (Deininger & Batzer, 1999).

O DNA repetido em tandem tem localização variada no genoma humano e a característica principal deste DNA está relacionada com o seu arranjo de repetições em seqüências altamente variáveis, sendo uma ferramenta de grande importância para o estudo da genética populacional e da genética forense (Gil *et al.,* 1985; Smouse & Chakraborty, 1986).

O DNA não-codificador repetido pode ser dividido em três classes distintas: DNA-satélite, minissatélite e microssatélite. Esta divisão está relacionada com a ocorrência de variação dos tamanhos médios das unidades de repetições nucleotídicas.

O DNA-satélite é caracterizado pela sua composição em longos blocos de DNA repetidos em tandem com localização próxima aos centrômeros, onde sua unidade de repetição é representada por seqüências simples ou moderadamente complexas, podendo variar de 5 a 250 pares de bases (Strachan & Read, 2010).

A classe dos DNA-minissatélites abrange seqüências de DNA altamente polimórficas (hiperváriaveis) repetidas em tandem de tamanho moderado denominadas VNTR – “Variable Number of Tandem Repeats”. Estas seqüências podem ser encontradas em regiões dispersas, ao longo do genoma nuclear humano, estando suas unidades organizadas em blocos de 8 a 100 pares de bases (Jeffreys *et al.*, 1985).

As primeiras seqüências VNTR descritas foram empregadas para auxiliar no estudo do mapeamento do genoma humano (Botstein *et al.,* 1980; Wyman *et al*., 1980). No entanto, com a descoberta de novas regiões hipervariáveis por Allec Jeffreys e colaboradores (1985), a proposição do uso destas para fins de identificação humana foi apresentada, sendo esta metodologia capaz de individualizar um perfil genético a partir da análise de um conjunto destes marcadores. A metodologia de análise está baseada no uso de sondas moleculares capazes de detectar seqüências VNTR no genoma humano digeridas por enzimas de restrição (Jeffreys *et al.*, 1985).

Ao contrário do DNA-minissatélite, os DNA-microssatélites são caracterizados por apresentarem seqüências repetitivas em tandem de tamanho curto, conhecidas como STR – “Short Tandem Repeats”, podendo também ser encontradas por todo o genoma. São formadas por unidades de repetição cujas seqüências nucleotídicas apresentam de 2 a 6 pares de bases, formando blocos de repetição de 50 a 500 pb (Silva, 2004a).

A partir dos anos 90, a técnica de identificação humana através de marcadores VNTRs foi substituída por um novo sistema de tipagem baseado na amplificação por PCR de marcadores microssatélites ou STR em cromossomos autossomos, tornando-se uma poderosa ferramenta de medida da variabilidade da estrutura do DNA em humanos, resultando em grandes avanços científicos e tecnológicos para a ciência forense (Jeffreys, 1987).

A utilização de regiões STR para identificação por DNA através da PCR foi um grande avanço científico, visto que tal metodologia utiliza quantidades mínimas de DNA, ou até mesmo amostras altamente degradadas para tipagem genética, ao contrário dos marcadores VNTR, que apesar do alto grau de polimorfismo requerem DNA em maiores quantidades e altamente preservado (Pena, 1995).

Além do seu uso na identificação humana, a utilização de marcadores STR teve grande destaque na pesquisa e diagnóstico de doenças como na área do câncer, fibrose cística (Morral & Estivill, 1992) e doença de Huntington (Weber *et al.,* 1992).

A análise individual de um único *locus* STR (monoplex) não fornece um poder de discriminação adequado para a análise forense, visto que tal parâmetro é dependente do número de *loci* analisados no sistema de tipagem por DNA utilizado. Daí a necessidade de sistemas multiplex, onde a junção de pares de *primers*

determina a amplificação de vários marcadores genéticos simultaneamente em uma única reação de PCR.

A obtenção de uma grande quantidade de dados em estudos de genética populacional e forense é um procedimento de grande importância para obtenção das freqüências alélicas e para a construção de bases de dados. Este fato, aliado à necessidade de reduzir a quantidade de amostra de DNA utilizada na análise, quando esta é limitada (principalmente nos casos forenses), levou ao desenvolvimento de sistemas capazes de amplificar simultaneamente diversos *loci* STR (Budowle *et al.,* 2000).

Os primeiros sistemas multiplex para tipagem humana por DNA foram constituídos por STR localizados nos cromossomos autossomos, trazendo grandes benefícios nas investigações de vinculo genético e nas investigações criminais. Atualmente, existem diferentes sistemas multiplex disponíveis comercialmente, capazes de amplificar até 15 STR autossômicos, permitindo alcançar um poder de discriminação elevado e adequado para a resolução da maioria dos casos forenses.

Sistemas multiplex contendo marcadores STR do cromossomo Y (Y-STR) foram também desenvolvidos, sendo estes bastante utilizados em análises de evidências de crimes sexuais, onde em muitos casos a amostra biológica do acusado encontra-se misturada à da suposta vítima. A aplicação destes marcadores também se relaciona com o propósito de determinação de parentesco entre indivíduos do sexo masculino e no estudo de populações (Góes *et al.*, 2005; Kayser, 2003; Domingues *et al.*, 2007).

Recentemente, o uso de marcadores STR do cromossomo X (X-STR) tem se tornado muito comum na prática forense, sendo uma ferramenta complementar à análise com STR autossômicos, ganhando grande destaque e tendo sua utilidade comprovada (Poetsch *et al.*, 2005).

A vantagem na utilização desses marcadores está nos casos de ausência paterna envolvendo filhas. Nesses casos, o DNA de familiares como tios, irmãs, meio-irmãs e avó paterna, que compartilham alelos do mesmo cromossomo X, é utilizado, apresentando uma maior eficiência se comparado aos STR autossômicos e mostrando-se altamente informativos (Edelmann *et al.*, 2004). Adicionalmente a esses estudos, os X-STR podem ser utilizados com eficiência em casos de parentesco quando a questão envolve irmãs que foram separadas ainda quando crianças, tendo o mesmo suposto pai e nos casos de investigação de maternidade

(Shin *et al.*, 2004). Apesar de a maternidade poder ser investigada por análise da seqüência do DNA mitocondrial, este procedimento é mais custoso e nem sempre se atinge o nível de discriminação necessário em casos forenses.

Análise de ligação

O estudo da análise de ligação está relacionado com a localização e distanciamento entre os *loci* no cromossomo. Assim, a freqüência de recombinação entre os *loci* é uma função da distância entre os mesmos. Quanto mais separados estiverem esses *loci*, mais freqüentemente irá ocorrer a recombinação entre eles. Por outro, lado se eles estiverem muito próximos, a recombinação será rara, configurando o fenômeno conhecido como desequilíbrio de ligação e, nestes casos, os marcadores associados devem ser tratados como haplótipos e não como marcadores independentes. Os haplótipos tendem a permanecerem intactos ao longo das gerações, se não houver mutações (Korf, 2008).

Em relação a sua aplicação em estudos evolutivos, os X-STR são considerados ferramentas importantes na reconstrução de árvores filogenéticas. Tal estratégia é possível devido ao fato de homens possuírem o seu conjunto de alelos na forma de haplótipo, sendo este transmitido ao longo de gerações sem sofrer recombinação (Schaffner, 2004).

O cromossomo X

O modo de transmissão do cromossomo X apresenta características particulares em estudos de genética populacional e também pela sua aplicação específica no campo forense, colocando-se em uma situação intermediária entre os genomas autossômicos e os uniparentais (cromossomo Y e DNA mitocondrial).

Em células humanas femininas o cromossomo X está presente como um par homólogo, semelhante aos autossômicos em muitos aspectos, ao contrário dos homens que possuem um único cromossomo X.

Devido ao fato de homens apresentarem um único cromossomo X, que não sofre recombinação genética, com exceção das regiões pseudo-autossômicas, o

padrão de transmissão confere aspectos importantes como, por exemplo, o fato de homens transmitirem todo o seu cromossomo X para suas filhas na forma de haplótipos, além da capacidade de a tipagem do cromossomo X de duas ou mais irmãs revelarem o haplótipo do cromossomo X paterno e parte do genótipo do X materno, conferindo a este cromossomo características únicas (Szibor *et al.*, 2003a).

O idiograma representativo do cromossomo X, com a distribuição de seus marcadores, bem como a localização física em Mbp (milhões de pares de base) e genética (em centimorgans – cM), a qual reflete a taxa de recombinação entre os mesmos, já foi estabelecido (Szibor, 2007). Tais informações são importantes, pois dizem respeito à aproximação ou distanciamento dessas regiões, o que possibilita a segregação conjunta ou independente dos marcadores.

A figura 2 representa o idiograma do cromossomo X, onde podemos observar a distribuição de alguns dos diversos X-STR utilizados na prática forense e em estudos populacionais. Adicionalmente, as informações disponíveis no apêndice A fornece informações relativas às localizações físicas e genéticas dos mesmos.

Figura 2 – Distribuição de marcadores STR ao longo do cromossomo X. Adaptado de www.chrx-str.org.

Condições para a utilização do cromossomo X na prática forense

Antes da utilização dos marcadores do cromossomo X na prática forense, um adequado planejamento deve ser desenvolvido, e etapas como seleção dos marcadores X-STR e padronização dos mesmos devem ser seguidas com vista ao aperfeiçoamento das condições de interpretações.

A seleção de marcadores STR está relacionada com o grau de polimorfismo de cada *locus*, e tal medida é feita com base na ocorrência de um maior ou menor número de alelos (heterozigosidade) para um dado *locus* em uma determinada população. Para a avaliação do grau de polimorfismo, é necessária a análise de parâmetros estatísticos objetivando a determinação do padrão de diversidade genética destes marcadores na população em estudo.

Por outro lado, a etapa de padronização é também um passo muito importante, pois nos fornece informações fundamentais relacionadas com alguns parâmetros de análise como, por exemplo, a temperatura de hibridização de cada *primer*, que devem ser bem próximas, concentração dos *primers*, concentração de DNA na reação de PCR e condições de tamponamento. A especificidade da amplificação de cada marcador e a compatibilidade dos *primers* no sistema multiplex, também são parâmetros de grande importância de modo a evitar a existência de interações entre os mesmos, com conseqüentes problemas na amplificação (Manta, 2008).

Além destes, a realização de testes de validação tem grande importância no sentindo de se estabelecer critérios de qualidade nas análises de amostras forenses e de vínculo genético.

Condições para a utilização dos marcadores X-STR em genética populacional

Para se determinar a variabilidade e a estrutura genética de uma população com base na utilização de marcadores STR, torna-se necessário investigar o padrão de heterozigosidade, ou polimorfismo, dos mesmos na população em estudo. Neste sentido, a construção de bases de dados de referência a partir dos dados de freqüências alélicas e haplotípicas de uma população específica é de grande

importância, por possibilitar a avaliação do índice de diversidade genética dos marcadores STR, estabelecendo, conseqüentemente, o perfil genético da população estudada (Chakraborty & Kidd, 1991).

Estudos recentes foram desenvolvidos no sentido de se verificar o potencial de sistemas multiplex de STR do cromossomo X em diferentes populações bem como analisar o perfil da diversidade genética nas mesmas (Bini *et al.*, 2005; Gomes *et al.*, 2007a; Silva, 2010).

Após a coleta de dados de freqüências alélicas e haplotípicas, torna-se importante, também, a análise de alguns parâmetros estatísticos fundamentais como, por exemplo, o desequilíbrio de ligação, que está relacionada com a ocorrência de segregação conjunta ou independente entre marcadores devido a localizações próximas entre si, bem como a análise do equilíbrio na população segundo o teorema de Hardy-Weinberg. Para a aplicação na prática forense, recomenda-se a utilização de marcadores com a transmissão alélica independente intra e entre os *loci*, evitando-se desta forma a segregação conjunta de marcadores para gerações seguintes como uma unidade única, (Balazs *et al*, 1990).

Em adição a esse princípio, a verificação de parâmetros de interesse forense, como, por exemplo, Poder de Discriminação feminino e masculino (PD_F e PD_M), que representa a probabilidade de dois indivíduos não relacionados geneticamente serem diferenciados pelo multiplex X-STR utilizado, quando escolhidos ao acaso na população; Conteúdo de Informação de Polimorfismo de cada marcador (PIC), que está relacionado com o número de repetições da unidade básica de cada *loci* STR; e Chance Média de Exclusão - *Mean Exclusion Chance* - (MEC) e Poder de Exclusão (PE), referente ao percentual de indivíduos cuja identidade ou paternidade biológica seria excluída, com base nos marcadores utilizados, também devem ser considerados no estudo.

Desenvolvimento de multiplex do cromossomo X

O estudo de marcadores genéticos autossômicos, do cromossomo Y e do DNA mitocondrial é uma prática generalizada em genética forense, permitindo resolver com eficiência a grande maioria das questões que se colocam. No entanto,

em alguns casos particulares, os marcadores do cromossomo X constituem uma poderosa ferramenta de informação, e tem o potencial de complementar eficientemente a análise de outros marcadores genéticos.

A utilização de marcadores STR do cromossomo X iniciou-se durante o século passado através de estudos no campo da genética clínica, correlacionando a herança do cromossomo X e doenças com a base molecular ainda pouco conhecida, como a hemofilia, distrofia muscular de Duchenne e deficiência de G6PD - Glicose-6-fosfato desidrogenase (Szibor, 2007).

Um dos primeiros marcadores estudados corresponde ao gene presente em humanos denominado ARA (Receptor Andrógeno – *Androgen Receptor*), (Desmarais *et al.*, 1998). Um segundo marcador de identificação muito utilizado, após sua descoberta, corresponde ao gene HPRTB (Fosforribosil Transferase Humana - *Human Phosphoribosyl Transferase*) [Mertens *et al.,* 1999]. Com o seqüenciamento do cromossomo X humano (Ross *et al.*, 2005), foram descritas diversas outras seqüências de STR no cromossomo X, correspondendo a 26 trímeros e 90 tetrâmeros repetitivos (Szibor *et al.*, 2003b).

Os multiplex descritos inicialmente compreendiam triplex (Wiegand *et al.*, 2003), quadruplex (Athanasiadou *et al.*, 2003; Lee *et al.*, 2004) e pentaplex (Zarrabeitia *et al.*, 2002a; Shin *et al.*, 2004). Com o interesse maior em relação ao uso desses marcadores, foi estudado um número maior de STR em populações diversas, como a coreana (Shin *et al.*, 2005), a italiana (Robino *et al.*, 2006), paquistanesa (Tariq, *in press*) e brasileira (Tavares *et al.*, 2008; Martins *et al*., 2009; Laura *et al*., 2010; Silva *et al.*, 2010).

Recentemente, um decaplex foi desenvolvido e utilizado em diferentes populações, inclusive brasileira, através de um estudo colaborativo liderado pelo grupo GHEP-ISFG (Grupo de Línguas Espanhola e Portuguesa da Sociedade Internacional de Genética Forense). Neste estudo, diferentes populações das regiões Ibéricas e latinas americanas foram estudadas com base em suas freqüências alélicas (Gusmão *et al.,* 2009).

O desenvolvimento e validação de novos sistemas de tipagem vêm sendo implementado no sentido de reduzir custos e disponibilizar informações não observadas em sistemas comerciais. Atualmente, os únicos sistemas comerciais para tipagem e identificação humana através do cromossomo X existentes no mercado são o Mentype® Argus X-8 e o Mentype® Argus X-12.

O sistema multiplex Argus X-8 é formado por oito marcadores X-STR e contém dados populacionais referentes às populações européia, asiática e africana (Becker, 2008). Já o sistema Argus X-12, formado por doze marcadores X-STR, foi recentemente disponibilizado comercialmente e estudos populacionais ainda estão sendo desenvolvidos (Bentayebi *et al., in press*; Edelmann *et al., in press*; Horvath *et al., in press*; Zeng *et al., in press*). Daí a importância de se desenvolver uma base de dados no Brasil, visto que o sistema multiplex LDD (X-STR) Decaplex desenvolvido neste estudo contém cinco marcadores em comum aos sistemas Argus X-8 e Argus X-12.

No âmbito deste trabalho foi desenvolvido um sistema multiplex capaz de amplificar dez X-STR numa única reação de PCR.

Validação de sistemas de tipagem STR

Antes de ser utilizado como um sistema de identificação humana com aplicação em análises de vinculação genética e forense, uma série de estudos deve ser realizada e alguns parâmetros de fundamental importância devem ser avaliados em um processo conhecido como validação (Micka *et al.*, 1999). Tais estudos estão relacionados com a sensibilidade do sistema multiplex, a presença de artefatos na reação de PCR, estudo de reprodutibilidade em diferentes amostras e a capacidade de discriminação do sistema multiplex quando submetido a um padrão de misturas de distintas amostras de DNA.

Os testes de sensibilidade estão relacionados com a determinação da quantidade mínima de DNA necessária para a amplificação de todos os *loci* STR que constituem o sistema de tipagem (Silva *et al.*, 2004b).

O aparecimento de artefatos pode representar um sério problema durante a análise de regiões STR e, conseqüentemente, prejudicar a identificação dos produtos da PCR (tipagem genética). A presença destes artefatos está relacionada à falhas durante o processo de amplificação, ocasionando a adição ou perda de uma unidade de repetição no material amplificado, gerando, conseqüentemente, picos ou bandas de baixa intensidade denominadas *stutter,* que pode erroneamente ser interpretado como um alelo real. Tal fenômeno pode ser minimizado através de

ajustes nas temperaturas de extensão dos *primers* durante a PCR (Jobling *et al.*, 2004).

Em um laboratório de análises forense, a fonte de DNA disponível para uma análise de vinculação genética, pode ser proveniente de diferentes tipos de amostras biológicas. Neste sentido, o estudo de reprodutibilidade torna-se importante por possibilitar a análise das diferentes fontes de materiais biológicos que podem ser amplificadas satisfatoriamente como sistema multiplex utilizado.

A determinação da tipagem genética de um perfil biológico obtido a partir de uma amostra ou evidência biológica com um suposto padrão de misturas de DNA é de grande importância para a análise forense, principalmente nos casos de crimes sexuais, uma vez que a amostra de DNA a ser analisada, proveniente da vítima, pode vir a estar misturada com o DNA de um suposto agressor. Neste sentido torna-se necessária a determinação da proporção a partir da qual a identificação do DNA de um dos componentes da mistura é perdida (Willot & Crosse, 1986).

Assim, a aplicação de marcadores X-STR marcadores na prática forense e na genética de populações requer a seleção de *loci* polimórficos, o aprimoramento de metodologias de genotipagem que permitam, simultaneamente, a amplificação de vários marcadores em uma única reação de PCR. Além do estudo de validação e construção de bases de dados de interesse populacional e forense para a população em questão (Kohlrausch *et al.*, 2005).

1 **OBJETIVOS**

1.1 **Objetivo gerall**

Desenvolvimento e validação de sistema multiplex de regiões STR do cromossomo X humano e estudo populacional.

1.2 **Objetivos específicos**

1. Desenvolver sistema multiplex capaz de amplificar em uma única reação de PCR um conjunto de 10 *loci* STR do cromossomo X;

2. Validar o sistema multiplex para uso em identificação humana:
   - Determinar a quantidade mínima de DNA para a amplificação simultânea dos 10 *loci* X-STR selecionados;
   - Avaliar a reprodutibilidade da tipagem dos alelos para os marcadores STR em diferentes fontes de amostras biológicas;
   - Analisar a presença de artefatos caracterizados como bandas *Stutter*;
   - Discriminar perfis genotípicos a partir da mistura de DNA.

3. Realizar estudo populacional:
   - Utilizar o sistema multiplex em referência para a genotipagem de indivíduos autodenominados afrodescendentes nascidos no Rio de Janeiro;
   - Estimar as freqüências alélicas na amostra populacional;
   - Construir banco de haplótipos X-STR masculinos;
   - Aplicar testes estatísticos para avaliar: a ocorrência de segregação independente/dependente entre os diferentes *loci* em estudo, diversidade genética, equilíbrio das freqüências genotípicas em uma amostra de população feminina segundo Hardy-Weinberg; e cálculos de parâmetros de análises forenses e investigação de vínculos genéticos.

4. Elevar o poder de discriminação dos sistemas correntes de genotipagem com aplicação dirigida à complementação (distância genética).

2 **MATERIAL E MÉTODOS**

2.1 **Amostras biológicas**

Na rotina forense, diferentes tipos de amostras biológicas podem ser utilizados como fonte de DNA genômico**,** e a coleta e estocagem das amostras são etapas de grande importância para o processo de identificação humana por DNA. Diferentes amostras biológicas como sangue periférico, dente, saliva, urina, sêmen, sangue em substrato sólido (FTA) e cabelo (com bulbo) foram utilizadas para amplificação com o LDD (X-STR) Decaplex desenvolvido nesse trabalho.

2.1.1 Detalhamento das amostras biológicas

Foram selecionadas para o estudo amostras biológicas de 170 indivíduos (68 homens e 102 mulheres), nascidos na cidade do Rio de Janeiro, autodenominados afrodescendentes, e não relacionados geneticamente. Todos participaram de perícias realizadas no Laboratório de Diagnósticos por DNA (LDD) da Universidade do Estado do Rio de Janeiro (UERJ), e consentiram na utilização de suas amostras biológicas, de forma não identificada, para a realização de estudos científicos.

A autorização foi dada mediante preenchimento de questionário em concordância com a utilização de seus materiais biológicos nas análises efetuadas pelo laboratório, sendo a identificação dos indivíduos mantida em sigilo.

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

2.1.2 Detalhamento das amostras biológicas utilizadas no estudo de validação do sistema multiplex

Diferentes fontes de amostras biológicas foram utilizadas no estudo de validação do sistema multiplex.

A amostra de material dentário utilizado neste trabalho foi proveniente da base de dados de exames de identificações civis, criminais e de vínculo genético do Laboratório de Diagnósticos por DNA, sendo esta escolhida aleatoriamente. A base de dados do laboratório envolve amostras de restos mortais, evidências criminais e suas respectivas referências. Já as amostras de saliva (mucosa oral), urina, sêmen, sangue (FTA) e cabelo (bulbo) foram obtidas a partir de um indivíduo saudável.

Na extração de DNA de dentes o material (1 a 2 dentes) foi previamente limpo com água, hipoclorito de sódio comercial (água sanitária) e etanol. Posteriormente, os dentes foram mantidos em solução de EDTA até que o esmalte dentário estivesse amolecido.

A amostra de mucosa oral (saliva) foi coletada com o auxílio de suab e mantida a -20 ºC até a extração do DNA. A amostra de sangue coletada foi transferida para um cartão de papel FTA, permanecendo nas mesmas condições de temperatura descritas acima, bem como a amostra de sêmen.

Alíquota de 1 mL de amostra de urina foi transferida para tubo de 1,5 mL e centrifugada por 10 min (Centrifuga Centimicro 242 - Fanem) a 1.200 x g. Ao precipitado foi adicionado o volume de 500 µL de tampão PBS sendo novamente centrifugado. Este procedimento foi repetido por duas vezes e o precipitado final foi armazenado a -20 ºC até a extração do DNA.

2.2 **Extração de DNA**

2.2.1 Extração por Salting-out

O método de extração de DNA por *Salting-out* (Miller, 1988) consistiu na utilização de aproximadamente 2 mL de sangue adicionado a 10 mL de tampão de lise TLH. Sendo a mistura homogeneizada manualmente e centrifugada por 5 min a 100 x g (Centrifuga Excelsa modelo: 206BL - Fanem) e, em seguida, descartado o

sobrenadante. Ao precipitado de células era adicionado 10 mL tampão TLH, repetindo-se a centrifugação. Após a remoção do sobrenadante, o precipitado de células era ressuspenso em 500 µl de tampão NLB seguindo-se a adição de 15 µL de SDS 10% e 40 µL de proteinase K 20 mg/mL e homogeneizado manualmente. Após a homogeneização, a preparação era mantida por 15-18 horas a 37º C ou por 2 horas a 56º C. Após a incubação eram adicionados 150 µL de NaCl 6M, sendo a preparação agitada vigorosamente manualmente, por aproximadamente 30 seg. Em seguida o material era centrifugado por 10 min, a 905 x g e o sobrenadante transferido para um novo tubo tipo falcon. Ao sobrenadante eram adicionados 2,5 mL de etanol absoluto, com agitação manual em movimentos circulares para precipitação dos ácidos nucleicos. O álcool era descartado e o precipitado ressuspenso em volumes entre 100 a 500 µL de tampão TE, dependendo da quantidade de DNA obtido.

2.2.2 Extração por Chelex

A extração do DNA através do método Chelex (Lareu *et al*., 1994) foi baseada na utilização de aproximadamente 10 µL da amostra de sangue, sendo esta transferida para tubo de micro centrifuga de 1,5 mL e adicionados 1 mL de água milliQ autoclavada. A mistura era homogeneizada suavemente e incubada por 30 min à temperatura ambiente, e em seguida centrifugada durante 5 min (Centrifuga Centimicro 242 - Fanem) a 1.200 x g. O sobrenadante era então descartado e ao precipitado era adicionado 0,5 mL de água milliQ autoclavada, sendo novamente homogeneizado e deixado à temperatura ambiente por 30 min. O material era centrifugado novamente por 5 min à 1.200 x g, sendo desprezado o sobrenadante. Em seguida, era adicionado 300 µL da solução Chelex-100 a 5% em água milliQ. A mistura era incubada a 56^0 C por 20 min e posteriormente a 95^0 C por 8 min. Após a adição de 1 mL de TE, o Chelex era sedimentado por centrifugação por 3 min a 12000 x g, sendo a solução de DNA armazenada a 4^0 C.

2.2.3 Extração por método orgânico/ fenol:clorofórmio:álcool isoamílico / filtro microcon®

A extração de DNA das amostras de saliva, urina, sêmen, sangue (FTA) e cabelo (bulbo) foram feitas através da transferência de 150 µL do material coletado, quando em estado líquido, cerca de dez fios de cabelo com a extremidade proximal bulbar, e de 0,5 mm^2 de secção de material coletado e fixado em substrato sólido cartão FTA para tubos de 1,5 mL. A estes, eram adicionados 500 µL de tampão de lise SEB, 15 de proteinase K (10 mg/mL) e incubados a 56^0 C por 16 horas. Após o período de digestão, foram adicionados 500 µL de fenol:clorofórmio:álcool isoamílico na proporção 25:24:1 (Life Technologies) e os tubos foram agitados e centrifugados por 5 min a 1.200 x g (Centrifuga Centimicro 212 - Fanem). A fase superior era transferida para um filtro microcon-100® (Amicon) e posteriormente centrifugada por 10 min para a concentração do DNA. Ao filtro, eram adicionados 500 µL de água milliQ autoclavada, sendo este centrifugado por 10 min a 1.200 x g. Este procedimento era repetido duas vezes. Ao filtro, era adicionado novamente água milliQ, em volume de 50 µL. Este foi invertido e transferido para um novo tubo, com posterior centrifugação a 1.200 x g por 5 min, com a finalidade de recuperar o DNA purificado retido na membrana.

Na extração de DNA de dentes o material era previamente limpo com água, hipoclorito de sódio comercial (água sanitária) e etanol. Posteriormente, os dentes eram mantidos em solução de EDTA até que o esmalte dentário estivesse amolecido. Em seguida, os dentes eram lavados com água destilada e centrifugados por 10 min a 12000 x g. Este procedimento era repetido por três vezes consecutivas. Finalmente os dentes eram secos e cortados longitudinalmente com o auxílio de alicate e bisturi, com o objetivo de expor a raiz dentária.

Ao material dentário eram adicionados 3 mL de Tampão de lise SEB e 40 µL proteinase K (0,3 mg/mL), sendo este incubado a 56^0 C por 16 horas. O DNA era extraído pela adição de um volume de fenol:clorofórmio:álcool isoamílico na proporção 25:24:1, na mesma proporção de volume de sobrenadante digerido. O tubo era agitado moderadamente e centrifugado por 10 min a 1.200 x g. Em seguida, a fase aquosa era transferida para um novo tubo, repetindo-se a operação por mais duas vezes. À fase aquosa superior adicionava-se o mesmo volume de clorofórmio:álcool isoamílico na proporção 24:1. O tubo foi então agitado

moderadamente e centrifugado por 10 min a 1.200 x g. O líquido sobrenadante era transferido para um novo tubo, repetindo-se a operação por mais duas vezes. A fase aquosa era transferida para uma coluna microcon-100 ®. Em seguida a coluna era centrifugada a 1.500 x g por 20 min, sendo o sobrenadante descartado. Posteriormente, a coluna era lavada três vezes com água milliQ estéril nas mesmas condições de centrifugação descrita anteriormente. Finalmente, a coluna era invertida e transferida para um novo tubo, sendo adicionado o volume de 50 µL de tampão TE com posterior centrifugação por um período de 10 min a 3.500 x g. A solução de DNA eluída da coluna era estocada a - 20^0 C até a sua utilização.

2.3 **Determinação da concentração de DNA**

A concentração do DNA extraído das amostras biológicas utilizadas no teste de reprodutibilidade e das amostras utilizadas nos testes de padrão misturas de DNAs era determinada por espectrofotometria. Os equipamentos utilizados para a leitura foram os espectrofotômetros modelo GeneQuant (Life Technologies), a 260nm, utilizando uma cubeta de quartzo com 0,5 mm de comprimento óptico, e o modelo NanoDrop 2000 (Thermo scientific), que utiliza um sistema de retenção e medição do volume da amostra diretamente em um pedestal óptico com um espectro de 190 a 840 nm.

2.4 **Seleção dos *loci***

A escolha dos *loci* X-STR utilizados nesse trabalho está relacionada com o grau de polimorfismo de cada marcador, sendo este parâmetro de grande importância para a construção de um multiplex com eficiente poder de discriminação e de exclusão, necessários para análise de vinculação genética e identificação humana por DNA.

As regiões que compõe o LDD (X-STR) Decaplex estão representadas na tabela 1, onde estão apresentadas, também, as respectivas unidades de repetição, fluoróforos, comprimentos de onda de excitação e emissão dos mesmos e

referências utilizadas para esse estudo. Já as seqüências de cada par de *primers* podem ser visualizadas na tabela 2.

Tabela 1 - *Loci* utilizados na construção do LDD (X-STR) Decaplex com respectivas unidades de repetição, fluoróforos, comprimentos de onda de excitação e emissão dos fluoróforos e referências bibliográficas.

Locus	**Repetição**	**Fluoróforo**	**λMax de emissão/Cor/ λMax de excitação**	**Referência**
DXS6807	$(GATA)_X$	6-FAM	500nm/azul/650nm	Poetsch *et al.*, 2005
DXS8378	$(CTAT)_X$	6-FAM	500nm/azul/650nm	Poetsch *et al.,* 2005
DXS7132	$(TCTA)_X$	6-FAM	500nm/azul/650nm	Becker *et al.,* 2008
DXS10074	$(AAGA)_X$	NED	546nm/amarelo/575nm	Becker *et al.,* 2008
DXS7424	$(TAA)_X$	6-FAM	500nm/azul/650nm	Poetsch *et al.*, 2005
DXS7133	$(ATAG)_X$	6-FAM	500nm/azul/650nm	Poetsch *et al.*, 2005
GATA172D05	$(TAGA)_X$	NED	546nm/amarelo/575nm	Gusmão *et al.*, 2008
DXS10101	$(AAAG)_X$	NED	546nm/amarelo/575nm	Becker *et al.*, 2008
DXS8377	$(AGA)_X$	VIC	538nm/verde/554nm	Poetsch *et al.*, 2005
DXS7423	$(TCCA)_X$	VIC	538nm/verde/554nm	chrx-str.org

Tabela 2 - *Loci* utilizados na construção do LDD (X-STR) Decaplex e respectivas seqüências de *primers*.

Locus	**Seqüência dos *primers* (5'- 3')**
DXS6807	GAGCAATGATCTCATTTGCA
	AAGTAAACATGTATAGGAAAAAGCT
DXS8378	CACAGGAGGTTTGACCTGTT
	AACTGAGATGGTGCCACTGA
DXS7132	TGATTAGGAATATCAAAGGCAAA
	CTTCTCTGGTTCTCTAGCTCACAT
DXS10074	TTATTATTTTGCCTACCTCTCTGAGC
	ATTTTCTCTCTGTCTTAGCTCCCATA
DXS7424	CTGCTTGAGTCCAGGAATTCAA
	GAACACGCACATTTGAGAACATA
DXS7133	GCTTCCTTAGATGGCATTCA
	CTTCCAAGAATCAGAAGTCTCC
GATA172D05	TAGTGGTGATGGTTGCACAG
	ATAATTGAAAGCCCGGATTC
DXS10101	TGTTTCAAAATAATATGGGAGTTGG
	TCTTTAATCTCTCACAGCCCCTAC
DXS8377	CACTTCATGGCTTACCACAG
	GACCTTTGGAAAGCTAGTGT
DXS7423	GTCTTCCTGTCATCTCCCAAC
	TAGCTTAGCGCCTGGCACATA

2.5 **Padronização da concentração dos *primers* no LDD (X-STR) Decaplex**

Diferentes versões do sistema multiplex foram preparadas com variáveis concentrações dos pares de *primers*, objetivando a amplificação equilibrada dos *loci* STR. Um ajuste adequado das concentrações de cada marcador foi estabelecido, chegando-se a concentração ideal final de 0,2 µM para todos os dez marcadores, no sistema aqui nomeado como LDD (X-STR) Decaplex. As reações de PCR foram desenvolvidas nos seguintes termocicladores: GeneAmp 2400, GeneAmp 9700 e Veriti (Applied Biosystems).

