

ALUISIO TRINDADE FILHO

**CARACTERIZAÇÃO GENÉTICA DA POPULAÇÃO DO DISTRITO FEDERAL
COM BASE EM MARCADORES STR DO CROMOSSOMO X**

BRASÍLIA 2010

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE CIÊNCIAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA SAÚDE**

ALUISIO TRINDADE FILHO

**CARACTERIZAÇÃO GENÉTICA DA POPULAÇÃO DO DISTRITO FEDERAL
COM BASE EM MARCADORES STR DO CROMOSSOMO X**

Dissertação apresentada como requisito parcial para obtenção do Título de Mestre em Ciências da Saúde pelo Programa de Pós-Graduação em Ciências da Saúde da Universidade de Brasília.

Orientador: Prof^ª. Dr^ª. Silviene Fabiana de Oliveira

BRASÍLIA-2010

ALUISIO TRINDADE FILHO

**CARACTERIZAÇÃO GENÉTICA DA POPULAÇÃO DO DISTRITO FEDERAL
COM BASE EM MARCADORES STR DO CROMOSSOMO X**

Dissertação apresentada como requisito parcial para
obtenção do Título de Mestre em Ciências da Saúde
pelo Programa de Pós-Graduação em Ciências da
Saúde da Universidade de Brasília.

Aprovado em 13 de outubro de 2010

BANCA EXAMINADORA

Prof^a. Dr^a. Silviene Fabiana de Oliveira (Presidente)
Universidade de Brasília

Prof. Dr. Aguinaldo Luiz Simões
Universidade de São Paulo – Ribeirão Preto

Prof^a. Dr^a. Maria de Nazaré Klautau Guimarães
Universidade de Brasília

*Dedico este trabalho,
Aos meus queridos pais Aluisio e Terezinha (in memoriam) pelo amor e
dedicação incondicionais; a minha esposa Regina, a luz que me ilumina
e aponta meu caminho e aos meus filhos Bernardo e Helena, meus
amigos e meus ídolos.*

AGRADECIMENTOS

Ao Prof. Dr. Celso Teixeira Mendes Junior, do Departamento de Química da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto - Universidade de São Paulo, pela gentileza e generosidade com que me acolheu em seu Serviço e pelo auxílio essencial para a realização das análises estatísticas referentes aos diversos parâmetros populacionais e forenses.

Aos servidores do Instituto de Pesquisa de DNA Forense que contribuíram para que este projeto pudesse ter êxito, em especial a Adriano Augusto Gouveia de Macedo, Ana Flávia Moreira, Célia Maria Marques de Santana, Evanilda Francisca de Oliveira, Givânia Caldas Oliveira Lopes, José de Souza Santos, Sandra Martins Dias e Sandra Yandecy Lucena Veiga que, com abnegação e desprendimento, me ajudaram na genotipagem das amostras biológicas que compõem minha amostra populacional.

A minha esposa Regina e aos meus filhos Bernardo e Helena, pela confiança em meu trabalho e pela riqueza de nossas interlocuções que constituem um estímulo constante para que eu busque o meu melhor.

À Prof^a Dra. Silviene Fabiana de Oliveira, minha orientadora, pela importante contribuição a minha formação e pela acolhida institucional deste projeto.

RESUMO

O cromossomo X apresenta um modo peculiar de herança. Na transmissão paterna, comporta-se como uniparental enquanto que na transmissão materna, como biparental. Esta excentricidade torna os marcadores genéticos situados nesse cromossomo particularmente úteis para determinadas análises no âmbito da genética forense. Adicionalmente, lhe empresta características que podem ser exploradas em estudos de genética de populações para verificar que semelhanças e diferenças existem entre a história contada por eles e a contada pelos demais marcadores. A incorporação dos marcadores do cromossomo X na genética forense tem uma história recente e vem ainda sendo consolidada, razão pela qual se verifica ainda uma relativa escassez de dados sobre estes marcadores nas populações, seja brasileira ou de outros países. Neste projeto, 205 homens e 210 mulheres nascidos no Distrito Federal foram analisados com 10 marcadores STRs localizados no cromossomo X humano (DXS8378, DXS9898, DXS7133, GATA31E08, GATA172D05, DXS7423, DXS6809, DXS7132, DXS9902 e DXS6789), compondo o sistema de amplificação multiplex denominado “X-Decaplex”, desenvolvido pelo Grupo de Espanhol-Português da *International Society for Forensic Genetics* (ISFG). Todos os indivíduos foram considerados não relacionados geneticamente, de acordo com o modo de herança do cromossomo X. As frequências alélicas, parâmetros populacionais, parâmetros forenses e desequilíbrio de ligação foram estimados, assim como a constituição genética da população considerando os três grupos parentais principais – europeus, africanos e ameríndios. Comparação com outras populações brasileiras analisadas para os marcadores em questão foram também realizadas. O STR mais informativo foi o DXS6809 e o menos informativo, o DXS8378. O Poder Médio de Discriminação do sistema como um todo para mulheres (PMD_F) foi de 0,99999999996 e para homens (PMD_M), 0,9999995. A Probabilidade Média de Exclusão Cumulativa para o Trio mãe filha e suposto pai ($PMEC_T$) e para o Duo filha e suposto pai ($PMEC_D$), considerando os *loci* agrupados, foi de 0,99999795 e 0,99988927, respectivamente. Não foram observados desvios significativos no equilíbrio da Hardy-Weinberg testado na amostra feminina, bem como não se observou Desequilíbrio de Ligação na

amostra masculina. Estes resultados habilitam o uso do Banco de Dados das frequências alélicas dos dez STRs do cromossomo X obtidas para a população do Distrito Federal, como base de dados para cálculos em análises forenses em casos envolvendo indivíduos desta população. Análises comparativas entre a população do Distrito Federal e as populações dos estados do Mato Grosso do Sul, Paraná, Rio de Janeiro e São Paulo sugerem uma maior diversidade genética para a população do Distrito Federal, provavelmente refletindo a constituição de sua população, resultado do afluxo de contingentes populacionais de todo o Brasil. Análises de distância genética com base em valores de G_{ST} mostraram diferenças significativas entre o Distrito Federal e os estados do Mato Grosso do Sul, Paraná e Rio de Janeiro. As proporções de mistura étnica da população do Distrito Federal, considerando como parentais as populações européia, africana e ameríndia, demonstraram um predomínio da européia (61,9%), seguida da africana (27%). Estes dados não diferem significativamente dos resultados observados com marcadores autossômicos na mesma população, sugerindo que ambas as classes de marcadores contam a mesma história genética, em que pese o fato dos marcadores do cromossomo X refletirem mais a herança materna que a paterna e que, para o Brasil, estudos com marcadores uniparentais terem demonstrado consistentemente uma influência maior para a parental européia na linhagem paterna, com as africana e ameríndia predominando para a parental materna. A inclusão dos dez CrX STRs em quatro casos de paternidade deficiente analisados previamente com uma bateria de 21 ou 22 STRs autossômicos demonstrou um grande impacto na valoração estatística quando a reconstituição do genótipo do suposto pai foi realizada mediante análise de uma filha biológica mas um impacto mínimo quando a reconstituição foi realizada com a suposta avó.

ABSTRACT

The X chromosome has a peculiar mode of inheritance. In paternal transmission, its behavior is uniparental, while in maternal transmission, it is biparental. Because of this eccentricity, the markers located in this chromosome are gradually being introduced into forensic genetics, particularly in deficiency paternity cases, where the alleged father is absent. This unique form of inheritance can be explored by population genetics in order to verify the extent of similarities and differences between the stories told by the X-chromosome markers and those told by other genetic markers. The inclusion of X chromosome markers in forensic genetics has a recent history and is still being consolidated, which is why there is still a relative paucity of data on these markers in populations, both in Brazil and in other countries. In this project, 205 men and 210 women (625 alleles) born in the Federal District were analyzed with 10 STR markers located on the human X chromosome (DXS8378, DXS9898, DXS7133, GATA31E08, GATA172D05, DXS7423, DXS6809, DXS7132, DXS9902 and DXS6789) composing the multiplex amplification system "X-Decaplex", developed by the Spanish-Portuguese Group of the International Society for Forensic Genetics (ISFG). All individuals were considered unrelated according to the X chromosome mode of inheritance. Allele frequencies, forensic parameters, population parameters and linkage disequilibrium were estimated as was the genetic make-up of the population considering the three main parental groups - Europeans, Africans and Amerindians. Comparison with other Brazilian populations analyzed for the markers in question were also performed. DXS6809 was the most informative STR and DXS8378, the least. The Mean Power of Discrimination of the system as a whole for women (PDC_F) was 0.99999999996 and for men (PDC_M) 0.9999995. The Mean Probability Cumulative of Exclusion for the trio mother-daughter-alleged father ($PMEC_T$) was 0.99999795 and that for the duo daughter-alleged father ($PMEC_D$) was 0.99988927, considering the loci grouped. There were no significant deviations from the Hardy-Weinberg expectations in the female sample tested and there was no observed Linkage Disequilibrium observed in the male sample. These results enable the allelic

frequencies of the ten STRs of X chromosome obtained for Federal District population to be used as a database for forensic analysis calculations in cases involving individuals from this population. Comparative analysis between the population of the Federal District and those of the states of Mato Grosso do Sul, Parana, Rio de Janeiro and São Paulo suggests a greater genetic diversity in the population of the Federal District, probably reflecting the make-up of its population, a result of the influx of population groups from the whole country. Analyses of genetic distance based on G_{ST} values showed significant differences between the population of the Federal District and those of the states of Mato Grosso do Sul, Parana and Rio de Janeiro. The proportions of ethnic mix in the population of the Federal District, considering parental populations as European, African and Amerindian, showed a predominance of European (61.9%), followed by African (27%). These data do not differ significantly from the results observed with autosomal markers in the same population, suggesting that both classes of markers have the same genetic history, despite the fact that the X chromosome markers reflect more maternal inheritance than paternal and that for Brazil, studies using uniparental markers have consistently shown a greater influence of Europe for paternal parental lineage, with African and American predominance for the maternal parent. The inclusion of the ten CRX STRs in four deficient paternity cases previously analyzed with a battery of 21 or 22 autosomal STRs showed a statistically significant impact on valuation when the reconstitution of the alleged father's genotype was performed by analysis of a biological child but a minimal impact when the supposed grandmother replaced the supposed father.

LISTA DE FIGURAS

Figura 1 - Ideograma apresentando a localização dos STRs do X-Decaplex ao longo do cromossomo X.....	48
Figura 2 - Eletroferograma de uma amostra de origem feminina	49
Figura 3 – Eletroferograma de uma amostra de origem masculina.....	49
Figura 4 – Dendrograma do tipo UPGMA representativo do relacionamento genético das cinco populações brasileiras e das três parentais avaliado com base nos dez CrX STRs, com os respectivos valores de <i>bootstrap</i>	80

LISTA DE TABELAS

Tabela 1 - Estimativa da composição da população brasileira em três épocas distintas com base no critério cor-da-pele.....	39
Tabela 2 - Composição da população brasileira e do Distrito Federal no ano 2000 segundo auto-denominação de cor da pele.	40
Tabela 3 – Localização e estrutura dos dez CrX STRs que compõem o multiplex X-Decaplex, sequência dos iniciadores e respectivos fluoróforos.....	45
Tabela 4- Parâmetros de ciclagem para amplificação dos X-STR via PCR.....	46
Tabela 5 - Distribuição das frequências alélicas para os 10 CrX STRs na amostra masculina da população do Distrito Federal	63
Tabela 6 - Distribuição das frequências alélicas para os 10 CrX STRs na amostra feminina da população do Distrito Federal	64
Tabela 7 - Distribuição das frequências alélicas combinadas de homens e mulheres para os 10 CrX STRs na população do Distrito Federal	65
Tabela 8 – Número de alelos por <i>locus</i> observados nas cinco populações e estimados após o processo de rarefação.....	66
Tabela 9 – Alelos privados, frequência e população onde foram observados ..	67
Tabela 10 – Número de alelos privados observados por população e estimados após o processo de rarefação	68
Tabela 11 - Parâmetros estatísticos populacionais e forenses na população do Distrito Federal analisada com os 10 CrX STRs	70
Tabela 12 - Média e desvio padrão da Heterozigose observada, Poder de Discriminação e Probabilidade Média de Exclusão obtidas com AS-STR e CrX STR.	71
Tabela 13- Teste exato para análise de desequilíbrio de ligação entre os marcadores CrX STR analisados aos pares.....	74
Tabela 14 - Teste exato para análise de desequilíbrio de ligação entre os marcadores AS STR analisados aos pares.....	75