2.6 **Amplificação dos STR no sistema LDD (X-STR) Decaplex**

Para a amplificação por PCR dos *loci* alvo dos pares de *primers* reunidos no LDD (X-STR) Decaplex, foi utilizado o QIAGEN® Multiplex PCR Kit, com tampão QIAGEN® multiplex PCR master mix 1x, mix dos *primers* na concentração de 0,2 µM e 1-2ng de DNA genômico em um volume final de 6,5 a 10 µL. As reações da PCR foram preparadas contendo amostras de DNA da linhagem celular feminina 9947A (Promega®) na concentração de 2 ng/µL, utilizada como controle positivo das reações. As condições termocíclicas utilizadas nas reações de amplificação estão apresentadas na tabela 3.

Tabela 3 – Condições termocíclicas utilizadas para amplificação do LDD (X-STR) Decaplex.

Temperatura	Tempo		Referência bibliográfica
95º C	15 min		
94º C	30 seg		
		10	
60º C	90 seg	**ciclos**	
72º C	60 seg		Gusmão *et al*, 2008
94º C	30 seg		
		20	
58º C	90 seg	**ciclos**	
72º C	60 seg		
60º C	60 min		

2.7 **Eletroforese Capilar**

A separação dos produtos da PCR foi realizada através de eletroforese capilar no seqüenciador automático ABI Prism 3130 Avant Genetic Analyser Instrument (Applied Biosystems).

2.7.1 Preparação das amostras (produtos da PCR)

Para a separação dos produtos da PCR, alíquotas de 1 µL desse produto foram adicionadas a mistura contendo 8,8 µL de formamida (Hi-Di Formamide, Applied Biosystems) e 0,2 µL de LIZ 500 Size Standard (Applied Biosystem) em cada poço da placa do seqüenciador. Em seguida, a placa contendo as amostras era desnaturada por 5 min a 95^0 C e posteriormente mantida em banho de gelo por 5 min ou até o início da eletroforese capilar.

2.7.2 Eletroforese capilar e detecção

Para a eletroforese foi utilizado o capilar Avant Capillary Array 36 cm (Applied Biosystems) preenchido com o polímero 3130 POP-7^{TM} (Performance Optimized Polymer) [Applied Biosystems]. Nos procedimentos de amplificação de multiplex através de eletroforese capilar, os *primers* são marcados com diferentes fluoróforos que, quando excitado por uma fonte de alta energia, emitem fluorescência em diferentes comprimentos de onda. O resultado obtido é representado por um eletroferograma onde podemos observar picos resultantes do processo de emissão, detecção e transformação do sinal luminoso, que são visualizados em unidades relativas de fluorescência (RFU).

Levando-se em consideração o padrão de herança do cromossomo X, nas amostras femininas a presença de dois picos representa a heterozigosidade para o marcador específico. Já a presença de um único pico é relativa a homozigosidade do marcador, visto que nas mulheres o cromossomo X aparece em dose dupla. Por sua vez, os homens por apresentarem um único cromossomo X, apresentam apenas um único pico.

2.7.3 Análise dos produtos da PCR

Para a análise dos fragmentos e tipagem dos alelos deste trabalho, foi utilizado o *Software* Gene Mapper v 3.01 (Applied Biosystems).

2.7.4 Construção dos bins

Para a genotipagem automática dos alelos no *Software* Gene Mapper, os bins para o sistema LDD (X-STR) Decaplex foram desenvolvidos contendo informações referentes a cada *l*ocus X-STR, suas respectivas faixa de alelos e tamanho em pares de bases. Tal ferramenta é de grande importância por facilitar a genotipagem dos alelos de forma rápida e segura, diminuindo conseqüentemente a chance de erros de análise. Para tal, foram selecionados aproximadamente 50 indivíduos

pertencentes ao banco de dados e, com o auxílio do DNA controle 9947A, o tamanho de cada alelo em pb foi determinado para cada *locus*. Um valor médio foi calculado para cada alelo, com um desvio padrão de 0,5. Arquivos em formato de bloco de notas foram criados contendo informações dos bins referentes a cada alelos, seus respectivos tamanhos em pb, fluoróforos e unidades de repetição (tri ou tetranucleotídicas). Finalmente, os arquivos foram importados para o *software* Gene Mapper para o processamento das análises.

2.7.5 Sequenciamento

Previamente à reação de sequenciamento, os fragmentos eram amplificados em reação monoplex nas mesmas condições utilizadas para o multiplex e com uma concentração de 2 µM. Os produtos de amplificação eram analisados no seqüenciador automático ABI Prism 3130 e posteriormente submetidos ao *Software* Gene Mapper v 3.01 (Applied Biosystems), conforme anteriormente descrito, para análise da qualidade dos produtos amplificados.

Em seguida, 5 µl do produto de amplificação eram purificados com 2 µl de ExoSAP-IT® (USB Corporation) sendo incubado a 37°C por 15 min e posteriormente, 85°C por 15 min. Os produtos purificados eram depois submetidos a uma reação de sequenciamento utilizando o tampão Big Dye Terminator Cycle Sequencing Ready Reaction Kit v3.0 (Applied Biosystems). A reação era realizada em microtubos no volume final de 5 µl, contendo 2,5 µl do produto amplificado purificado, 0,5 µl do *primer* a 2 µM e 2 µl da solução de reação (kit 1:1). Cada amostra era seqüenciada em duas reações diferentes (uma contendo *primer* para a síntese de DNA na direção 5´-3´ "*forward*" e outra na direção 3'-5' "*reverse*"), cada uma contendo um dos pares de *primer* usado na reação de PCR. As reações eram desenvolvidas nos termoclicladores GeneAmp 2400, GeneAmp 9700 e Veriti (Applied Biosystems) nas seguintes condições: 96^0 C por 15 seg, 50^0 C por 1 seg, 60^0 C por 2 min por 25 ciclos.

Após a reação de sequenciamento, a amostra era purificada em resina Sephadex G-50 (Sambrook *et al.,* 1989). Em tal metodologia 750 µl de Sephadex (solução de 10g em 150 mL) era aplicada em uma coluna de purificação própria.

Esta era colocada dentro de um tubo 1,5 mL e posteriormente centrifugada por 5.000 x g por 3 min. O filtrado era descartado e a coluna centrifugada novamente. O volume total do sequenciamento era adicionado à coluna Sephadex, sendo esta centrifugada novamente a 5.000 x g por 3 min.

Finalmente, aos produtos de sequenciamento purificados, eram adicionados 8 µl de formamida (Hi-Di Formamide, Applied Biosystems), e desta mistura, 12 µl eram aplicados em cada poço da placa do seqüenciador. As amostras eram então submetidas à eletroforese capilar em condições desnaturantes já descrita acima, utilizando o seqüenciador automático ABI Prism 3130 preenchido com polímero POP-7^{TM} (Applied Biosystems), e analisadas com o *software* Sequencing Analysis 3.7 (Applied Biosystems).

2.8 **Estratégias de validação**

2.8.1 Determinação do grau de sensibilidade em amostra de DNA

Este estudo objetivou a análise da quantidade mínima de DNA necessária para a amplificação e tipagem dos alelos referentes aos dez *loci* X-STR do LDD (X-STR) Decaplex. Para a determinação de tal parâmetro, foram realizadas reações de PCR nas mesmas condições descritas anteriormente (ver item 5.3 Material e Métodos), utilizando como amostra referência o DNA padrão 9947A em diferentes quantidades. As quantidades de DNA utilizadas foram: 10ng, 5ng, 2ng, 1ng, 0,5ng, 0,25ng, 0,125ng e 0,0625ng. A análise dos produtos da PCR foi realizada por eletroforese capilar e estão descritas no item 3.6.

2.8.2 Estudo de reprodutibilidade das tipagens utilizando o LDD (X-STR) Decaplex

Amostras de DNA de diferentes fontes biológicas foram coletadas e utilizadas neste estudo, tais como: dente, saliva, sangue, cabelo (bulbo), urina e sêmen. Após a extração realizada pelo método orgânico (fenol), os DNA foram amplificados pelo sistema de tipagem comercial *Identifiler*® (Applied Biosystem) contendo marcadores

dos cromossomos autossomos e, posteriormente, amplificadas pelo LDD (X-STR) Decaplex. A análise dos produtos da PCR foi realizada por eletroforese capilar e estão descritas no item 3.6.

2.8.3 Análise das bandas *stutter*

Nesta etapa do estudo, DNA de 15 amostras biológicas foram amplificados com o LDD (X-STR) Decaplex e submetidos à eletroforese capilar no seqüenciador automático ABI3130 (Applied Biosystem).

As alturas dos picos indicam a intensidade relativa de fluorescência emitida, estando, portanto, relacionadas ao processo de emissão, detecção e transformação do sinal luminoso e são medidas em RFU (ver item 3.6.2 Material e Métodos). Para este estudo, a razão correspondente entre a altura do pico do alelo e a altura correspondente com o pico *stutter* foi determinada para cada marcador X-STR.

2.8.4 Análise do padrão de misturas de DNA

A determinação de perfis biológicos obtidos a partir de uma amostra contendo mistura de DNA foi realizada a partir dos materiais biológicos extraídos de dois indivíduos, um do sexo feminino e um do sexo masculino. As amostras de DNA foram extraídas pelo método salting-out (ver item 3.2.1). Para a realização deste estudo, os DNA dos respectivos indivíduos foram misturados e amplificados, em uma única reação de PCR, em diferentes proporções, tais como 10:1ng; 5:1ng; 2,5ng:1; 1:1ng; 1:2,5ng; 1:5ng; e 1:10ng. A análise foi realizada inicialmente com o sistema LDD (X-STR) Decaplex e, posteriormente, com o sistema *Identifiler*.

2.9 **Análise estatística**

As freqüências alélicas e haplotípicas, estimativa da diversidade gênica, teste de Equilíbrio de Hardy-Weinberg (HWE), Heterosigosidade observada (Ho) e

Heterosigozidade esperada (He), análise do desequilíbrio de ligação e comparação entre as sub-populações masculina e feminina (teste exato de diferenciação) foram realizadas através do *software* Arlequin versão 3.01 (Excoffier *et al*, 2005). O cálculo de diversidade haplotípica foi realizado conforme Nei, 1987.

Conteúdo de Informação Polimórfica (PIC) e Poder de Exclusão (PE) foram calculados através do programa PowerStats v1.2 (Promega Corporation). A Chance Média de Exclusão de trios envolvendo filhas - *Mean Exclusion Chance* – *(*MECt) e duos pai/filha (MECd), Poder de Discriminação entre mulheres (PDm) e homens (PDh) foram calculados através das fórmulas citadas por Desmarais e colaboradores, 1998. Quando se realizaram análises estatísticas envolvendo múltiplos testes de significância independente, aplicou-se a correção de Bonferroni (1936), dividindo o nível de significância de cada teste (0,05) pelo número de testes realizados.

No estudo de caso, foram calculados alguns parâmetros de paternidade como: Índice de Vinculação (IV) que reflete quantas vezes é mais provável que a pessoa que está sendo analisada seja o pai biológico do que outro individuo selecionado ao acaso, sendo calculado pelo inverso da probabilidade de vinculação; O Índice Cumulativo de Vinculação (ICV) que é calculado através do produto dos IV de cada *locus* individualmente (Hart & Clarck, 1997); a Probabilidade de Vinculação (PV), que está baseado na teoria de Bayes (NCRII report,1996); o Poder de Exclusão (PE) referente ao percentual de indivíduos que teria a paternidade biológica excluída, sendo considerada a ocorrência de incompatibilidade genética (Chakraborty *et al*, 1999); o Poder de Exclusão acumulado (PEac) que é calculado através do produto do PE de cada *locus* individualmente; a Probabilidade de Coincidência (PC) referente a probabilidade de um mesmo genótipo ser originado de mais de uma pessoa (Hart & Clarck, 1997) e a Razão de Verossimilhança (RV) que indica a raridade do perfil genotípico de um indivíduo, sendo esta, a medida da força da prova considerando a hipótese de que dois perfis apresentam a mesma origem (Weir, 1996).

2.10 **Lista de soluções e tampões utilizados**

2.10.1 Extração de DNA Genômico

Tampão de lise de hemáceas (TLH)	
NH_4CL	155 mM
$KHCO_3$	10 mM
EDTA	1 mM

Tampão TE (pH 7,6)	
Tris-Cl pH 7,6	10 mM
EDTA pH 8,0	1 mM

Tampão de Lise de núcleo (NLB)	
Tris-Cl pH 8	10 mM
NaCl	0,4 mM
EDTA	2 mM

Tampão fosfato PBS (pH 7,4)	
NH_4CL	155 mM
$KHCO_3$	10 mM
NaCl	136 mM

Tampão de extração SEB (pH 8)	
EDTA pH 8,0	10 mM
NaCl	100 mM
Tris-Cl pH 8,0	10 mM
SDS	2%

2.10.2 Amplificação da Polimerase em cadeia

Multiplex Qiagen Master Mix 2x
HotStar Taq^{R} DNA Polimerase
Multiplex PCR Buffer pH 8,7
dNTP Mix

3 **RESULTADOS**

Este trabalho teve como um dos objetivos produzir um sistema multiplex que promovesse a amplificação simultânea de múltiplos *loci* do cromossomo X humano com aplicações em genética populacional e forense. Para isto, foram selecionados dez *loci* com elevados graus de polimorfismos, conforme verificado em estudos anteriores (Szibor, 2007; Becker *et al.,* 2008).

Com o sistema multiplex, que denominamos de LDD (X-STR) Decaplex, foi realizada a genotipagem de 170 indivíduos nascidos na cidade do Rio de Janeiro, autodenominados afrodescendentes, participantes de investigações de vínculos genéticos, conforme item 3.1.1, Material e Métodos.

Um estudo de validação objetivando a utilização do sistema multiplex LDD (X-STR) Decaplex em análises forenses foi desenvolvido. Para isso, testes de sensibilidade do sistema multiplex, produção de artefatos pela reação da PCR, estudo de reprodutibilidade e discriminação de perfis genéticos em amostras de misturas de DNAs foram realizados.

3.1 **Desenvolvimento de sistema multiplex**

Etapas caracterizadas pela produção inicial de um sistema hexaplex (Tavares *et al.*, 2008) e outro octaplex se relacionam com o desenvolvimento da versão atual reunindo dez marcadores X-STR, foco deste estudo. A construção do sistema LDD (X-STR) Decaplex objetivou a amplificação simultânea de *loci* X-STR altamente polimórficos, padronização das concentrações molares de utilização bem como a qualidade dos produtos de amplificação.

Com o sistema LDD (X-STR) Decaplex foram realizadas amplificações a partir de alíquotas dos DNA dos indivíduos nas condições mostradas na tabela 4. Na figura 3 é mostrado o eletroferograma relativo à amplificação de 2ng do DNA controle positivo 9947A. Os bins desenvolvidos para o sistema LDD (X-STR) Decaplex (ver item 3.7.4 – Material e Métodos), referentes aos alelos de cada *locus* do multiplex, são representados pelas barras verticais na cor cinza.

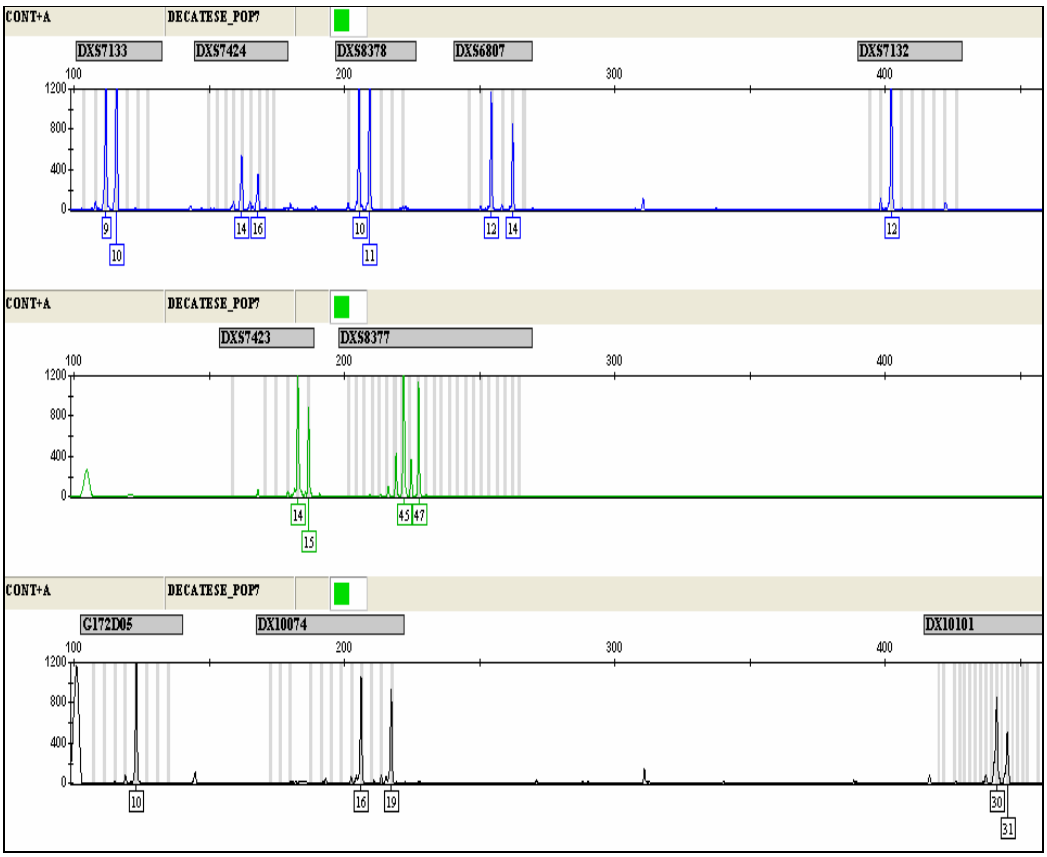


Figura 3- Alelos X-STR a partir de DNA controle 9947 utilizando o LDD (X-STR) Decaplex.

Os pares de *primers* relacionados com a amplificação dos dez *loci* que compõe o LDD (X-STR) Decaplex foram marcados com três fluoróforos diferentes: 6-FAM, VIC e NED.

No eixo das abscissas (X) são apresentados os produtos de amplificação (alelos) em pb, que encontram-se também nomeados de acordo com nomenclatura internacional.

No eixo das ordenadas (Y) são mostradas em RFU as quantidades de produtos de amplificação (alelos) gerados em cada *loci*.

3.2 **Construção do banco de freqüências alélicas**

Após a determinação dos tamanhos, em pares de bases, utilizando o *software* Gene Mapper, os alelos amplificados em cada uma das regiões X-STR foram tabulados com a finalidade de identificar os alelos presentes em cada *locus*, estimar as suas freqüências, bem como avaliar parâmetros estatísticos, forenses e de vinculação genética.

Alelos amplificados pelo sistema LDD (X-STR) Decaplex podem ser visualizados nas figuras 4 e 5, nas quais são mostrados os produtos das amplificações de amostras de um DNA feminino e outro masculino, respectivamente. Verifica-se que os picos representativos dos produtos amplificados (alelos X-STR) apresentam intensidades de fluorescência na faixa entre 800 a 1.200 RFU. Um total de 272 alelos foram tipados em 102 mulheres e 68 homens, cujos perfis genéticos e haplotípicos são apresentados nos apêndices B e C, respectivamente.

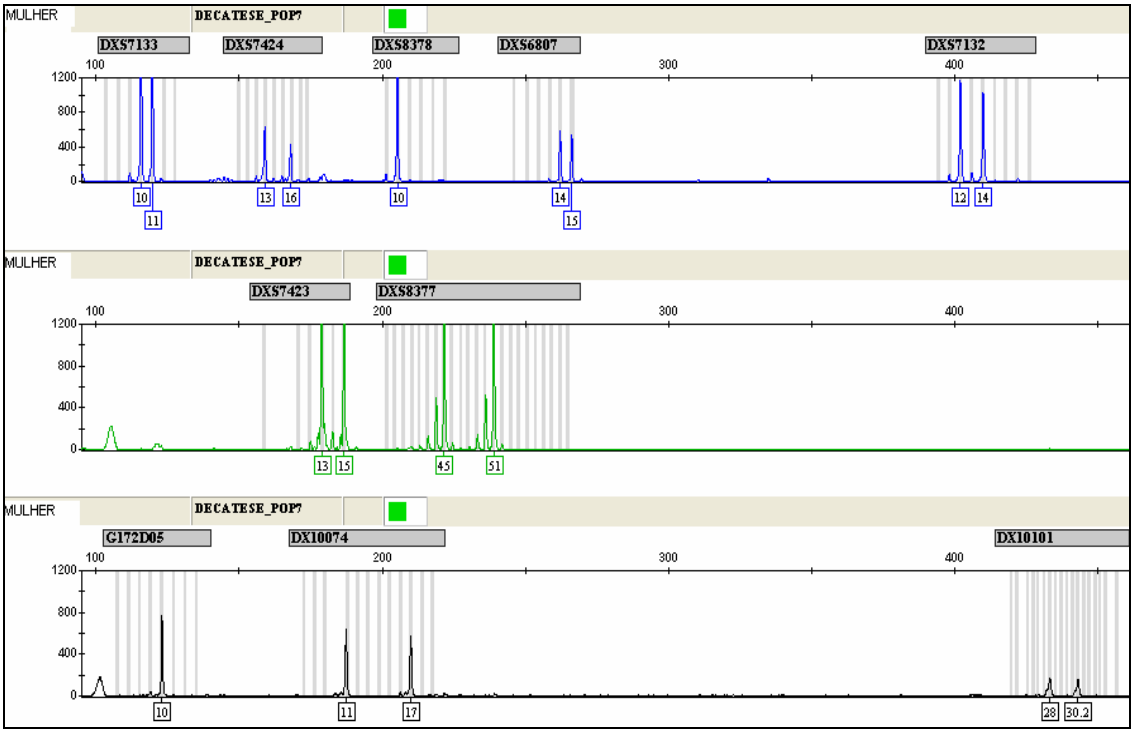


Figura 4- Alelos X-STR amplificados a partir de DNA de um indivíduo afrodescendente, do sexo feminino, do Rio de Janeiro utilizando o LDD (X-STR) Decaplex.

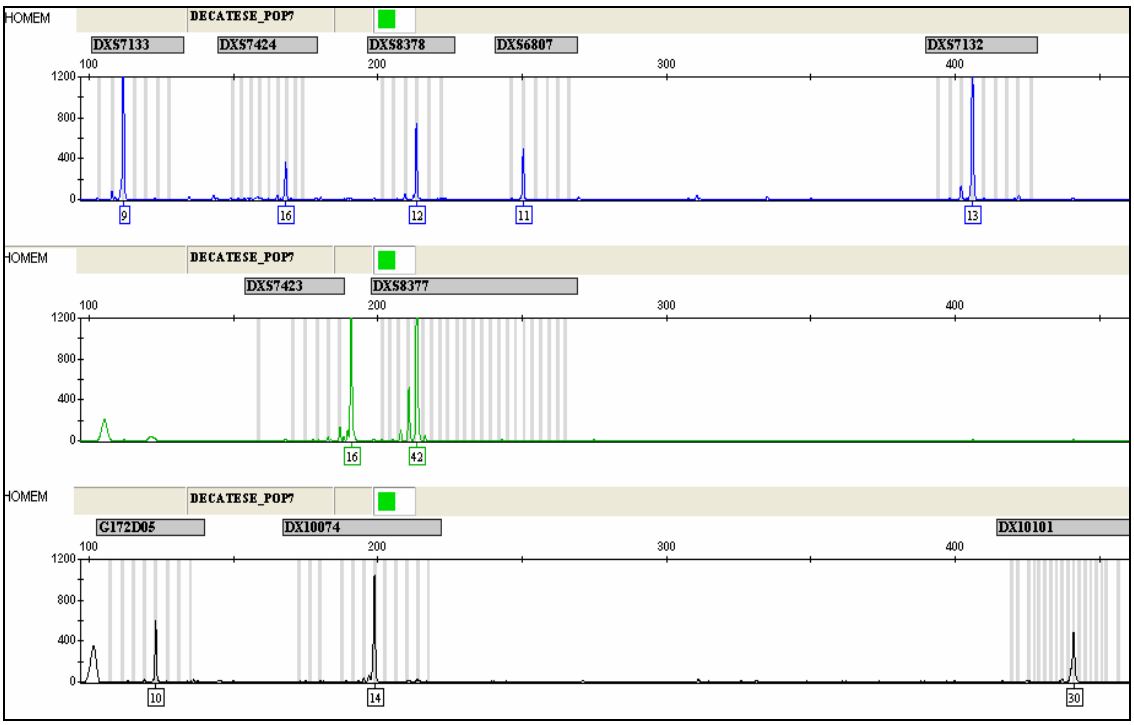


Figura 5- Alelos X-STR amplificados a partir de DNA de um indivíduo afrodescendente, do sexo masculino, do Rio de Janeiro utilizando o LDD (X-STR) Decaplex.

3.2.1 Estimativas das freqüências alélicas

A distribuição das freqüências dos alelos verificados, em cada uma das regiões microssatélites do LDD (X-STR) Decaplex, no conjunto dos indivíduos autodenominados afrodescendentes nascidos no Rio de Janeiro, que constituem o presente trabalho, é apresentada nas tabelas 4, 5 e 6. As mesmas referem-se às freqüências alélicas observadas para homens (n=68), mulheres (n= 102) e para a amostra total (n= 272), respectivamente. As freqüências alélicas foram utilizadas como base de dados para os cálculos dos parâmetros estatísticos e de interesse forense.

Através da análise da tabela 6, pode ser observado que o alelo 11 foi o mais freqüente para o *locus* DXS7133, representando 56,62% dos alelos genotipados. Para o *locus* DXS7424, o alelo 15 foi o mais freqüente (freqüência de 22,06%). Para os *loci* DXS8378 e DXS6807, o alelo 11 foi o mais freqüente na amostra genotipada, apresentando valores de 33,09% e 48,16%, respectivamente. O alelo 14 foi o mais freqüente nos *loci* DXS7132 e DXS7423, representando 28,68% e 41,54%, respectivamente. No *locus* DXS8377, os alelos 46 e 49 foram os mais freqüentes (freqüência de 10,66%). O Alelo 9 foi o mais frequente para o *locus* GATA172D05, apresentando valor de 23,9%. O *locus* DXS10074 apresentou uma freqüência de 18,38% para o seu alelo mais freqüente (alelo 16). E para o *locus* DXS10101, os alelos 31 e 32 foram os mais freqüentes, representando 12% para cada um dos alelos genotipados.

Com relação ao intervalo de variação dos alelos na amostra populacional estudada (tabela 9), este variou de 9 a 15 o *locus* DXS7133 (7 alelos), de 9 a 18 para o *locus* DXS7424 (10 alelos), de 9 a 13 para o *locus* DXS8378 (5 alelos), de 11 a 17 para o *locus* DXS6807 (7 alelos), de 10 a 18 para o *locus* DXS7132 (9 alelos), de 8 a 17 para o *locus* DXS7423 (7 alelos), de 39 a 59 para o *locus* DXS8377 (21 alelos), para o *locus* GATA172D05 o intervalo foi de 6 a 12 (7 alelos), para o *locus* DXS10074 foi de 7 a 20 (14 alelos) e por último, para o *locus* DXS10101 o intervalo de variação dos alelos foi de 24.2 a 35 (18 alelos) .

Tabela 4 – Distribuição das freqüências dos alelos verificados nos dez *loci* X-STR estudados na fração masculina da amostra populacional constituída por indivíduos do Rio de Janeiro autodenominados afrodescendentes.

DXS7133		**DXS7424**		**DXS8378**		**DXS6807**		**DXS7132**	
alelo	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*
9	0,1765	**11**	0,1029	**9**	0,0147	**11**	0,500	**12**	0,1324
10	0,1765	**12**	0,0882	**10**	0,2794	**12**	0,0147	**13**	0,3382
11	0,5441	**13**	0,2500	**11**	0,3529	**13**	0,0588	**14**	0,2206
12	0,0735	**14**	0,1471	**12**	0,3088	**14**	0,0294	**15**	0,2206
13	0,0294	**15**	0,2941	**13**	0,0441	**15**	0,2794	**16**	0,0441
		16	0,1176			**16**	0,1176	**17**	0,0294
								18	0,0147

DXS7423		**DXS8377**		**Gata172D05**		**DXS10074**		**DXS10101**	
alelo	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*
13	0,0735	**40**	0,0294	**6**	0,2794	**7**	0,1029	**24.2**	0,1029
14	0,3676	**41**	0,0294	**7**	0,0294	**8**	0,0588	**26.2**	0,0147
15	0,4412	**42**	0,0294	**8**	0,1912	**11**	0,0147	**28**	0,0441
16	0,1029	**43**	0,0441	**9**	0,2794	**12**	0,0735	**28.2**	0,0294
17	0,0147	**44**	0,0147	**10**	0,0882	**13**	0,0294	**29**	0,1029
		45	0,0294	**11**	0,0882	**14**	0,1471	**29.2**	0,0441
		46	0,1618	**12**	0,0441	**15**	0,1912	**30**	0,0735
		47	0,0882			**16**	0,1912	**30.2**	0,0735
		48	0,0882			**17**	0,0735	**31**	0,1324
		49	0,1029			**18**	0,0882	**31.2**	0,0588
		50	0,0882			**19**	0,0294	**32**	0,1176
		51	0,0735					**32.2**	0,0588
		52	0,0735					**33**	0,1029
		53	0,0294					**34**	0,0294
		54	0,0441					**35**	0,0147
		55	0,0294						
		56	0,0294						
		58	0,0147						

Tabela 5 - Distribuição das freqüências dos alelos verificados nos dez *loci* X-STR estudados na fração feminina da amostra populacional constituída por indivíduos do Rio de Janeiro autodenominados afrodescendentes.

DXS7133		**DXS7424**		**DXS8378**		**DXS6807**		**DXS7132**	
alelo	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*
9	0,1912	**9**	0,0049	**9**	0,0196	**11**	0,4755	**10**	0,0049
10	0,1176	**10**	0,0147	**10**	0,2941	**12**	0,0098	**11**	0,0196
11	0,5735	**11**	0,0637	**11**	0,3235	**13**	0,0294	**12**	0,1373
12	0,0833	**12**	0,0931	**12**	0,3284	**14**	0,2353	**13**	0,2108
13	0,0245	**13**	0,2059	**13**	0,0343	**15**	0,2010	**14**	0,3088
14	0,0049	**14**	0,2108			**16**	0,0343	**15**	0,2402
15	0,0049	**15**	0,1961			**17**	0,0147	**16**	0,0441
		16	0,1422					**17**	0,0196
		17	0,0490					**18**	0,0147
		18	0,0196						

DXS7423		**DXS8377**		**Gata172D05**		**DXS10074**		**DXS10101**	
alelo	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*
8	0,0098	**39**	0,0196	**6**	0,2206	**7**	0,0833	**24.2**	0,0049
12	0,0098	**40**	0,0049	**7**	0,0588	**8**	0,0245	**25.2**	0,0049
13	0,0784	**41**	0,0245	**8**	0,1667	**10**	0,0098	**26**	0,0147
14	0,4314	**42**	0,0637	**9**	0,2255	**11**	0,0294	**26.2**	0,0147
15	0,3578	**43**	0,0294	**10**	0,2010	**12**	0,1078	**27**	0,0294
16	0,0833	**44**	0,0784	**11**	0,1029	**13**	0,0441	**28**	0,0441
17	0,0294	**45**	0,0637	**12**	0,0245	**14**	0,1029	**28.2**	0,0245
		46	0,0882			**15**	0,1471	**29**	0,1078
		47	0,0833			**16**	0,1814	**29.2**	0,0784
		48	0,0686			**16.2**	0,0098	**30**	0,1029
		49	0,1078			**17**	0,1324	**30.2**	0,0735
		50	0,0931			**18**	0,0980	**31**	0,1275
		51	0,0882			**19**	0,0245	**31.2**	0,0588
		52	0,0588			**20**	0,0049	**32**	0,1225
		53	0,0343					**32.2**	0,0539
		54	0,0196					**33**	0,1176
		55	0,0441					**34**	0,0147
		56	0,0147					**35**	0,0049
		57	0,0049						
		58	0,0049						
		59	0,0049						

Tabela 6 – Distribuição das freqüências dos alelos verificados nos dez *loci* X-STR estudados na fração total da amostra populacional constituída por indivíduos do Rio de Janeiro autodenominados afrodescendentes.

DXS7133		**DXS7424**		**DXS8378**		**DXS6807**		**DXS7132**	
alelo	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*
9	0,1875	**9**	0,0037	**9**	0,0184	**11**	0,4816	**10**	0,0037
10	0,1324	**10**	0,0110	**10**	0,2904	**12**	0,0110	**11**	0,0147
11	0,5662	**11**	0,0735	**11**	0,3309	**13**	0,0368	**12**	0,1360
12	0,0809	**12**	0,0919	**12**	0,3235	**14**	0,1838	**13**	0,2426
13	0,0257	**13**	0,2169	**13**	0,0368	**15**	0,2206	**14**	0,2868
14	0,0037	**14**	0,1949			**16**	0,0551	**15**	0,2353
15	0,0037	**15**	0,2206			**17**	0,0110	**16**	0,0441
		16	0,1360					**17**	0,0221
		17	0,0368					**18**	0,0147
		18	0,0147						

DXS7423		**DXS8377**		**Gata172D05**		**DXS10074**		**DXS10101**	
alelo	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*	*alelo*	*freq.*
8	0,0074	**39**	0,0147	**6**	0,2353	**7**	0,0882	**24.2**	0,0294
12	0,0074	**40**	0,0110	**7**	0,0515	**8**	0,0331	**25.2**	0,0037
13	0,0772	**41**	0,0257	**8**	0,1728	**10**	0,0074	**26**	0,0110
14	0,4154	**42**	0,0551	**9**	0,2390	**11**	0,0257	**26.2**	0,0147
15	0,3787	**43**	0,0331	**10**	0,1728	**12**	0,0993	**27**	0,0221
16	0,0882	**44**	0,0625	**11**	0,0993	**13**	0,0404	**28**	0,0441
17	0,0257	**45**	0,0551	**12**	0,0294	**14**	0,1140	**28.2**	0,0257
		46	0,1066			**15**	0,1581	**29**	0,1066
		47	0,0846			**16**	0,1838	**29.2**	0,0699
		48	0,0735			**16.2**	0,0074	**30**	0,0956
		49	0,1066			**17**	0,1176	**30.2**	0,0735
		50	0,0919			**18**	0,0956	**31**	0,1287
		51	0,0846			**19**	0,0257	**31.2**	0,0588
		52	0,0625			**20**	0,0037	**32**	0,1213
		53	0,0331					**32.2**	0,0551
		54	0,0257					**33**	0,1140
		55	0,0404					**34**	0,0184
		56	0,0184					**35**	0,0074
		57	0,0037						
		58	0,0074						
		59	0,0037						

3.2.2 Confirmação da ocorrência de alelo microvariante

Reações de sequenciamento do DNA foram realizadas para a confirmação do alelo microvariante 24.2 referente ao *locus* DXS10101 (ver item 3.7.5, Material e Métodos). Para controle da reação, foi também seqüenciado um alelo não microvariante, no caso, o 26. Através da análise dos resultados utilizando *software* Seqncing Analysis, o alelo 26 do *locus* DXS10101 continha uma seqüência composta por 26 repetições AAAG, e apresentava adicionalmente na posição inicial quatro timinas residuais, enquanto o alelo 24.2 era formado por 24 repetições AAAG e seis timinas residuais, demonstrando-se a inserção de duas timinas (figuras 6 e 7). De acordo com os nossos resultados, foi confirmado ser um alelo microvariante o produto de amplificação nomeado de 24.2.

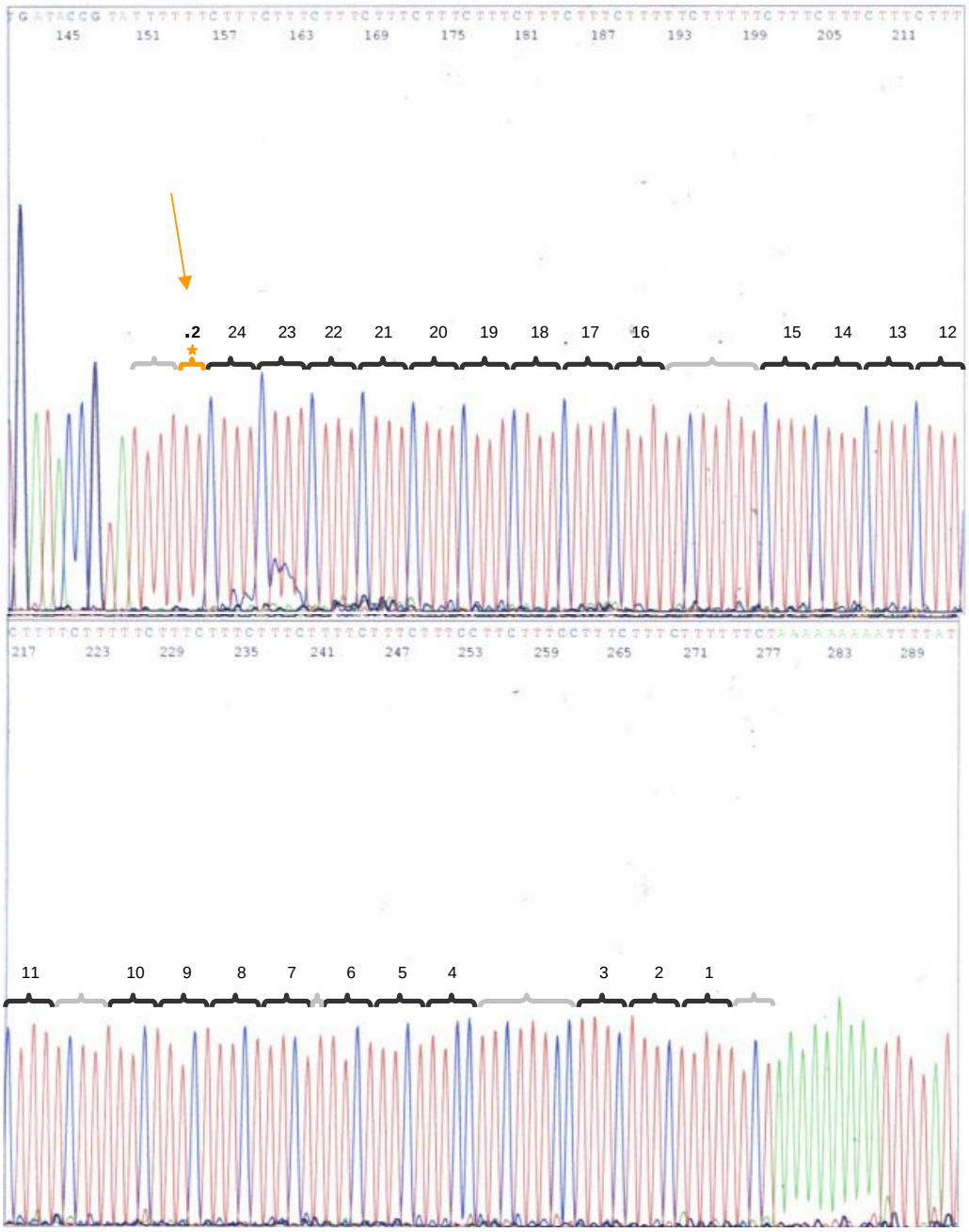


Figura 6- Resultado da reação de sequenciamento do fragmento 24.2 do *locus* DXS10101 (orientação 3´-5').

Legenda: T = Timina; A = Adenina; G = Guanina; C= Citosina; As chaves em preto representam a seqüência repetitiva **AAAG**; A chave/seta em amarelo sinaliza a inserção das bases microvariantes.

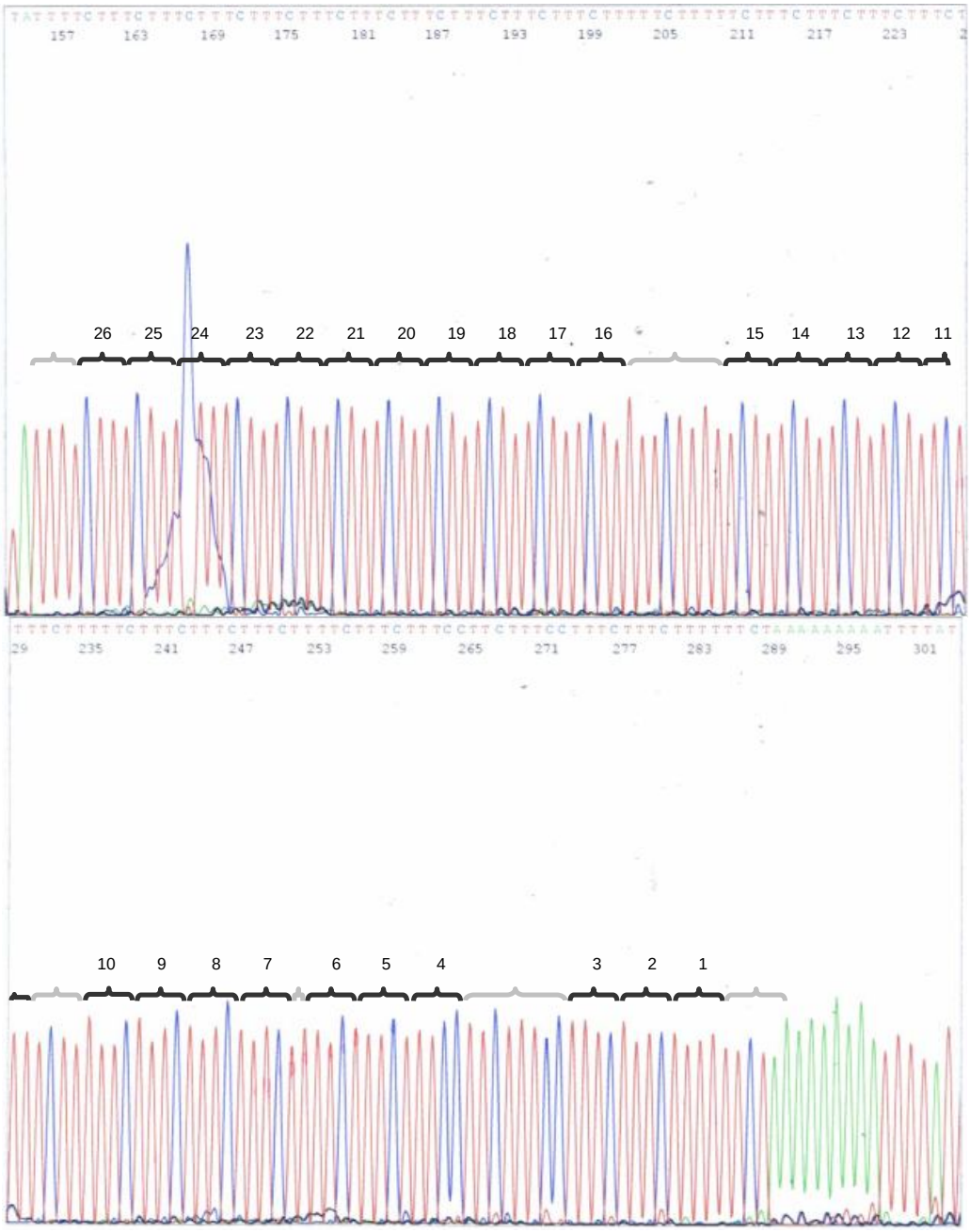


Figura 7- Resultado da reação de sequenciamento do fragmento 26 do *locus* DXS10101 (orientação 3´-5').

Legenda: T = Timina; A = Adenina; G = Guanina; C= Citosina; As chaves em preto representam a seqüência repetitiva **AAAG**.

3.2.3 Verificação das freqüências haplotípicas

As freqüências de haplótipos formados pelos alelos amplificados, em cada um dos dez *loci* do LDD (X-STR) Decaplex, foi estimada com base na genotipagem dos indivíduos do sexo masculino, cuja transmissão alélica ocorre na forma de um haplótipo simples devido ao fato de apresentarem um único cromossomo X.

Através da análise dos resultados em *software* Gene Mapper, pôde-se verificar a ocorrência de haplótipos únicos. No apêndice C, são apresentados os diferentes haplótipos resultantes da análise genética de 68 homens para os dez *loci* X-STR. A freqüência de cada haplótipo é 0,014706 e a diversidade haplotípica 0,985293882.

3.2.4 Análise de parâmetros estatísticos

Os resultados dos testes estatísticos realizados com base nas freqüências dos alelos tipados nos dez *loci* X-STR são mostrados na tabela 7. Os dados foram utilizados para a análise da heterosigozidade observada, heterosigozidade esperada, do equilíbrio de Hardy-Weinberg e da diversidade genética.

Os valores do parâmetro heterozigosidade observada equivalem à proporção observada de indivíduos heterozigotos para um dado *locus* na população em estudo. Estes valores variaram entre 0,5686 para o *locus* DXS6807 e 0,9509 para o *locus* DXS8377. Já os valores de heterozigosidade esperada, que mede a proporção esperada de indivíduos heterozigotos em um determinado *locus* com base modelo de Hardy-Weinberg, variaram de 0,6750 para o *locus* DXS7423 e 0,9336 para o *locus* DXS8377.

A análise do equilíbrio de Hardy-Weinberg foi realizada com o objetivo de verificar se as freqüências alélicas e genotípicas, relativas às genotipagens dos indivíduos da amostra populacional estudada nos dez *loci* STR, estavam de acordo com as proporções esperadas para uma população em equilíbrio. Para tal verificação, foi realizado o teste estatístico de ajustamento do qui-quadrado através do *software* arlequin.

Para um determinado *locus* se apresentar em equilíbrio de Hardy-Weinberg, os testes devem mostrar valores de *p* maiores que 0,05, confirmando que as proporções alélicas e genotípicas observadas encontram-se em conformidade com as proporções genotípicas esperadas, caracterizando, conseqüentemente, a independência de transmissão dos alelos do *locus* em referência. Os valores observados inicialmente revelaram que os *loci* DXS6807, DXS7423, GATA172D05 e DXS10101 apresentaram desvio do equilíbrio de Hardy-Weinberg. Utilizando a correção de Bonferroni, que é aplicada em análises estatísticas envolvendo testes com valores de $p \leq 0,05$, foram obtidos valores de *p* corrigidos, indicando que os referidos *loci* encontravam-se em equilíbrio segundo Hardy-Weinberg.

O teste exato de diferenciação realiza a análise entre as sub-populações masculina e feminina, levando-se em consideração as freqüências alélicas obtidas para as mesmas. Valores menores que 0,05 indicam significância de distinção no *locus* analisado. A comparação entre as duas sub-populações mostrou diferença significante apenas para o *locus* DXS6807 (valor de *p* de 0,00037). Mantendo-se a diferenciação como significativa mesmo após a correção de Bonferroni. Para os demais *loci,* não foram observadas diferenças significativas entre as frações masculina e feminina da amostra populacional em estudo.

A diversidade gênica foi calculada com base nas freqüências de todas as formas alélicas verificadas para o conjunto dos dez *loci* estudados na população. A análise da diversidade gênica permite a comparação dos níveis de heterozigosidade dentro da população. Esse parâmetro leva em consideração o grau de polimorfismo de cada *locus*, e conseqüentemente, o seu poder de discriminação na amostra estudada. Através dos dados obtidos, pôde-se verificar que os índices de diversidade variaram entre 0,6218 para o *locus* DXS7133, e 0,9327 para o *locus* DXS8377.

Tabela 7 - Parâmetros estatísticos para os *loci* do sistema LDD (X-STR) Decaplex estudados na amostra populacional do Rio de Janeiro autodenominada afrodescendente.

Marcador	HO	HE	EHW	Teste Exato	Diversidade Gênica
DXS7133	0.5882	0.6161	0.69848	0.8814	0.6218
DXS7424	0.8333	0.8429	0.50837	0.2921	0.8353
DXS8378	0.6960	0.7028	0.74582	1.0000	0.7024
DXS6807	0.5686	0.6791	0.01207*****	0.00037*****	0.6835
DXS7132	0.7647	0.7846	0.64853	0.5782	0.7850
DXS7423	0.6078	0.6750	0.04393*****	0.8689	0.6720
DXS8377	0.9509	0.9336	0.23251	0.6108	0.9327
GATA172D05	0.7549	0.8217	0.00479*****	0.2883	0.8175
DXS10074	0.8333	0.8893	0.28905	0.9096	0.8867
DXS10101	0.8627	0.9151	0.01095*****	0.2145	0.9171

HO= Heterozigosidade Observada; **HE=** Heterozigosidade Esperada; **EHW=** Equilíbrio de Hardy-Weinberg; ***** indica valores de *p* onde foram aplicadas correção de Bonferroni para análises múltiplas.

3.2.5 Análise do desequilíbrio de ligação

O teste exato para análise do desequilíbrio de ligação foi realizado para todos os pares de *loci* envolvidos neste estudo (tabela 8). Com exceção do par DXS6807-8377, os resultados não se revelaram consistentes para afirmar se que ocorra uma associação entre os pares de *loci*, demonstrando, portanto, encontrarem-se os mesmos em equilíbrio de ligação ($p \geq 0,05$).

Após a aplicação da correção de Bonferroni para análises múltiplas, foi obtido um valor de *p* ajustado para o par DXS6807-8377 ($p = 0,001111$), comprovando a não associação gamética. A correção de Bonferroni é feita com o objetivo de manter

o nível de significância em 0,05. Neste teste estatístico foi considerado 0,05/n, sendo n o número de testes realizados igual a 45 (um para cada par de *loci*).

Tabela 8 – Análise do desequilíbrio de ligação para os pares de marcadores na população masculina do Rio de janeiro autodenominada afrodescendente.

Pares de *loci*	Valor de *p* ± dp
7133-7424	0.577327 ± 0.004527
7133-8378	0.445941 ± 0.002721
7424-8378	0.771881 ± 0.002670
7133-6807	0.510792 ± 0.003025
7424-6807	0.914455 ± 0.002297
8378-6807	0.330594 ± 0.002379
7133-7132	0.642871 ± 0.003445
7424-7132	0.484455 ± 0.004426
8378-7132	0.387228 ± 0.002818
6807-7132	0.183564 ± 0.002164
7133-7423	0.345446 ± 0.002925
7424-7423	0.156040 ± 0.002506
8378-7423	0.603267 ± 0.003857
6807-7423	0.407624 ± 0.002363
7132-7423	0.200990 ± 0.002657
7133-8377	0.160297 ± 0.002991
7424-8377	0.361287 ± 0.002017
8378-8377	0.383465 ± 0.003083
6807-8377	0.038812***** ± 0.001621
7132-8377	0.416040 ± 0.001428
7423-8377	0.418416 ± 0.003987
7133-GATA172D05	0.273168 ± 0.002256
7424-GATA172D05	0.458317 ± 0.003402
8378-GATA172D05	0.667129 ± 0.002894
6807-GATA172D05	0.910495 ± 0.001635
7132-GATA172D05	0.874257 ± 0.001917
7423-GATA172D05	0.882574 ± 0.001935
8377-GATA172D05	0.858416 ± 0.001121

7133-10074	0.899208 ± 0.001544
7424-10074	0.868614 ± 0.002342
8378-10074	0.930396 ± 0.001709
6807-10074	0.199505 ± 0.002936
7132-10074	0.499307 ± 0.001804
7423-10074	0.377327 ± 0.002417
8377-10074	0.204059 ± 0.001924
GATA172D05-10074	0.997327 ± 0.000442
7133-10101	0.507624 ± 0.002814
7424-10101	0.191485 ± 0.001618
8378-10101	0.258713 ± 0.002395
6807-10101	0.620693 ± 0.001733
7132-10101	0.608515 ± 0.002017
7423-10101	0.752574 ± 0.002336
8377-10101	0.822871 ± 0.001004
GATA172D05-10101	0.196733 ± 0.002100
10074-10101	0.418614 ± 0.001443

p* =** valor obtido com o cálculo do teste exato; **dp=** desvio padrão; ** indica valor de *p* onde foi aplicada correção de Bonferroni para análises múltiplas.

3.2.6 Parâmetros forenses

Investigações de vínculo genético e forense requerem a análise de diversos parâmetros estatísticos de interesse forense objetivando a correta caracterização de um perfil genético. Tendo como base a análise de marcadores STR do cromossomo X, parâmetros como: a Chance Média de Exclusão em trios (MECt), relativo à investigação de paternidade envolvendo uma criança, sua genitora e o suposto pai, a Chance Média de Exclusão em duos (MECd), relacionado com a investigação de paternidade envolvendo uma criança e o seu suposto pai, o Poder de Discriminação (PD) em mulheres e homens, o Conteúdo de Informação de Polimorfismo (PIC) e o Poder de Exclusão (PE) são de fundamental importância. A tabela 9 mostra esses valores para os dez *loci* STR alvos do sistema de tipagem LDD (X-STR) Decaplex.

O *locus* DXS7133 apresentou o menor valor de MEC trios (0,6048), o menor valor de MEC duos (0,43), o menor PD no componente masculino (0,6195) e o menor PIC, com valor de 0,5732.

O *locus* DXS8377 apresentou o maior valor de MEC em trios envolvendo pai, mãe e filha (0,9248), o maior valor de MEC em duos pai e filha (0,8644), o maior PD no componente feminino (0,9905) e no componente masculino (0,9293) e o maior PIC, com valor de 0,9245. Além de apresentar o maior valor para o PE (0,9002).

Na tabela 10 pode ainda ser observado que o *locus* DXS7423 apresentou o menor poder de discriminação no componente feminino da amostra populacional em estudo, com valor de 0,8319. E que o *locus* DXS6807 apresentou o menor PE, com valor de 0,2549.

Os valores de MEC e de PD total ou acumulado tende a 100% quando são considerados dez ou mais *loci.* A comparação entre os valores apresentados na tabela mostrou que o valor de MEC obtido para trios foi de 99,9999819% e de MEC para duos de 99,998541%, respectivamente. O PD em mulheres e homens demonstrou índices acumulados de 99,99999999995% e 99,9999942%, respectivamente.

Tabela 9- Parâmetros estatísticos de avaliação da eficiência forense para os *loci* do sistema LDD (X-STR) Decaplex na amostra populacional do Rio de Janeiro autodenominada afrodescendente.

Marcador	MECt	MECd	PDm	PDh	PIC	PE
DXS7133	0,6048	0,43	0,8405	0,6195	0,5732	0,2769
DXS7424	0,8106	0,6966	0,9502	0,8323	0,8185	0,6623
DXS8378	0,6397	0,4942	0,8498	0,6998	0,6391	0,4222
DXS6807	0,6365	0,4908	0,8537	0,6809	0,6265	0,2549
DXS7132	0,7483	0,6178	0,9187	0,7821	0,7467	0,5182
DXS7423	0,6107	0,4661	0,8319	0,6695	0,6151	0,3004
DXS8377	0,9248	0,8644	0,9905	0,9293	0,9245	0,9002
GATA172D05	0,7882	0,667	0,9394	0,8144	0,7919	0,5182
DXS10074	0,8722	0,7828	0,9752	0,8834	0,8713	0,6559
DXS10101	0,9072	0,8362	0,9861	0,9137	0,8539	0,6794
Acumulado	0,999999819	0,99998541	0,9999999999995	0,999999942		

MECt= Chance Média de Exclusão em trios pai/mãe/filha; MECd= Chance Média de Exclusão em duos pai/filha; PDm= Poder de Discriminação na amostra feminina; PDh= Poder de discriminação na amostra masculina; PIC= Conteúdo de Informação de Polimorfismo e PE = Poder de Exclusão.

3.3 **Estratégias de validação**

3.3.1 Determinação do grau de sensibilidade

A quantidade de produto gerado na reação de PCR é dependente da quantidade de DNA necessária para a eficiente amplificação de todos os alelos que constituem o perfil genético de um indivíduo. Visando verificar a quantidade mínima de DNA que se relacionava com a amplificação dirigida pelo sistema LDD (X-STR) Decaplex dos alelos de todos os dez *loci* X-STR, diferentes reações de PCR contendo DNA 9947A nas quantidades de 10ng, 5ng, 2ng, 1ng, 0,5ng, 0,25ng, 0,125ng e 0,0625ng foram procedidas.

Nas figuras 8 e 9, observa-se que todos os alelos foram detectados quando se aplicou 10ng e 5ng de DNA às reações de PCR, tendo as quantidades de produtos variado de 4.000 para 1.800 RFU, respectivamente.

Na figura 10, é observada amplificação de todos os alelos para os dez *loci* utilizados quando se aplicou 2ng de DNA às reações, porém verifica-se uma queda acentuada nas quantidades dos produtos relativos aos *loci* DXS7424 e DXS6807 (aproximadamente 200 RFU).

Nas reações de PCR que receberam 1ng e 0,5ng de DNA (figuras 11 e 12), com exceção do *locus* DXS6807 para o qual não foi verificado o alelo 14, os produtos relativos aos demais *loci* foram obtidos à semelhança das reações que receberam 2 ng de DNA.

Na figura 13, onde são mostrados os resultados relativos à reação de PCR contendo 0,25ng de DNA, verifica-se a ausência de amplificação no *locus* DXS6807 e uma queda nas quantidades de produtos relativos aos *loci* DXS7133 (alelo 9), DXS7424 (alelo 16), GATA172D05 (alelo 10) e DXS10101 (alelo 31). Para tais regiões, as alturas dos picos variaram de 25 a 80 RFU.

Nas figuras 14 e 15, observam-se resultados relativos às reações de PCR que receberam 0,125ng e 0,0625ng de DNA, respectivamente. Quando se aplicou 0,125ng, com exceção do *locus* DXS8377, onde apenas um alelo foi amplificado, não foi observado nenhum produto de amplificação para os demais *loci*. A aplicação de 0,0625ng de DNA à reação de PCR, resultou na completa uma ausência de amplificação para os dez *loci* X-STR estudados.

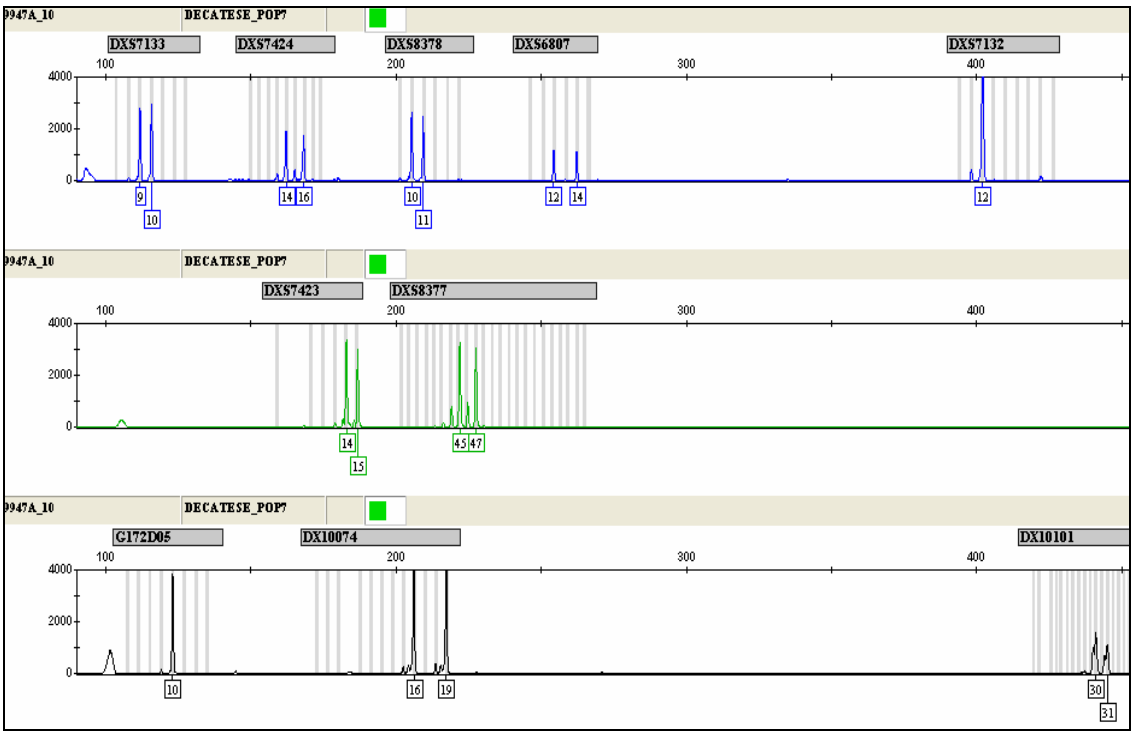


Figura 8 – Perfil alélico gerado pela amplificação de 10ng de DNA 9947A com o sistema LDD (X-STR) Decaplex.

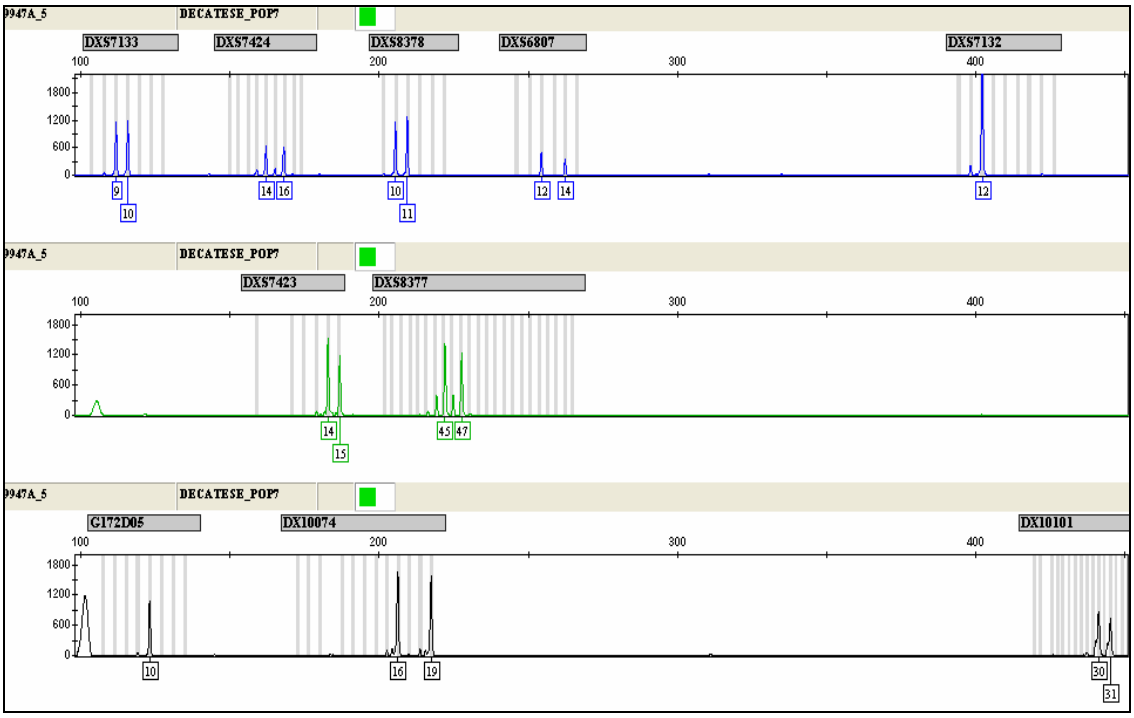


Figura 9 – Perfil alélico gerado pela amplificação de 5ng de DNA 9947A com o sistema LDD (X-STR) Decaplex.

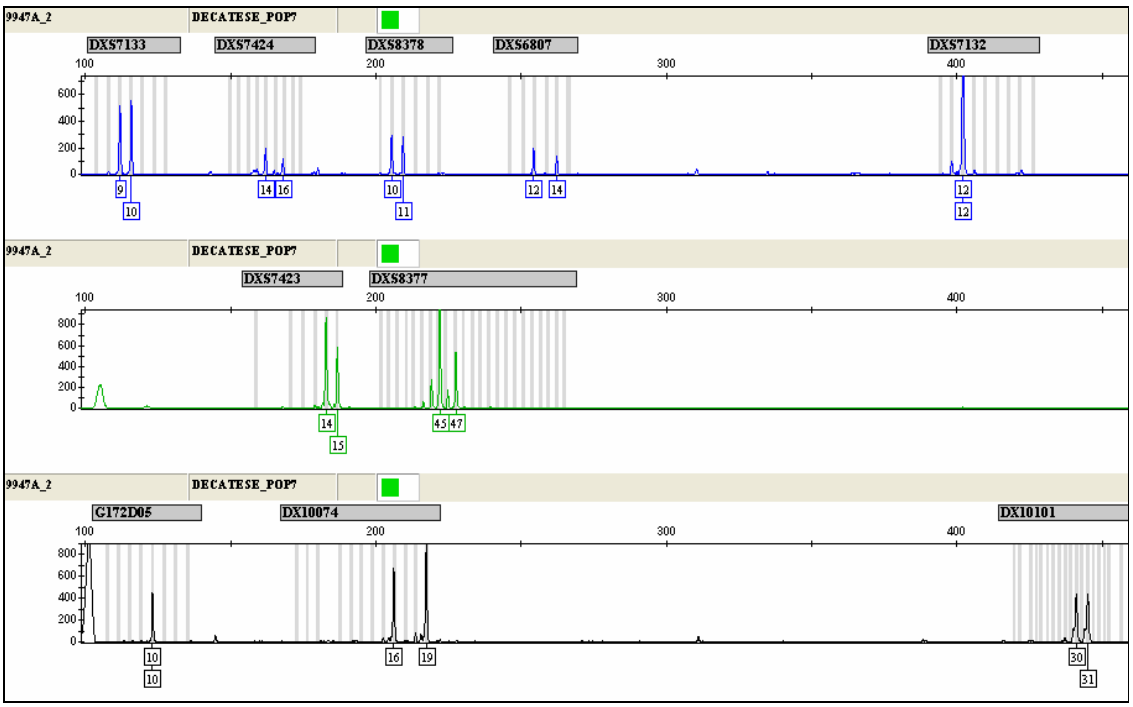


Figura 10 – Perfil alélico gerado pela amplificação de 2ng de DNA 9947A com o sistema LDD (X-STR) Decaplex.

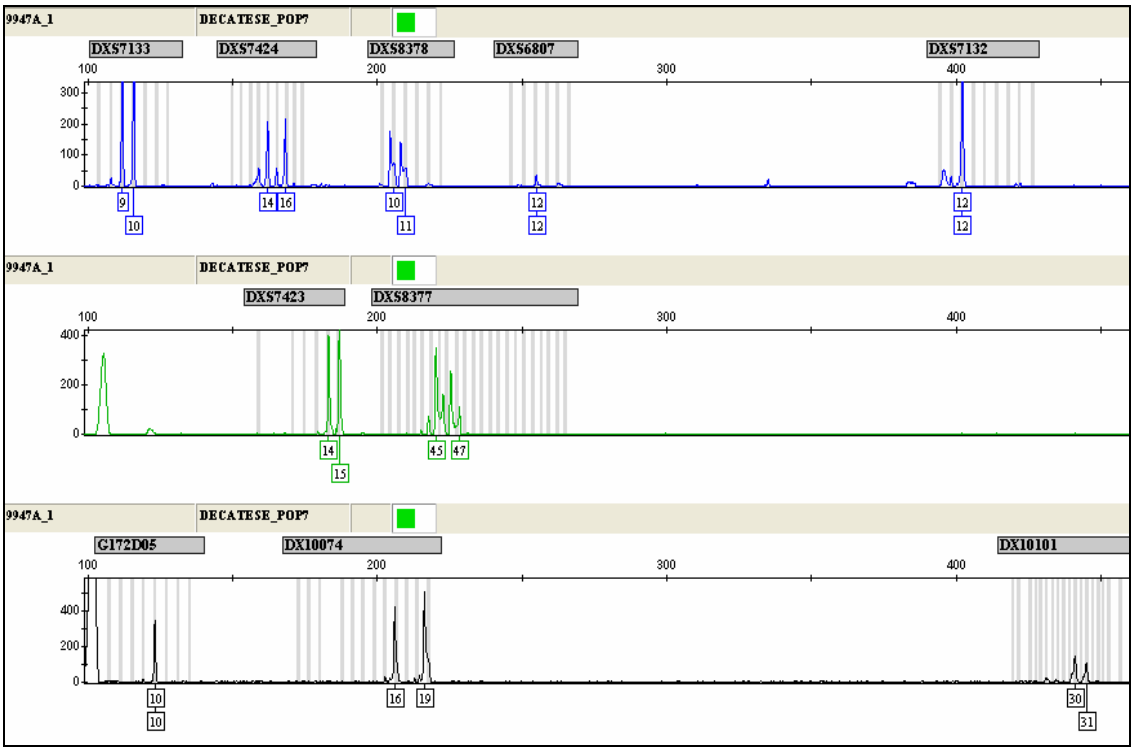


Figura 11 – Perfil alélico gerado pela amplificação de 1ng de DNA 9947A com o sistema LDD (X-STR) Decaplex.