Tabela 15 - Teste exato para análise de desequilíbrio de ligação entre os marcadores Y STR analisados aos pares	76
Tabela 16 - Estimativa de contribuição genética ancestral na composição atual de cinco populações brasileiras com base na análise dos CrX STRs.	77
Tabela 17 – Estimativa de mistura ancestral para a população do Distrito Federal: comparação dos resultados obtidos com base em análises por CrX STRs e por AS STRs.....	79
Tabela 18 – Matriz de distâncias genéticas com base nos dez CrX STRs entre os pares de populações medidas pelos valores de G_{ST}	81
Tabela 19 – Índices de Paternidade (IP) e Probabilidades de Paternidade (PP) obtidos com STRs autossômicos, STRs do cromossomo X e para o conjunto das duas classes de marcadores.	82

LISTA DE ABREVIATURAS E SÍMBOLOS

AS STR	<i>Short Tandem Repeat</i> Autossômico
CrX STR	<i>Short Tandem Repeat</i> do cromossomo X
cM	centi-Morgan
CODIS	<i>Combined DNA Index System</i>
DF	Distrito Federal
DNA	Ácido Desoxirribonucléico
dNTP	Desoxinucleotídeo trifosfato
EDTA	Ácido Etilenodiaminotetraacético
g	Gramma
HCl	Ácido Clorídrico
He	Heterozigose esperada
Ho	Heterozigose observada
HW	Equilíbrio de Hardy-Weinberg
IBD	Idêntico por Descendência
IBGE	Instituto Brasileiro de Geografia e Estatística
IP	Índice de Paternidade
IPDNA	Instituto de Pesquisa de DNA Forense
ISFG	<i>International Society for Forensic Genetics</i>
kb	Kilobase
DL	Desequilíbrio de ligação
mb	Megabase
MgCl ²	Cloreto de Magnésio
mL	Mililitro
mM	Milimol
MS	Mato Grosso do Sul
NaOH	Hidróxido de Sódio
pb	Pares de Base
PCR	<i>Polymerase Chain Reaction</i>
PD _F	Poder de Discriminação feminino
PD _M	Poder de Discriminação masculino
PDC _F	Poder Médio de Discriminação Feminino

PDC _M	Poder Médio de Discriminação Masculino
PME _D	Probabilidade Média de Exclusão em Duo
PME _T	Probabilidade Média de Exclusão em Trio
PMEC _D	Probabilidade Cumulativa Média de Exclusão para Duo
PMEC _T	Probabilidade Cumulativa Média de Exclusão para Trio
PP	Probabilidade de Paternidade
PR	Paraná
RFLP	<i>Restriction Fragment Length Polymorphism</i>
RJ	Rio de Janeiro
SNP	<i>Single Nucleotide Polymorphism</i>
SP	São Paulo
STR	<i>Short Tandem Repeat</i>
TRIS	<i>tris (hydroxymethyl) aminomethane</i>
VNTR	<i>Variable Number of Tandem Repeat</i>
Y STR	<i>Short Tandem Repeat</i> do cromossomo Y
°C	Graus Celcius
μL	Microlitro
μM	Micromol
θ	Teta

SUMÁRIO

1 INTRODUÇÃO	16
2 REVISÃO BIBLIOGRÁFICA	18
2.1 IDENTIFICAÇÃO HUMANA.....	18
2.1.1 Identificação humana em bases populacionais	18
2.1.2 Identificação humana em bases individuais.....	19
2.2 A GENÉTICA FORENSE E OS MARCADORES GENÉTICOS DO TIPO STRs – EVOLUÇÃO HISTÓRICA.....	21
2.3 O CROMOSSOMO X.....	26
2.3.1 Uso Forense dos marcadores STR do Cromossomo X	28
2.3.2 Os CrX STRs nos exames de paternidade cíveis	33
2.3.3 Os CrX STR em outros exames de vínculo genético	35
2.3.4 Os CrX STRs em análise de vestígios criminais.....	36
2.3.5 Aberrações cromossômicas e marcadores do cromossomo X	37
2.4 BANCO DE DADOS DE FREQUÊNCIA ALÉLICA	37
2.5. ORIGEM DAS POPULAÇÕES BRASILEIRA E DO DISTRITO FEDERAL .	38
3 OBJETIVOS	41
4. MATERIAL E MÉTODOS	42
4.1 CARACTERIZAÇÃO DA AMOSTRA	42
4.1.1 Amostra para análise dos marcadores do cromossomo X	42
4.1.2 Amostra para análise dos marcadores dos cromossomos autossômicos	42
4.1.3 Amostra para análise dos marcadores do cromossomo Y	43
4.2 COLETA DAS AMOSTRAS	43
4.3 PROCEDIMENTOS LABORATORIAIS	43
4.3.1 Extração do DNA	43
4.3.2 Marcadores genéticos	44
4.3.3 Reação de PCR.....	46
4.3.4 Detecção dos produtos de Amplificação e Assinalamento alélico.....	47
4.4. OUTRAS AMOSTRAS POPULACIONAIS.....	50
4.5 ANÁLISES ESTATÍSTICAS.....	50
4.5.1 Frequências Alélicas e Heterozigose.....	51

4.5.2 Aderência ao equilíbrio de Hardy-Weinberg.....	52
4.5.3 Poder de Discriminação.....	52
4.5.4. – Probabilidade Média de Exclusão.....	53
4.5.5 Comparação dos parâmetros Heterozigose, Poder de Discriminação e Probabilidade de Exclusão entre CrX STRs e AS STRs.....	53
4.5.6 Estatísticas F e Diversidade Inter-populacional.....	54
4.5.7 Diversidade Alélica.....	55
4.5.8 Equilíbrio de Ligação Genotípica.....	55
4.5.9 Estimativas de contribuição parental na composição atual da população.....	57
4.5.10 Distâncias Genéticas e Dendrograma.....	58
4.5.11 Uso dos CrX STRs na resolução de casos reais de paternidade deficientes	60
5 RESULTADOS E DISCUSSÃO.....	62
5.1 DISTRIBUIÇÃO DAS FREQUÊNCIAS ALÉLICAS.....	62
5.2 PARÂMETROS POPULACIONAIS E FORENSES.....	68
5.3 ANÁLISE DE MISTURA GENÉTICA E DIVERSIDADE INTER- POPULACIONAL.....	77
5.4 IMPACTO DA INTRODUÇÃO DE DADOS DO X- DECAPLEX EM QUATRO CASOS DE PATERNIDADE DEFICIENTES.....	81
7 REFERÊNCIAS BIBLIOGRÁFICAS.....	86

1 INTRODUÇÃO

A Genética Forense, ramo da genética humana que visa identificar a origem de material biológico, individualizar pessoas ou estabelecer vínculo genético entre elas, experimentou uma radical transformação a partir de meados da década de 1980, face à introdução de metodologias de exploração direta do DNA, o que acabou por substituir os denominados marcadores clássicos (proteínas, isoenzimas e grupos sanguíneos). Apesar destes marcadores terem sido utilizados desde o início do século passado, eles apresentam um baixo grau de variabilidade, oferecendo desta forma um reduzido poder discriminatório para identificação humana (1,2). Por outro lado, alguns desses marcadores são tecido-específicos, além de usualmente serem encontrados em quantidade e qualidade insuficientes em vestígios criminais para que se possa obter algum resultado.

A exploração direta de regiões polimórficas do DNA teve início com as análises por RFLP - *Restriction Fragment Length Polymorphism* ou “Polimorfismo de Comprimento de Fragmentos de Restrição” (3). Quase simultaneamente foi criada nos Estados Unidos a revolucionária metodologia de amplificação seletiva *in vitro* de sequências definidas do DNA, denominada PCR (*Polymerase Chain Reaction* ou “Reação em Cadeia de Polimerase” (4). Na genética forense, as técnicas mais utilizadas atualmente exploram o polimorfismo caracterizado por sequências de bases nitrogenadas que se repetem *in tandem* nas regiões satélite do DNA, regiões estas não codificantes, dispersas aleatoriamente no genoma humano. Inicialmente, longas sequências com motivos variando da ordem de algumas dezenas de pares de bases, denominadas minissatélites, eram o alvo das análises. Posteriormente, sequências menores, contendo de dois a dez pares de bases, denominadas microssatélites ou *Short Tandem Repeat* (STR), passaram a ser utilizadas com esta finalidade, por se adequarem melhor à amplificação pelo método da PCR (5). Pelo seu potencial de gerar perfis genéticos com elevados índices de discriminação e pela grande reprodutibilidade dos resultados, vultosos investimentos em pesquisa foram e vêm sendo feitos ao redor do mundo na técnica da PCR, no aprimoramento dos procedimentos laboratoriais, na criação de insumos e equipamentos e no aprofundamento das análises estatísticas.

O Brasil logo se inseriu neste contexto, inicialmente com laboratórios privados voltados para análises de paternidade cível e, em seguida, com as instituições públicas, focadas nas análises de material biológico oriundo de investigações criminais. A justificativa para tal iniciativa deve-se à relevância social da prova pericial judicial, a qual se constitui em um dos pilares da investigação criminal atual (6)

Considerando os marcadores nucleares, desde o início os STRs localizados nos cromossomos autossômicos têm preponderado na genética forense, sendo utilizados nas análises de vestígios de cenário de crime, identificação de cadáveres e testes de paternidade, dentre outros. Posteriormente foram introduzidos marcadores situados no cromossomo Y, usados com frequência em casos deficientes de paternidade com o filho do sexo masculino e em crimes sexuais com mistura de material masculino/feminino (7), ou mesmo com espermatozóide não detectado em esfregaços vaginais (8).

Atualmente muitas pesquisas têm se voltado para marcadores STRs situados no cromossomo X (CrX STR). Embora seu uso ainda esteja pouco difundido na prática forense, eles têm demonstrado ser de grande valia em situações específicas. Neste contexto, pode-se citar seu emprego na ampliação das análises nos casos complexos de investigação de paternidade cíveis, no qual o suposto pai não pode ser testado, principalmente por já ter falecido, com o potencial de complementar as análises realizadas com marcadores autossômicos, marcadores do cromossomo Y e mitocondrial (9,10).

Este trabalho procurou caracterizar a população do Distrito Federal para 10 marcadores STRs do cromossomo X (DXS8378, DXS9898, DXS7133, GATA31E08, GATA172D05, DXS7423, DXS6809, DXS7132, DXS9902, e DXS6789), para os quais poucas populações foram já estudadas (11), mediante a análise de parâmetros populacionais e forenses. Adicionalmente foram feitos estudos comparativos com dados relativos a populações de quatro outros estados brasileiros (Mato Grosso do Sul, Paraná, Rio de Janeiro e São Paulo), bem como análise de mistura genética considerando que uma população brasileira qualquer seja tri-híbrida, composta das parentais européia, africana e ameríndia.

2 REVISÃO BIBLIOGRÁFICA

2.1 IDENTIFICAÇÃO HUMANA

A identificação humana pode ser vista sob dois aspectos: identificação de grupos populacionais e identificação de indivíduos. Ambas as abordagens, quando vistas em uma perspectiva histórica, remetem a um passado remoto.

2.1.1 Identificação humana em bases populacionais

O estudo das “raças” humanas data da antiguidade. A existência de diferenças entre grupos populacionais de origens geográficas distintas deve ter sido uma constatação dos primeiros viajantes de longas distâncias. O grego Heródoto (século V a.C.) descreveu o nome, localização geográfica, hábitos e aparência física de muitos povos, a maioria mediterrânea (12), razão pela qual é considerado não só o pai da história como também da antropologia (13). Para o naturalista romano Plínio, o velho (século I a.C.), o sol, por estar mais próximo da África, queimava seus habitantes, deixando-os com a pele enegrecida e os cabelos escuros e encaracolados, enquanto que os povos do norte, vivendo bem mais distantes do sol, tinham pele clara e cabelos lisos e amarelos (14). J. F. Blumenbach, considerado o pai da antropologia física, na sua tese de doutorado em medicina considerou a espécie humana como única, com cinco variantes, de acordo com a cor da pele: Caucásiana, Mongol, Etíope (incluindo os demais africanos), Americana e Malaia, incluindo nesta última os povos das ilhas do sudeste asiático e da Oceania conhecida à época (12). Com o passar do tempo, a tentativa de classificação da espécie humana passou a ser feita com o emprego de métodos com base em princípios científicos. Nos dias atuais, muita pesquisa está sendo produzida com fins de identificar a origem geográfica dos povos, a proximidade entre eles, sob vários pontos de vista, bem como o grau de mistura genética na composição de uma

determinada população. A resposta a estas questões baseia-se em estudos de arqueologia, de linguística e da genética.