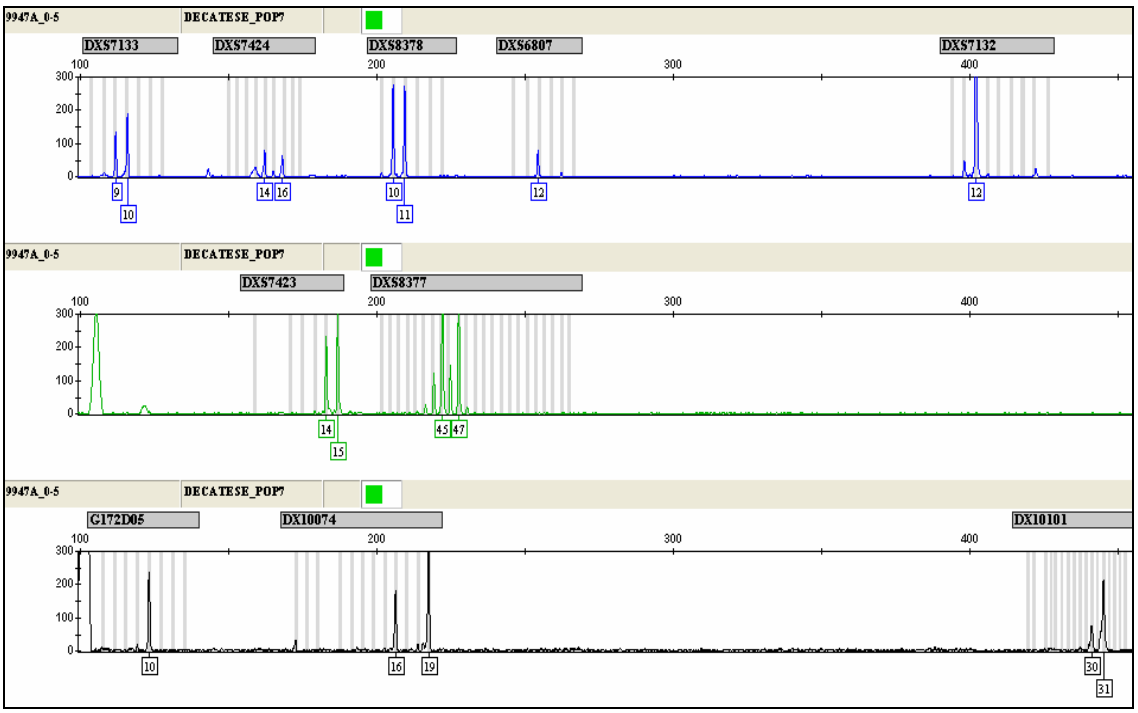


Figura 12 – Perfil alélico gerado pela amplificação de 0,5ng de DNA 9947A com o sistema LDD (X-STR) Decaplex.

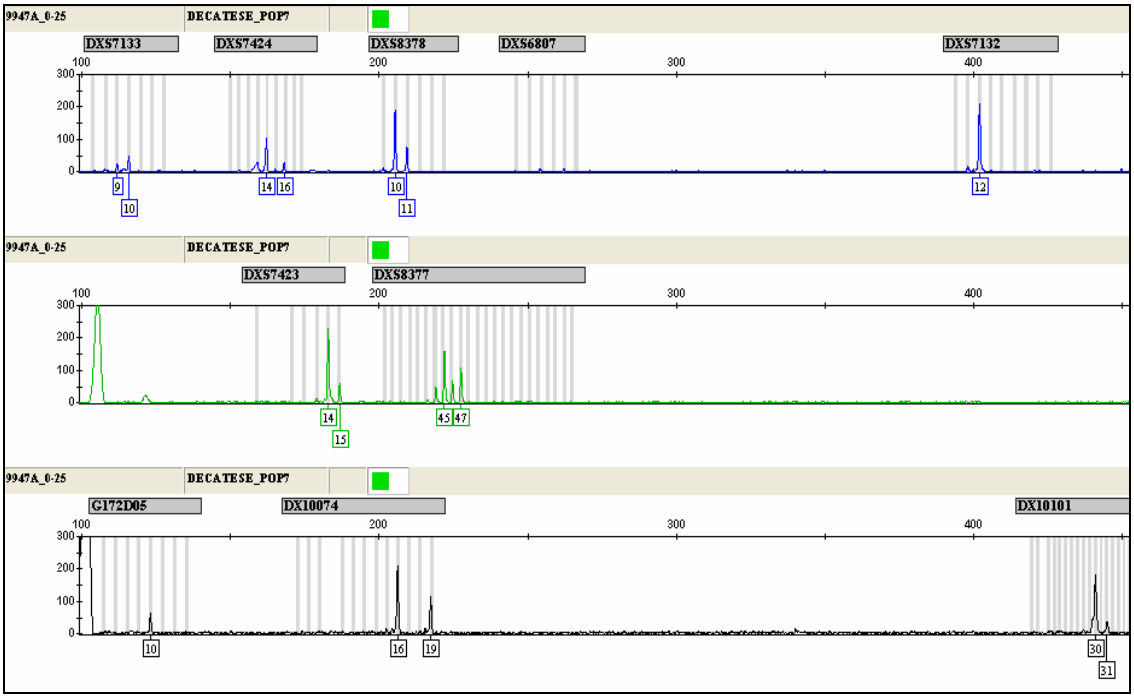


Figura 13 – Perfil alélico gerado pela amplificação de 0,25ng de DNA 9947A com o sistema LDD (X-STR) Decaplex.

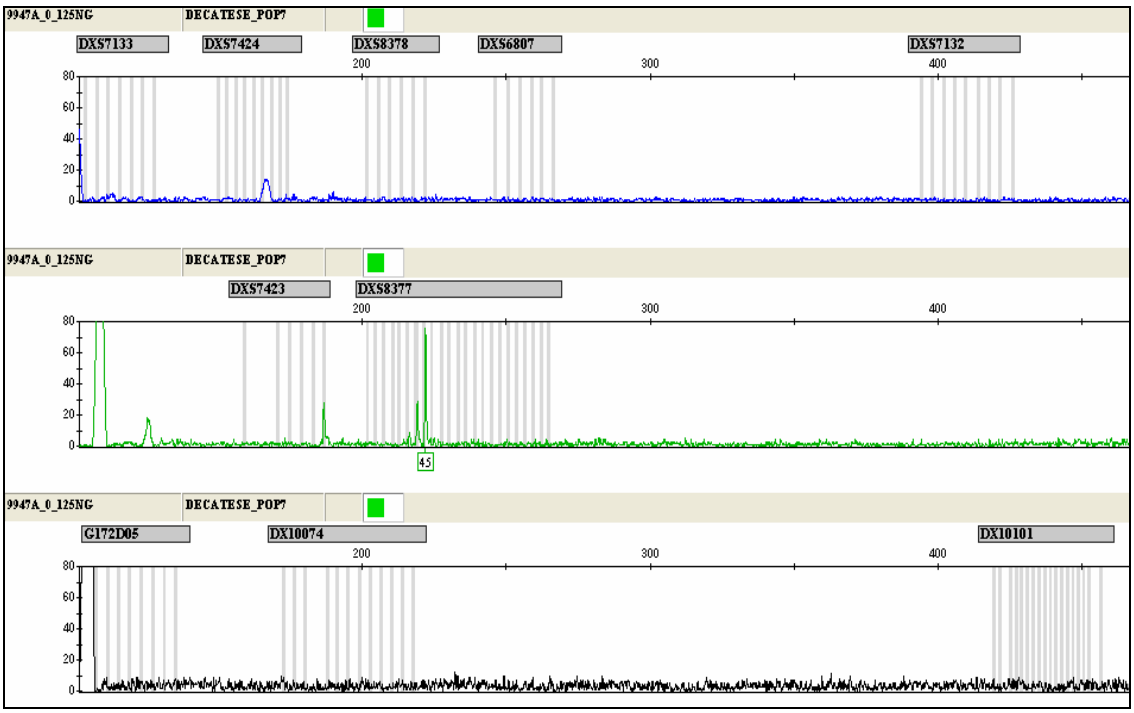


Figura 14 – Perfil alélico gerado pela amplificação de 0,125ng de DNA 9947A com o sistema LDD (X-STR) Decaplex.

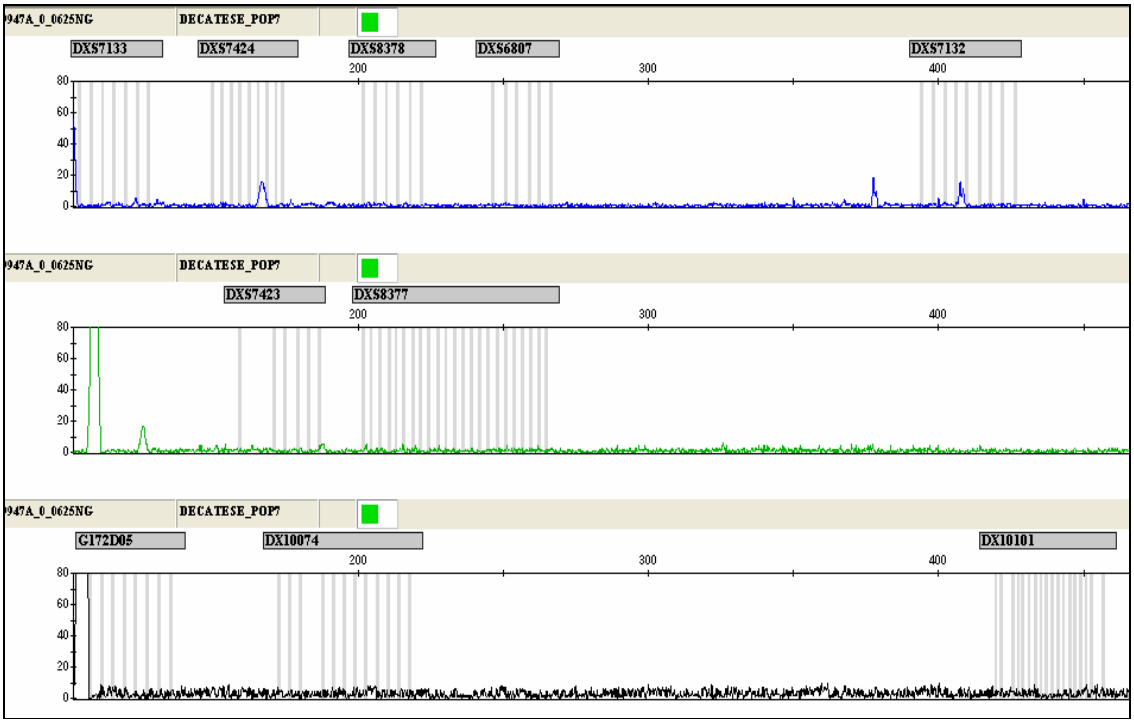


Figura 15 – Perfil alélico gerado pela amplificação de 0,0625ng de DNA 9947A com o sistema LDD (X-STR) Decaplex.

3.3.2 Análise de reprodutibilidade da tipagem de alelos X-STR a partir de diferentes amostras biológicas

Nesta etapa, diferentes amostras biológicas foram utilizadas como fonte de DNA genômico. Uma delas consistiu de dente proveniente do acervo de amostras de restos mortais do Laboratório de Diagnósticos por DNA. As amostras de saliva (mucosa oral), cabelo (bulbo), sangue (FTA), urina e sêmen foram provenientes de um mesmo indivíduo saudável. As extrações de DNA foram realizadas pelo método orgânico (Ver item 3.2 – Material e Métodos). A concentração dos DNAs nas reações de amplificação por PCR com o sistema LDD (X-STR) Decaplex foi 5ng na reação.

Inicialmente, foi utilizado o sistema de tipagem comercial *Identifiler*®, formado por iniciadores para quinze *loci* STR autossômicos e posteriormente, o sistema LDD (X-STR) Decaplex. Após a amplificação por PCR, os produtos foram separados por eletroforese capilar em seqüenciador automático 3130 e analisados conforme descrito no item 3.7.3 (Material e Métodos).

A amplificação inicial com sistema *Identifiler* teve como objetivo verificar a qualidade das preparações de DNA através da amplificação dirigida por um sistema comercial bem estabelecido técnica e cientificamente, confirmando-se o sexo do doador bem como padronizar as condições ideais de utilização das amostras de DNA. Para cada preparação de DNA, foram observados produtos relativos a todos os *loci* do sistema *Identifiler*, inclusive da amelogenina, confirmando o perfil masculino do doador. Para o sistema LDD (X-STR) Decaplex, semelhante padrão de amplificação foi observado nas dez regiões estudadas para as diferentes amostras de DNA. As quantidades de produtos obtidas variaram de, aproximadamente, 800 a 4000 RFU (figuras 16-27).

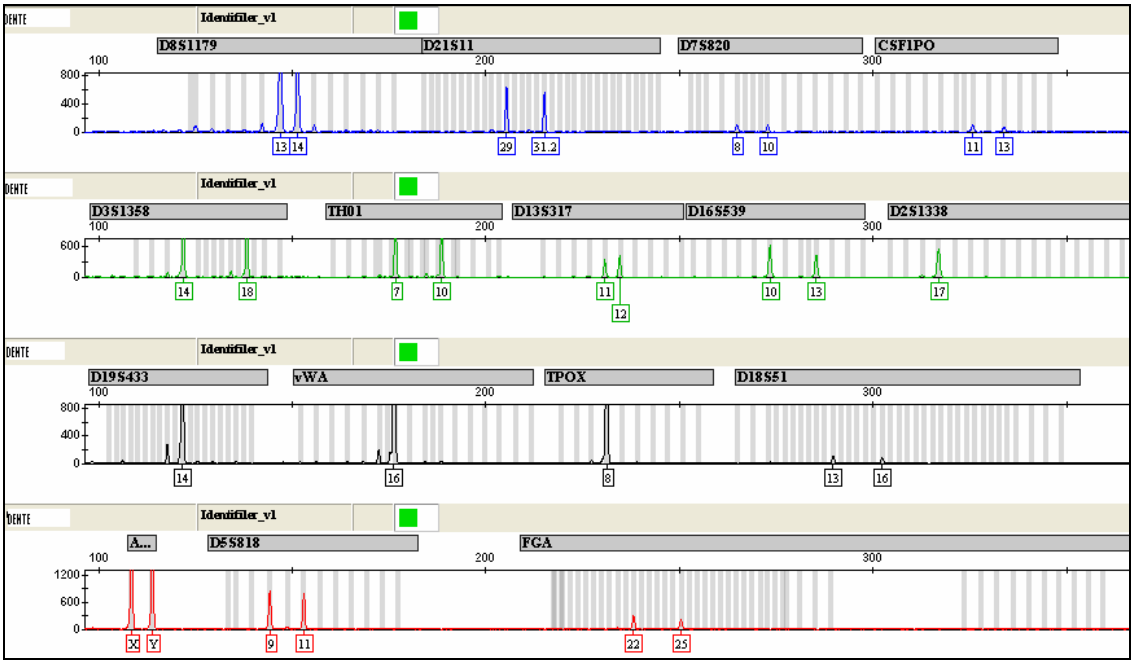


Figura 16 - Perfil alélico gerado pela amplificação de 5ng de DNA de dente com o sistema *Identifiler*.

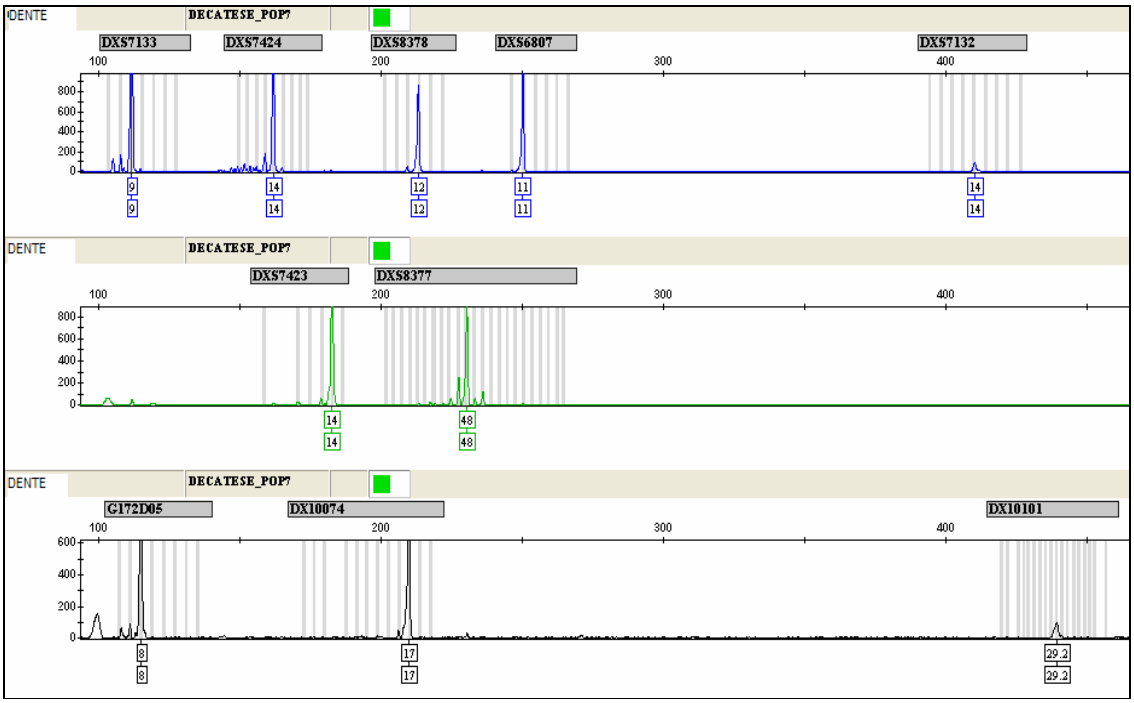


Figura 17 Perfil alélico gerado pela amplificação de 5ng de DNA de dente com o sistema LDD (X-STR) Decaplex.

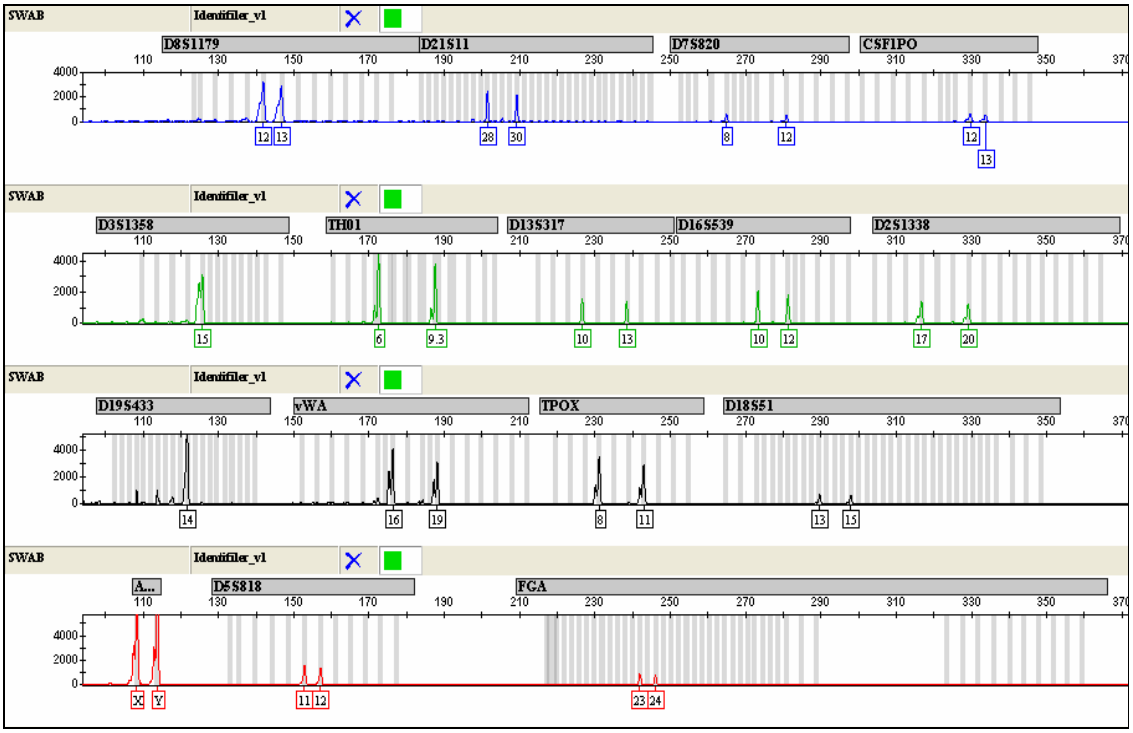


Figura 18 - Perfil alélico gerado pela amplificação de 5ng de DNA de saliva (mucosa oral) com o sistema *Identifiler*.

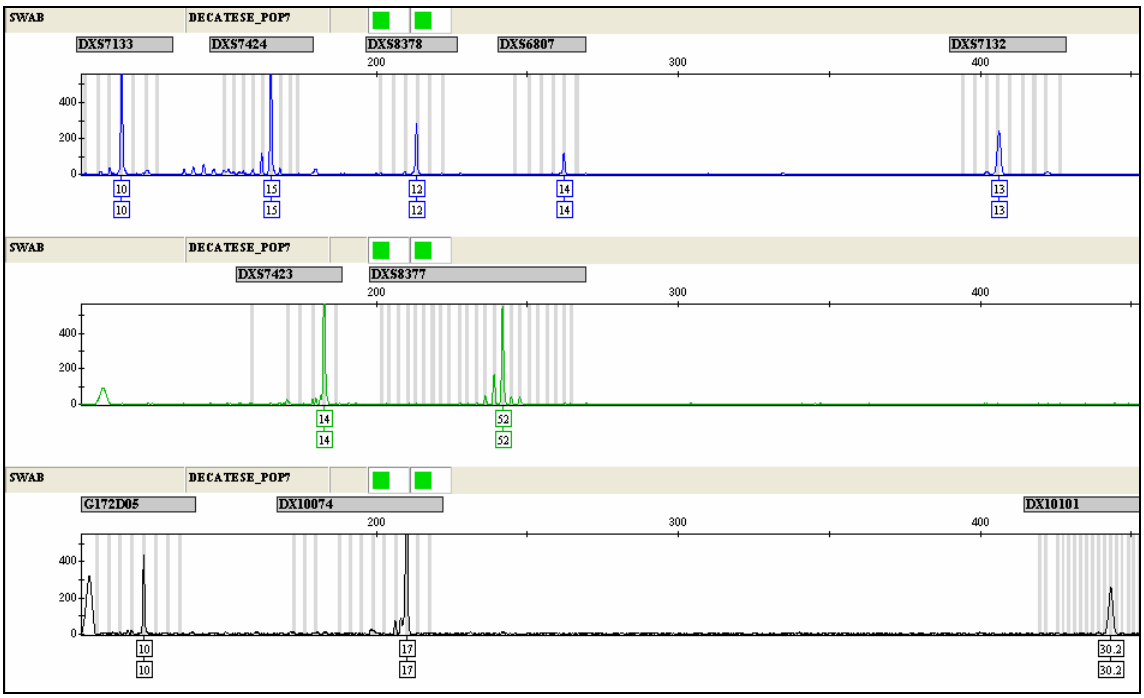


Figura 19 - Perfil alélico gerado pela amplificação de 5ng de DNA de saliva (mucosa oral) com o sistema LDD (X-STR) Decaplex.

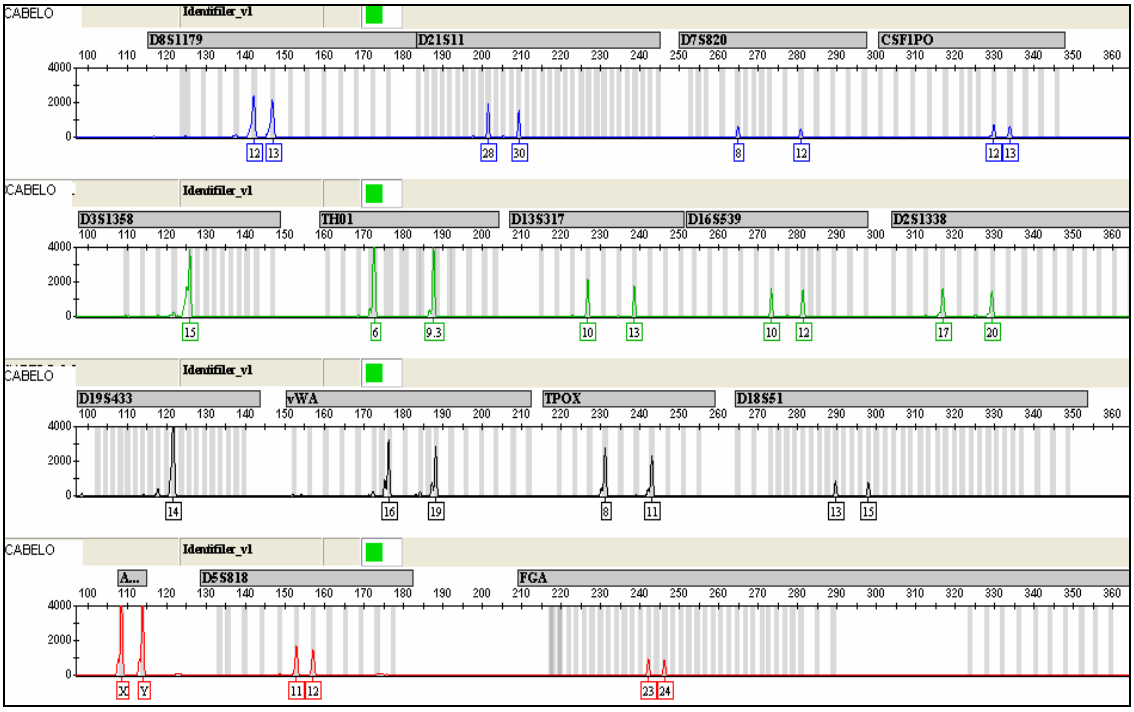


Figura 20 - Perfil alélico gerado pela amplificação de 5ng de DNA de cabelo (bulbo) com o sistema *Identifiler*.

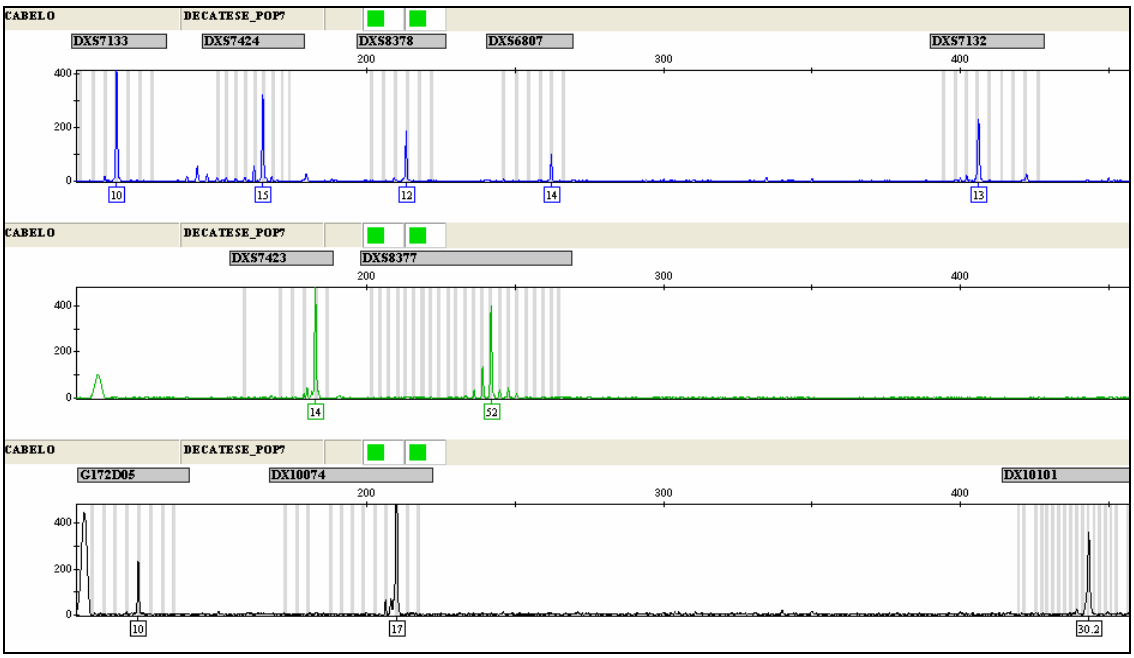


Figura 21 - Perfil alélico gerado pela amplificação de 5ng de DNA de cabelo (bulbo) com o sistema LDD (X-STR) Decaplex.

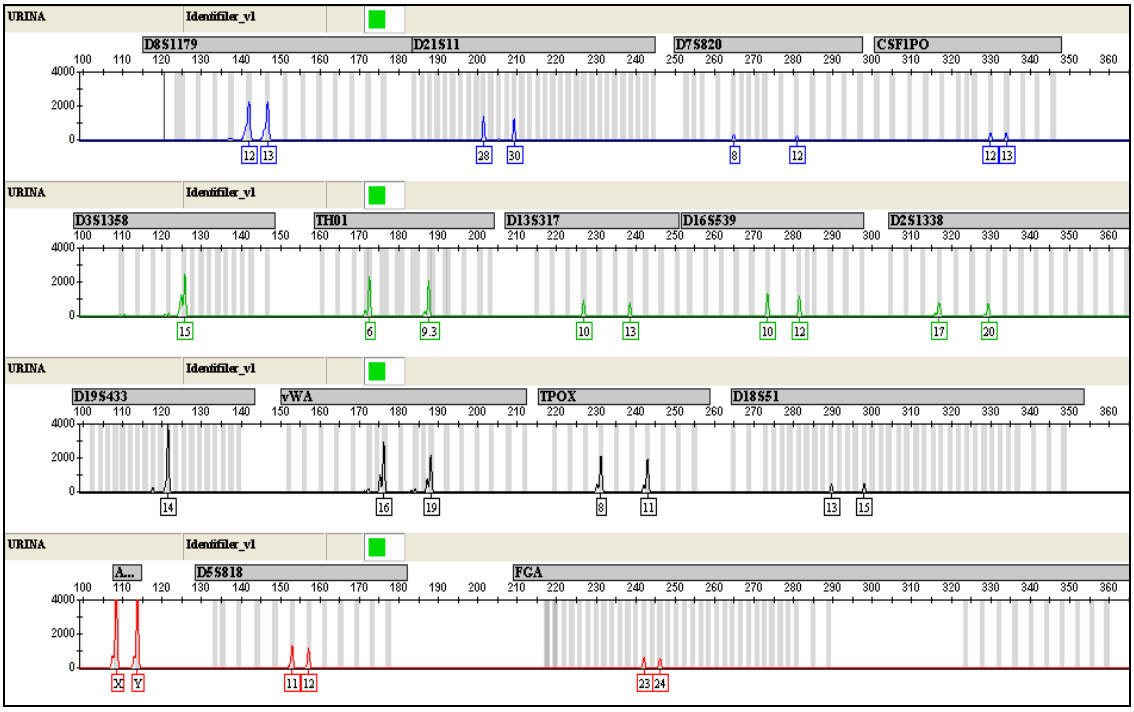


Figura 22 - Perfil alélico gerado pela amplificação de 5ng de DNA de urina com o sistema *Identifiler*.

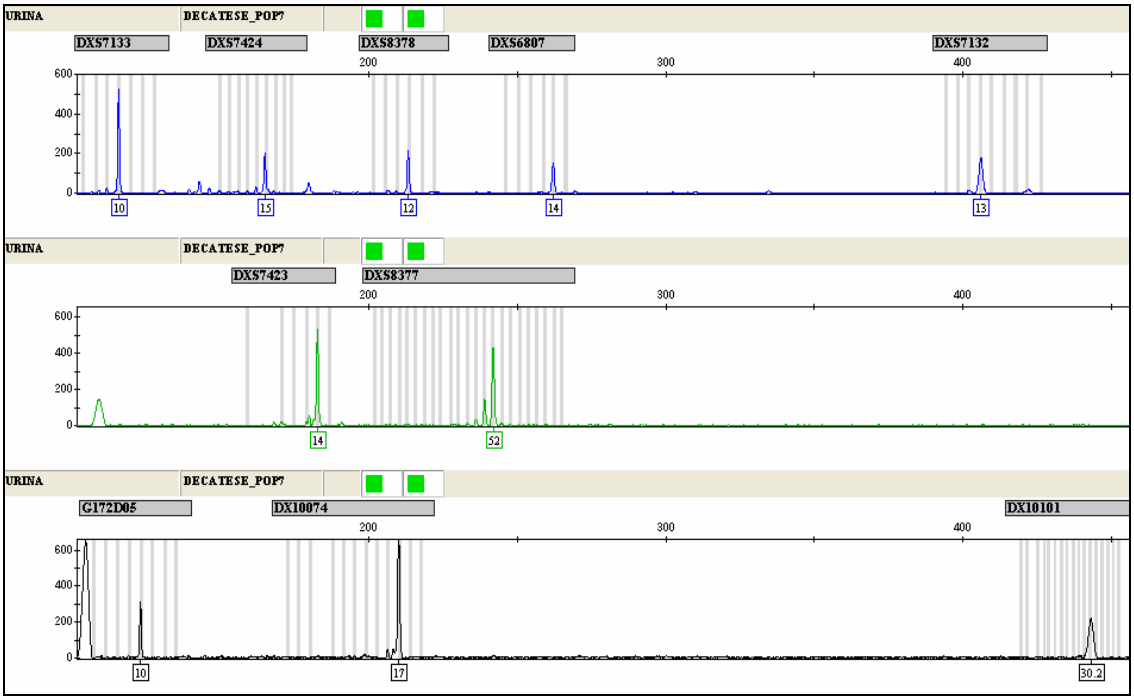


Figura 23 - Perfil alélico gerado pela amplificação de 5ng de DNA de urina com o sistema LDD (X-STR) Decaplex.