A contribuição da genética nesta área é dada principalmente por meio dos estudos de marcadores genéticos de diversas categorias, como mutações pontuais e inserção/deleção. Os marcadores genéticos do tipo STR também se prestam bem a este propósito, em especial os resultantes de mutações localizadas em regiões não codificantes, sem que sobre eles sejam exercidas pressões seletivas relevantes (15).

2.1.2 Identificação humana em bases individuais

A bíblia nos conta que certa vez, duas mulheres procuraram o rei Salomão de Israel para que resolvesse uma disputa de maternidade de uma criança recém-nascida. Cada uma se dizia a mãe da criança. O rei, dispondo à época apenas do seu discernimento, propôs dividir a criança ao meio com uma espada e entregar a cada mulher uma metade. Uma delas aceitou. A outra, no entanto, caiu aos pés do rei aos prantos e implorou para que desse a criança a qualquer pessoa, mas que a deixasse viver. O rei decidiu então que a segunda era a verdadeira mãe. Este deve ser o mais antigo caso de que se tem relato de uma decisão judicial envolvendo uma disputa de filiação (16) e, portanto, de identificação humana individual. A identificação humana em bases individuais pertence ao campo da medicina legal e da criminalística. Nela se inserem as disputas de filiação cíveis, bem como a vasta gama de casos criminais, envolvendo identificação de criminoso ou vítima, de pessoa desaparecida, confronto entre vestígios de um ou de vários delitos, todos com a finalidade de fornecer provas técnicas robustas para o auxílio da justiça.

Os métodos utilizados na tentativa de identificação humana individual variaram muito ao longo do tempo. Segundo os registros, os métodos mais antigos careciam de qualquer base científica, servindo ao duplo propósito de identificação e aplicação de punição àquele que se queria identificar (2). Era o caso do uso sistemático do ferrete e da mutilação, prática pelas quais se costumava marcar indivíduos infratores com ferro em brasa ou promover amputação de orelhas, dedos, mãos, língua (17). Como se pode depreender, esta prática era restrita a indivíduos condenados por alguma infração penal.

A necessidade de identificação de grandes contingentes populacionais levou ao desenvolvimento de métodos que se baseavam no registro de algumas características morfológicas, ainda que com um razoável grau de imprecisão, destacando-se o assinalamento sucinto de algumas destas características, a fotografia simples e o retrato falado, todos eles ainda de uso corrente nos dias atuais. Com o tempo evoluíram para a adoção de métodos mais elaborados e já com base científica, que caíram em desuso principalmente pela sua pouca praticidade. O primeiro e mais notório desses sistemas, o antropométrico de Bertillon, foi introduzido em 1882 em Paris e utilizava uma série de características consideradas fixas desde a idade adulta até a morte (17).

No final do século XIX, o sistema de classificação a partir da constituição anatômica dos desenhos papilares nas faces palmares das extremidades das falanges, conhecido como impressões digitais, começou a ser utilizado e hoje é um dos pilares da identificação tanto cível quanto criminal. Os registros existentes do uso oficial desse método para identificação pessoal datam de 1901, pelo serviço de identificação civil da cidade de Nova York/Estados Unidos (1). No Brasil, o método foi oficialmente implantado em 1903. A aceitação das impressões digitais para identificação humana é feita pela presunção de que a probabilidade da existência de dois indivíduos com impressões digitais idênticas é infinitamente pequena (17).

Embora o estudo de impressões digitais apresente-se como uma ferramenta poderosa para identificação humana, capaz de individualizar pessoas e levar à autoria de crimes, esse método apresenta limitações na seara criminal. Encontrar fragmentos em locais de crime em quantidade e qualidade mínimas para identificação nem sempre é possível. Nesses locais, os vestígios biológicos (sangue, saliva, espermatozoides, pêlos, entre outros) são encontrados com mais frequência e a identificação de sua origem, se da vítima ou do suspeito, pode ser crucial no curso de uma investigação, conduzindo eventualmente à identificação do autor de um delito.

Nesse enfoque, foram descritas as provas genéticas empregadas para fins de identificação humana, as quais deram um grande impulso ao campo da investigação de paternidade. Didaticamente podem-se dividi-las em provas genéticas pré-mendelianas, baseadas na semelhança fisionômica entre indivíduos aparentados, principalmente entre pais e filhos, e as provas genéticas mendelianas. Estas últimas são classificadas em não-sanguíneas, com base em características tais como

morfologia do lóbulo da orelha, redemoinhos do cabelo, braquidactilia, cor dos olhos, cor da pele e sensibilidade gustativa à feniltiouréia; e sanguíneas, baseadas estas em marcadores genéticos (18).

A história da utilização dos marcadores genéticos aplicados às ciências forenses iniciou-se com a utilização do sistema de grupos sanguíneos ABO e do fator Rh a partir da análise de mancha de sangue. Nos anos que se seguiram, vários outros sistemas sanguíneos protéicos foram sendo descobertos e logo incorporados à área da identificação humana. Tais sistemas, cuja descoberta teve início há pouco mais de 100 anos, dominaram o cenário por quase todo o século XX. Sua substituição começou a ocorrer em meados da década de 80 daquele século, com a incorporação das técnicas de genotipagem do material genético diretamente. Além disso, como em todas as células nucleadas existe DNA genômico, tornou-se possível a análise de um amplo espectro de amostras biológicas.

2.2 A GENÉTICA FORENSE E OS MARCADORES GENÉTICOS DO TIPO STRs – EVOLUÇÃO HISTÓRICA

Nos dias atuais, e já por quase duas décadas, os STRs (*short tandem repeat*) vêm sendo largamente usado no mapeamento genético, bem como se constituindo nos principais marcadores biológicos com fins de identificação humana, comportando-se como o esteio da genética forense atual. Esta, com base em princípios científicos sólidos, teve início no alvorecer do século 20, quando Karl Landsteiner descreveu o sistema sanguíneo ABO e observou que indivíduos podiam ser agrupados de acordo com seu tipo sanguíneo (19). Descrito como um sistema polimórfico com herança mendeliana, apresenta o alelo “I” recessivo em relação aos alelos “A” e “B”. estes codominantes entre si. Em 1915 foi descrito pela primeira vez por Leone Lattes o uso do sistema ABO na resolução de casos de paternidade (20). Sistemas sanguíneos eritrocitários identificados posteriormente, como ABH, Rh, MNSs, Duffy, Kell, Kidd, GmKm aumentaram o número de marcadores imunológicos utilizados na tentativa de individualização de manchas de sangue e outros vestígios biológicos. Além destes, diversos marcadores enzimáticos foram identificados e incorporados na prática forense, como a adenilato deaminase (ADA), anidrase

carbônica II (CA-II), fosfatase ácida eritrocitária (EAP), haptoglobina, transferrina, glioxalase I (GLO), esterase D (EsD), fosfoglucomutase (PGM), 6 fosfogluconato desidrogenase, componente específico de grupo (GC), dentre outros. Nos anos 50 foi descrito o sistema de histocompatibilidade humano, o HLA (*human leucocyte antigen*), composto por vários *loci*, quatro dos quais usados para análise de variabilidade humana (HLA-A, HLA-B, HLA-C e HLA-D).

A probabilidade de uma mancha de sangue se originar de um indivíduo tendo uma combinação particular de genótipos é determinada pelo produto da distribuição na população dos alelos dos marcadores genéticos empregados, desde que se segreguem independentemente. À medida que o número de marcadores utilizados aumenta, o compartilhamento por dois indivíduos de um conjunto de genótipos resultante do acaso diminui. Usados em conjunto, os sistemas descritos eram capazes de produzir perfis individuais com razoável poder de discriminação (20). Estes marcadores, no entanto, apresentavam algumas restrições ao uso em larga escala. Em razão do baixo nível de variabilidade apresentado, oferecendo reduzido poder discriminatório para identificação humana, exceção feita ao sistema HLA, a utilização destes marcadores era decisiva nos casos em que o suposto pai podia ser excluído (1,2), porém menos importante quando não se caracterizava uma exclusão, em razão dos índices estatísticos nem sempre elevados que eram obtidos com os cálculos estatísticos. Desvantagens adicionais eram o fato de necessitarem de uma quantidade substancial de material biológico para a sua caracterização (20), o que apenas eventualmente ocorre em cenas de crime, bem como não estarem presentes em todos os tecidos e líquidos orgânicos. Estas peculiaridades limitavam bastante o seu uso na maioria dos casos forenses, permitindo o seu emprego em larga escala apenas nas disputas de paternidade cível.

Com a descrição de regiões do genoma humano altamente variáveis e o desenvolvimento de tecnologia para seu acesso, as restrições citadas foram minimizadas e uma revolução foi produzida na análise genética forense humana (21). São regiões compostas de sequências repetitivas de DNA (22), encontrando-se agrupadas ou dispersas ao longo do genoma, formando sítios altamente variáveis.

A utilização da genotipagem de regiões genômicas via acesso direto ao material genético teve início em meados da década de 1980 e rapidamente substituiu os marcadores protéicos como ferramenta da genética na identificação humana. O pioneirismo coube a Sir Alec Jeffreys, da Universidade de Leicester,

Reino Unido, ao resolver um caso de imigração, determinando vínculo genético entre mãe e filho (3) e ao identificar Colin Pitchfork como o indivíduo que violentou e matou duas jovens de 15 anos (15).

O marcador genético aqui descrito localiza-se geralmente em regiões genômicas não codificadoras e comumente são evolutivamente neutros, isto é, não sujeitos a pressões seletivas. Desta forma, mutações nessa região se acumulam ao longo do tempo. Estas regiões se caracterizam por sequências ou blocos de DNA que se repetem *in tandem*, denominadas motivo. De acordo com a quantidade de pares de bases de cada unidade de repetição, a sequência é denominada minissatélite (VNTR – *variable number of tandem repeat*), ou microssatélite (STR – *short tandem repeat*).

A variabilidade observada nessas regiões resulta da existência de diferentes alelos cuja variação diz respeito à diferença no número de repetições, originadas por recombinação desigual ou pelo deslizamento da fita de DNA durante a replicação, levando ao acréscimo ou ao decréscimo de sequências repetidas na fita filha (23,24). Ocorrendo em células germinativas, estas alterações resultam em mutações transmitidas à prole, com uma frequência superior à que ocorre em regiões genômicas de sequência simples (25).

A metodologia empregada de início para a exploração dessa variabilidade denomina-se RFLP (*restriction fragment length polymorphism*) e consiste da digestão de DNA obtido da amostra a ser analisada com endonucleases de restrição que cortam o DNA em fragmentos de diversos tamanhos, variando de 1 a 20 kb. Os fragmentos são então submetidos à eletroforese e separados por tamanho, sendo visualizados por sondas *multiloci* que hibridizam com vários minissatélites ao mesmo tempo. A análise dos resultados é complexa, não se sabendo ao certo quais *loci* estão sendo analisados, bem como o grau de independência entre os mesmos. O desenvolvimento de sondas unilocais contornou este problema, tornando possível a análise considerando a segregação mendeliana (26).

Durante alguns anos, essa foi a metodologia dominante, inclusive no Brasil, em testes de paternidade (27). Sendo uma técnica lenta e dependente de uma quantidade razoável de DNA genômico, encontra uma grande limitação na análise de amostras forenses, nas quais frequentemente a quantidade de material biológico disponível é muito pequena. Por estas razões, quase não se encontra mais em uso,

tendo sido rapidamente suplantada pela análise dos STRs pela metodologia da PCR.

Os STRs, ou microssatélites, são caracterizados por sequências repetitivas de dois a dez pares de bases, já tendo sido descritos alguns milhares no genoma humano (28). Representando cerca de 3% do genoma humano, encontram-se dispersos por todo o material genético, ocorrendo em média um a cada 10.000 pares de bases (29). Os STRs apresentam larga aplicabilidade no mapeamento genético, sendo muito utilizados na identificação de genes envolvidos em doenças. Na genética forense, começaram a ser usados no início dos anos 1990 (30,31), suplantando rapidamente a utilização de outros marcadores. Atualmente, a vasta maioria dos casos forenses envolve análises de STRs.

O uso em larga escala dos STRs como marcadores genéticos coincidiu com a difusão da tecnologia da PCR. Idealizada em meados dos anos 1980 pelo químico norte-americano Karen Mullis (32), a PCR permite a amplificação exponencial de pequenas quantidades de DNA, por meio de variação na temperatura a que o DNA genômico é submetido, constituindo ciclos que se repetem por um número médio de 30 vezes. Ao final, as sequências do DNA previamente escolhidas são multiplicadas por um fator de até um bilhão. Enquanto os métodos previamente discutidos requeriam cerca de 500 ng de DNA íntegro para a obtenção de resultados, na reação de PCR, quantidades tão pequenas quanto 0,5 a 1,0 ng são tipicamente utilizadas, podendo ser obtido produto de amplificação com 250 pg ou menos, ou mesmo com material de uma única célula (33).