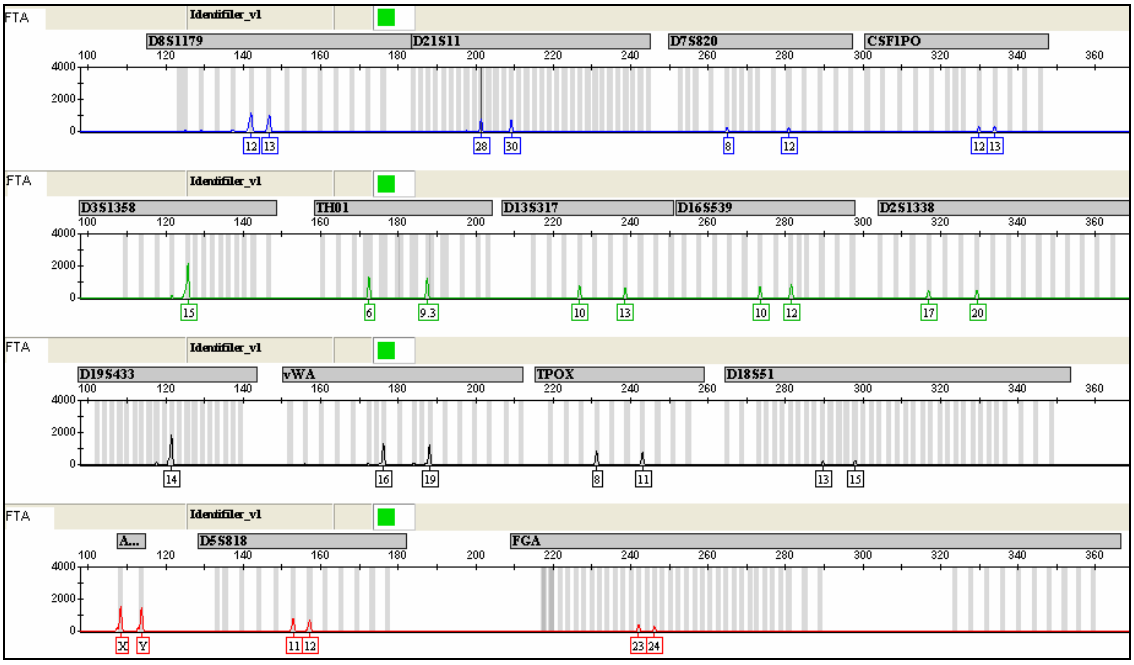


Figura 24 - Perfil alélico gerado pela amplificação de 5ng de DNA de sangue (FTA) com o sistema *Identifiler*.

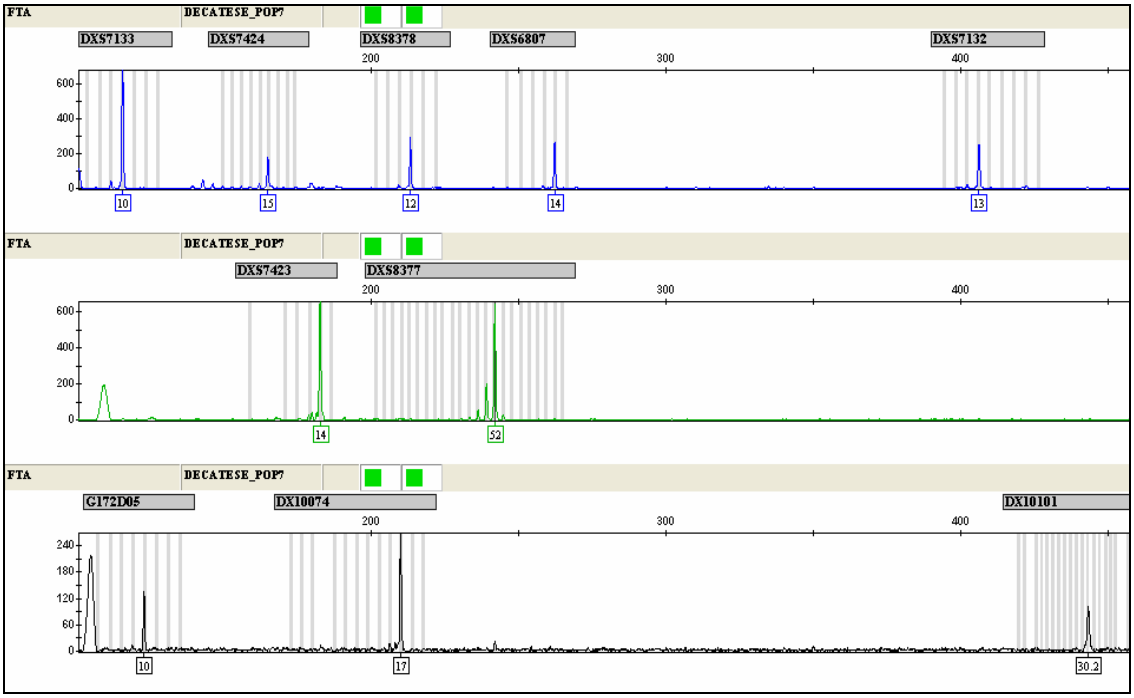


Figura 25 - Perfil alélico gerado pela amplificação de 5ng de DNA de sangue (FTA) com o sistema LDD (X-STR) Decaplex.

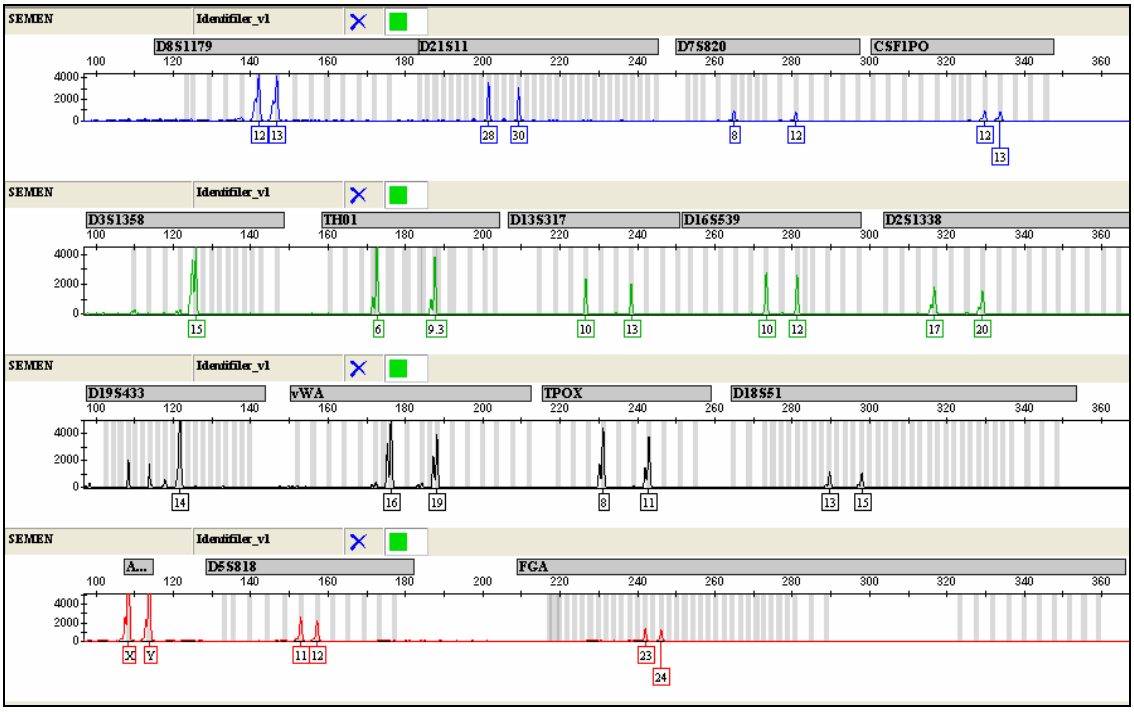


Figura 26 - Perfil alélico gerado pela amplificação de 5ng DNA de sêmen com o sistema *Identifiler.*

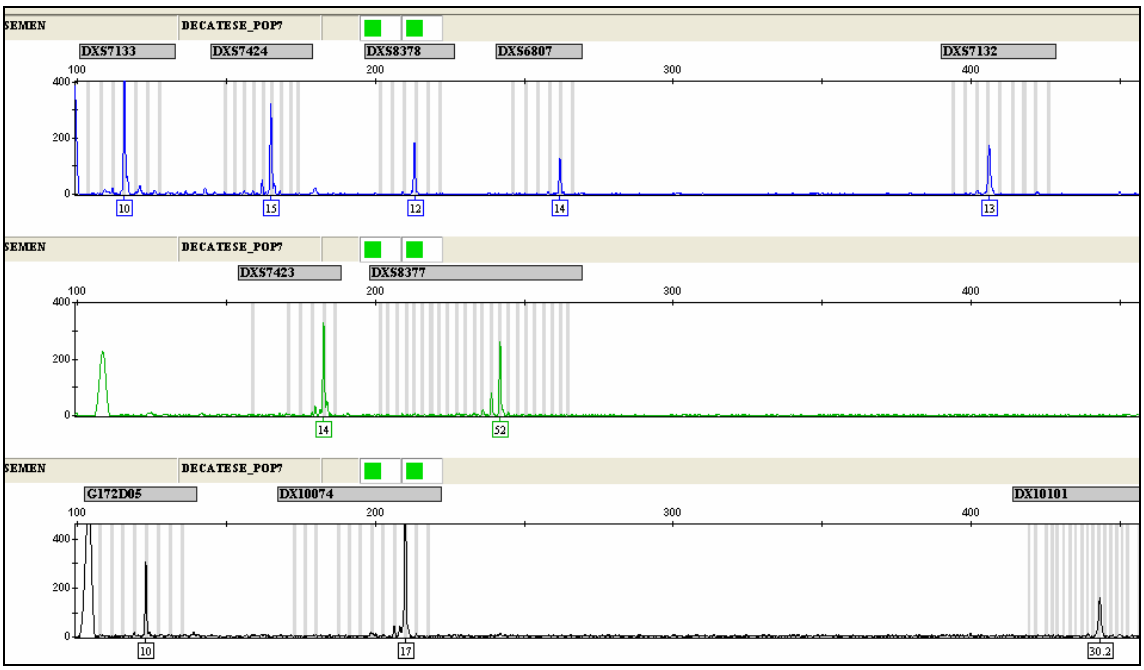


Figura 27 - Perfil alélico gerado pela amplificação de 5ng de DNA de sêmen com o sistema LDD (X-STR) Decaplex.

3.3.3 Análises de bandas *stutter*

A análise de bandas *stutter* foi realizada através da amplificação de quinze alíquotas de preparações de DNA, extraídos por *chelex,* com o sistema LDD (X-STR) Decaplex. Os produtos de PCR foram submetidos à eletroforese capilar em seqüenciador automático 3130 e genotipados com *software* Gene Mapper (conforme item 3.7.3 – Material e Métodos).

Quando cabível, os alelos e suas respectivas bandas *stutter* eram detectados e analisados quanto às suas quantidades (alturas dos respectivos picos), medidas em RFU. A relação entre as alturas do alelo e do seu *stutter* foi avaliada e expressa em percentual dividindo-se o valor relativo à banda *stutter* por aquele do alelo real.

Na figura 28 é mostrado histograma com as relações entre as intensidades das bandas *stutter* e seus respectivos alelos verdadeiros nos *loci* DXS7133, DXS8378, DXS7132, DXS7423 e DXS8377 do sistema LDD (X-STR) Decaplex. Não foram observadas bandas *stutter* nos *loci* DXS7424, DXS6807, GATA172D05, DXS10074 e DXS10101. O *locus* DXS7133 apresentou a menor relação entre bandas *stutter* e alelo real com valor de 6,81%, seguido do *locus* DXS8378 com percentual de 6,91%, DXS7423 com 8,15% e DXS7132 com 9,38%. O percentual de bandas *stutter* foi mais elevado para o *locus* DXS8377, observando-se um valor de 28,2%.

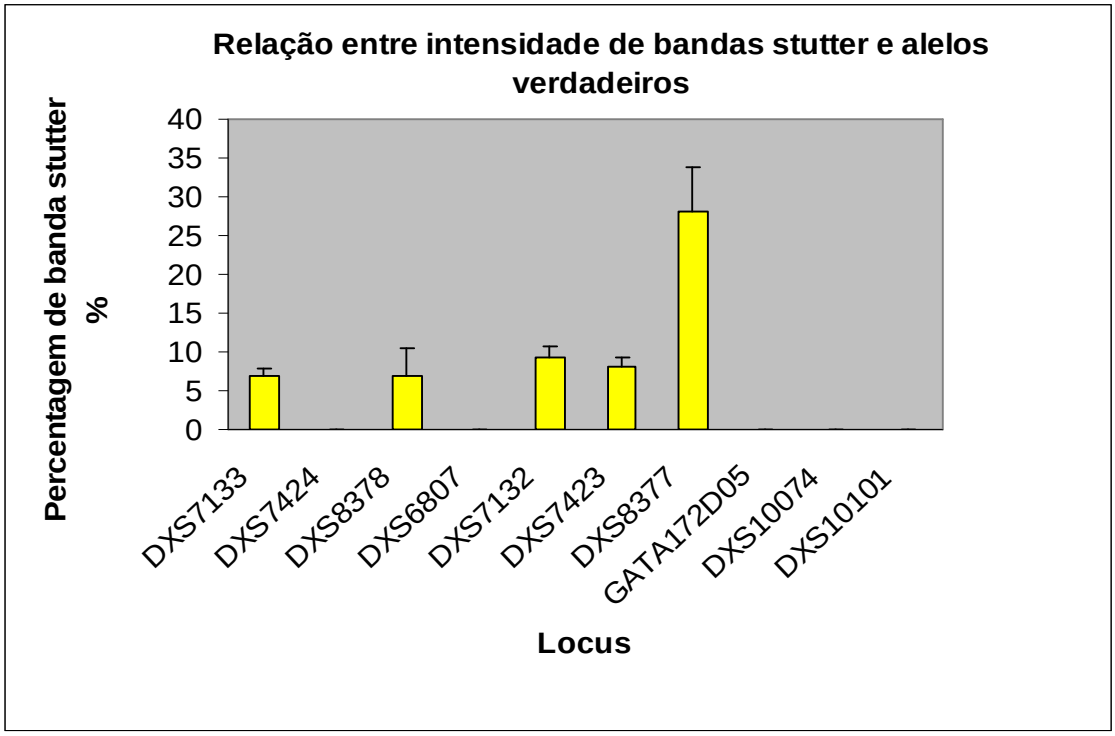


Figura 28 – Relação entre as intensidades de bandas *stutter* e alelos verdadeiros verificados nos *loci* do sistema LDD (X-STR) Decaplex. As barras verticais representam o desvio padrão referente a cada marcador.

3.3.4 Análises de misturas

Para a verificação da eficiência do sistema LDD (X-STR) Decaplex em discriminar perfis genéticos em misturas, alíquotas de soluções formadas por diferentes proporções de um DNA feminino e outro masculino foram amplificadas com os sistemas *Identifiler* e LDD (X-STR) Decaplex. As proporções de DNA feminino e masculino utilizadas foram: A - 10:1ng; B - 5:1ng; C - 2,5ng:1; D - 1:1ng; E - 1:2,5ng; F - 1:5ng; e G - 1:10ng.

Neste estudo, buscaram-se verificar, em cada *locus*, quais das proporções das misturas de DNA se relacionavam com a produção de todos os alelos dos dois indivíduos com qualidade compatível com as suas inequívocas genotipagens, ou seja, se havia uma relação a partir da qual não se verificaria a amplificação do conjunto completo de marcadores que caracterizam um dos genomas, por exemplo.

Os resultados podem ser visualizados na figura 29 (A - J). Os histogramas mostram as quantidades de produtos (alelos) amplificados, medidas em RFU, nas diferentes proporções de DNA (feminino: masculino). Dos quinze marcadores do sistema *Identifiler*, quatro foram selecionadas para visualização em histogramas (figuras 29.L- 29.O). A seleção dos mesmos foi baseada na não sobreposição dos alelos da amostra feminina em relação à masculina.

Com exceção do *locus* DXS10101 na proporção B - 5:1ng, os dois genomas foram amplificados nos demais *loci* do LDD (X-STR) decaplex. No sistema *Identifiler*, para a região D7S820 verificou-se amplificação parcial nas proporções D (1:1ng), E (1:10ng) e F (1:5ng), observando-se maiores quantidades de produtos relacionados ao DNA masculino.

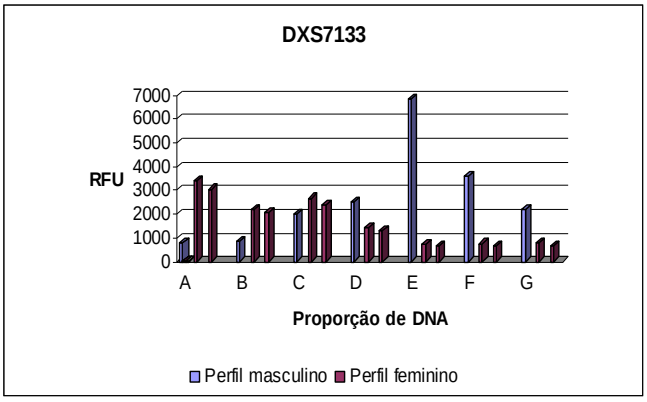


Figura 29.A – Padrão de amplificação no *locus* DXS7133

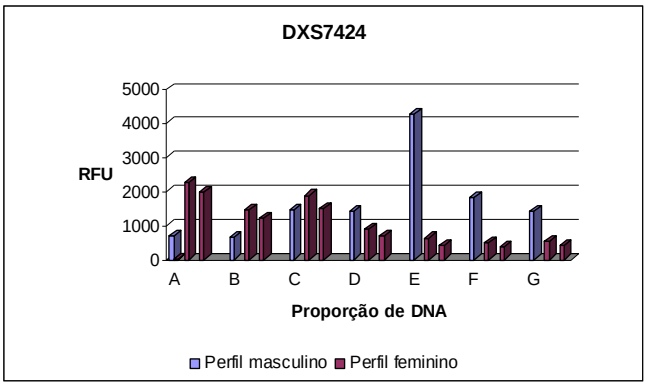


Figura 29.B – Padrão de amplificação no *locus* DXS7424

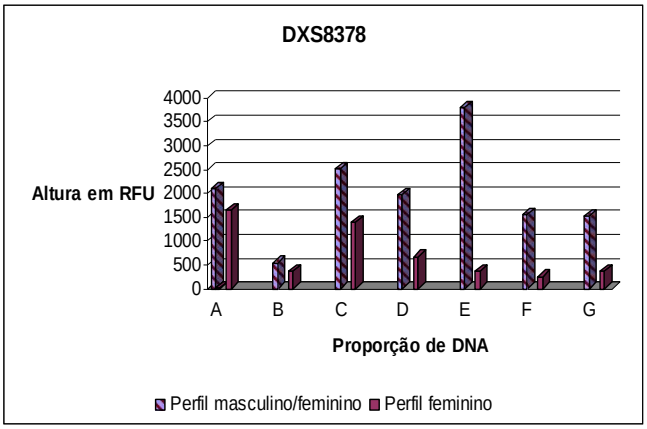


Figura 29.C – Padrão de amplificação no *locus* DXS8378

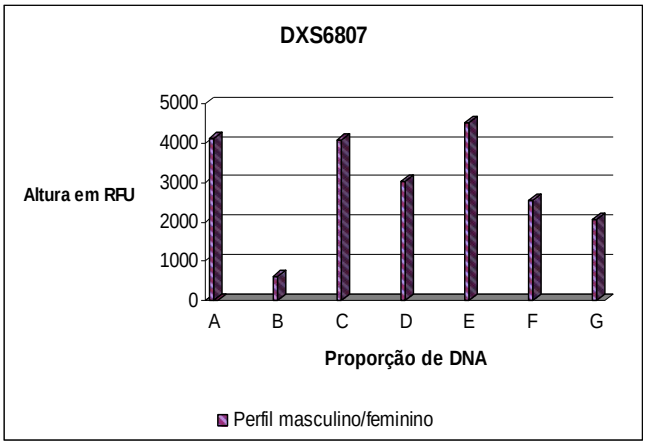


Figura 29.D – Padrão de amplificação no *locus* DXS6807

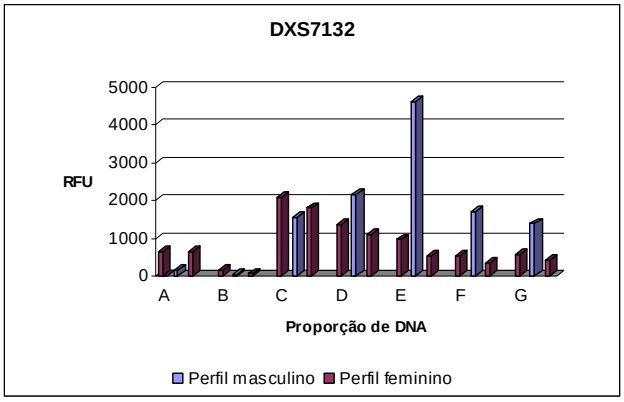


Figura 29.E – Padrão de amplificação no *locus* DXS7132

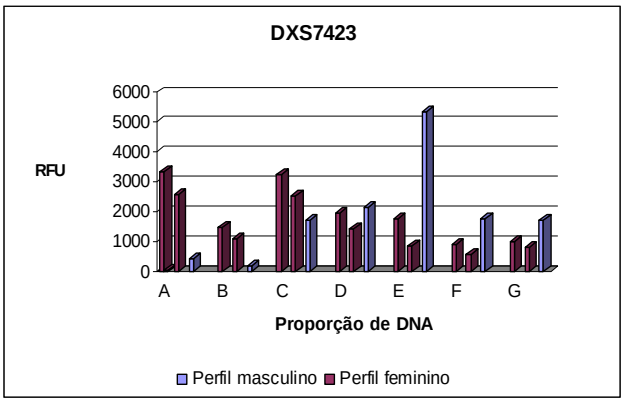


Figura 29.F – Padrão de amplificação no *locus* DXS7423

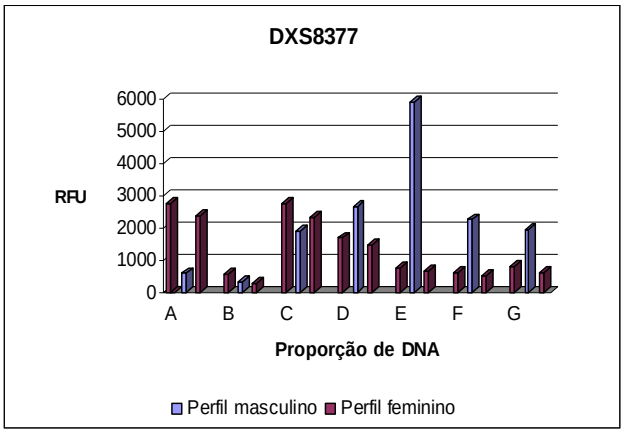


Figura 29.G – Padrão de amplificação no *locus* DXS8377

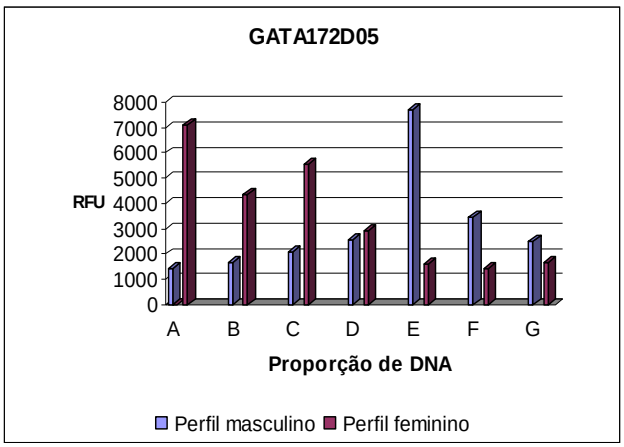


Figura 29.H – Padrão de amplificação no *locus* GATA172D05

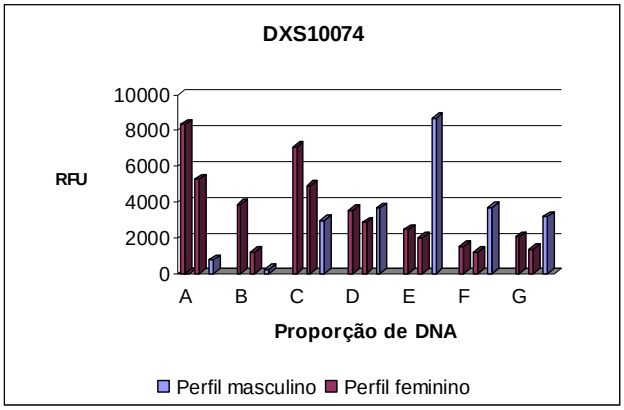


Figura 29.I – Padrão de amplificação no *locus* DXS10074

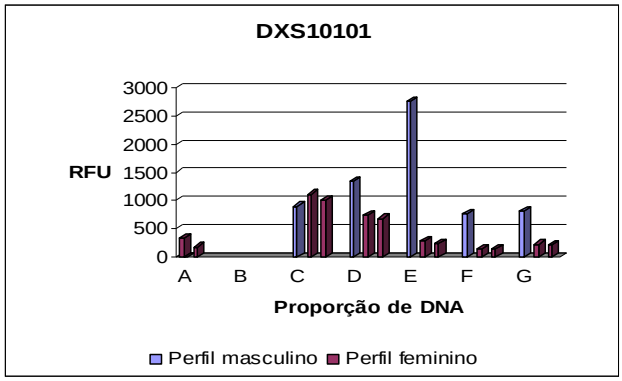


Figura 29.J – Padrão de amplificação no *locus* DXS10101

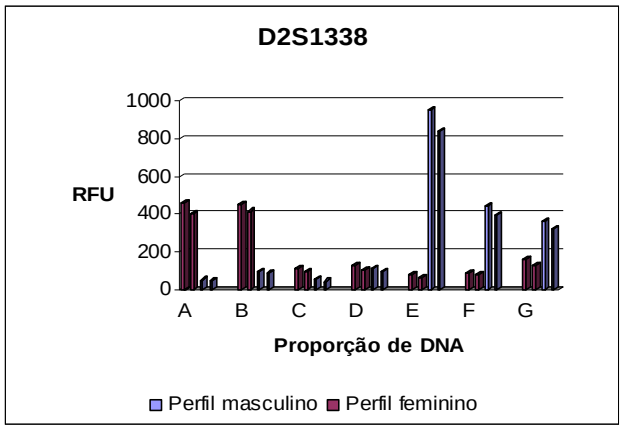


Figura 29.L – Padrão de amplificação para o *locus* autossômico D2S1338

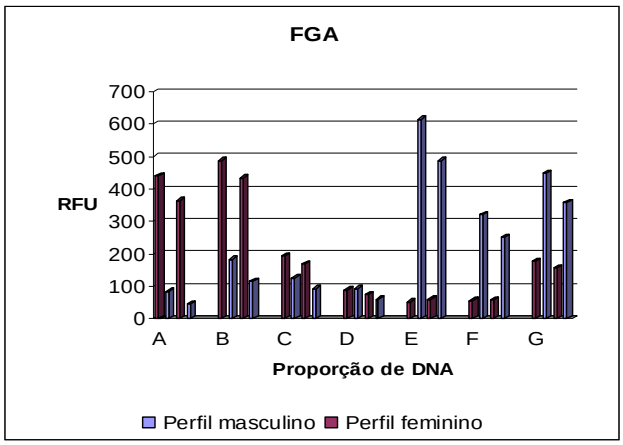


Figura 29.M – Padrão de amplificação no *locus* FGA

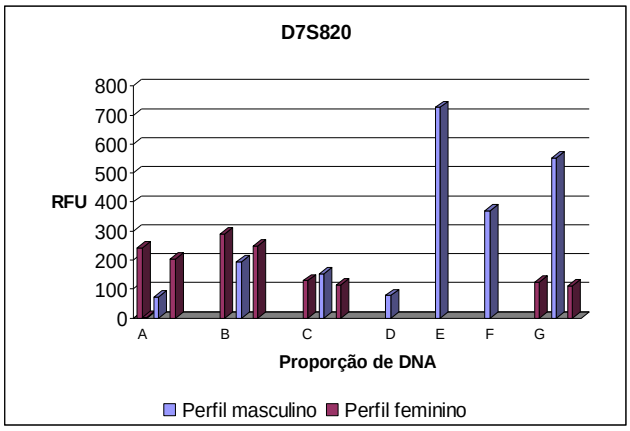


Figura 29.N – Padrão de amplificação no *locus* D7S820

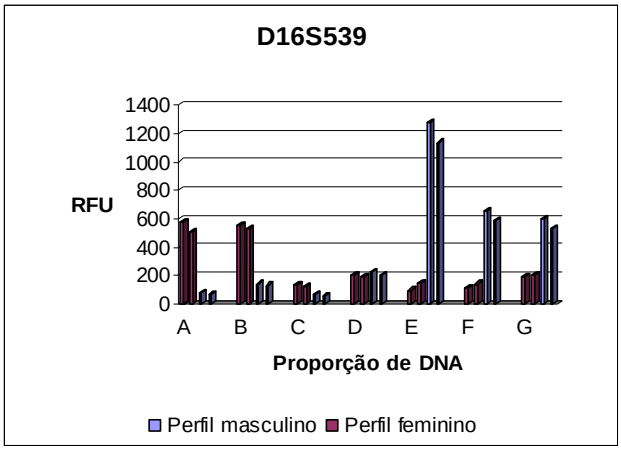


Figura 29.O – Padrão de amplificação no *locus* D16S539

Legenda: Quantificação de produtos de amplificação por PCR (alelos). Alíquotas de uma mistura de DNA feminino e masculino nas proporções 10:1 **(A)**; 5:1 **(B)**; 2,5:1 **(C)**; 1:1 **(D)**;1:10 **(E)**; 1:5 **(F)**;1: 2,5 **(G)** foram amplificadas por PCR com os sistemas LDD (X-STR) Decaplex (29.A a 29.J) e *Identifiler* (29.L a 29.O).

3.4 **Estudo de casos**

Os STR do cromossomo X são ferramentas úteis na investigação de relações de parentesco, principalmente em investigações de paternidade na ausência do suposto pai, em que a criança em questão é do sexo feminino. Nesses casos, na dependência dos seus vínculos genéticos com o investigado ausente, os cromossomos X de familiares do suposto pai podem ser analisados.

3.4.1 Reconstrução de perfil genético do investigado a partir de suas filhas

Em pesquisas de vinculação genética na ausência do suposto pai, a reconstrução do perfil genético do investigado através da genotipagem de suas filhas biológicas, e de suas respectivas progenitoras, deve ser realizada. Na figura 30, são apresentados heredogramas representativos de três meio-irmãs paternas. A partir da análise de múltiplos *loci* X-STR de suas filhas biológicas, é possível reconstruir o haplótipo X-STR suposto pai, com o intuito de verificar a coincidência, ou não, deste com o haplótipo X paterno da suposta filha, o que permitirá incluir a cerca da hipótese da partilha de paternidade.

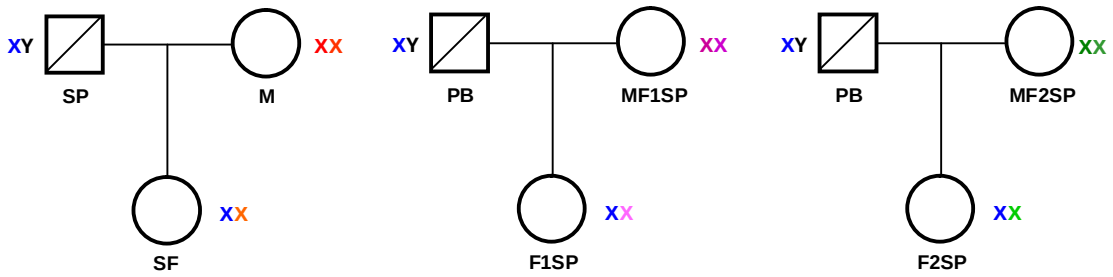


Figura 30- Heredogramas representativos de três meio-irmãs paternas.
Legenda- Mãe (M); Suposta Filha **(SF)**; Suposto Pai Investigado (**SP)**; Pai Biológico de F1SP e F2SP **(PB)**; Mãe da Filha 1 do Suposto Pai **(MF1SP)**; Filha 1 do Suposto Pai **(F1SP)**; Mãe da Filha 2 do Suposto Pai **(MF2SP)**; Filha 2 do Suposto Pai **(F2SP)**.

3.4.2 Pesquisa de vínculo genético a partir da suposta avó paterna

Quando da ausência do suposto pai, a genotipagem da suposta filha, da sua genitora e da mãe do suposto pai, em marcadores STR do cromossomo X, possibilitará incluir a respeito da ocorrência, ou não, de vínculo genético paterno entre as supostas neta e avó. O perfil genético da suposta avó é de relevante importância uma vez que o único cromossomo X do suposto pai está contido no seu genoma (figura 31).

Na tabela 10, é apresentado um caso típico de investigação de paternidade com a participação de uma menina, sua genitora e a mãe do suposto pai. Parâmetros estatísticos tais como o Índice de Vinculação por *locus* (IV), o índice Cumulativo de Vinculação (ICV), a Probabilidade de Vinculação e o Poder de Exclusão (PE) foram calculados. Estes parâmetros são calculados com base nas freqüências dos alelos paternos e nos perfis genéticos dos demais indivíduos, para cada *locus*, a partir de um banco de dados representativo da distribuição frequencial dos alelos polimórficos na população (item 3.9 – Material e Métodos).

Os resultados possibilitaram o cálculo de um ICV de 40.337.671,37, o qual expressa uma probabilidade de certeza superior a 99,99999% a favor de a genitora do investigado ser a avó paterna biológica da criança. O poder de exclusão associado a tal investigação genética foi verificado ser superior a 99,7220%.

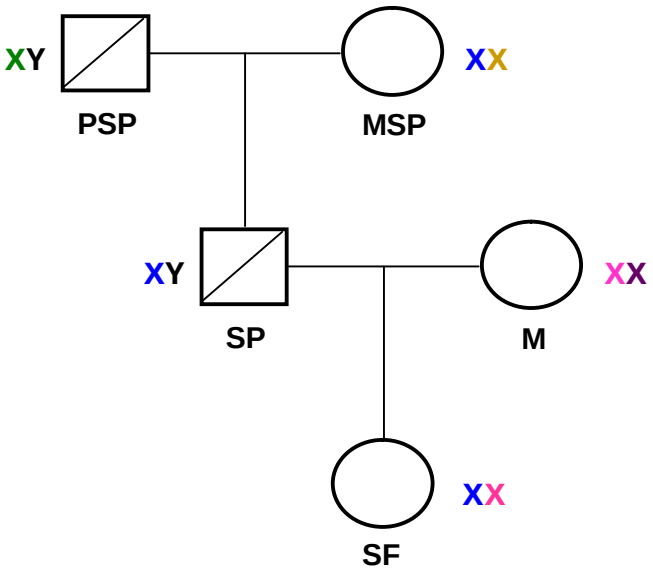


Figura 31 - Heredograma representativo de vínculo genético paterno entre neta e avó paterna.

Pai do Suposto Pai **(PSP)**; Mãe do Suposto Pai **(MSP)**; Suposto Pai **(SP)**; Mãe **(M)**; Suposta Filha **(SF).**

Tabela 10 - Parâmetros estatísticos em investigação de vínculo genético paterno com a participação de uma suposta filha, sua genitora e da mãe do suposto pai utilizando o sistema LDD (X-STR).

Locus	**M**		**SF**		**MSP**		**freq. alelo paterno**	**IV**	**PE**
DXS7133	11	12	10	11	10	11	0,132352941	3,77777778	0,090884516
DXS7424	14	17	14	16	13	16	0,136029412	3,67567568	0,447718426
DXS8378	10	11	11	11	10	11	0,330882353	1,51111111	0,447718426
DXS6807	14	14	11	14	11	11	0,481617647	2,07633588	0,268720264
DXS7132	13	14	11	13	11	14	0,014705882	34	0,508704585
DXS7423	16	16	14	16	14	15	0,415441176	1,20353982	0,386042928
DXS8377	45	49	45	56	52	56	0,018382353	27,2	0,963573205
G172D05	11	12	11	11	9	11	0,099264706	5,03703704	0,81132407
DX10074	8	16	8	11	11	13	0,025735294	22,6695684	0,920753136
DX10101	28.2	31	31	31.2	28	31.2	0,058823529	8,5	0,885813149
								ICV	40.337.671,37
								PV	0,999999975
								PEac	0,997220304

Legenda – Perfil X-STR da mãe da suposta filha **(M);** Perfil X-STR da suposta filha **(SF)**; Perfil X-STR da mãe do suposto pai **(MSP)**; **IV:** Índice de Vinculação no *locus*; **ICV:** Índice Cumulativo de Vinculação; **PV:** Probabilidade de Vínculo; **PE:** Poder de Exclusão do *locus***; PEac:** Poder de Exclusão acumulado.

Nota: Os alelos paternos da suposta filha, ou não excluídos de serem paternos, encontram-se sublinhados na SF e na MSP.

3.4.3 Parâmetros estatísticos na análise forense

No âmbito do DNA forense, o perfil genético de uma amostra questionada, um vestígio ou evidência biológica pode ser comparado com aqueles de indivíduos referências com o intuito de determinar a sua origem. Neste sentido, foram realizados testes estatísticos para avaliar os seguintes parâmetros forenses: a Probabilidade de Coincidência (PC) e a Razão de Verossimilhança (RV), item 3.9 – Material e Métodos. O perfil genético da genitora do investigado na investigação genética apresentada na tabela 10, se relaciona com uma Probabilidade de Coincidência de 1,23 x 10^{-15}. A raridade do perfil genotípico, ou Razão de Verossimilhança (RV), é 1 em 8,11 x 10^{+14}, (tabela 11).

Tabela 11 - Parâmetros estatísticos na análise forense calculado com base na freqüência alélica dos *loci* STR do sistema LDD (X-STR) Decaplex, conforme mostrado na tabela 6.

Locus	**MSP**		**freq. do genótipo**	
DXS7133	10	11	0,149870242	
DXS7424	13	16	0,05901276	
DXS8378	10	11	0,19220372	
DXS6807	11	11	0,231955558	
DXS7132	11	14	0,008434256	
DXS7423	14	15	0,314635597	
DXS8377	52	56	0,002297794	
G172D05	9	11	0,04744269	
DX10074	11	13	0,0020815	
DX10101	28	31.2	0,005190311	
			PC	1,232E-15
			RV	8,114E+14

Legenda- PC: Probabilidade de Coincidência, **RV**: Razão de Verossimilhança.

Com a finalidade de comparar a eficiência dos sistemas LDD (X-STR) Decaplex e *Identifiler*, os parâmetros estatísticos de ICV, PV, PEacu, PC e RV, utilizados em casos de investigações criminais e de vínculo genético foram, também, calculados para o sistema constituído por marcadores autossômicos.

A genotipagem das partes envolvidas na investigação de vínculo genético paterno esquematizado na figura 31 com o sistema *Identifiler*, constituído por quinze *loci* STR autossômicos, se associam com resultados de ICV de 1.221,81, que expressa uma probabilidade de certeza de 99,9182%, e com um poder de exclusão acumulado de 99,9999%. A probabilidade de coincidência e a razão de verossimilhança foram de $5,69 \times 10^{-21}$ e $1,75 \times 10^{20}$, respectivamente (tabela 12).

Tabela 12 - Parâmetros estatísticos de vínculo genético e de interesse forense calculado com base nos sistemas de tipagem LDD (X-STR) Decaplex e *Identifiler*.

	ICV	**PV**	**PEac**	**PC**	**RV**
Decaplex	40.337.671,37	99,9999975	99,7220304	1,232E-15	8,114E+14
Identifiler	1.221,81	99,918220	99,9999	5,69533E-21	1,75583E+20

Legenda - **ICV**: Índice Cumulativo de Vinculação; **PV:** Probabilidade de Vínculo; **PEac**: Poder de Exclusão acumulado; **PC:** Probabilidade de Coincidência; **RV**: Razão de Verossimilhança.

4 Discussão

4.1 Construção do sistema LDD (X-STR) Decaplex

O desenvolvimento de sistemas multiplex para identificação humana contendo marcadores STR do cromossomo X, como o sistema LDD (X-STR) Decaplex, possibilitam a rápida construção de bases de dados de freqüências alélicas e haplotípicas para vários *loci* X-STR simultaneamente e, a partir destas, a obtenção de dados estatísticos de interesse para estudos de vínculos genéticos e de investigações criminais.

Neste trabalho, os marcadores X-STR que compõe o sistema LDD (X-STR) Decaplex foram selecionados entre os mais polimórficos observados nos sistemas de tipagem Decaplex do GHEP (Gusmão *et al.*, 2009) e Octaplex do Argus (Becker *et al*., 2008), tendo sido seus graus de polimorfismos e parâmetros forenses avaliados em estudos com diferentes populações. O fato de selecionarmos marcadores já investigados em outras populações e/ou incluídos em outros sistemas de genotipagem traz como vantagem adicional a possibilidade de se realizar estudos comparativos entre a população do Rio de Janeiro e outras populações mundiais.

As freqüências alélicas obtidas para a amostra populacional analisada, bem como os demais parâmetros estatísticos, foram comparados com dados genéticos disponíveis a partir de estudos de diferentes populações com distintos sistemas multiplex X-STR, incluindo o sistema de amplificação comercial Argus-X8 (tabela 13).

Nas etapas iniciais do presente estudo, dois sistemas multiplex constituídos por seis *loci* X-STR (Hexaplex) e por oito *loci* X-STR (Octaplex) foram desenvolvidos no Laboratório de Diagnósticos por DNA - UERJ. Em sua versão final, o LDD (X-STR) Decaplex incorpora dez *loci* STR, cujos produtos de amplificação gerados a partir de uma única reação de PCR são analisados por eletroforese capilar em seqüenciador automático.

Tabela 13 – *Loci* X-STR estudados em diferentes populações mundiais.

População	***Locus* estudado**	***n***	**Referência**
Portugal	DXS7133, DXS8378, DXS7132, GATA172D05, DXS7423	636	Gusmão *et al.*, 2009
Espanha	DXS7133, DXS8378, DXS7132, GATA172D05, DXS7423	279	Gusmão *et al.*, 2009
Argentina	DXS7133, DXS8378, DXS7132, GATA172D05, DXS7423	574	Gusmão *et al.*, 2009
Costa Rica	DXS7133, DXS8378, DXS7132, GATA172D05, DXS7423	627	Gusmão *et al.*, 2009
Colômbia	DXS7133, DXS8378, DXS7132, GATA172D05, DXS7423	342	Gusmão *et al.*, 2009
Rio de Janeiro	DXS7133, DXS8378, DXS7132, GATA172D05, DXS7423	165	Gusmão *et al.*, 2009
São Paulo	DXS7133, DXS8378, DXS7132, GATA172D05, DXS7423	306	Gusmão *et al.*, 2009
Mato Grosso do Sul	DXS7133, DXS8378, DXS7132, GATA172D05, DXS7423	300	Gusmão *et al.*, 2009
Uganda	DXS7133, DXS7423	393	Gomes *et al.*, 2009
Amazonas	DXS7133	466	Rodrigues *et al.*, 2008
Gana	DXS7424, DXS6807	357	Poetsch *et al*., 2009
Rio de Janeiro	DXS7424, DXS6807	397	Tavares *et al*., 2008
Japão	DXS8378, DXS7132, DXS7423, DXS10074, DXS10101	675	Tamura *et al.*, 2009
Itália	DXS8378, DXS7132, DXS7423, DXS10074, DXS10101	135	Cerri *et al*., 2008
Gana	DXS8378, DXS7132, DXS7423, DXS10074, DXS10101	290	Thiele *et al*., 2008
Angola	DXS7423	237	Gomes *et al*., 2007a
Estados Unidos	DXS8377	130	Gomes *et al*., 2007b
Espanha	DXS8377	290	Aler *et al*., 2007
Portugal	DXS8377	347	Pereira *et al*., 2007

4.2 **Construção do banco de freqüência de alelos X-STR**

O conhecimento das freqüências de alelos STR é de grande importância para estudos populacionais e de genética forense, sendo utilizadas em cálculos estatísticos relativos à investigações criminais, ou de vínculo genético, para a caracterização, por exemplo, da raridade de um perfil genético em população.

4.2.1 Estimativas das freqüências alélicas

As freqüências alélicas dos marcadores genéticos incluídos no sistema LDD (X-STR) Decaplex foram estimadas a partir de uma amostra constituída por 170 indivíduos nascidos no Rio de Janeiro autodenominados afrodescendentes. Nas tabelas 4, 5 e 6 são mostradas as distribuições das freqüências alélicas em homens e mulheres, bem como na amostra total (n= 272).

No *locus* DXS7133, os alelos 9, 10 e 11 foram os mais freqüentes, à semelhança do verificado nas populações de Portugal, Espanha, Argentina, Costa Rica, Colômbia, Rio de Janeiro (miscigenada), São Paulo e Mato Grosso do Sul, (Gusmão *et al.*, 2009), bem como em uma amostra populacional de Uganda (Gomes *et al.*, 2009). O alelo 15, com freqüência de 0,37% nos indivíduos afrodescendentes do Rio de Janeiro, não foi observado em nenhuma das amostras populacionais estudadas por Gusmão e colaboradores (2009), sendo, entretanto, verificado em uma amostra populacional da região amazônica, com freqüência de 0,2% (Rodrigues *et al.*, 2008).

No *locus* DXS7424, os alelos 13, 14 e 15 apresentaram freqüências elevadas, e tal padrão de distribuição foi verificado na população de Gana (Poetsch *et al.*, 2009). Na amostra populacional miscigenada do Rio de Janeiro (Tavares *et al.*, 2008), os alelos mais freqüentes foram 14, 15 e 16. Na amostra populacional alvo deste estudo, o alelo 16 foi o quarto mais freqüente.

Os alelos 10, 11 e 12 foram os mais freqüentes do *locus* DXS8378, resultado também verificado nas populações de Portugal, Espanha, Argentina, Costa Rica, Colômbia, Rio de Janeiro (miscigenada), São Paulo e Mato Grosso do Sul (Gusmão

et al., 2009). O mesmo padrão de distribuição frequencial foi observado em estudos populacionais realizados com o sistema de amplificação comercial Argus-X8, como Japão (Tamura *et al.*, 2009), Itália (Cerri *et al*., 2008) e Gana (Thiele *et al*., 2008).

No *locus* DXS6807, os alelos mais freqüentes foram 11, 14 e 15. Estes também foram os de maiores freqüências nas amostras populacionais miscigenada do Rio de Janeiro (Tavares *et al.,* 2008) e de Gana (Poetsch *et al.*, 2009).

Os alelos 13, 14 e 15 foram os mais freqüentes no *locus* DXS7132, sendo o mesmo observado nas populações de Portugal, Espanha, Argentina, Costa Rica, Colômbia, Rio de Janeiro (miscigenada), São Paulo e Mato Grosso do Sul (Gusmão *et al.*, 2009). Estudos populacionais no Japão (Tamura *et al.*, 2009), Itália (Cerri *et al*., 2008) e Gana (Thiele *et al*., 2008), nos quais os indivíduos foram genotipados com multiplex comercial Argus-X8, também apresentaram um padrão similar de distribuição alélica.

No *locus* DXS7423, observou-se freqüências elevadas para os alelos 14, 15 e 16, sendo estas equivalentes àquelas populações de Portugal, Espanha, Argentina, Costa Rica, Colômbia, Rio de Janeiro, São Paulo e Mato Grosso do Sul (Gusmão *et al.*, 2009), Japão (Tamura *et al.*, 2009), Itália (Cerri *et al*., 2008) e Gana (Thiele *et al*., 2008). O alelo 8, verificado em 0,74% dos indivíduos da amostra em estudo, não foi observado em nenhuma das populações analisadas com o sistema Argus-X8. Sendo o mesmo observado apenas na população do Mato Grosso do Sul (Gusmão *et al.*, 2009), Uganda (Gomes *et al*., 2009) e Angola (Gomes *et al*., 2007a).

No *locus* DXS8377, os alelos 46, 49 e 50 apresentaram freqüências elevadas, sendo o mesmo padrão observado em estudos com grupos populacionais dos Estados Unidos (Gomes *et al*., 2007b) e Espanha (Aler *et al*., 2007). Os alelos 49, 50 e 51 foram os mais freqüentes em uma amostra populacional de Portugal (Pereira *et al*., 2007).

No *locus* GATA172D05, os alelos mais freqüentes foram 8, 9 e 10. Na amostra populacional do Rio de Janeiro (miscigenada) e Mato Grosso do Sul foi observado o mesmo padrão. Em Portugal, Espanha, Argentina Costa Rica e Colômbia, os mais freqüentes foram os alelos 8, 10 e 11, enquanto na população de São Paulo os alelos 9, 10 e 11 foram os mais observados (Gusmão *et al.*, 2009).

A distribuição alélica no *locus* DXS10074 mostrou freqüências mais elevadas para os alelos 15, 16 e 17, sendo o mesmo padrão também observado nas

populações do Japão (Tamura *et al.*, 2009), Itália (Cerri *et al*., 2008) e Gana (Thiele *et al*., 2008).

No *locus* DXS10101, os alelos mais freqüentes foram 31, 32 e 33, à semelhança do verificado na população de Gana (Thiele *et al*., 2008). As freqüências destes alelos são próximas daquelas observadas em estudos com indivíduos do Japão (Tamura *et al.*, 2009) e Itália (Cerri *et al*., 2008). Na amostra populacional em estudo, no *locus* DXS10101 foi verificada a ocorrência do alelo 24.2 com freqüência de 2,94%, o qual não foi observado nas amostras populacionais do Japão e Itália, sendo, entretanto verificado na população de Gana, com uma freqüência de 1,69%.

Em função da não ocorrência do alelo 24.2 nas populações anteriormente estudadas, com exceção de Gana, realizou-se o seqüenciamento do referido alelo com a finalidade de confirmar as suas estruturas e tamanho. O alelo 26, sem a unidade microvariante, também foi seqüenciado como padrão de referência. Através dos resultados mostrados nas figuras 6 e 7, pode ser observada a presença de quatro timinas para os alelos 24.2 e 26, além da presença de mais duas timinas para o primeiro alelo. Tais observações confirmaram a ocorrência do alelo microvariante 24.2.

Tal alelo foi observado com freqüência significativa apenas na amostra da população do Rio de Janeiro constituída por afrodescendentes e na população de Gana. Esses resultados sugerem que o alelo 24.2 pode estar associado à herança africana. No entanto, um maior número de estudos em amostras de origem africana deve ser realizado no sentido de confirmar tal observação.

Em algumas regiões X-STR foram observadas variações nas freqüências dos alelos entre a população aqui estudada e aquelas com as quais ela foi comparada com, por exemplo, nos *loci* DXS7133, DXS7423, DXS8377, GATA172D05 e DXS10101. No entanto, foi observado que praticamente todos os alelos genotipados na amostra em estudo estão presentes em diferentes populações de distintos continentes (Gusmão *et al.*, 2009; Poetsch *et al*., 2009; Tamura *et al.*, 2009; Cerri *et al*., 2008; Thiele *et al*., 2008; Gomes *et al*., 2007A; Gomes *et al*., 2007b; Aler *et al*., 2007; Pereira *et al*., 2007). De acordo com Chakraborty e colaboradores, tal padrão mostra que as variações genéticas entre as populações são apenas devidas às diferenças das freqüências dos alelos (Chakraborty *et al*., 1999).

4.2.2 Estimativas das freqüências haplotípicas

Nos indivíduos do sexo masculino o cromossomo X é único, sendo ainda o fenômeno da recombinação genética ausente. Devido a estes fatos, os marcadores genéticos no cromossomo X dos homens constituem haplótipos que, sem alterações, são integralmente transmitidos para suas filhas.

Com o objetivo de estimar as freqüências haplotípicas relativas aos dez *loci* que formam o LDD (X-STR) Decaplex foram genotipados 68 homens não relacionados geneticamente. Verificou-se que todos os indivíduos estudados apresentam haplótipos individualizados. A freqüência haplotípica decorrente da genotipagem dos 68 indivíduos foi 0,014706 e a diversidade haplotípica 0,985293882. Resultado similar foi observado quando um outro decaplex foi utilizado em estudos com uma população de Portugal (Pereira *et al*., 2007).

A diversidade haplotípica é um parâmetro estatístico que representa a probabilidade de se selecionar dois indivíduos ao acaso na população e estes apresentarem haplótipos diferentes para um determinado conjunto de marcadores. Assim, podemos inferir que a probabilidade de dois homens não aparentados da população afrodescendente do Rio de Janeiro apresentarem o mesmo haplótipo é praticamente zero, demonstrando a alta capacidade discriminativa do sistema LDD (X-STR) decaplex.

4.2.3 Análise de parâmetros estatísticos

Após a determinação das freqüências alélicas dos marcadores X-STR na população estudada, foi realizado o cálculo de parâmetros estatísticos relacionados com análises de vinculação genética e forense. È sabido que o grau de polimorfismo dos *loci* que compõem um multiplex está diretamente relacionado com a obtenção de índices satisfatórios para tais parâmetros.

Nos resultados apresentados na tabela 7, pode ser verificado que as taxas de heterozigosidade observadas não diferem significativamente daquelas esperadas para todos os *loci* analisados. Além disso, foi verificado que, em geral, os *loci* com um número maior de alelos cujas freqüências não mostraram marcante

predominância de um, ou mais alelos, na população apresentaram taxas de heterozigosidade relativamente maior como, por exemplo, os *loci* DXS7424, DXS7132, DXS8377, GATA172D05, DXS10074 e DXS10101. Por outro lado, os *loci* DXS7133, DXS8378, DXS6807, DXS7423 que apresentam um número menor de alelos com predominância de alguns alelos na população, apresentaram taxas de heterozigosidade menor, demonstrando a relação direta entre as freqüências alélicas e o grau de heterozigosidade. Estes resultados concordam com aqueles observados em outras populações como as de Portugal, Espanha, Argentina, Costa Rica, Colômbia, Rio de Janeiro (miscigenada), São Paulo e Mato Grosso do Sul (Gusmão *et al.*, 2009).

Segundo Szibor e colaboradores (2003a), uma das propriedades que deve ser avaliada em um marcador com vista a aplicação na prática forense é o equilíbrio de Hardy-Weinberg. Se uma população está em equilíbrio de Hardy-Weinberg, presume-se que esta é infinita, que não há eventos de mutação e seleção, que os cruzamentos são aleatórios e que o fluxo gênico não é capaz de alterar a composição alélica desta população. Assim, através das freqüências alélicas obtidas é possível determinar a proporção dos diferentes genótipos da população (Beiguelman, 1996).

O equilíbrio de Hard-Weinberg foi testado na parcela feminina da amostra populacional em estudo, sendo inicialmente observados desvios do equilíbrio nas regiões DXS6807, DXS7423, GATA172D05 e DXS10101. Com a aplicação da correção de Bonferroni, foi verificado que todos os *loci* encontravam-se em equilíbrio mostrando-se os mesmos apropriados para cálculos de probabilidades genotípicas, representando um importante parâmetro em investigações de vínculos genéticos e forenses. Em análises realizadas para a verificação do equilíbrio de Hardy-Weinberg nas diferentes populações utilizadas como base de comparação, foi observado desvio do equilíbrio apenas para o *locus* DXS7423 na população da Espanha (Gusmão *et al.*, 2009). O valor obtido, de 0,00178, se manteve significativo mesmo após correção de Bonferroni.

Considerando o teste exato de diferenciação entre as sub-populações masculina e feminina, observamos um valor de *p* significativo apenas para o *locus* DXS6807. Tal resultado sugere a necessidade de se analisar a distribuição das freqüências dos alelos para as sub-populações masculina e feminina ao longo de mais gerações para a confirmação de eventual diferenciação entre as mesmas. Para

os outros nove *loci* não foram observadas diferenças significativas quando as duas sub-populações foram comparadas (p≥0,05). A ocorrência de diferenças significativas em diferentes marcadores indicaria a necessidade de usar dados populacionais específicos de cada uma das sub-populações. A ausência de diferenciação entre homens e mulheres conforme verificado para os *loci* do LDD (X-STR) Decaplex permite considerar uma base de dados única para estudos com essa população.

Este resultado já era esperado uma vez que, na mulher, ocorre recombinação para a sua filha entre os cromossomos X herdados por herança materna e paterna. Neste tocante, outros estudos também mostraram a inexistência de diferenças significativas entre os gêneros masculinos e femininos, como Robino e colaboradores (2006) em estudo com a população italiana, Tavares e colaboradores (2008) na amostra populacional do Rio de Janeiro (miscigenada) e Zarrabeitia e colaboradores (2006) na população espanhola.

O presente estudo demonstrou que os marcadores X-STR selecionados para a realização deste trabalho apresentaram uma elevada diversidade gênica na população afrodescentente do estado do Rio de Janeiro, com valores variando de 0,6218 para o *locus* DXS7133 a 0,9327 para o *locus* DXS8377. Evidenciando que os *loci* que compõe o LDD (X-STR) Decaplex são altamente informativos e, portanto, úteis em aplicações forenses como estudos de identificação individual e testes de parentesco.

4.2.4 Análise do desequilíbrio de ligação

A análise de desequilíbrio de ligação, ou associação entre marcadores, é importante por possibilitar a verificação de possíveis ocorrências de segregação preferencial entre *loci* gaméticos. No presente trabalho foram estudados marcadores localizados exclusivamente em cromossomos X, tornando-se necessário comprovar a não associação entre os mesmos.

Em 2004, Schaffner realizou estudos relevantes relacionados com o cromossomo X, e uma deles se relaciona aos intervalos extensos de desequilíbrio de ligação (Schaffner 2004). Uma vez que a taxa de recombinação no cromossomo X é mais baixa do que nos cromossomos autossomos, devido ao fato deste

recombinar somente nas mulheres, poder-se-ia esperar uma maior associação entre *loci* do cromossomo X em relação aos autossômicos.

Embora os *loci* DXS8377 e DXS7423 encontrem-se muito próximos no cromossomo X, (com localização de 149.3 e 149.4 Mb respectivamente – Ver apêndice A), não foi verificada associação entre eles. Estudos anteriores já haviam mostrado a não associação entre esse par de marcadores genéticos do cromossomo X (Zarrabeitia *et al.,* 2002b; Lee *et al*., 2004). Este resultado pode ser explicado pelo fato de que apesar de os *loci* se situarem muito próximos, a pequena taxa de recombinação entre eles poderia promover, ao longo de gerações, uma importante variabilidade genética, reestabelecendo-se o equilíbrio de ligação.

Por outro lado, no nosso estudo foi obtido um valor significativo para o par DXS6807-DXS8377 (tabela 8), indicando a presença de associação entre este par de marcadores. No entanto, a distancia física entre estes *loci* é de aproximadamente 144 Mb (dados do NCBI - National Center for Biotechnology Information – Apêndice A). E após correção de Bonferroni, os resultados dos testes de associação para este par de marcadores não foram consistentes para a população em estudo.

Em síntese, não é de se esperar que os *loci* presente no LDD (X-STR) Decaplex sejam transmitidos de forma associada, podendo estes, consequentemente, serem considerados *loci* independentes.

4.2.5 Parâmetros forenses

Os parâmetros estatísticos na análise forense calculados com base na freqüência dos alelos observados para o LDD (X-STR) Decaplex na amostra populacional estudada demonstraram que o multiplex é altamente discriminativo, confirmando a possibilidade da utilização do mesmo em análises forenses (tabela 9).

Os valores elevados obtidos para a Chance Média de Exclusão acumulada para o LDD (X-STR) Decaplex, variando de 99,998541% a 99,9999819%, confirmam a utilidade deste sistema em testes de paternidade envolvendo filhas, onde estão disponíveis duos pai/filha ou em trios completos, bem como em investigações de maternidade de indivíduos do sexo masculino. Os valores de MEC para trios e duos verificados em diversas populações mundiais mostraram-se bem próximos daqueles observados para a amostra populacional avaliada no presente estudo. Quando trios

e duos foram analisados, os valores de MEC foram, respectivamente, de 99,999% e 99,996% na amostra populacional da Espanha (Aler *et al*., 2007), 99,99994% e 99,996% para Portugal (Pereira *et al*., 2007), 99,9997% e 999862% na amostra de Uganda (Gomes *et al.*, 2009) e de 99,9997% e 99,98% para Gana (Poetsch *et al*., 2009).

O Poder de Discriminação (PD) corresponde à probabilidade de dois indivíduos selecionados ao acaso na população apresentarem genótipos diferentes. Neste trabalho, o PD acumulado foi de 99,99999999995% no componente feminino da amostra populacional em estudo e de 99,9999942% no componente masculino. Portanto, a chance de duas mulheres, ou de dois homens, apresentarem idênticos perfis genéticos para o conjunto de X-STR considerado é de 0,0000000000005 e 0,000000058.

O PD acumulado observado em outras populações mundiais mostra que para a população de Portugal, este variou de 99,9999999996% no componente feminino e 99,99999% para o componente masculino, quando utilizando um sistema contendo dez *loci* X-STR (Pereira *et al*., 2007). Para a população de Gana, um sistema multiplex contendo onze regiões X-STR forneceu um PD para mulheres e homens de 99,999999% e 99,9999%, respectivamente (Poetsch *et al*., 2009). Tais resultados confirmam o potencial do sistema LDD (X-STR) Decaplex em estudos de identificação humana.

O elevado grau de polimorfismo dos dez *loci* que constituem o LDD (X-STR) Decaplex foi demonstrado através da análise das distribuições das suas freqüências alélicas, assumindo como medidas de análise do polimorfisrmo parâmetros como a taxa de heterozigosidade e a diversidade gênica. Nos resultados obtidos não foram demonstradas diferenças significativas entre as taxas de heterozigosidade observada e esperada, e os valores elevados de diversidade genética, evidenciam a eficiência do sistema, mostrando-se este altamente informativo em análises forenses e para identificação individual.

4.3 **Estratégias de validação**

Estudos de validação vêm sendo realizados com o objetivo de se identificar e determinar diferentes aspectos relacionados com procedimentos metodológicos, direcionados à otimização de técnicas aplicadas em análises forense que utilizam como fonte de estudo evidências ou amostras biológicas (Acar *et al*., 2009). Neste sentido, é de fundamental importância um estudo minucioso de alguns aspectos como especificidade da amplificação, estabilidade, sensibilidade do multiplex, reprodutibilidade do método, parâmetros da PCR, presença de artefatos e misturas de diferentes amostras de DNA.

Neste trabalho, foram realizados alguns dos estudos relatados acima relacionados com os procedimentos de validação objetivando o aperfeiçoamento e aplicação do LDD (X-STR) Decaplex em investigações forenses e de vinculação genética.

4.3.1 Efeito da quantidade de DNA na eficiência do sistema multiplex

A sensibilidade relacionada com a quantidade mínima de DNA necessária à amplificação e detecção dos alelos dos dez *loci* incluídos no LDD (X-STR) Decaplex foi avaliada através da adição de 0,0625 a 10ng de DNA a diferentes reações de PCR (figuras 8 a 15). O DNA utilizado para tal análise foi o padrão 9947A.

Os resultados mostraram produtos de amplificação para todos os *loci* do multiplex quando 10 a 2ng de DNA eram adicionados às reações de PCR, sendo ainda detectados alelos quando 1 a 0,25ng de DNA eram aplicados às reações. É importante destacar que mesmo nas reações contendo 0,25ng de DNA, onde houve a perda dos alelos do *locus* DXS6807, as outras nove regiões tipadas são suficientes para a discriminação de um perfil biológico. Tal informação é de grande relevância para as análises de amostras criminais, que muitas vezes se encontram altamente degradadas e com pouca quantidade de DNA.

Os resultados obtidos nesta análise estão em concordância com aqueles observados em outros estudos, onde se verificou que a faixa ótima de amplificação e

detecção de alelos de um sistema multiplex contendo dez marcadores variava entre 0,75 e 2ng de DNA (Levedakou *et al.*, 2002).

Por outro lado, se uma quantidade elevada de DNA é adicionada às reações de PCR, um excesso de produto de amplificação pode ser gerado. A adição de quantidades elevadas de DNA às reações de PCR pode dificultar a análise, principalmente, de regiões que apresentam bandas *stutter*, onde esta pode ser erroneamente confundida com um alelo real, gerando conseqüentemente falsos resultados. Como apresentado na figura 8, o resultado obtido pela adição de 10ng de DNA na reação de PCR gerou bandas *stutter* na região DXS8377 de aproximadamente 1.000 RFU.

4.3.2 Estudo de reprodutibilidade do LDD (X-STR) Decaplex

Este estudo teve como objetivo avaliar a aplicabilidade do sistema LDD (X-STR) na rotina forense através da verificação da reprodutibilidade dos resultados quando da genotipagem de distintas amostras biológicas. Dente, mucosa oral, bulbo capilar, sangue, urina e sêmen foram utilizadas fontes de DNA. Tais amostras foram amplificadas, adicionalmente, com o sistema autossômico *Identifiler.*

Em relação ao tipo de material empregado neste estudo, o elemento dentário era considerado a amostra forense mais crítica por se tratar de uma amostra exumada. Tais amostras são consideradas de difícil análise devido à presença de ácidos húmicos e fúlvicos, do solo, que parecem inibir a reação de PCR (Góes, 2003).

Em nossos experimentos, observou-se amplificação de todos os alelos de todas as regiões alvo do sistema LDD (X-STR) Decaplex desenvolvido no presente estudo, bem como para o sistema *Identifiler*. Tal fato contribui para demonstrar a utilidade do sistema decaplex para amplificação de diferentes amostras biológicas, incluindo amostras provenientes de restos mortais, com qualidade dos produtos de amplificação tão satisfatória quanto aquela obtida com um sistema de amplificação comercial. Para a análise dos produtos gerados através das reações de PCR, observou-se que as quantidades de DNA, empregadas nas diferentes reações, foram suficientes para a amplificação de todas as regiões STR de ambos os sistemas de tipagem.