A rápida difusão do emprego dos STRs decorreu principalmente das seguintes razões:

- Tamanhos dos fragmentos gerados pela PCR tipicamente entre 100 e 400 pb, comparados com alguns kb para os VNTRs, tornando-os de fácil amplificação pela tecnologia da PCR;
 - Baixos níveis de artefatos no produto de amplificação, principalmente quando os STRs contem quatro ou cinco pares de bases na unidade de repetição;
 - Grande quantidade de alelos por *locus*, gerando alto poder de discriminação;
 - Fácil separação dos alelos quando em heterozigose, mesmo com uma unidade de repetição ou menos de diferença;
-

-
- Passível de amplificação mesmo em material biológico degradado;
 - Passível de ser usado em sistemas multiplexes, isto é, a genotipagem simultânea de diversos *loci*

No início foram usados basicamente os STRs localizados nos cromossomos autossômicos. Para tanto, a escolha do painel de STRs utilizados mais rotineiramente foi feita de forma a que os marcadores estivessem preferivelmente localizados em cromossomos diferentes ou, se no mesmo cromossomo, a uma distância entre os marcadores que garantisse a segregação independente, ou seja, 50 centi-Morgan (cM) ou mais. Em 1992 foi descrito o primeiro STR situado fora dos cromossomos autossômicos. Nessa ocasião foi descrito o marcador DYS19 situado no cromossomo Y (34). O valor da descrição de marcadores situados no cromossomo Y na identificação humana vem do fato de estarem presentes apenas em homens, sendo transmitidos diretamente de pai para filho, de forma que todos os indivíduos do sexo masculino que compartilham a mesma linhagem patrilínea apresentam o mesmo haplótipo, salvo os casos de ocorrência de mutação. Também analisados em sistemas multiplexes, os STRs do cromossomo Y vêm sendo usados como ferramenta auxiliar aos autossômicos, em casos selecionados, principalmente de paternidade deficientes ou de violência sexual contra mulher. Uma restrição inerente aos STRs do cromossomo Y é que apenas linhagens, e não indivíduos, podem ser identificadas, já que os homens que compartilham a mesma linha paterna apresentam o mesmo perfil genético.

O cromossomo X foi o último a ter seu polimorfismo explorado tanto na genética forense, quanto em estudos de genética de populações, a despeito de conter mais informação genética que o cromossomo Y (37). Embora o início da caracterização de STRs do cromossomo X seja contemporâneo ao dos STRs autossômicos (AS STRs) e os do cromossomo Y (Y STRs) (35), o seu uso efetivo só começou a ser difundido há pouco mais de 10 anos (36). Devido ao seu curioso modo de herança, às vezes uniparental e haplotípica, às vezes biparental mediante recombinação e segregação, o cromossomo X apresenta peculiaridades exclusivas a serem exploradas em genética de populações (37). Na genética forense, em que pese sua utilização ser ainda pouco difundida, o foco nos seus marcadores vem aumentando rapidamente nos últimos anos (11,38,39,40,41,42,43).

2.3 O CROMOSSOMO X

Em algumas espécies animais, como por exemplo o gafanhoto, as fêmeas possuem um cromossomo a mais que os machos. O cromossomo extra, presente em outras espécies de insetos, é denominado cromossomo X. Em outras espécies, as fêmeas possuem dois cromossomos X, sendo ditas XX e os machos, apenas um, designados como sendo X0, onde o "0" indica ausência de um cromossomo (44). Há outro sistema de diferenciação sexual denominado ZW, comum em aves, onde o sexo heterogamético é a fêmea e não o macho (45).

Em muitos outros animais, incluindo os humanos, machos e fêmeas têm o mesmo número de cromossomos. Netas espécies, o dimorfismo sexual é dependente ainda de outro cromossomo, que nos humanos é o cromossomo Y, presente apenas nos machos. Descobertos no início do século 20 graças aos trabalhos dos citologistas americanos C. E. McClung, N. M. Stevens, W. S. Sutton e E. B. Wilson (46), o X e o Y são denominados de cromossomos sexuais por estarem diretamente relacionados com a definição do sexo cromossômico.

As células da maioria dos homens contêm um cromossomo X e um Y. Estes não formam um verdadeiro par de homólogos, de forma que a maioria das regiões dos dois cromossomos são hemizigotas em machos. Já as fêmeas, por possuírem dois cromossomos X, apresentam um par de homólogos, os quais se comportam como se fosse um par autossômico (47). A maior diferença é que, como mecanismo de compensação de dose, apenas uma cópia do cromossomo X é ativa em cada célula, sendo a cópia adicional inativada, conforme proposto por Lyon (48), formando o corpúsculo de Barr.

Os cromossomos X e Y de humano, bem como o de outros animais, evoluíram provavelmente de um par autossômico (49). Esta diferenciação teve início por volta de 240 a 320 milhões de anos atrás, logo após a divergência entre as linhagens mamífera e aviária. Estima-se que o processo tenha passado por quatro fases distintas, cada uma delas suprimindo a recombinação em uma determinada região dos cromossomos, partindo da extremidade do braço longo para a extremidade do braço curto (50). Atualmente ainda ocorre troca de material genético entre os dois cromossomos sexuais durante a meiose. As regiões de homologia localizam-se nas extremidades de ambos os cromossomos e são denominadas

pseudo-autossômicas, devido ao fato de que os *loci* aí localizados são herdados à maneira de qualquer outro *locus* autossômico (51).

Durante o período evolutivo, o cromossomo Y foi perdendo sequências e genes, de forma a apresentar atualmente um padrão único de sequências repetitivas (52,53). Contrastando com este padrão de evolução, o cromossomo X conservou suas características autossômicas, sendo fisicamente o mais estável cromossomo, ao menos entre os mamíferos placentários (54). Uma característica que distingue o cromossomo X dos demais é o seu modo peculiar de herança. As mulheres, possuindo duas cópias, e os homens, apenas uma, faz com que dois terços da história do cromossomo X se passem nas mulheres. Por esta razão, a taxa de mutação no cromossomo X é esperada ser menor que nos autossômicos e no cromossomo Y, já que as mutações gaméticas por substituições nucleotídicas ocorrem menos frequentemente em mulheres que em homens, devido ao maior número de mitoses na linha germinativa masculina (55,56,57,58). Uma forma de efetuar esta medição tem sido analisar regiões homólogas dos cromossomos X e Y e comparar com uma sequência ancestral inferida. Estudos em humanos e outros primatas têm confirmado uma taxa de mutação maior na linha germinativa masculina (58,59,60,61,62).

Por conta do modo de herança, a diversidade genética no cromossomo X é menor que entre os demais cromossomos, sendo menor também a divergência interespécies (37). Um amplo estudo comparando sequências genéticas demonstrou que entre humanos e chimpanzés, a divergência é de 83% da que foi observada para os cromossomos autossômicos (63). A diversidade genética é influenciada também pelo tamanho efetivo da população. Como no homem falta a segunda cópia do cromossomo X, o tamanho efetivo da população é de apenas 3/4 da dos autossômicos. A soma destes dois efeitos faz com que a diversidade esperada para o cromossomo X, medida principalmente pela taxa de heterozigose, situe-se abaixo da observada para os autossômicos. Embora baixa quando comparada com autossômicos, sua diversidade é cerca de duas vezes maior que a do cromossomo Y, sendo suficiente para fornecer uma razoável quantidade de sítios informativos, da ordem de uma a dez substituições por 1000 pb (37). Outro efeito do menor tamanho populacional é o aumento da deriva genética para o cromossomo X, resultando em um nível de sub-estruturação populacional maior, quando comparado com os

autossômicos; ou seja, as populações devem diferir mais quando vistas sob a ótica do cromossomo X que dos autossômicos (37).

Em humanos, recombinação do cromossomo X ocorre apenas nas mulheres, uma vez que os homens possuem apenas uma cópia. Desta forma, apenas cerca de 2/3 dos cromossomos X se recombinam a cada geração, como já demonstrado experimentalmente (64). Como resultado, espera-se um desequilíbrio de ligação (DL) para marcadores do cromossomo X maior que para os autossômicos (37).

2.3.1 Uso Forense dos marcadores STR do Cromossomo X

A idéia fundamental para o emprego dos marcadores do cromossomo X na prática forense derivou do conhecimento adquirido no campo da genética clínica (47). Muitas doenças e características tais como hemofilia, distrofia muscular de Duchenne, deficiência de G6PD e Daltonismo, apresentam um modo de herança tipicamente ligado ao X. Se um homem portador de uma doença ligada ao cromossomo X é fértil, todas suas filhas serão portadoras daquela característica, transmitida à metade da próxima geração. Nesta próxima geração, metade das filhas será igualmente portadora do alelo e metade dos filhos exibirá o quadro clínico, por serem hemizigotos para o cromossomo X. Este estado de hemizigose masculina facilita sobremaneira os estudos de ligação entre marcadores no cromossomo X. O conhecimento deste contexto é importante não só para a genética clínica, mas também para as análises de vínculo genético na prática forense.

O primeiro passo para a inclusão do cromossomo X na genética forense ocorreu com a descrição do grupo sanguíneo Xga na década de 60 (65) Mann *et al.* 1962). Com o advento da tecnologia da PCR, muitos STRs foram sendo descritos, alguns deles no cromossomo X. Os primeiros de que se tem conhecimento foram o HPRTB (66), HUMARA (67) e DXS981 (35), descritos na primeira metade da década de 1990. Ainda na década de 90 foram descritas fórmulas para calcular os parâmetros forenses, levando em consideração o modo de herança único do cromossomo X (68), as quais foram revistas posteriormente (9) Szibor *et al.*, 2003). No entanto, a inclusão dos marcadores do cromossomo X na identificação humana só começou a ser disseminada já na presente década (9,37,69). Embora mais de

uma centena de STRs do cromossomo X tenha sido já descrita, Szibor (<http://www.chrx-str.org/> Acesso em 28 abr. 2010) lista apenas 55 como estando em uso na identificação humana.

Embora ainda pouco utilizados na prática forense, a demanda pela incorporação dos marcadores CrX STR vem sendo rapidamente incrementada. O Grupo Espanhol-Português da *International Society for Forensic Genetics* (GEP-ISFG) relatou que entre 2006 e 2008 o número de laboratórios que reportaram resultados baseados em análises de CrX STR no exercício de controle de qualidade anual (www.gep-isfg.org – acesso em 12 abr. 2010) aumentou de 11 para 24.

Os marcadores STR do cromossomo X, no que diz respeito ao modo de herança, diferem de todos os outros localizados nos demais cromossomos, bem como do DNA mitocondrial. O cromossomo X se comporta de uma maneira mista, englobando características dos cromossomos autossômicos, do Y e do DNA mitocondrial. Isto porque as mulheres, apresentando um par de homólogos do X, o transmite a todos os seus descendentes, mediante recombinação e segregação alélica, se comportando neste aspecto, até certo ponto, como os cromossomos autossômicos. Desta forma tanto os filhos quanto as filhas herdaram uma cópia do material genético do cromossomo X materno, após recombinação meiótica. Já os homens, possuidores de apenas um cromossomo X, o transmite na forma de haplótipo, sem recombinação, e apenas para a prole feminina. A herança dos STRs do cromossomo X se processa ora como marcador biparental, ora como uniparental, a depender do sexo do genitor. Desta forma, o conteúdo dos CrX STRs das mulheres resulta de uma herança haplotípica do pai e uma herança por recombinação e segregação da mãe. Já os homens, com apenas um cromossomo X e, portanto, hemizigotos, herdaram este X apenas da mãe, mediante recombinação e segregação.

Outro aspecto a ser considerado no uso dos STRs do cromossomo X com fins forenses é o fato de estarem localizados em um mesmo cromossomo, cuja sequência recentemente elucidada demonstra conter cerca de 155 Mb (70). Esse fato implica em certos cuidados quando das análises estatísticas, devendo-se ter um conhecimento exato da localização genética de cada marcador no cromossomo. Os marcadores autossômicos, preponderantes na resolução dos casos forenses, estão localizados, via de regra, em cromossomos diferentes ou, se em um mesmo cromossomo, a uma distância tal que permita recombinação meiótica com uma

frequência de 0,5. Desta forma, a valoração estatística, quando há compatibilidade entre os genótipos das amostras testadas, é feita com base na frequência dos alelos genotipados e pela regra do produto das Razões de Verossimilhança (razão entre duas probabilidades mutuamente excludentes para explicar um mesmo conjunto de dados) de cada um dos marcadores utilizados (20). Já a análise de STRs localizados em um mesmo cromossomo, como é o caso do X, requer o conhecimento da extensão da ligação e do desequilíbrio de ligação entre eles (9).

Ligação em genética se refere à associação por herança de alelos de dois ou mais marcadores, além do que seria esperado em bases aleatórias (71). A abordagem clássica para estudo de ligação se baseia em análise de heredograma (9). No caso do cromossomo X, o ideal é genotipar uma série de mães e dois ou mais filhos do sexo masculino de cada uma delas (72). Como os homens são hemizigotos e apresentam herança apenas materna, a tarefa de obter a frequência de recombinação e, portanto, o grau de ligação, é particularmente eficiente para o cromossomo X. Com esta estratégia se obtêm o mapa genético, com a distância que separa os marcadores no cromossomo, dada em cM. Outra abordagem denominada mapeamento por radiação híbrida (RH) permite acessar diretamente a distância entre marcadores. Por este método, DNA de um painel de células somáticas híbridas, criadas pela fusão de células de hamster não radioativas com linfoblastos humanos marcados com isótopo radioativo, é submetido à amplificação por PCR. Análises estatísticas da presença ou ausência de produtos específicos da PCR permitem determinar a distância física entre marcadores conhecidos e novos marcadores. Vale salientar que nem sempre as distâncias genética e física são coincidentes (73), devendo o RH ser validado por outras investigações.

Marcadores genéticos ligados geneticamente são sintênicos, isto é, residem em um mesmo cromossomo. A ligação pode ser mais ou menos intensa, a depender da frequência de recombinação meiótica. A recombinação é um fenômeno que ocorre a uma taxa que depende fundamentalmente da distância em pares de base que separa duas regiões de um mesmo cromossomo. Como regra, uma distância física de um Mb corresponde a uma distância genética de um cM, devendo ocorrer uma recombinação a cada 100 meioses (9). Já para uma distância de 50 Mb, equivalente a 50 cM, o esperado é que ocorra recombinação com uma frequência de 0,5. Desta forma, marcadores localizados a distâncias menores que 50 cM devem apresentar algum grau de ligação entre eles, sendo que quanto menor a distância

mais forte tenderá a ser a ligação. Para marcadores com forte DL, as frequências genotípicas não podem ser obtidas pela simples multiplicação das frequências gênicas, devendo ser estabelecidas mediante análise populacional (9).

Outro aspecto a ser levado em consideração no processo de recombinação é que ela parece não ocorrer de forma uniforme ao longo do genoma, não sendo completamente entendida atualmente na espécie humana (47). A taxa de recombinação ocorre de forma diferenciada entre os dois sexos, tendendo a ser maior nas regiões próximas dos telômeros nos homens e mais alta próximo ao centrômero nas mulheres (64). Nos homens, a distribuição da recombinação pode ser estudada diretamente pela tipagem de um grande número de espermatozoides. Pesquisas usando esta estratégia têm demonstrado a existência de locais preferenciais, ou “*hot spots*”, onde a frequência de recombinação se processa a uma taxa maior que o esperado, enquanto outras regiões são menos propensas à troca de material genético entre cromossomos homólogos na meiose, não se sabendo ainda quão universal ou variável este fenômeno ocorre ao longo de todo o genoma (74,75,76). Uma plataforma óbvia para tais pesquisas em mulheres é o uso de marcadores do cromossomo X, por ser este o único local em que recombinação ocorre em mulheres, na ausência de recombinação nos homens (37).

Um aspecto adicional com relação ao uso forense dos CrX STRs diz respeito ao desequilíbrio de ligação, o qual se refere à ocorrência, na população, de uma frequência maior de uma determinada combinação entre alelos de dois genes, do que a esperada pelo produto de suas frequências individuais (47). Em outras palavras, desequilíbrio de ligação é a medida estatística de associação entre alelos de diferentes *loci*, em uma população. Ele é modulado principalmente pela distância que separa dois marcadores em um cromossomo, mas não é uma função apenas desta variável, existindo outras influências não muito bem compreendidas atualmente (41). Estruturação populacional, seleção, deriva genética e fluxo migratório são citados como fatores que influenciam na extensão do DL (77,78).

Para um painel de marcadores, o estudo de DL em uma dada amostra populacional é feito mediante análise dos mesmos aos pares, com o objetivo de verificar a associação preferencial de alelos de um marcador com alelos do outro marcador. A presença de DL é estabelecida pela observação de desvio significativo das frequências haplotípicas observadas em relação ao que seria esperado pelo produto das frequências alélicas (9).

Em razão das altas taxas de mutação, os STRs tendem a ser muito menos afetados por DL do que os polimorfismos pontuais do tipo SNP, contribuindo para diluir o efeito do desequilíbrio de ligação (41). No entanto, DL pode ocorrer entre marcadores localizados muito próximos em um cromossomo, devendo ser devidamente avaliado o seu efeito, antes de serem usados na prática forense (47,79). Analisando três CrX STRs (DXS6801–DXS6809–DXS6789) com distâncias inferiores a cinco cM em uma população alemã, foi observado forte desequilíbrio de ligação entre os três pares de marcadores (72). Por outro lado, segundo (41), para marcadores distanciados mais de cinco cM no cromossomo X, DL não costuma ocorrer.

Em um estudo brasileiro (78), 11 microssatélites, a maioria localizada no braço longo do cromossomo X, foram investigados com vistas a avaliar a extensão do desequilíbrio de ligação entre os *loci*. A população estudada era composta de dois grupos: uma amostra de 200 indivíduos não relacionados residentes no Rio Grande do Sul e um conjunto de amostras de 131 ameríndios igualmente não relacionados provenientes de tribos localizadas no Brasil, Argentina e Paraguai. Foram realizadas análises de pares de marcadores e aplicada a correção de Holm-Sydak. Os resultados demonstraram que DL estatisticamente significativo ($p > 0,001$) foi observado em apenas um de 55 pares (DXS6789-GATA31E08) na população do Rio Grande do Sul, e em três de 55 pares (DXS9895-GATA31E08, DXS7130-GATA31E08 e DXS9895-DXS7423) na população ameríndia. Estruturação populacional originada possivelmente por endogamia, casamentos consanguíneos e eventual efeito fundador foram evocados para explicar uma maior DL entre ameríndios.

Em uma análise de uma amostra da população alemã composta de 210 homens, utilizando um painel de 16 STRs do cromossomo X, foi encontrada uma DL significativa em apenas um par de marcadores: DXS101 e DXS7424 (39). Já em um estudo colaborativo conduzido pelo GEP, um sistema multiplex composto por dez CrX STR foi estudado em populações ibero-americanas de seis países (Portugal, Espanha, Brasil, Argentina, Colômbia e Costa Rica). No total foram genotipados 1492 homens. Para um nível de significância de 0,0011 obtido após a correção de Bonferroni, apenas uma entre 675 comparações realizadas (45 comparações por população) apresentou níveis compatíveis com DL entre os *loci* DXS7133-DXS6789 (11). Outros estudos populacionais têm obtido resultados semelhantes (41,80,81).

Sistemas multiplexes de amplificação de STRs do cromossomo X já foram desenvolvidos, como o X-Decaplex (11) Gusmão *et al.*, 2009) utilizado neste trabalho, com amplificação simultânea de dez CrX STRs e Mentype® Argus X-8 (Biotype® AG), composto de oito marcadores (DXS8378, HPRTB, DXS7423, DXS7132, DXS10134, DXS10074, DXS10101 e DXS10135). No Brasil, algumas populações já foram caracterizadas para marcadores STRs do cromossomo X, como exemplo, Mato Grosso do Sul, Paraná e Rio de Janeiro (11), São Paulo (11,82) Pará (83) e Rio Grande do Sul (78).

2.3.2 Os CrX STRs nos exames de paternidade cíveis

Atualmente, a principal indicação para o emprego forense dos CrX STR tem sido os casos complexos de paternidade, ditos casos deficientes, nos quais o suposto pai não pode ser testado, sendo substituído pelos seus parentes (79). Os casos envolvendo o clássico trio composto de mãe, criança e suposto pai são resolvidos usualmente utilizando apenas dados de marcadores autossômicos (9). Quando a mãe não está disponível e a relação testada é apenas pai/filha, pode ser necessária a inclusão dos marcadores CrX STR para a resolução do caso, especialmente quando o material biológico apresenta dificuldade para ser analisado, como nos casos de exumação (47). Ressalte-se que em apenas metade das disputas de paternidade tais marcadores podem ser usados, restringindo-se às situações nas quais o filho em disputa é do sexo feminino (40), tendo recebido, portanto, um cromossomo X paterno. No caso do filho ser do sexo masculino, não existe, no genótipo deste último, alelo dito IBD (*identical by descent* – idêntico por descendência), sendo, portanto, absolutamente não informativa a análise destes marcadores.

Se o suposto pai não pode ser testado, a informação relativa a seu haplótipo pode ser recuperada basicamente de duas maneiras: (a) por reconstituição do haplótipo paterno mediante a genotipagem genética de uma ou mais filhas biológicas e da genitora destas; (b) por estimativa do referido haplótipo, o qual pode ser feito mediante análise da suposta avó ou, na ausência desta, da sua prole masculina. A informação sobre a suposta avó pode ainda ser recuperada pela

genotipagem de sua prole feminina e do suposto avô (9), permitindo quais alelos foram transmitidos por um e por outro genitor do suposto pai.

Quando o haplótipo do suposto pai é individualizado por genotipagem do seu material biológico ou de filhas biológicas, a valoração estatística decorrente da eventual compatibilidade da paternidade deve levar em consideração o desequilíbrio de ligação entre os *loci* testados. Esta parece tender a não ser muito pronunciada entre os CrX STR. A recombinação alélica no cromossomo X na proporção de 2/3 dos indivíduos a cada geração bem como a alta taxa de mutação entre os microssatélites em geral são invocados como fatores determinantes para o pouco DL observado na maioria das populações testadas. Em todo caso, para que as frequências alélicas de cada *locus* analisado possam ser usadas para prever as frequências haplotípicas, o ideal é avaliar o grau de DL na própria população das quais as pessoas envolvidas fazem parte (11). A ausência de associação preferencial entre alelos dos vários marcadores testados permite levar em consideração todos eles no cálculo do Índice de Paternidade, definido como a razão entre as duas probabilidades mutuamente excludentes envolvidas para explicar a combinação dos genótipos observados nos indivíduos testados: a probabilidade do suposto pai testado ser efetivamente o pai biológico e a probabilidade de ser outro indivíduo qualquer, selecionado aleatoriamente na população.

Se o conjunto de alelos do suposto pai é estimado a partir do genótipo da suposta avó, a regra da multiplicação entre os *loci* estudados implica em uma abordagem mais cautelosa. A mulher, por ser portadora de dois cromossomos X, transmite aos seus descendentes cada um dos alelos de um determinado *locus* com uma probabilidade de 0,5. A valoração estatística pela multiplicação destas probabilidades para vários *loci* só pode ser efetuada caso a taxa de recombinação entre os *loci* se situe por volta de 0,5. Para marcadores ligados, a segregação não é independente e a regra da multiplicação implica em sérias distorções da frequência haplotípica observada, comparada à esperada pela multiplicação das frequências genotípicas. Em um estudo recente, (9) identificaram quatro grupos de ligação para o cromossomo X, sugerindo que marcadores pertencentes a um mesmo grupo sejam tratados como haplótipos. Desta forma, se o genótipo da suposta avó substitui o haplótipo do suposto pai nas análises, a regra da multiplicação só valeria para marcadores pertencentes a grupos de ligação distintos. Esta regra seria igualmente válida para as demandas envolvendo o duo mãe-filha.

2.3.3 Os CrX STR em outros exames de vínculo genético

Embora a grande indicação para a exploração dos CrX STRs seja os casos deficientes de disputas de paternidade, em algumas outras circunstâncias eles podem agregar informações valiosas aos resultados obtidos com os STR autossômicos, eventualmente se constituindo no fator decisivo para uma clara compreensão da situação. Em algumas disputas de paternidade envolvendo mais de um parente consanguíneo na condição de suposto pai, o poder de exclusão dos STRs fica consideravelmente reduzido. Nestas circunstâncias, os CrX STRs podem ser superiores aos AS STRs. Por exemplo, pai e filho não compartilham alelos idênticos por descendência (IBD). Quando ambos se encontram na condição de suposto pai, os CrX STRs serão mais informativos que os AS STRs. Irmãos por outro lado compartilham um alelo materno com uma probabilidade de 0,5. Esta corresponde à mesma probabilidade de compartilhar um alelo IBD de um *locus* autossômico (47).