4.3.3 Análises de bandas *stutter*

No desenvolvimento de sistemas multiplex para aplicação em identificação humana por DNA, o aparecimento de artefatos pode representar um sério problema para a correta definição dos alelos dentre os produtos de amplificação. Alguns artefatos são decorrentes de falhas durante o processo de amplificação e podem resultar em bandas de baixa intensidade denominadas *stutter* (figura 32), podendo ser confundidas como alelo verdadeiro (Daniels *et al*., 2004). Tal problema pode ser minimizado através da análise e determinação da relação entre as intensidades das bandas *stutter* e do alelo verdadeiro.

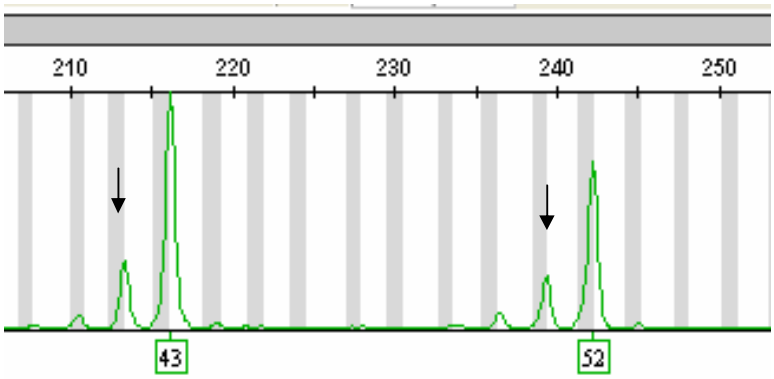


Figura 32 - Eletroferograma mostrando bandas *stutter* (setas) em relação ao alelo verdadeiro.

O objetivo deste estudo foi determinar a relação entre as intensidades de banda *stutter* e seus respectivos alelos reais para todos os *loci* do sistema LDD (X-STR) Decaplex. Tal análise foi realizada através da amplificação por PCR de quinze amostras de DNA, sendo os resultados mostrados na figura 28.

De acordo com os resultados apresentados, verifica-se que para os *loci* DXS7424, DXS6807, GATA172D05, DXS10074 e DXS10101 não foram observadas bandas *stutter*. Os percentuais mais elevados destes artefatos foram verificados para os *loci* DXS7132 (9,38%) e DXS8377 (28,2%).

A utilização de STR tetranucleotídico em sistemas de tipagem é recomendada devido ao fato de estes serem menos sujeitos à falhas no processo de amplificação,

prevenindo consequentemente o aparecimento das bandas *stutter* (Van Hoofstat *et al.,* 1998). No sistema LDD (X-STR) Decaplex os *loci* DXS7424 e DXS8377 apresentam unidades de repetição trinucleotídicas, os demais *loci* unidades de repetição tetranucleotídicas.

No presente estudo, foi verificado entre *loci* com repetições trinucleotídicas um maior percentual de *stutter* para aqueles com um maior número de unidades de repetição. Na figura 28, verificamos que o *locus* DXS8377 com até 21 unidades de repetição apresentou o maior percentual de *stutter* (28,2%). Enquanto que no *locus* DXS7424, com 10 unidades de repetição, não foram observados tais artefatos. A relação entre o aparecimento de *stutter* com o aumento do número de unidades de repetição foi também observada em um *locus* autossômico por Walsh e colaboradores (1996).

De acordo com Zarrabeitia e colaboradores (2002b), a "tendência" do *locus* DXS8377 a produzir bandas *stutter* pode, por vezes, dificultar a interpretação de alguns resultados, como por exemplo, nos casos de amostras contendo misturas de DNA. Como apresentado nos testes de sensibilidade, (ver item 4.3.1 – Resultados), a intensidade das bandas *stutter* torna-se acentuada com o aumento da concentração de DNA na amostra. Logo, a concentração de DNA é um fator crítico para a obtenção de resultados coerentes.

4.3.4 Padrão de mistura de DNA

Em investigações criminais relativas, por exemplo, a abuso sexual ou a partir de cenas de crime associadas a múltiplos assassinatos**,** é comum a ocorrência de evidências contendo mistura de DNA de duas ou mais pessoas. Nestes casos, cada rastro biológico encontrado na cena de um crime ou no corpo de uma vítima pode representar um vestígio, ou uma prova fundamental para elucidar questões criminais.

Em alguns casos de análise de mistura a aplicação de sistemas multiplex contendo marcadores do cromossomo X mostra-se mais eficiente que os sistemas autossômicos, principalmente na distinção do componente feminino do masculino. A exceção quando todos os *loci* observados forem homozigóticos (Szibor *et al.,* 2003a).

Esta etapa do estudo teve como objetivo verificar a proporção da mistura de DNA que se relacionavam com a produção de todos os alelos dos dois indivíduos em condições compatíveis com as suas genotipagens. Para tal, foram realizadas reações de PCR utilizando os sistemas LDD (X-STR) Decaplex e *Identifiler* em diferentes proporções de DNA feminino-masculino, tais como A - 10:1ng; B - 5:1ng; C - 2,5:1ng; D - 1:1ng; E - 1:2,5ng; F - 1:5ng; e G - 1:10ng

Quanto às intensidades dos alelos, os resultados mostrados nas figuras 29.A – 29.O indicam que para o sistema LDD (X-STR) Decaplex a proporção 2,5:1ng de DNA feminino-masculino foi a que se proporcionou os melhores resultados. Já com sistema autossômico, tal proporção foi a de 1:1ng. Esse resultado reflete uma maior eficiência do sistema LDD (X-STR) Decaplex em relação ao sistema *Identifiler* por possibilitar a amplificação dos alelos masculinos mesmo quando o DNA masculino se encontra em menor quantidade comparado com o feminino. Os resultados estão em concordância com os observados no estudo de validação do sistema comercial Argus X-8, (Acar *et al.*, 2009).

Quanto à parâmetro proporção mínima detectável de DNA masculino em relação ao DNA feminino, foi verificado que um excesso de dez vezes de DNA feminino (10ng) ainda permite a fiel genotipagem do DNA masculino (1ng) em ambos os sistemas de amplificação. Tal resultado expressa uma eficiência para o LDD (X-STR) Decaplex tão satisfatória quanto aquela obtida para o sistema de amplificação comercial *Identifiler*, quando quantidades mínimas de DNA são disponíveis para a determinação de perfis genéticos em evidências contendo mistura de amostras biológicas de distintos indivíduos.

4.4 **Estudo de casos**

O modo de transmissão do cromossomo X lhe confere características particulares aplicáveis em investigações de vínculos genéticos na ausência do suposto pai. O perfil genético do investigado pode ser obtido através da genotipagem de suas filhas biológicas, e de suas respectivas genitoras, inferindo-se o haplótipo do investigado a partir dos idênticos alelos paternos das suas filhas. Portanto, quando da investigação de relações de parentesco pai/filha, os

marcadores X-STR mostram-se altamente informativos, apresentando maior eficiência comparada aos STR autossômicos (Aquino *et al.,* 2009).

Quando o suposto pai não é excluído de ser o pai biológico da criança em questão um índice de vinculação entre as partes, que expressará a probabilidade de paternidade, deve ser calculado a partir da freqüência do haplótipo X-STR.

Parâmetros estatísticos relativos às análises de vinculação genética e forense foram calculados com base nas freqüências dos alelos de marcadores do sistema LDD (X-STR) Decaplex para um trio envolvendo mãe, criança e suposta avó paterna, visando a determinação da eficácia do sistema na prática forense. Com o intuito de comparar as eficiências do sistema LDD (X-STR) Decaplex e de um sistema autossômico largamente utilizado na prática forense, os parâmetros foram também calculados para o *Identifiler*.

Neste contexto, o índice de vinculação é obtido pela relação entre a hipótese de se observar na população os genótipos da mãe, da criança e da suposta avó paterna, sendo a última efetivamente a avó paterna sobre a hipótese de outra pessoa que não a testada ser a avó paterna. Com base nos resultados apresentados na tabela 10, pode-se observar um valor de ICV de 1 em 40.337.671,37, com uma probabilidade de vinculação (PV) superior a 99,99999% a favor de a suposta avó paterna testada ser a avó biológica da criança em questão em relação a qualquer outro indivíduo na população não relacionado geneticamente a ela. O poder de exclusão acumulado para os dez *loci* analisados neste caso é de 99,7220%.

Na tabela 12, é possível comparar os valores obtidos para o sistema LDD (X-STR) Decaplex com os obtidos para o *Identifiler*. O valor de ICV para o LDD (X-STR) Decaplex foi de 40.337.671,37 enquanto o obtido para o *Identifiler* foi 1.221,81, demonstrando a eficiência do sistema desenvolvido em investigações de vínculo genético como a referida.

Adicionalmente, a Probabilidade de Vínculo genético verificada para o sistema LDD (X-STR) Decaplex foi de 99,99999% enquanto que para o sistema *Identifiler* foi de 99,9182%, demonstrando que o conjunto de *loci* utilizado no presente estudo é altamente informativo, podendo ser utilizado diretamente em investigações de vinculação genética ou em complementação a outros sistemas, como o *Identifiler*.

Por outro lado, o sistema *Identifiler* apresentou um maior Poder de Exclusão (PE) acumulado (99,9999%) comparado ao LDD (X-STR) Decaplex (99,7220%). A

obtenção de um maior PE acumulado para o *Identifiler* se relaciona com as diferenças de número e graus de polimorfismos dos *loci* que constituem cada um dos sistemas.

A probabilidade de ocorrer coincidências entre perfis genéticos (PC) e a raridade de um perfil genético obtido pelo cálculo da Razão de Verossimilhança (RV) observados para o LDD (X-STR) Decaplex foram tão significativos quanto aqueles obtidos para o *Identifiler,* e indicam ser praticamente impossível uma completa coincidência entre perfis genéticos quando dez *loci* X-STR são analisados.

Pesquisas de vinculação genética como a abordada é muito comum na rotina dos laboratórios forenses, a maioria dos quais as realizam através da análise exclusiva de *loci* STR autossômicos. Como demonstrado, a incorporação de marcadores STR do cromossomo X nas suas rotinas elevará significativamente a possibilidade de se obter resultados conclusivos.

Em investigações de paternidade, tendo em vista que o pai biológico transmite o seu único cromossomo X para todas a(s) sua(s) filha(s), o índice de vinculação pode ser calculado com base na freqüência na população do haplótipo transmitido pelo suposto pai para a suposta filha. Considerando tal situação, a freqüência de cada um dos diferentes haplótipos definidos para os 68 indivíduos genotipados com o LDD (X-STR) Decaplex de 0,014706% se relacionaria com um índice de paternidade de 68, indicando que tal dado poderia apenas complementar genotipagens realizadas com sistemas autossômicos. Alternativamente, maior número de indivíduos do sexo masculino poderia ser genotipado como forma de se definir inclusões de paternidade envolvendo filhas apenas com a utilização de *loci* STR do cromossomo X.

5 CONCLUSÃO

1- O desenvolvimento do sistema LDD (X-STR) Decaplex possibilitou a amplificação eficiente de dez marcadores X-STR em uma única reação de PCR, permitindo a genotipagem e caracterização genética de uma amostra populacional do estado do Rio de Janeiro autodenominada afrodescendente, de uma forma rápida e utilizando um baixo volume de amostra.

2- Os marcadores selecionados para este estudo mostraram-se altamente polimórficos na amostra populacional estudada, apresentando uma elevada diversidade genética.

3- Através da análise do desequilíbrio de ligação, não foi observada a associação entre os marcadores na população estudada, podendo estes *loci* ser tratados como independentes para a análise forense.

4- Os parâmetros de avaliação da eficiência forense mostraram que esses marcadores são altamente informativos e, portanto, úteis em aplicações forenses. Os elevados valores de poder de discriminação acumulado para o conjunto de *loci* STR confirmam a potencialidade do LDD (X-STR) Decaplex em estudos de identificação; as elevadas chances de exclusão em trios e duos comprovam a utilidade deste sistema em casos de investigação de parentesco biológico.

5- O limite de sensibilidade do LDD (X-STR) Decaplex em condições padrão é em torno de 0,5ng. A eficiência do sistema foi demonstrada também na reprodutibilidade da genotipagem de diferentes amostras biológicas, incluindo as de restos mortais, bem como para a tipagem de misturas onde co-existam duas amostras biológicas.

REFERÊNCIAS

1- Acar E, Bulbul O, Rayimoglu, Shahzad M S, Argac D, Altuncul H, et al. Optimização and validation studies of the Mentype Argus X-8 kit for paternity cases. Forensic Sci Int: Genetics. 2009; 2:47-48.

2- Aler M, Diz OS, Gomes I, Gisbert M, Carracedo A, Amorim A, et al. Genetic data of 10 X-STRs in a Spanish population sample. Forensic Sci Int. 2007; 173: 193–196.

3- Aquino J, Peixe C, Silva D, Tavares C, Carvalho EF. A X-chromosome STR hexaplex as a powerful tool in deficiency paternity cases. Forensic Sci Int: Genetics. 2009; 2: 45–46.

4- Athanasiadou D, Bellinghausen BS, Rittner C, Alt KW, Schneider PM. Development of a quadruplex PCR system for the genetic analysis of X-chromosomal STR loci. Int Congress Series. 2003; 1239: 311-314.

5- Avery OT, Macleod C, Mccarty M. Studies on the chemical nature of the substance inducing transformation of Pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus type III. J. Exper. Med. 1944; 79 (2): 137-158.

6- Balazs I, Baird ML, McElfres K, Udey J. Hardy-Weinberg equilibrium analysis of DNA polymorphic *loci* that exhibit a continuos distribution of fragment size. In H. Polesky and W. Mayr, Eds. Advances in Forensic Haemogenetics 3. New York: Springer-Verlag. 1990; 71-74.

7- Becker D, Rodig H, Augustin C, Edelmann J, Gotz F, Hering S, et al. Population genetic evaluation of eight X-chromosomal short tandem repeat *loci* using Mentype Argus X-8 PCR amplification kit. Forensic Sci Int: Genetics. 2008; 2: 69-74.

8- Beiguelman B. Curso Prático de Bioestatística. 4ª ed. Sociedade Brasileira de Genética: Ribeirão Preto, SP, Brasil, 1996.

9- Bentayebi K, Picornell A, Bouabdeallah M, Castro JA, Aboukhalid R, Squalli D, et al. Genetic diversity of 12 X-chromosomal short tandem repeats in the Moroccan population. Forensic Sci Int: Genetics. *in press.*

10- Bini C, Ceccardi S, Ferri G, Pelotti S, Ferri G, Alù M, et al. Development of a heptaplex PCR system to analyse X-chromosome STR loci from five italian population samples a collaborative study. Forensic Sci Int. 2005; 153: 231-236.

11- Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilità. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze. 1936; 8:3-62.

12- Botstein D, White RL, Skolnick M, Davis RW. Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. Am J Hum Genet. 1980; 32:314-331.

13- Budowle B, Chakraborty R, Carmondy G Monson, KL. Source attribution of a Forensic DNA profile. Forensic Sci communications. 2000; 2(3):1-6.

14- Butler JM. Fundamentals of forensic DNA typing. Edition Elsevier Academic Press; 2009.

15- Cerri N, Verzeletti A, Gasparini F, Poglio A, Mazzeo E, Ferrari F. Population data for 8 X-chromosome STR loci in a population sample from Northern Italy and from the Sardinia island. Forensic Sci Int. 2008; 1: 173-175.

16- Chakraborty R, Kidd K. The utility of DNA typing in forensic work. Science. 1991; 254: 1735-1739.

17- _____, Stivers DN, Su B, Zohong Y, Budowle B. The utility of short tandem repeat *loci* beyond human identification: Implications for development of new DNA typing systems. Eletrophoresis. 1999; 20: 1682-1696.

18- Daniels DL, Hall AM, Ballantyne J. SWGDAM. Developmental Validation of a 19-*Locus* Y-STR System for Forensic Casework. J.Forensic Sci. 2004; 49(4): 668-683.

19- Deininger PL, Batzer MA. Alu Repeats and Human Disease. Molecular Genetics and Metabolism. 1999; 67: 183–193.

20- Desmarais D, Zhong Y, Chakraborty R, Perreault C, Busque L. Development of a highly polymorfic STR marker for identity testig purposes at the human androgen receptor gene (HUMARA). J Forensic Sci. 1998; 43: 1046-1049.

21- Domingues PM, Gusmão L, Silva DA, Amorim A, Rinaldo WP, Carvalho EF. Sub-Saharian África descendents in Rio de Janeiro (Brazil): population and mutation data for 12 Y-STR loci. Int J Legal Med. 2007; 12: 238-241.

22- Edelmann J, Lessig R, Klintschar M, Szibor R. Advantages of X-chromossomal microsatellites in deficiency paternity testing: presentation of cases. ICS. 2004; 1261: 257-259.

23- _____, Bonengel SL, Naue J, Hering S. X-chromosomal haplotype frequencies of four linkage groups using the Investigator Argus X-12 Kit. Forensic Sci Int: Genetics. *in press.*

24- Excoffier L, Laval G, Schneider S. Arlequin version 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1. 2005; 47-50.

25- Fowler JC, Burgoyne LA, Scott AC, Harding HW. Repetitive deoxyribonucleic acid (DNA) and human genome variation – A concise review relevant to forensic biology. J Forensic Sci. 1988; 33: 1111-1126.

26- Gill P, Jeffreys AJ, Werrett DJ. Forensic application of DNA fingerprints. Nature. 1985; 318: 577-579.

27- Góes ACS. Tipagem humana por DNA: otimização de condições para análise post mortem [Tese doutorado]. Rio de Janeiro, RJ: Universidade do Estado do Rio de Janeiro, 2003; 257 pp.

28- _____, Carvalho EF, Gomes I, Silva DA, Gil EHF, Amorim A, et al. Population and mutation analysis of 17 Y-STR loci from Rio de Janeiro (Brazil). Int J Legal Med. 2005; 119: 70-76.

29- Gomes I, Alves C, Maxzud K. Analysis of 10 X-STRs in three african populations. Forensic Sci Int: Genetics. 2007a; 1: 208-211.

30- _____, Prinz M, Pereira R, Meyers C, Mikulasovich RS, Amorim A, et al. Genetic analysis of three US population groups using an X-chromosomal STR decaplex. Int J Legal Med. 2007b; 121:198-203.

31- _____, Pereira V, Gomes V, Prata MJ, Pinto N, Carracedo A, et al. The Karimojong from Uganda: Genetic characterization using an X-STR decaplex system. Forensic Sci Int: Genetics. 2009; 3: e127–128.

32- Gusmão L, Alves C, Sánchez P, Zarrabeitia MT, Abovich MA, Aragon I, et al. Results of the GEP-ISFG collaborative study on an X-STR Decaplex. Forensic Sci Int: Genetics. 2008; 1: 677-679.

33- _____, Diz PS, Alves C, Gomes I, Zarrabeitia MT, Abovich M, et al. A GEP collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations. Int J Legal Med. 2009; 123: 227-234.

34- Hartl DL & Clark AG. Principles of Population Genetics 3ª. ed. Massachussets: Sinauer Associates. 1997.

35- Horvath G, Zalan A, Kis Z, Pamjav H. A genetic study of 12 X-STR loci in the Hungarian population. Forensic Sci Int: Genetics. *in press.*

36- Jeffreys AJ, Wilson V, Thein S. Hipervariable “minisatellite” regions in human DNA. Nature. 1985; 316: 76-79.

37- _____. Highly variable minisatellites and DNA fingerprints. Biochem SocTrans. 1987; 15: 309-317.

38- Jobling M, Hurles M, Tyler SC. Human Evolutionary Genetics. Garland Science 2004; 125-137.

39- Kayser M. The Human Y Chromosome - An Introduction to Genetics and Applications. Forensic Sci Review. 2003; 15: 77-89.

40- Kohlrausch FB, Callegari-Jacques SM, Tsuneto LT, Petzl-Erler ML, Hill K, Hurtado AM. Geography influences microsatellite polymorphism diversity in Amerindians. Am J Phys Antropol. 2005; 126(4) 463-470.

41- Korf BR. Genética humana e genômica. 3ª ed. Rio de Janeiro: Guanabara Koogan; 2008.

42- Lareu MV, Phillips CP, Carracedo A, Lincoln PJ, Syndercombe C D, Thomson JA. Investigation of the STR locus HUMTH01 using PCR and two electrophoresis formats: UK and Galician Caucasian population surveys and usefulness in paternity investigations. Forensic Sci Int.1994; 66: 41-52.

43- Laura M. Cainé MS, Zarrabeitia MT, Riancho JA, Pinheiro M. Genetic data of a Brazilian population sample (SantaCatarina) using an X-STR decaplex. Forensic J Legal Med. 2010; 17: 272-274.

44- Lee HY, Park MJ, Jeong CK, Lee SY, Yoo JE, Chung U, et al. Genetics characteristics and population study of 4 X-chromosomal STRs in Koreans: evidence for a null allele at DXS9898. Int J Legal Med. 2004; 118: 355-360.

45- Lesko LJ, Salerno RA, Spear BB, Anderson DC, Anderson T, Brazell C. Pharmacogenetics and Pharmacogenomics in Drug Development and Regulatory Decision Making: Report of the First FDA-PWG-PhRMA-DruSafe Workshop. J Clin Pharmacol. 2003; 43: 342-358.

46- Levedakou NE, Freeman AD, Budzynski JM. Characterization anda validation studies of PowerPlex 2.1, a nine-locus short tandem repeat (STR) multiplex system and a penta D monoplex. J Forensic Science 2002; 47(4): 1-16.

47- Manta FSN. Desenvolvimento e validação de sistemas multiplex de tipagem para identificação humana por DNA [Dissertação de mestrado]. Rio de Janeiro, RJ: Universidade do Estado do Rio de Janeiro, 2008; 128p.

48- Martins JA, Costa JC, Paneto GG, Gusmão L, Diz PS, Carracedo A, et al. Genetic data of 10 X-chromosomal loci in Vitória population (Espírito Santo State, Brazil). Forensic Sci Int: Genetics. 2009; 2: 394–395.

49- Mertens G, Gielis M, Mommers N, Mularoni A, Lamartine J, Heylen H, et al. et al. Mutation of the repeat number of the HPRTB locus and structure of rare intermediate alleles. Int J Legal Med. 1999; 112: 192-194.

50- Micka KA, Amiott EA, Hochenberry TL, Sprecher CJ, Lins AM, Rabbach DR, et al. validation of a nine–*locus* and a four-*locus* fluorescent STR multiplex system. J.Forensic Sci 1999; 44(6): 1243-1257.

51- Miller S. A simples Salting out procedure for extraction DNA from human nucleated cells. Nucl. Acids Res. 1988; 16 (3): 1215.

52- Monaghanl MT, Balke M, Gregory TR, Vogler AP. DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. Phil. Trans. R. Soc. B. 2005; 360: 1925–1933.

53- Morral N, Estivill X. Multiplex PCR amplification of three microsatellites within the CFTR gene. Genomics. 1992; 13:1362-1364.

54- National Research Council. The Evaluation of Forensic DNA Evidence. National Academy Press 1996.

55- Nei, M. Molecular Evolutionary Genetics. Columbia University Press, New York, NY. 1987.

56- Pasternak JJ. Uma introdução à genética molecular humana: mecanismos das doenças hereditárias. 2ª ed. Rio de Janeiro. Guanabara Koogan, 2007.

57- Pena SDJ. Pitfalls of paternity testing based solely on PCR typing of minisatellites and microsatellites. Am J Hum. Genet. 1995; 56: 1503-1506.

58- Pereira R, Gomes I, Amorim A, Gusmão L. Genetic diversity of 10 X chromosome STRs in northern Portugal. 2007; 12:192–197.

59- Poetsch M, Petersmann H, Repenning A, Lignitz E. Development of two pentaplex systems with X-chromosomal STR loci and their allele frequencies in a northeast German population. Forensic Sci Int. 2005; 155: 71-76.

60- _____, Mostaqim DE, Tschentscher F, Edmund NLB, Timmann C, Horstmann RD, et al. Allele frequencies of 11 X-chromosomal loci in a population sample from Ghana. Int J Legal Med. 2009; 123: 81-83.

61- Robino C, Giolitti A, Gino S, Torre C. Development of two multiplex PCR systems for the analysis of 12 X-chromosomal STR loci in a northwestern Italian population sample. Int J Legal Med. 2006; 120: 315-318.

62- Rodrigues EMR, Leite FPN, Hutz MH, Palha TJBF, Santos AKCR, Santos SEBS. A multiplex PCR for 11 X chromosome STR markers and population data from a Brazilian Amazon Region. Forensic Sci Int: Genetics. 2008; 2: 154–158.

63- Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, et al. The DNA sequence of the human X chromosome. Nature. 2005; 434: 325-337.

64- Sambrook J, Fritsch EF, Maniats T. Molecular cloning: a laboratory manual. 2ª ed; vol 2, 1989.

65- Schaffner SF. The X Chromosome in population genetics. Nature Reviews Genetics. 2004; 5: 43-48.

66- Shin KJ, Kwon BK, Lee SS, Yoo JE, Park MJ, Chung U, et al. Five highly informative X- chromosomal STRs in Koreans. Int J Legal Med. 2004; 118: 37-40.

67- Shin SH, Yu JS, Park SW, Min GS, Chung KW. Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. Forensic Sci Int. 2005; 147: 35-41.

68- Silva DA. Metodologia para análise forense por DNA: Estudo de validação para 14 loci STR e construção de banco de dados [Tese de doutorado]. Rio de Janeiro, RJ: Universidade do Estado do Rio de Janeiro, 2004a; 300 pp.

69- _____, Góes ACS, Carvalho JJ, Carvalho EF. The usefulness of DNA extracted from vaginal smear slides in forensic casework - sperm DNA typing in rape cases. São Paulo Medical Journal. 2004b; 122: 34-38.

70- Silva IHF, Barbosa AG, Azevedo DA, Diz OS, Gusmão L, Tavares CC, et al. An X-chromosome pentaplex in two linkage groups: Haplotype data in Alagoas and Rio de Janeiro populations from Brazil. Forensic Sci Int: Genetics. 2010; 4: 95–100.

71- Smouse P, Chakraborty R. The use of restriction fragment length polymorphisms in paternity analysis. Am J Hum Genet. 1986; 38: 918-939.

72- Strachan T, Read A. Human molecular genetics. 4ª ed. Nova Iorque: Garland Science; 2010.

73- Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D. Use of X-linked markers for forensic purposes. Int J Legal Med. 2003a; 117: 67-74.

74- _____, Plate I, Edelmann J, Hering S, Kuhlisch E, Michael M., et al. Chromosome X haplotyping in deficiency paternity testing principles and case report. ICS. 2003b; 1239: 815-820.

75- Szibor R. X-chromosomal markers: Past, present and future. Forensic Sci Int: Genetics. 2007; 1: 93–99.

76- _____, Edelmann J, Hering S. The Forensic ChrX Research Group. Atualizada em 2009. Acesso em 2010. Disponível em http://www.chrx-str.org/.

77- Tamura A, Tsutsumi H, Hara M, Takada A, Saito K, Suzuki K, et al. Genetic studies of eight X-STRs in a Japanese population. Legal Medicine. 2009; 11: 451–452.

78- Tariq MA, Ullah O, Riazuddin SA, Riazuddin S. Allele frequency distribution of 13 X-chromosomal STR loci in Pakistani population. Int J Legal Med. *In press.*

79- Tavares CC, Gusmão L, Domingues C, Domingues P, Silva DA., Aquino J, et al. Population data for six X-chromosome STR loci in a Rio de Janeiro (Brazil) sample: Usefulness in forensic casework. Forensic Sci Int: Genetics. 2008; 1:164–166.

80- Thiele K, Loffler S, Loffler J, Gunthner F, Nitschke K, Edelmann J, et al. Population data of eight X-chromosomal STR markers in Ewe individuals from Ghana. Forensic Sci Int: Genetics. 2008; 1: 167–169.

81- Van Hoofstat DE, Deforce DL, Millecamps RE, Brochez VM, Van Geldre EG, Vergauwe A, et al. Population genetic study of short tandem repeat loci in the Belgian population, using capillary electrophoresis. Electrophoresis. 1998; 19 (5): 719-22.

82- Walsh PS, Fildes NJ, Reynolds R. Sequence analysis and characterization of stutter products at the tetranucleotide repeat *locus* vWA. Nucleic Acids Res. 1996; 24(14): 2807-2812.

83- Watson JD, Crick FHC. A Molecular structure of nucleic acids – A structure for deoxyribonucleicbacid. Nature. 1953; 171 (4356): 737-738.

84- Weber JL, Riess O, Wolff G, Andrew S, Collins C, Graham R, et al. Delineation of a 50 kilobases segment containing the recombination site in a sporadic case of Huntington´s disease. Nature genetics. 1992; 2:216-222.

85- Weir B. S. The second National Research Council Report on Forensic DNA Evidence. Am J. Hum. Genet. 1996; 59: 497-500.

86- Wiegand P, Berger B, Edelmann J Parson W. Population genetic comparisons of three X-chromosomal STRs. Int J Legal Med. 2003; 117: 62-66.

87- Willott GM, Crosse MA. The detection of spermatozoa in the mouth. J. Forensic Sci. 1986; 26: 125-128.

88- Wyman AR, White R. A highly polymorphic locus in human DNA. Proc. Natl. Acad. Sci. EUA. 1980; 77: 6754-6758.

89- Zarrabeitia MT, Amigo T, Sañudo C, Zarrabeitia A, Lamuño DG, Riancho JA. A new pentaplex system to study short tandem repeat markers of forensic interest on X-chromosome. Forensic Sci Int. 2002a; 129: 85-89.

90- _____, Amigo T, Sañudo C, Pancorbo MM, Riancho JA. Sequence structure and population data of two X-linked markers: DXS7423 and DXS8377. Int J Legal Med. 2002b; 116: 368–371.

91- _____, Alonso A, Martin J, Miguel A, Gonzalez G, Juan C et al. Study of six X linked tetranucleotide microsatellites: population data from five Spanish regions. Int J Legal Med. 2006; 120:147–150.

92- Zeng X, Ren Z, Chen J, Lv D, Tong D, Chen H, et al. Genetic polymorphisms of twelve X-chromosomal STR loci in Chinese Han population from Guangdong Province. Forensic Sci Int: Genetics. *in press.*

APÊNDICE A - Marcadores X-STR descritos e sua localização no cromossomo. A localização dos marcadores foi obtida a partir das bases de dados NCBI “National Center for Biotechnology Information” e Rutgers Map. Adaptado de www.chrx-str.org.

Loci	**Localização física (Mb) NCBI**	**Localização genética (cM) [Rutgers Map V.2]**
DXS9900	1.307	0.31
DXS6807	4.753	14.76
DXS9895	7.387	17.09
DXS9906	7.391	17.10
DXS10148	9.198	19.84
DXS10135	9.199	20.03
DXS8378	9.330	20.21
DXS9902	15.234	32.32
DXS6795	23.254	44.24
DXS9907	32.010	55.32
DXS6810	42.804	75.12
GATA144D04	44.898	78.96
DXS10076	48.194	85.04
DXS10077	48.202	86.06
DXS10078	48.207	85.07
DXS10161	55.999	89.67
DXS10160	56.506	89.86
DXS10159	56.766	90.01
DXS10162	61.800	90.65
DXS10163	62.000	90.66
DXS10164	62.161	90.66
DXS10165	63.994	90.73
DXS7132	64.572	90.75
DXS10079	66.632	90.82
HumARA	66.682	90.81

DXS10074	66.894	90.83
DXS10075	66.915	90.83
DXS981 (STRX1)	68.114	92.81
DXS6800	78.567	97.49
DXS6803	86.318	99.40
DXS9898	87.682	101.29
DXS9905	88.749	103.39
DXS6801	92.378	106.08
DXS6809	94.825	108.12
DXS6789	95.336	108.47
DXS6799	95.336	110.71
DXS7424	100.505	115.25
DXS101	101.300	116.15
DXS6797	107.368	117.74
DXS7133	108.928	118.18
DXS6804	111.999	122.32
GATA172D05	113.061	124.36
DXS7130	118.084	130.28
GATA165B12	120.706	136.18
DXS10103	133.246	149.37
HPRTB	133.443	149.66
DXS10101	133.482	149.75
GATA31E08	140.062	160.54
DXS9908 (DXS7127)	142.769	169.87
DXS8377	149.310	183.66
DXS10146	149.335	183.72
DXS10134	149.401	183.96
DXS10147	149.410	184.01
DXS7423	149.460	184.19
DXS10011	150.939	188.70

APÊNDICE B - Perfis genéticos constituídos pelos alelos tipados após amplificação, por PCR, com o LDD (X-STR) Decaplex, de alíquotas dos DNA de 102 mulheres afrodescendentes do Rio de Janeiro, conforme autodenominação.