Quando um aborto legal é praticado, o material biológico examinado se constitui, via de regra, de uma pequena quantidade de restos fetais e uterinos, nem sempre passíveis de dissecação e identificação dos seus constituintes. Em tais casos, o exame genético costuma apresentar uma mistura do DNA fetal e materno. Se a gravidez resultou supostamente de incesto e o feto é do sexo masculino, os marcadores do cromossomo Y podem ser utilizados eficientemente para auxiliar os autossômicos. Para os fetos do sexo feminino, os marcadores autossômicos e do cromossomo X são as duas únicas ferramentas possíveis, com o poder de exclusão dos CrX STRs sendo superior ao dos AS STRs. A prova da paternidade, no entanto, é menos fácil de ser obtida com os CrX STRs. Isto porque se um homem engravida a própria filha, o haplótipo do feto estará necessariamente contido no genótipo da mãe, não sendo possível afirmar que no material analisado existia DNA fetal (47). Nestes casos não são observados mais de dois alelos por *locus*, mesmo havendo mistura de material fetal e materno.

Nos casos de identificação de cadáver, o material biológico nem sempre é de boa qualidade, especialmente se são encontrados apenas restos ósseos. Devido à relativa incerteza da paternidade nos tempos atuais, a identificação é mais confiável se feita com a linha materna. Nestas circunstâncias, os marcadores do cromossomo

X se constituem em uma alternativa valiosa. Quando a relação mãe-filha é testada, os CrX STR se equivalem aos STR autossômicos. O vínculo mãe-filho, por outro lado, é mais eficientemente avaliado pelos marcadores do X que pelos autossômicos. O poder de exclusão nesta circunstância se equivale ao da relação pai-filha (9).

2.3.4 Os CrX STRs em análise de vestígios criminais

Quando vestígios de um delito são examinados para se identificar a fonte biológica, os marcadores do cromossomo X na maioria das vezes são menos informativos que os autossômicos. O Poder de Discriminação (PD) dos CrX STRs variam a depender do sexo, sendo igual aos AS STRs quando vestígios de origem feminina são comparados com indivíduo do sexo feminino e inferior a estes se a origem do vestígio e o indivíduo testado são do sexo masculino. Este último cenário se deve ao fato de no sexo masculino ser examinado apenas um alelo por *locus* (9).

No caso de um vestígio apresentar mistura de material genético de indivíduos de ambos os sexos, a procura da fração masculina pelos CrX STRs não apresenta vantagem. Isto porque a chance dos alelos masculinos estarem incluídos no componente feminino é maior que para os autossômicos, em decorrência da hemizigose masculina. Se a situação se inverte e o perfil procurado é o feminino, o mesmo estado hemizigoto masculino para os marcadores do cromossomo X se transforma em uma vantagem, quando comparado com os autossômicos. Isto porque a completa inclusão do genótipo feminino no haplótipo masculino só se daria na rara eventualidade da mulher ser homozigota para todos os *loci* estudados (9).

Com o intuito de otimizar a performance dos CrX STRs em análise de vestígios com material biológico intensamente degradado, novos iniciadores vêm sendo desenhados de forma a gerar *amplicons* de menor comprimento (84), à semelhança dos mini STRs autossômicos.

2.3.5 Aberrações cromossômicas e marcadores do cromossomo X

Nas células de mulheres sem aberrações cromossômicas sexuais, o cromossomo X está presente como um par de homólogos. Para um dado casal fértil, a presença de irregularidades gonossômicas pode ser negligenciada, já que a sua presença, via de regra, se associa a infertilidade. Porém, tais aberrações podem estar presentes em cariótipos dos descendentes de forma inesperada e eventualmente não detectada (9). Exemplo disso é a Síndrome de Turner, cujo cariótipo é 45,X. Dessa forma, sempre que a análise de vários marcadores ligados ao X se mostrarem homozigotos deve-se pensar na possibilidade de se estar diante de uma mulher com cariótipo 45,X ou uma mulher com insensibilidade androgênica e cariótipo 46,XY (9). Na forma como alteram as análises de vínculo genético, tais aberrações se equivalem às dissomias uniparentais autossômicas (85). Outra condição aberrante é a Síndrome de Klinefelter, na qual indivíduos masculinos apresentam cariótipo 47,XXY ou 48,XXXYY. Diante de um homem com genótipo heterozigoto para os CrX STRs, deve-se pensar na ocorrência dessa síndrome genética.

Independente de qual destas alterações esteja presente, elas interferem na acurácia dos testes, fazendo com que os marcadores do cromossomo X percam a validade em análises de vínculo genético (47). Aqui vale a observação de que qualquer alteração cromossômica ou de outra natureza que se configure em uma característica desfavorável descoberta fortuitamente em um exame de identificação humana, deve ser vista pela ótica da confidencialidade, só devendo ser revelada ao indivíduo afetado mediante solicitação expressa deste ou de seu representante legal (9).

2.4 BANCO DE DADOS DE FREQUÊNCIA ALÉLICA

Um Banco de Dados de referência para cálculo de frequência alélica é constituído a partir da genotipagem de uma amostra de indivíduos selecionados em uma população. Idealmente deveria ser estruturado com base em uma amostragem

aleatória simples ou cientificamente estruturada da população. No entanto, por ser difícil de ser implementada, dispendiosa e pouco prática, esta estratégia pouco é utilizada. De fato, a maioria dos bancos é composta de amostras de conveniência, de fontes como bancos de sangue, laboratórios de exames de paternidade, centros de aconselhamento genético, funcionários de laboratórios e indivíduos acusados de crimes (86). É necessário, no entanto, que o banco seja acreditado teoricamente mediante a análise de parâmetros como Equilíbrio de Hardy-Weinberg, Equilíbrio de Ligação, Poder de Discriminação, Poder de Exclusão e Heterozigose.

Bancos de Dados costumam ser usados sem que a amostragem tenha saído da população onde ele é aplicado. Parte dessa realidade deve-se ao fato de que muitas populações ainda não dispõem de tais bancos de dados. Por outro lado, estudos mostram que uma população qualquer no mundo apresenta 85% da variabilidade genética para marcadores autossômicos, que contam a história das migrações masculina e feminina. Diferenças entre populações de um mesmo continente contribuem para outros 6% e apenas 9 a 13% da variação genética diferencia populações de diferentes continentes (87). No entanto, idealmente, os cálculos estatísticos em genética forense devem ter por base um banco de dados que reflita o real balanço entre os alelos da população onde ele será aplicado (86).

2.5. ORIGEM DAS POPULAÇÕES BRASILEIRA E DO DISTRITO FEDERAL

O Brasil é um país de composição étnica bastante diversa. Fazendo parte do continente americano, foi dos últimos rincões do planeta a ser povoado por humanos. Desde um período compreendido entre 11 e 15 mil anos atrás, habitava por quase todo o país uma população com características físicas homogêneas, compondo o grupo ameríndio (88). No início do século 16 os europeus portugueses aqui aportaram e pouco depois iniciaram a colonização. Naquela época o território atual do Brasil era habitado por cerca de 2.000.000 de indígenas. Desde então, este povo foi submetido a um drástico declínio demográfico devido a conflitos com os colonizadores europeus e doenças para as quais ele não estava adaptado (89), além da amalgamação com os novos povos que aqui chegaram. Para compor força de trabalho escrava, em meados do mesmo século os portugueses iniciaram a

migração forçada oriunda da África subsaariana, política que se estendeu por 300 anos. Durante todo o período colonial, outros europeus também para cá migraram, culminando com a vinda de grandes contingentes de alemães, italianos e espanhóis entre os séculos XIX e XX (90), após o declínio da política escravagista de africanos e seus descendentes, fixando-se principalmente no sudeste e sul do país. Por fim, na primeira metade do século passado outra onda migratória digna de registro foi de contingentes asiáticos, principalmente japoneses. Dos migrantes que colonizaram o Brasil, aproximadamente 58% eram europeus, 40% africanos e 2% asiáticos (91). Em decorrência deste padrão de colonização, a população brasileira é uma das mais heterogêneas no mundo, resultando basicamente da mistura de Europeus (principalmente portugueses), Africanos e Ameríndios nativos. Esta contribuição tri-étnica da população brasileira não se deu de forma uniforme, havendo predomínio de ameríndios na região norte, africanos no nordeste e europeus no sul (92). A Tabela 1 apresenta a variação estimada da composição étnica brasileira em 480 anos (93).

Tabela 1 - Estimativa da composição da população brasileira em três épocas distintas com base no critério cor-da-pele.

Cor-da-pele	1500	1890	1980
Indígena	2.000.000	440.000	200.000
Branco	-	6.302.000	65.000.000
Mistos	-	6.000.000	46.000.000
Negros	-	2.000.000	7.000.000
Asiáticos recentes	-	-	733.000

O Distrito Federal, sede atual do governo brasileiro, existe há 50 anos. Instalado em uma área pouco povoada do estado de Goiás, é a menor unidade da federação brasileira e a que tem a menor quantidade de habitantes. Segundo dados do IBGE, em 2007 sua população era de cerca de 2.450.000 habitantes (94). A constituição desta população é o resultado predominante de fluxos migratórios oriundos de todo o país, principalmente da Região Nordeste, seguida da Sudeste e Centro Oeste (95).

A Tabela 2 apresenta a distribuição das populações brasileira e do Distrito Federal segundo os critérios cor da pele/ancestralidade adotados pelo IBGE, de acordo com o censo do ano 2000. A metodologia utilizada para a classificação dos recenseados foi a autodeterminação.

Tabela 2 - Composição da população brasileira e do Distrito Federal no ano 2000 segundo auto-denominação de cor da pele.

Cor da pele	Brasil	%	Distrito Federal	%
Branca	91.298.042	53,74	1.008.199	49,15
Preta	10.554.336	6,21	98.462	4,80
Parda	65.318.092	38,45	918.305	44,77
Amarela	761.583	0,45	7.996	0,39
Indígena	734.127	0,43	7.154	0,35
Sem declaração	1.206.675	0,71	11.029	0,54

Fonte – IBGE censo 2000

3 OBJETIVOS

Este trabalho teve como objetivo geral caracterizar uma população brasileira do Distrito Federal para marcadores genéticos do tipo STRs situados no cromossomo X humano. Para tanto, os objetivos específicos foram:

1. Analisar a distribuição das frequências alélicas de 10 STRs situados no cromossomo X: DXS8378, DXS9898, DXS7133, GATA172D05, GATA31E08, DXS7423, DXS6809, DXS7132, DXS9902 e DXS6789;
2. Analisar parâmetros de interesse da genética forense a partir desses dados;
3. Analisar parâmetros populacionais e compará-los com os dados referentes a outras populações brasileiras;
4. Estimar o padrão de desequilíbrio de ligação entre os marcadores STRs situados no cromossomo X e comparar com os resultados obtidos para um conjunto de marcadores STRs autossômicos não sintênicos e um conjunto de marcadores STRs localizados no cromossomo Y;
5. Estimar a contribuição genética das populações parentais européia, africana e ameríndia para a formação da população do Distrito Federal do ponto de vista dos marcadores STR do cromossomo X e comparar os dados com os resultados obtidos para populações de outras unidades da federação;
6. Comparar as estimativas de contribuições parentais para a formação da população do Distrito Federal pela análise dos STRs do cromossomo X, com as estimativas alcançadas com base em marcadores autossômicos;
7. Analisar o impacto produzido pelos dez STRs do cromossomo X na valoração estatística em casos de paternidade deficientes.

Este projeto é parte integrante do projeto de pesquisa “Análise de marcadores de DNA Bialélicos e STRs do Cromossomo X em populações humanas no Centro-Oeste (Brasil)” aprovado pelo Conselho de Ética em Pesquisa da Faculdade de Medicina da UNB em 28 de setembro de 2005 sob o número CEP-FM 066/2005.

4. MATERIAL E MÉTODOS

4.1 CARACTERIZAÇÃO DA AMOSTRA

Todos os indivíduos cujas amostras biológicas foram utilizadas para as tipagens genéticas haviam cedido seu material para a realização de exames genéticos de paternidade civil ou exames na esfera criminal. A coleta foi precedida da assinatura, por parte do doador ou do seu representante legal, de um termo de consentimento denominado “Declaração de Doação Voluntária”.

4.1.1 Amostra para análise dos marcadores do cromossomo X

Um total de 415 indivíduos nascidos no Distrito Federal, sendo 205 homens e 210 mulheres, integraram a amostra utilizada nas análises dos STRs do cromossomo X. As pessoas foram selecionados de 331 casos de vínculo genético de paternidade cível, analisados no Instituto de Pesquisa de DNA Forense da Polícia Civil do Distrito Federal (IPDNA), compreendendo o período de outubro de 2007 a janeiro de 2010. Apenas indivíduos não relacionados geneticamente considerando o cromossomo X foram selecionados para compor a amostra. Observando este critério, em 84 protocolos pesquisados, mais de um indivíduo foi selecionado.