CÓDIGO	DXS7133	DXS7424	DXS8378	DXS6807	DXS7132	DXS7423	DXS8377	GATA172	DXS10074	DXS10101
M1	11	14	10	11	14	13	50	9	16.2	31
	11	15	12	14	15	14	51	9	17	31.2
M2	9	17	12	11	12	15	44	10	16	29
	11	18	12	11	17	15	52	10	16	32
M3	11	14	10	15	13	14	47	6	15	31.2
	12	14	11	15	17	14	53	10	16	33
M4	11	13	9	14	12	14	47	6	12	29
	11	16	11	14	15	14	50	10	15	33
M5	11	15	10	11	14	14	42	9	12	30
	11	17	12	15	14	15	42	9	15	33
M6	11	14	10	14	13	15	45	8	15	30
	13	14	11	15	15	15	51	8	15	32
M7	9	15	10	11	13	13	42	6	12	29
	10	15	10	15	15	15	51	9	14	32
M8	10	15	12	14	12	14	41	8	12	32.2
	11	15	12	14	14	14	51	8	15	32.2
M9	9	14	11	11	14	14	48	6	12	28.2
	11	15	12	11	15	14	56	10	17	29
M10	9	13	9	11	11	14	48	6	8	24.2
	12	16	12	11	14	15	52	6	11	29
M11	11	15	11	13	12	13	49	9	12	31.2
	12	17	12	14	14	14	55	10	18	32
M12	9	13	11	11	14	14	45	10	8	32
	11	18	12	11	15	14	56	11	18	33
M13	9	13	10	15	12	17	45	7	15	32.2

	11	14	12	15	12	17	51	12	16	32.2
M14	9	11	10	14	12	13	45	8	12	28
	11	12	11	15	15	14	52	9	12	29
M15	11	14	10	11	16	15	41	6	17	30.2
	11	15	12	15	18	15	48	8	17	30.2
M16	10	15	11	11	12	14	42	7	14	29
	11	16	12	11	14	14	44	9	14	31
M17	10	13	10	11	14	14	50	10	16	32
	11	16	11	15	15	14	54	12	18	32
M18	11	11	11	11	14	14	48	9	14	31
	11	13	11	11	14	14	52	9	14	32
M19	10	9	10	14	15	15	42	6	12	33
	11	13	11	14	18	16	47	6	15	33
M20	11	12	12	11	15	13	45	6	16	28
	11	13	12	11	15	15	52	6	17	30.2
M21	9	15	9	15	13	14	43	8	12	27
	11	15	12	15	15	15	55	10	16	30
M22	11	11	10	11	13	14	45	6	12	31
	11	13	11	11	13	15	50	10	14	31.2
M23	11	11	10	11	14	14	39	10	16	29
	11	15	11	15	16	15	46	11	17	31
M24	9	13	11	11	11	13	47	6	16	26
	12	14	11	11	16	15	50	10	18	26
M25	11	13	11	11	13	14	47	6	7	28.2
	11	14	12	14	13	16	51	9	15	33
M26	11	14	11	11	15	14	41	6	12	28
	11	16	12	11	16	14	48	10	13	30
M27	11	14	10	11	13	13	46	6	17	28.2
	11	18	11	11	14	15	50	10	17	29
M28	10	14	12	11	14	14	48	9	12	29
	11	15	12	11	15	15	51	10	19	32.2
M29	10	15	10	14	14	14	48	6	15	29.2
	12	16	11	14	14	15	54	9	16	33
M30	9	13	10	11	15	14	39	9	7	30.2
	10	15	12	14	15	15	54	11	17	31

M31	9	12	11	11	13	15	49	8	17	29
	10	14	12	11	14	15	52	8	18	30.2
M32	11	10	11	11	12	8	50	9	7	31
	12	11	11	16	14	16	58	12	16	32.2
M33	11	13	10	14	15	14	42	9	15	26.2
	11	14	12	14	15	15	47	9	15	31
M34	11	12	12	11	13	15	46	7	7	29
	13	13	12	14	14	16	48	8	14	33
M35	11	12	10	11	12	15	39	6	11	30
	11	13	12	14	15	15	54	8	15	30
M36	11	14	11	11	12	15	47	6	16	29.2
	11	15	12	15	15	15	53	6	17	31.2
M37	11	14	10	11	12	15	47	6	8	32
	11	14	12	15	13	16	51	9	15	33
M38	11	12	10	12	13	14	44	8	13	31
	12	16	13	16	13	15	49	11	15	33
M39	9	14	11	11	14	14	48	9	16	30
	11	15	11	14	16	15	50	11	17	30.2
M40	10	14	12	11	13	13	46	9	14	25.2
	11	16	12	15	14	15	49	11	15	30
M41	9	12	10	11	15	14	43	6	16.2	29
	11	15	12	14	16	15	45	10	18	32.2
M42	9	16	10	11	12	14	46	6	16	30
	12	18	11	11	16	15	47	10	18	34
M43	9	13	10	11	12	14	42	8	16	29
	11	13	12	11	14	16	49	10	17	32
M44	11	12	10	14	14	14	50	6	16	26.2
	11	12	12	17	15	14	55	7	18	27
M45	11	13	11	14	12	14	46	9	10	31
	12	14	12	14	13	15	49	10	15	32
M46	11	12	10	11	13	12	40	9	11	28.2
	11	13	10	11	14	13	52	9	12	30
M47	11	10	10	11	13	13	55	9	14	29
	12	16	11	17	14	15	55	10	15	31
M48	9	13	11	11	14	14	45	10	17	29.2

	10	15	12	16	14	16	48	12	18	32
M49	11	13	11	11	12	8	46	9	12	31
	11	15	12	15	14	15	50	11	14	32
M50	11	14	11	11	14	14	46	9	15	26
	11	16	11	12	14	14	49	9	15	29
M51	9	13	10	11	12	15	47	8	16	30
	11	16	10	15	13	15	47	8	17	30.2
M52	11	11	11	11	15	14	42	10	12	29.2
	11	16	13	15	15	16	51	11	18	30
M53	11	15	11	11	14	15	44	6	16	29
	11	15	13	11	15	15	46	9	16	32
M54	11	15	11	14	14	14	47	6	11	31.2
	11	15	11	15	15	16	48	8	13	34
M55	11	13	10	14	12	14	49	8	7	32
	15	17	10	14	18	17	49	11	19	32
M56	9	13	10	11	13	14	44	8	14	28
	9	15	11	13	14	16	49	10	16	29
M57	9	14	10	14	13	14	42	8	16	31
	9	15	11	14	13	14	55	10	19	32
M58	11	12	10	11	14	15	52	6	15	31
	11	13	11	15	14	15	56	9	16	32
M59	11	12	10	11	13	15	45	9	15	29
	11	16	12	15	14	16	49	9	16	32
M60	9	15	10	14	14	14	44	10	14	29.2
	11	17	12	14	15	15	51	11	18	30
M61	9	15	12	11	13	14	44	10	17	30
	9	17	12	11	15	14	46	11	18	32
M62	10	14	11	11	14	15	47	8	13	30.2
	10	14	11	11	15	15	55	8	20	31
M63	9	10	10	11	14	15	43	6	16	27
	9	14	10	14	15	15	49	10	19	30
M64	12	14	11	11	14	14	49	7	16	30
	14	17	12	11	15	15	51	8	16	31.2
M65	9	14	10	11	14	14	41	10	14	30
	11	15	11	15	14	14	52	12	16	31

M66	9	15	10	11	13	14	46	11	17	33
	11	16	10	11	13	15	50	11	17	33
M67	9	13	10	11	14	15	49	7	7	27
	11	14	11	16	17	15	59	9	18	30.2
M68	10	13	12	14	13	14	42	6	16	30.2
	11	14	13	14	13	14	50	9	17	31.2
M69	11	13	11	11	14	14	44	8	12	33
	13	17	13	11	14	17	51	10	14	33
M70	9	13	10	14	12	14	50	10	15	31
	11	15	12	15	14	14	52	11	18	31.2
M71	9	12	10	11	13	15	46	6	8	28
	11	15	12	15	16	16	50	8	17	29
M72	11	11	10	11	14	14	50	10	7	31
	12	16	10	11	15	16	51	11	19	34
M73	11	13	11	11	13	14	48	6	7	32.2
	12	13	13	14	15	15	53	8	7	33
M74	10	16	11	15	15	14	44	10	11	31.2
	10	16	12	15	15	17	49	11	16	32
M75	10	11	10	14	13	14	44	7	15	28
	11	14	11	14	15	14	45	11	16	31.2
M76	11	11	10	11	13	14	42	7	7	33
	11	17	11	14	13	16	42	11	17	33
M77	11	13	12	14	11	13	47	8	7	31
	11	14	12	15	15	15	50	10	12	32
M78	10	13	11	11	12	13	46	6	16	29.2
	11	14	12	16	14	14	47	9	16	29.2
M79	9	14	11	11	12	14	42	8	15	29.2
	11	15	11	11	15	15	51	8	18	31
M80	11	15	11	11	12	14	39	9	7	30
	11	16	12	16	13	15	53	11	13	31
M81	11	13	10	11	14	13	49	6	10	30
	11	14	12	15	15	15	55	6	13	35
M82	11	13	10	15	11	14	44	6	7	29
	11	15	12	15	13	15	47	10	18	33
M83	9	11	12	13	10	12	52	6	12	31

	10	13	12	14	13	15	53	6	17	32.2
M84	9	16	10	15	13	14	41	8	13	28.2
	11	16	12	15	14	15	44	9	17	32.2
M85	9	14	10	11	15	13	43	6	7	29.2
	10	15	11	15	16	14	44	10	16	29.2
M86	11	13	11	11	14	15	49	6	14	32
	11	14	11	14	14	15	50	9	14	33
M87	11	16	9	11	12	14	46	7	15	28
	11	16	11	15	14	15	49	8	16	28
M88	10	13	12	11	14	14	44	9	8	29.2
	11	17	12	14	17	14	49	10	17	31
M89	9	14	10	11	13	16	49	6	14	29.2
	9	16	11	15	15	16	57	10	14	31
M90	11	12	10	11	15	14	44	6	15	32
	12	13	12	11	15	14	46	8	15	33
M91	11	12	12	11	12	15	48	9	17	27
	13	15	12	14	15	15	53	9	18	29.2
M92	10	13	10	14	12	13	45	10	11	28
	11	16	10	15	14	15	51	10	17	30.2
M93	10	14	11	14	13	13	45	6	12	32.2
	13	15	12	15	13	17	53	8	13	33
M94	9	13	10	11	14	15	44	7	12	30.2
	11	14	11	15	15	15	48	10	16	31.2
M95	9	12	11	13	12	14	43	8	7	30.2
	11	16	11	13	14	15	51	11	15	32
M96	11	11	10	11	13	14	45	9	12	29.2
	11	13	12	15	14	15	46	9	18	30
M97	11	12	10	14	14	14	50	6	16	26.2
	11	12	12	17	15	14	55	7	18	27
M98	11	16	11	11	12	14	46	6	15	29.2
	11	16	12	11	13	15	49	7	17	31
M99	11	14	10	11	14	14	49	9	7	29.2
	12	15	12	14	15	14	52	9	14	30.2
M100	11	12	12	14	14	14	43	9	14	29
	12	14	12	16	15	14	50	10	18	33

M101	10	11	11	11	13	14	47	6	13	30
	12	14	11	13	14	14	51	9	14	33
M102	11	11	12	11	12	15	46	8	7	30.2
	11	16	13	15	15	16	51	11	16	31

Legenda: M= Código de cada indivíduo do sexo feminino analisado.

APÊNDICE C – Haplótipos formados pelos alelos tipados após amplificação, por PCR, com o LDD (X-STR) Decaplex, de alíquotas dos DNA de 68 homens afrodescendentes do Rio de Janeiro, conforme autodenominação.

Hp	DXS7133	DXS7424	DXS8378	DXS6807	DXS7132	DXS7423	DXS8377	GATA172	DXS10074	DXS10101	n
Hp1	9	16	10	11	13	14	51	12	16	31.2	1
Hp2	11	13	11	11	13	15	48	8	14	28	1
Hp3	13	12	12	11	12	16	45	9	7	32.2	1
Hp4	12	16	10	15	13	14	49	8	12	26.2	1
Hp5	11	13	13	15	13	16	48	6	18	29	1
Hp6	11	14	12	15	12	14	46	6	17	32.2	1
Hp7	11	15	13	16	15	14	56	6	7	32	1
Hp8	11	14	10	11	14	15	50	9	17	28	1
Hp9	11	12	11	12	15	14	53	10	18	35	1
Hp10	9	13	12	16	15	15	49	9	15	29	1
Hp11	11	13	10	16	15	15	43	8	12	32.2	1
Hp12	11	13	10	15	14	15	46	9	8	31	1
Hp13	10	13	10	11	14	15	54	6	14	34	1
Hp14	11	16	12	11	13	15	40	6	18	33	1
Hp15	11	12	13	11	13	14	41	8	16	31	1
Hp16	11	13	11	11	14	15	40	8	17	29	1
Hp17	13	15	12	15	13	14	58	9	16	33	1
Hp18	11	11	10	11	13	14	46	7	16	29.2	1
Hp19	11	11	11	15	14	14	48	7	14	24.2	1
Hp20	11	15	11	11	14	14	41	9	16	31	1
Hp21	9	16	12	11	13	16	42	10	14	30	1
Hp22	11	15	10	11	15	15	48	6	14	30.2	1
Hp23	10	15	11	11	12	15	47	10	12	31	1
Hp24	10	15	12	15	16	14	48	10	16	28	1
Hp25	10	15	12	11	13	14	46	12	17	30.2	1

Hp26	10	13	11	13	12	15	55	6	7	30	1
Hp27	11	15	11	11	15	15	43	9	14	32	1
Hp28	11	14	12	15	13	14	42	9	7	28.2	1
Hp29	11	13	12	11	13	15	46	9	15	33	1
Hp30	12	12	10	13	12	15	47	9	15	24.2	1
Hp31	11	14	10	13	14	15	49	8	7	30	1
Hp32	9	14	11	11	13	15	53	6	12	32	1
Hp33	10	16	12	11	12	14	47	10	15	24.2	1
Hp34	11	12	11	15	14	16	52	9	15	24.2	1
Hp35	11	11	12	15	12	15	48	6	12	33	1
Hp36	11	15	11	11	14	15	50	12	15	29	1
Hp37	11	15	10	11	13	15	50	9	15	24.2	1
Hp38	9	16	11	11	12	13	50	6	16	30	1
Hp39	11	11	10	15	13	14	46	8	14	29.2	1
Hp40	11	13	12	16	15	14	51	9	13	32.2	1
Hp41	9	15	12	15	15	16	46	9	19	32	1
Hp42	11	15	9	14	15	15	49	8	7	29.2	1
Hp43	11	16	10	16	16	17	49	11	19	31	1
Hp44	9	13	10	16	14	15	52	11	15	33	1
Hp45	9	13	10	15	13	15	43	11	18	31	1
Hp46	11	12	11	11	13	14	44	6	8	28.2	1
Hp47	12	15	11	13	15	15	56	11	16	24.2	1
Hp48	10	15	12	11	14	13	47	9	14	29	1
Hp49	11	14	10	11	15	13	54	9	16	31	1
Hp50	11	15	12	15	18	15	46	8	16	33	1
Hp51	11	14	12	11	15	14	54	11	17	29	1
Hp52	9	11	12	11	17	15	50	8	15	34	1
Hp53	11	13	11	11	13	14	52	8	8	33	1
Hp54	10	13	12	11	13	14	51	6	16	32	1
Hp55	12	14	11	11	15	14	55	8	8	31.2	1
Hp56	11	14	11	15	14	15	47	9	15	32	1
Hp57	10	13	11	11	17	15	51	6	14	29	1
Hp58	11	16	10	15	16	16	52	6	16	30.2	1
Hp59	10	13	12	15	15	15	46	6	7	31.2	1
Hp60	9	14	11	15	13	15	49	10	16	31.2	1

Hp61	11	13	10	11	13	14	46	11	14	24.2	1
Hp62	11	11	11	16	14	14	47	9	15	32	1
Hp63	10	11	10	11	14	13	51	6	11	30	1
Hp64	9	15	11	16	12	13	50	8	18	31	1
Hp65	12	15	11	11	13	16	46	6	15	30.2	1
Hp66	9	15	11	11	14	14	49	6	15	30.2	1
Hp67	10	15	12	14	13	14	45	9	18	32	1
Hp68	11	15	11	15	15	15	52	6	13	31	1

Legenda: Hp= haplótipo de cada indivíduo do sexo masculino analisado; **n**= número de observações de cada haplótipo.

APÊNDICE D – Population data for six X-chromosome STR loci in a Rio de Janeiro (Brazil) sample: Usefulness in forensic casework.

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Forensic Science International: Genetics Supplement Series 1 (2008) 164–166
FSI GENETICS SUPPLEMENT SERIES
www.elsevier.com/locate/FSIGSS

Research article

Population data for six X-chromosome STR loci in a Rio de Janeiro (Brazil) sample: Usefulness in forensic casework

Celso C. Tavares [a,b], Leonor Gusmão [c], Cristiane Domingues [a], Patrícia Domingues [a], Dayse A. Silva [a,d], Juliana Aquino [a], Carla Peixe [a], António Amorim [c,e], Elizeu F. de Carvalho [a,*]

[a] *Lab. Diagnósticos por DNA, DECB, IBRAG, Univ. do Estado do Rio de Janeiro, Brazil*
[b] *Departamento de Bio. Cel. e Genética, IBRAG, Univ. do Estado do Rio de Janeiro, Brazil*
[c] *IPATIMUP, Instituto de Patologia e Imunologia Molecular da Univ. do Porto, Portugal*
[d] *Departamento de Patologia e Labs., FCM, Univ. do Estado do Rio de Janeiro, Brazil*
[e] *Faculdade de Ciências, Universidade do Porto, Portugal*

Received 21 August 2007; accepted 10 October 2007

Abstract

This study presents data for the X-chromosome STR loci DXS7133, DXS7424, DXS8378, DXS6807, DXS7423 and DXS8377. In order to establish a database, unrelated individuals (males and females) from Rio de Janeiro were typed for the above loci. No significant differences were observed between allele frequencies in male and female samples (non-differentiation exact P values ≥ 0.156). Hardy–Weinberg equilibrium was tested in the female sample and no significant deviations were found. All six markers have shown to be highly polymorphic in our sample with gene diversities varying between 0.6797 for DXS7133, and 0.9260 for DXS8377. Pairwise linkage disequilibrium analysis did not allow discharging a possible association between DXS7133 and DXS7424 alleles in Rio de Janeiro population. Parameters of forensic interest, like PD_M, PD_F, Het_{obs}, Het_{exp}, were calculated for each locus. The high discrimination power estimated in both males and females, as well as mean exclusion chance in father/daughter duos and in father/mother/daughter trios, demonstrates the usefulness of these six markers in forensic investigation.
© 2008 Elsevier Ireland Ltd. All rights reserved.

Keywords: X-chromosome; STR; Database; Rio de Janeiro; Brazil

1. Introduction

X-chromosome STR markers have been used in supplementation to autosomal and Y-chromosomal STRs markers in forensic analysis [1]. They can be used to establish kinship, especially in deficiency of paternity cases when female descendents are under investigation. Men transmit to his daughters one X-chromosome they received from their mothers. Therefore, in the absence of the father, paternal grandmother and uncles can be used to infer the father's profile. In these cases, we can also use the information obtained from sisters and half-sisters who share the X-chromosome haplotypes with females under investigation. X-STRs tend to present a mean exclusion chance higher than that of autosomal STRs [2].

The aim of this study was to describe the polymorphism of six X-chromosome STR loci, involving three X-chromosome linkage groups I, II and IV (DXS7133, DXS7424, DXS8378, DXS6807, DXS7423 and DXS8377) in Rio de Janeiro population. DXS8377 revealed to be the most polymorphic among these markers.

In order to evaluate the usefulness and potential applications of these six X-chromosome markers in forensic practice we perform HWE and LD tests, as well as genetic comparison of the two female and male samples. Statistical parameters with forensic relevance were calculated, such as PD_F, PD_M, PIC, MEC.

2. Materials and methods

Blood samples were collected from individuals living in Rio de Janeiro, Brazil. A total of 263 unrelated healthy donors (129 males and 134 females) were typed for DXS7133, DXS7424

* Corresponding author at: Rua São Francisco Xavier, 524, 20550-013, Rio de Janeiro, Brazil. Tel.: +55 21 25877902; fax: +55 21 25877662.
E-mail address: elizeufc@hotmail.com (E.F. de Carvalho).

1875-1768/$ – see front matter © 2008 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigss.2007.10.100

Table 1
Allele frequencies and forensic interest parameters

Allele	DXS7133	DXS7424	DXS8378	DXS6807	DXS7423	Allele	DXS8377
6	0.0025					39	0.0030
7				0.0030		40	0.0150
8	0.0025					41	0.0301
9	0.3224		0.0025			42	0.0481
10	0.1788	0.0076	0.3073			43	0.0602
11	0.4282	0.0277	0.3451	0.4698		44	0.0602
12	0.0504	0.0705	0.3023	0.0210	0.0060	45	0.0632
13	0.0025	0.1335	0.0403	0.0331	0.0512	46	0.0903
14	0.0126	0.1537	0.0025	0.2198	0.3433	47	0.1024
15		0.2242		0.1867	0.4216	48	0.1265
16		0.2645		0.0572	0.1385	49	0.0843
17		0.0982		0.0090	0.0361	50	0.0933
18		0.0176			0.0030	51	0.0602
20		0.0025				52	0.0783
						53	0.0240
						54	0.0271
						55	0.0180
						56	0.0120
						57	0.0030
n	397	397	397	332	332		332
PD_M	0.6797	0.8247	0.6952	0.6910	0.6811		0.9232
PD_F	0.8381	0.9455	0.8435	0.8614	0.8425		0.9888
MEC_T	0.6198	0.7995	0.6309	0.6479	0.6253		0.9179
MEC_D	0.4739	0.6823	0.4847	0.5032	0.4805		0.8532

and DXS8378 and 213 individuals (94 males and 119 females) for DXS6807, DXS7423 and DXS8377.

DNA was extracted using salting out method described by Miller et al. [3].

Multiplexes PCR reactions were performed using primers described by Poetsch et al. [4]. Primers were fluorescent labeled with fluorescein or hexamethyl rhodamin. PCR reaction was carried out in a volume of 10 μL containing 1–2 ng DNA, 0.5 μM of DXS7133 primers, 0.5 μM of DXS7424, 0.5 μM of DXS8378, 1 μM of DXS6807, 1 μM of DXS7423 and DXS8377, 200 μM of each dNTP, 1.5 mM $MgCl_2$, 1 unit Platinum Taq DNA Polymerase (Invitrogen, Corp.) and 1× buffer (20 mM Tris–HCl, 50 mM KCl, pH 8.4).

Thermocycling conditions consisted of an initial denaturation at 95 °C for 12 min; 30 cycles at 95 °C for 60 s, 60 °C for 60 s and 72 °C for 3 min; and a final extension step at 72 °C for 30 min. Alternatively it was used a different protocol consisting in an initial denaturation at 95 °C for 15 min; followed by 10 cycles of 30 s at 94 °C, 90 s at 60 °C, and 60 s at 72 °C; and 20 cycles of 30 s at 94 °C, 90 s at 58 °C, 60 s at 72 °C with a final incubation for 30 min at 72 °C [5].

Aliquots of 1 μL of PCR product were mixed with 9.5 μL formamide and 0.5 μL of ILS 600 size standard (Promega Corp., Madison, WI, USA) and separated by capillary electrophoresis on an ABI Prism 3100 Avant Genetic Analyzer instrument (AB, Applied Biosystems). Allelic sizes (bp) were determined by comparison with those from the cell line NA9947A [6].

Allele frequencies and gene diversity estimations, Hardy–Weinberg equilibrium test, linkage disequilibrium analysis of female sample, male and female samples comparison were performed using the ARLEQUIN Software Ver 3.01 [7]. Mean exclusion chance in trios involving daughters (MEC_T) and in father/daughter duos (MEC_D) as well as power of discrimination in females (PD_F) and in males (PD_M) were calculated using formulae according to Desmarais et al. [8].

3. Results and discussion

A Brazilian population sample of genetically unrelated individuals was studied for six X-chromosome markers: DXS7133, DXS7424, DXS8378, DXS6807, DXS7423 and DXS8377. Exact test of differentiation based on allele frequencies show no significant differences ($P = 0.431$ for DXS7133, $P = 0.773$ for DXS7424, $P = 0.193$ for DXS8378, $P = 0.157$ for DXS6807, $P = 0.379$ for DXS 7423 and $P = 0.329$ for DXS8377) between female and male samples. In view of this, samples were pooled and allele frequencies are presented in Table 1. Power of discrimination in males (PD_M) and in females (PD_F), and mean exclusion chance in trios (MEC_T) and in duos (MEC_D) involving daughters are indicated in Table 1. Hardy–Weinberg equilibrium was tested in the female sample and no deviations were observed at the six studied loci ($P = 0.6888$ for DXS7133, $P = 0.3073$ for DXS7424, $P = 0.3572$ for DXS8378, $P = 0.0593$ for DXS6807, $P = 0.3165$ for DXS7423 and $P = 0.1002$ for DXS8377).

Although DXS7133 and DXS7424 belong to the same linkage group [1], and thus can show association in some populations, in the male sample the exact test of linkage disequilibrium was only marginally significant ($P = 0.046360 \pm 0.000448$). Haplotype data can be obtained from the corresponding author on demand.

In accordance to other studies [5,9], DXS8377 was the most polymorphic marker in the Rio de Janeiro population.

Funding source

FAPERJ – Rio de Janeiro Government for Scientific Research Agency. The sponsor had no involvement in the development of the paper or decisions related to the paper.

Conflict of interest

None.

References

[1] R. Szibor, et al., Use of X-linked markers for forensic purposes, Int. J. Legal Med. 117 (2003) 67–74.

[2] J. Edelmann, et al., Validation of the DXS7424 and the linkage situation on the X-chromosome, Forensic Sci. Int. 125 (2002) 217–222.

[3] S.A. Miller, et al., A single salting out procedure for extracting DNA from human nucleated cells, Nucleic Acids Res. 16 (1988) 1215.

[4] M. Poetsch, et al., Development of two pentaplex systems with X-chromosomal STR loci and their allele frequencies in a northeast German population, Forensic Sci. Int. 155 (2005) 71–76.

[5] I. Gomes, et al., Genetic analysis of three US population groups using an X-chromosomal STR decaplex, Int. J. Legal Med. 121 (2007) 198–203.

[6] R. Szibor, et al., Cell line DNA typing in forensic genetics—the necessity of reliable standards, Forensic Sci. Int. 138 (2003) 37–43.

[7] L. Excoffier, et al., Arlequin ver. 3.01: an integrated software package for population genetics data analysis, Evol. Bioinform. Online 1 (2005) 47–50.

[8] D. Desmarais, et al., Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA), J. Forensic Sci. 43 (1998) 1046–1049.

[9] M.A. Moreno, et al., Validation of five X-chromosomal STRs DXS6800, DXS6807, DXS6798, DXS8377 and DXS7423 in an Antioquian population sample, Int. Congr. Ser. 1288 (2006) 295–297.

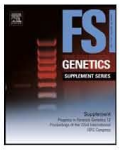
APÊNDICE E – A X-chromosome STR hexaplex as a powerful tool in deficiency paternity cases.

Forensic Science International: Genetics Supplement Series 2 (2009) 45–46

Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/FSIGSS



Research article

A X-chromosome STR hexaplex as a powerful tool in deficiency paternity cases

Juliana Aquino [a], Carla Peixe [a], Dayse Silva [a,b], Celso Tavares [a], Elizeu F. de Carvalho [a,*]

[a] DNA Diagnostic Laboratory (LDD), Biology Institute, State University of Rio de Janeiro, Brazil
[b] Pathology Department, Faculty of Medicine, State University of Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 26 August 2009
Accepted 27 August 2009

Keywords:
X-chromosome STR markers
Paternity
Deficiency cases
Likelihood ratio
Women

ABSTRACT

The X-chromosome short tandem repeat (STR) markers have been described as very adequate tools for solving deficiency paternity cases and kinship tests when women are involved. In the absence of the alleged father, presumed paternal relationship can be more efficiently investigated by using a set of six to ten X-STR markers compared to fifteen autosomal STR. For this study, we compared the usefulness of a X-STR hexaplex developed in our laboratory (DXS7133, DXS7424, DXS8378, DXS6807, DXS7423 and DXS8377) and the commercial kit Identifiler in solving deficiency paternities. We have worked on distinct groups of caseworks involving daughters, their mothers and presumed paternal grandmothers or putative half sisters and their respective mothers. The PCR products were separated by capillary electrophoresis and detected in an ABI Prism 3100. In the majority of the caseworks (>90%), the likelihood ratio (LR) obtained by using the X-STR hexaplex was higher than the LR value observed when the Identifiler kit was used for genotyping. The combination of the two STR typing systems was able to solve all the cases.

© 2009 Published by Elsevier Ireland Ltd.

1. Introduction

Autosomal short tandem repeat (STR) genotyping are the most frequent molecular tool used in DNA paternity testing [1]. However, for more than 70% of the kinship analysis autosomal STR loci do not produce conclusive results. In this context, X-chromosome STR markers are able to supplement the results of autosomal STR allowing likelihood ratios to be increased [2], especially when female offspring are under investigation since men transmit to his daughters the single X-chromosome they received from their mother [3].

The aim of this study was to compare results obtained by using fifteen autosomal STR kit with those supplemented by a X-chromosome STR hexaplex in kinship cases where putative paternal grandmothers or half sisters were genotyped in the absence of the alleged father.

2. Materials and methods

DNA of the blood samples was extracted by Chelex™ 100 resin beads (Biorad) [4]. DNA samples were profiled using Identifiler™ Kit (Applied Biosystem, Foster City, USA) and six unlinked chr-X markers, included in a X-hexaplex STR multiplex system.

2.1. X-STR amplification

PCR amplification was carried out using the Multiplex PCR Kit (Qiagen) at 1× Qiagen Multiplex PCR master mix, in a final volume of 10 μL containing 1–2 ng DNA. Final primers concentration in the reaction was 0.2 μM for all markers. Primers were fluorescent labeled with fluorescein (DXS7133, DXS7424, DXS6807 and DXS8378) and hexamethyl rhodamin (DXS7423 and DXS8377).

Thermocycling conditions consisted of an initial desnaturation at 95 °C for 12 min; 30 cycles at 95 °C for 60 s, 60 °C for 60 s and 72 °C for 3 min; and a final extension step at 72 °C for 30 min. Alternatively was used a scheme with an initial desnaturation at 95 °C for 15 min; followed by 10 cycles of 30 s at 94 °C, 90 s at 60 °C, and 60 s at 72 °C; and 20 cycles of 30 s at 94 °C, 90 s at 58 °C, 60 s at 72 °C with a final incubation for 30 min at 72 °C [2].

2.2. Analysis of PCR products

Aliquots of 1 μL of PCR product were mixed with 9.5 μL formamide and 0.5 μL of ILS 600 size standard (Promega Corp., Madison, WI, USA) and separated by capillary electrophoresis on an ABI Prism 3100 Avant Genetic Analyzer instrument with denaturing polymer 3100 POP-4™ (AB, Applied Biosystems, Foster City, CA 94404, USA) with GeneMapper ID 3.1 Analysis Software. Allelic sizes (bp) were determined by comparison with those from

* Corresponding author at: Laboratório de Diagnósticos por DNA, Rua São Francisco Xavier, 524, 20550-013, Rio de Janeiro, Brazil. Tel.: +55 21 2334 0594; fax: +55 21 2334 2183.
E-mail address: elizeufc@hotmail.com (E.F. de Carvalho).

1875-1768/$ – see front matter © 2009 Published by Elsevier Ireland Ltd.
doi:10.1016/j.fsigss.2009.08.183

Table 1
Likelihood ratios for distinct groups of kinship cases using the Identifiler (ID) and X-STR hexaplex.

	Case 1		Case 2		Case 3		Case 4	
LR	Identifiler (ID)	ID + hexaplex	Identifiler (ID)	ID + hexaplex	Identifiler (ID)	ID + hexaplex	Identifiler (ID)	ID + hexaplex
M, C and pPGM	2.37	505	13.12	85,458	3.66	434	12.00	3,839
M/MpHS, C and pHS	77.06	17,436	2.70	343	109.20	10,433	69.42	21,551
M, C, pHS and MpHS	29.00	6,128	97.11	37,637	74.00	39,433	1.04	155

M, mother; C, child; pPGM, putative paternal grandmother; pHS, putative half sister; MpHS, putative half sister mother.

the cell line NA9947A [5]. Likelihood ratios (LR) were computed with support of the alleles and male genotypes population database of Rio de Janeiro [6].

3. Results and discussion

The X-chromosome STR hexaplex was used in a population sample from Rio de Janeiro, Brazil [6]. The Accumulated Power of Discrimination and Power of Exclusion for the six loci were 0.99999 and 0.99975, respectively, demonstrating that the multiplex can be an excellent tool to complement the information given by the current commercial kits since the six loci in the multiplex revealed to be highly polymorphic to the studied population.

In this study, the Likelihood Ratios (LR) observed for 12 kinship cases were very low when only the Identifiler™ Kit was used (Table 1). After genotyping the DNA samples additionally with the X-chromosome STR hexaplex, the LR values increased from 100 to 6,500 times. Therefore, those results confirmed that the X-chromosome STR hexaplex developed in our laboratory is a powerful molecular tool for complement autosomal STR analysis in complex kinship cases.

Financial support

Financial support was granted by Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) and DNA Program - State University and Justice Court of Rio de Janeiro, Brazil.

Conflict of interest

None.

References

[1] H.A. Hammond, L. Jin, Y. Zhong, C.T. Caskey, R. Chakraborty, Evaluation of 13 short tandem repeat loci for use in personal identification applications, Am. J. Hum. Genet. 55 (1994) 175–189.
[2] D. Hatsch, C. Keyser, R. Hienne, L. Bertrand, Resolving paternity relationships using X-Chromosome STRs and bayesian networks, J. Forensic Sci. 52 (2007) 895–897.
[3] C. Toni, S. Presciuttini, I. Spinetti, A. Rocchi, R. Domenici, Usefulness of X-chromosome markers in resolving relationships: report of a court case involving presumed half sisters, Int. Congress Ser. 1288 (2006) 301–303.
[4] P.S. Walsh, D.A. Metzger, R. Higuchi, Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material, Biotechniques 10 (4) (1991) 506–513.
[5] D. Becker, H. Rodig, C. Augustin, J. Edelmann, F. Gotz, S. Hering, et al., Population genetic evaluation of eight X-chromosomal short tandem repeat loci using Mentype Argus X-8 PCR amplification kit, Forensic Sci. Int.: Genet. 2 (2008) 69–74.
[6] C. Tavares, L. Gusmão, C. Domingues, P. Domingues, D.A. Silva, J. Aquino, C. Peixe, A. Amorim, E.F. Carvalho, Population data for six X-Chromosome STR loci in a Rio de Janeiro (Brazil) sample: usefulness in forensic casework, Forensic Sci. Int. Genet.: Suppl. Ser. v1 (2008) 164–166.