4.1.2 Amostra para análise dos marcadores dos cromossomos autossômicos

Foram selecionados 102 indivíduos de casos rotineiros de análises de paternidade com o trio estabelecido, constituindo-se de filhos nos respectivos casos, cujas amostras biológicas foram coletadas no período de janeiro de 2008 a dezembro de 2009. Três critérios foram usados para a seleção da amostra: (1) compatibilidade da paternidade pretendida, (2) Índice de Paternidade superior a

10.000, (3)mãe, filho e suposto pai não compartilharem um mesmo genótipo heterozigoto em qualquer dos dez marcadores analisados.

4.1.3 Amostra para análise dos marcadores do cromossomo Y

Composta de amostras biológicas de 204 indivíduos não compartilhando a mesma linhagem patrilínea, selecionados de casos ordinários de paternidade e de amostras referência de casos criminais, coletadas no período de outubro de 2007 a fevereiro de 2010.

4.2 COLETA DAS AMOSTRAS

De cada doador foi coletada uma amostra de aproximadamente 3,0 mL de sangue periférico, em tubo contendo anticoagulante EDTA. Das crianças foram coletados esfregaços de células da mucosa oral em *swab* estéril. Dos supostos pais foram coletados adicionalmente esfregaços de mucosa oral. Durante a coleta foi indagado do doador se havia história de transfusão de sangue nos últimos 90 dias e/ou transplante de medula óssea. Após a coleta o material biológico foi estocado em freezer a -20°C até o momento da extração do DNA.

4.3 PROCEDIMENTOS LABORATORIAIS

4.3.1 Extração do DNA

O DNA das amostras biológicas foi extraído por um método rápido de extração com o emprego de NaOH/TRIS (96). No Anexo 1 estão apresentados os protocolos utilizados para extração de sangue e de *swab* bucal.

4.3.2 Marcadores genéticos

O DNA extraído de cada amostra destinada a análises com marcadores STR do cromossomo X humano foi amplificado pelo método da PCR para o sistema multiplex de amplificação do DNA denominado X-Decaplex (11), composto dos seguintes marcadores: DXS8378, DXS9898, DXS7133, GATA31E08, GATA172D05, DXS7423, DXS6809, DXS7132, DXS9902, e DXS6789. A amplificação era feita em uma única reação multiplex por amostra. A sequência de apresentação dos marcadores é a mesma de Gusmão *et al.* (11) e refere-se à ordem de leitura dos mesmos no eletroferograma. Todos os marcadores analisados apresentam unidade de repetição composta de 04 bases. Quatro microssatélites apresentam estruturas complexas, compostas de mais de um motivo de repetição e/ou sequências não repetitivas intercalantes (DXS9898, GATA31E08, DXS6809 e DXS6789). O STR DXS9902 foi descrito por Edelman *et al.* (39) como sendo uma sequência regular da estrutura de repetição GATA. O uso da estrutura TAGA neste *locus* resulta em uma unidade de repetição adicional. Seguindo orientações da ISFG, neste trabalho foi usada a última nomenclatura. A localização cromossômica, estrutura dos STRs, sequência dos iniciadores e fluoróforos utilizados na reação de amplificação estão listados na Tabela 3. A Figura 1 demonstra graficamente a localização dos dez STRs no cromossomo X.

As amostras destinadas a análise pelos marcadores autossômicos foram amplificadas com o sistema multiplex da amplificação Identifiler (AmpFISTR® Identifiler® PCR Amplification Kit - Applied Biosystems, CA, USA), com co-amplificação simultânea de 15 STRs autossômicos (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818 e FGA) e o marcador de dimorfismo sexual Amelogenina.

As amostras destinadas a análise pelos marcadores do cromossomo Y foram amplificadas com o sistema multiplex da amplificação Y-filer (AmpFISTR® Yfiler® PCR Amplification Kit - Applied Biosystems, CA, USA), com co-amplificação simultânea de 16 STRs.

Tabela 3 – Localização e estrutura dos dez CrX STRs que compõem o multiplex X-Decaplex, sequência dos iniciadores e respectivos fluoróforos.

STR	Localização	Sequência repetitiva	Sequência dos iniciadores (5' → 3') foward - reverse	Fluoróforo	Referência
DXS8378	Xp22.31	(CTAT) _n	TTAGGCAACCCGGTGGTCC ACAAGAACGAACTCCAACCTC	6-FAM	(97)
DXS9898	Xq21.33	(TATC) ₂ -ATC-(TATC) _n	CGAGCACACCTACAAAAGCTG TAGGCTCACCTCACTGAGCA	6-FAM	(98)
DXS7133	Xq22	(ATAG) _n	CACTTCCAAAAGGGGAAAAA ACTTGTA CT TGGTGGGAGGAA	6-FAM	(97)
GATA31E08*	Xq27	(AGGG) ₂₋₃ -(AGAT) _n	GCAAGGGGAGAAGGCTAGAA TCAGCTGACAGAGCACAGAGA	6-FAM	(69)
GATA172D05	Xq26-27	(TAGA) _n	TAGTGGTGATGGTTGCACAG ATAATTGAAAGCCCGGATTC	VIC	(97)
DXS7423	Xq27	(TCCA) _n	GTCTTCCTGTCATCTCCCAAC TAGCTTAGCGCCTGGCACATA	VIC	(42)
DXS6809	Xq21-33	(CTAT) _n -(ATCT) ₃ -N ₉ -(TATC) _n - (ATCT) _n -N ₁₀ -(ATCT) _n	TCCATCTTTCTCTGAACCTTCC TGCTTTAGGCTGATGTGAGG	VIC	(99)
DXS7132	Xq11.1- n ₁₂	(TCTA) _n	TCCCCTCTCATCTATCTGACTG CACTCCTGGTGCCAACTCT	NED	(97)
DXS9902*	Xp22.1	(TAGA) _n	CTGGGTGAAGAGAAGCAGGA GGCAATACACATTCATATCAGGA	NED	(39)
DXS6789	Xq21.3	(TATC) _n -(TATG) _n -(TATC) _n	CTTCATTATGTGCTGGGGTAAA ACCTCGTGATCATGTAAGTTGG	NED	(40)

* Par de iniciadores apresentado em Gusmão *et al.* (11).

4.3.3 Reação de PCR

A mistura de amplificação foi preparada em ambiente físico específico para tal no IPDNA, em capela de fluxo laminar, em um volume total de 9,0 µL por amostra, contendo 5,0 µL de água estéril, 1,0 µL de tampão AmpliTaq Gold (*Applied Biosystems*, Foster City, CA), 2,5 mM de MgCl₂, 0,5 µM de cada dNTP (GeneAmp dNTP Mix), 1,0 µL de 10x *primer* Mix X-Decaplex (0,2 µM de cada iniciador na reação de PCR) e 2,5 UI de AmpliTaq Gold DNA Polimerase (*Applied Biosystems*, Foster City, CA). O protocolo laboratorial utilizado encontra-se no Anexo 1.

Uma vez que a amplificação era feita por lote de amostras, a mistura era preparada em um microtubo de 0,5 mL e em seguida distribuída em microtubos de 0,2 mL, um para cada amostra, aos quais se adicionava 1,0 µL do DNA genômico extraído, que representa aproximadamente entre 0,5 e 2,0 ng de DNA. Um controle negativo contendo a mistura de amplificação e 1,0 µL de água estéril no lugar do DNA era adicionado. O controle positivo da amplificação foi feito com a linhagem celular 9947A (*Promega Corporation*), com perfil alélico conhecido.

A amplificação foi feita em dois modelos de termocicladores - GeneAmp 9600 (*Perkin-Elmer*) e GeneAmp 9700 (*Applied Biosystems*) -, previamente validados para genotipagem mediante análise de fragmentos, de acordo com os parâmetros descritos na Tabela 4.

Tabela 4- Parâmetros de ciclagem para amplificação dos X-STR via PCR

Temperatura	Tempo	Especificação
95 °C	15 minutos	Desnaturação Inicial
94 °C	30 segundos	10 ciclos
60 °C	90 segundos	
72 °C	60 segundos	
94 °C	30 segundos	20 ciclos
58 °C	90 segundos	
72 °C	60 segundos	
72°C	60 minutos	Extensão final

4.3.4 Detecção dos produtos de Amplificação e Assinalamento alélico

A separação e detecção alélica foi feita por Eletroforese Capilar em Analisador Automático de DNA ABI 3130 (Applied Biosystems), em condições desnaturantes, com polímero POP-4 (Applied Biosystems), de acordo com o protocolo apresentado no Anexo 1. Foram empregados dois equipamentos, um configurado para análise de 16 amostras simultâneas e outro configurado para análise de quatro amostras, a escolha sendo feita de acordo com a quantidade de amostras a serem analisadas em cada lote de amplificação. Os dados gerados foram coletados com a utilização do *software* Data Collection v.3.0. O tamanho dos fragmentos foi determinado por comparação com o padrão interno GeneScan™ – 500LIZ® Size Standard (Applied Biosystems), recorrendo ao *software* GeneMapper v3.2 (Applied Biosystems).

Para que fosse alcançada uma correta genotipagem das amostras, inicialmente foram amplificadas as linhagens genéticas 9947A (Applied Biosystems) e K562 (Promega Corporation), com composição alélica conhecida para os 10 marcadores do cromossomo X, além de sete amostras biológicas utilizadas nos exercícios dos anos 2005 e 2006 do Controle de Qualidade do Grupo Espanhol-Português da *International Society for Forensic Genetics* (GEP-ISFG), cujos perfis alélicos foram disponibilizados no endereço eletrônico <http://www.gep-isfg.org/ISFG/portada-idioma.php>, Acesso em 20 fev. 2009.

O tamanho, em pares de bases, obtido para cada um dos fragmentos gerados destas nove amostras padrão foi utilizado para criar as faixas (*Bin*) de localização de cada alelo no aplicativo GeneMapper v3.2. As faixas para os demais alelos já descritos na literatura para cada um dos marcadores e não presentes nas amostras referidas foram então ajustadas de acordo com a diferença, em pares de bases, em relação aos alelos das citadas amostras. Os ajustes foram feitos de forma individual para cada um dos dois analisadores automáticos de DNA, uma vez que as condições de corrida eletroforética variam de um equipamento para outro. As Figura 2 e Figura 3 apresentam eletroferogramas gerados pelo aplicativo GeneMapper v3.2, resultantes de análises de uma amostra feminina e outra masculina, respectivamente.

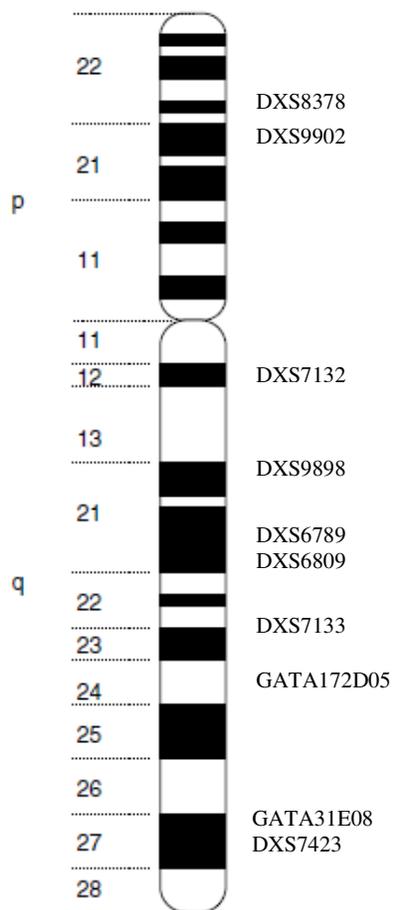


Figura 1 - Ideograma apresentando a localização dos STRs do X-Decaplex ao longo do cromossomo X.

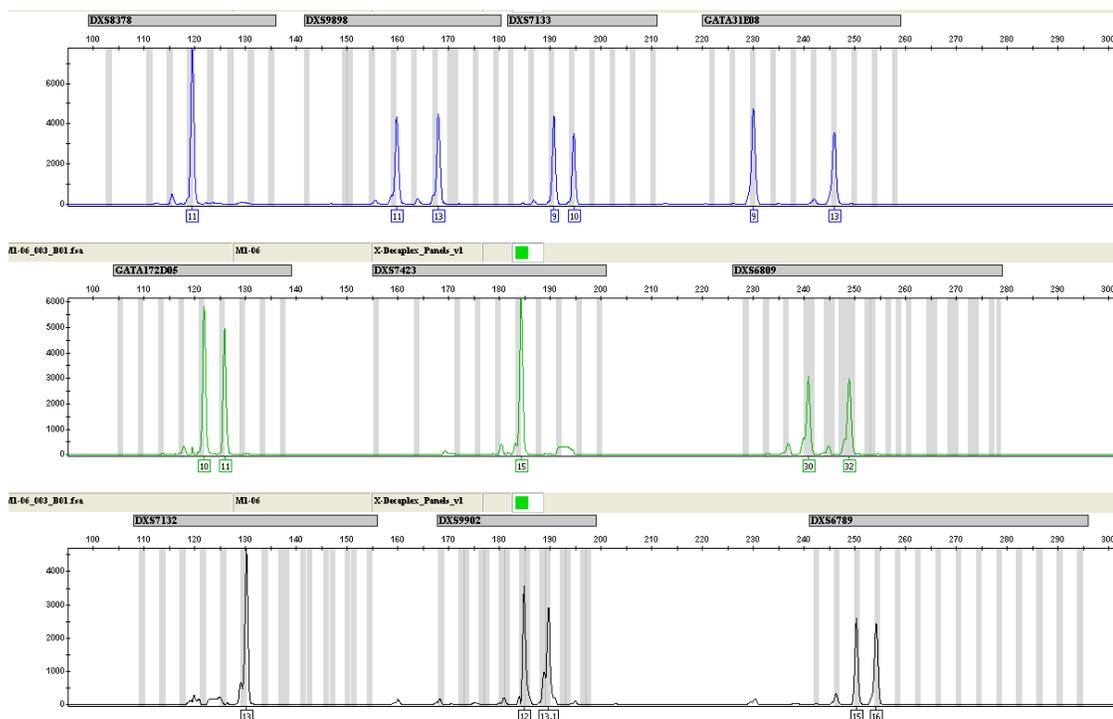


Figura 2 - Eletroferograma de uma amostra de origem feminina

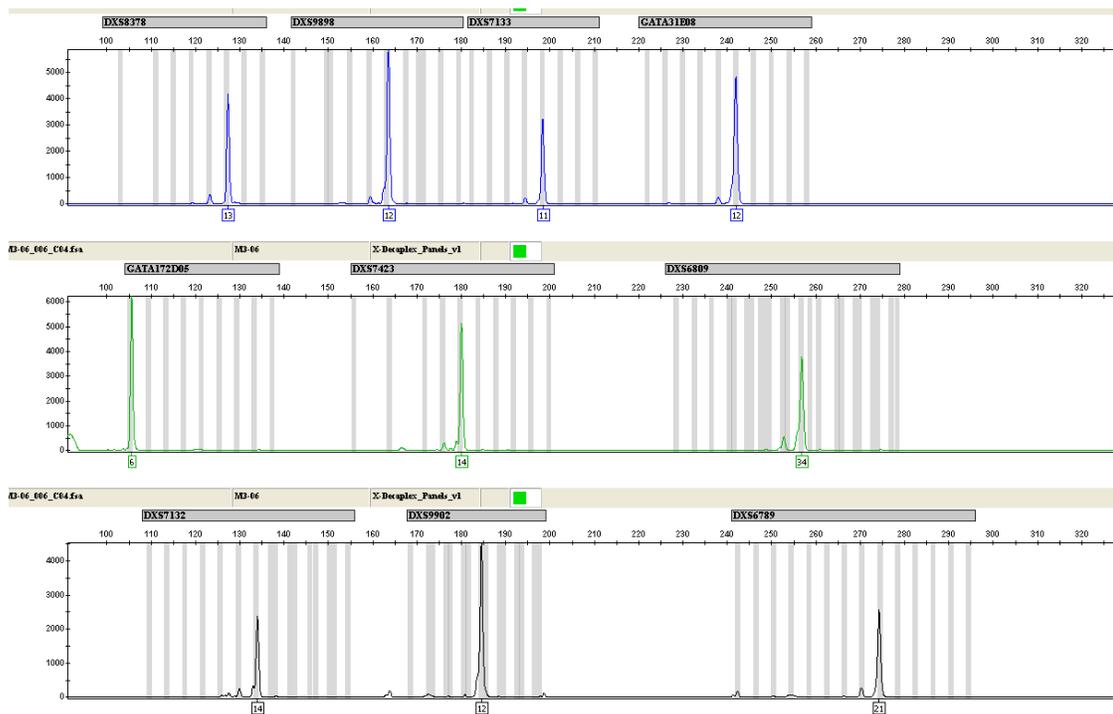


Figura 3 – Eletroferograma de uma amostra de origem masculina

4.4. OUTRAS AMOSTRAS POPULACIONAIS

Em um esforço colaborativo coordenado pelo Grupo Espanhol Português da *International Society for Forensic Genetics* (GEP-ISFG), amostras de 15 populações Íbero-americanas foram analisadas para os 10 CrX STRs aqui descritos (11). Quatro destas populações eram brasileiras: Mato Grosso do Sul (n=300), Paraná (n=309), Rio de Janeiro (n=165) e São Paulo (N=306). O “n” se refere ao número de cromossomos e não o número de indivíduos, considerando que homens possuem um cromossomo e mulheres, dois. Os parâmetros populacionais e forenses descritos para as quatro populações brasileiras (11) - frequências alélicas, heterozigose esperada, heterozigose observada, equilíbrio de Hardy-Weinberg, poder de exclusão médio para trio, poder de exclusão médio para duo, poder de discriminação em mulher e poder de discriminação em homem - foram utilizados para a realização de análises estatísticas comparativas com a população do Distrito Federal. Os genótipos das mulheres e haplótipos dos homens destas quatro populações serviram de base para a realização de análises comparativas quanto a distâncias genéticas e mistura étnica em relação à população objeto deste trabalho.

4.5 ANÁLISES ESTATÍSTICAS

As análises estatísticas foram realizadas com vistas a estimar os seguintes parâmetros de interesse tanto populacional quanto forense: frequências alélicas, heterozigose observada e esperada, aderência ao equilíbrio de Hardy-Weinberg, poder de discriminação, probabilidade de exclusão, estatística F, diversidade alélica, equilíbrio de ligação genotípica, estimativas de contribuição parental na constituição atual da população e distâncias genéticas.

4.5.1 Frequências Alélicas e Heterozigose

As frequências alélicas para cada um dos STRs em cada população foram estimadas por contagem direta, utilizando-se o programa Arlequin v.3.0 (100). A frequência do alelo “ i ” é dada pela fórmula:

$$p_i = \frac{M \cdot A_i Y + F \cdot A_i A_j + 2F \cdot A_i A_i}{M + 2F}$$

sendo que M corresponde ao número de homens, F ao número de mulheres, $A_i Y$ à frequência relativa de homens portadores do alelo i , $A_i A_i$ a frequência relativa de mulheres homozigotas para o alelo i e $A_i A_j$ a somatória da frequência relativa de todos os genótipos femininos portadores do alelo “ i ” exceto quando $i = j$.

O número médio de alelos por *locus* em cada população foi obtido pelo programa GDA v1.1 (101), fazendo-se a somatória do número de alelos observados em cada um dos *loci* analisados em uma dada população, dividindo-se o valor encontrado pelo número total de *loci*.

Uma vez que os homens são hemizigotos para os STRs do cromossomo X, a heterozigose observada (H_o) foi calculada apenas na amostra feminina, pelo mesmo programa, pelo método de contagem direta, a partir da seguinte equação:

$$H_o = \sum_{j \neq i} A_i A_j$$

A heterozigose média esperada (He) em cada população foi o parâmetro utilizado para a estimativa da diversidade genética intra-populacional, conforme a equação 8.6 apresentada em Nei (102), utilizando-se o programa GDA v1.1:

$$He = \sum_{j=1}^r h_j / r$$

sendo “ r ” o número de *loci* utilizados e h_j , de acordo com a equação 8.1 de Nei (102), a heterozigose esperada para cada *locus* na j -ésima população, estimada pela equação:

$$h = 1 - \sum_{i=1}^m p_i^2$$

sendo m o número de alelos e p_i a frequência do i -ésimo alelo.

4.5.2 Aderência ao equilíbrio de Hardy-Weinberg

Em uma população panmítica, a proporção entre indivíduos homocigotos e heterocigotos para um determinado marcador bi ou multi-alélico é determinada pelas frequências alélicas, segundo os postulados de Hardy-Weinberg. A aderência ao equilíbrio de Hardy-Weinberg foi estimada apenas na amostra feminina, de acordo com teste exato descrito por Zaykin *et al.* (103), implementado no programa GDA v1.1, escolhido por se adequar bem à análise de marcadores com grande número de alelos (VNTRs e STRs), em situações em que a amostra utilizada e/ou algumas frequências alélicas são pequenas.

Este teste usa de procedimentos aleatórios para obter uma aproximação da probabilidade exata real. Os valores encontrados correspondem à probabilidade de se observar dados com distribuição tão ou mais extrema do que a encontrada, assumindo-se que a população esteja em equilíbrio. A hipótese nula no caso é de que a união dos gametas ocorra de maneira aleatória.

4.5.3 Poder de Discriminação

O Poder de Discriminação (PD) se refere à probabilidade inerente a um marcador genético de que dois indivíduos selecionados aleatoriamente em uma população apresentem genótipos distintos. Este é um dos parâmetros utilizados para avaliar a robustez de um marcador genético para ser utilizado em exames forenses na esfera criminal. Diferentemente dos autossômicos, nos marcadores do cromossomo X esta probabilidade difere entre os sexos, uma vez que o homem apresenta apenas um cromossomo X e a mulher dois. Este parâmetro foi estimado pelas fórmulas apresentadas a seguir, a primeira para mulher (PD_F) e a segunda para homem (PD_M) (68):

$$PD_F = 1 - 2\left(\sum_i f_i^2\right)^2 + \sum_i f_i^4$$

$$PD_M = 1 - \sum_i f_i^2$$

Sendo que f_i se refere à frequência do i -ésimo alelo do marcador analisado.

4.5.4. – Probabilidade Média de Exclusão

A Probabilidade Média de Exclusão (PME) se constitui em um dos parâmetros forenses utilizados em uma investigação de paternidade. Ela fornece a probabilidade média que um marcador genético possui em excluir um indivíduo não relacionado ao pai biológico, escolhido ao acaso na população. A PME foi introduzida para marcadores autossômicos em trios envolvendo mãe, filho e suposto pai (104), no entanto, este parâmetro não se aplica quando marcadores do cromossomo X são testados. Para este caso, foram descritas expressões para o cálculo da PME em situações envolvendo o trio mãe, filha e suposto pai (PME_T) ou quando apenas o duo filha e suposto pai é testado (PME_D). Esta última também se adéqua para a análise do duo mãe-filho. As fórmulas, de acordo com (68) são:

$$PME_T = 1 - \sum_i f_i^2 + \sum_i f_i^4 - \left(\sum_{i < j} f_i^2 \right)^2$$

$$PME_D = 1 - 2 \sum_i f_i^2 + \sum_i f_i^3$$

Sendo que f_i se refere à frequência do i -ésimo alelo do marcador analisado.

4.5.5 Comparação dos parâmetros Heterozigose, Poder de Discriminação e Probabilidade de Exclusão entre CrX STRs e AS STRs

Os três parâmetros acima relacionados analisados para o conjunto dos dez CrX STRs foram comparados com os dados obtidos para um conjunto de 18 STRs autossômicos (D16S539, D7S820, D13S317, D5S818, CSF1PO, TPOX, TH01, vWA, F13A01, FESFPS, F13B, LPL, Penta E, D18S51, D21S11, D3S1358, FGA e D8S1179) analisados na mesma população (105). Para cada conjunto de

marcadores foi obtida uma média para cada um dos parâmetros analisados. Para os CrX STRs usou-se o Poder de Discriminação Feminino, por serem as mulheres dizigóticas, assim como qualquer indivíduo o é para os AS STRs. A Probabilidade Média de Exclusão para ambos os tipos de marcadores levou em consideração o conjunto mãe, filha e suposto pai.

4.5.6 Estatísticas F e Diversidade Inter-populacional

O G_{ST} (coeficiente de coancestralidade) descrito por Wright (106), refere-se à correlação de genes de diferentes indivíduos em uma mesma população, sendo uma estimativa da diversidade genética inter-populacional. Sua estimativa foi feita pelo método de Weir (107), utilizando o programa GDA v1.1.

Esse programa emprega o procedimento de *bootstrap* para estimar a significância dos valores de G_{ST} obtidos. Para isso, realiza um total de 10.000 replicações sobre o banco de dados original, utilizadas no cálculo dos valores de G_{ST} e estabelece intervalos de confiança de 95% e 99%, obtendo-se valores de significância aos níveis de 0,05 e 0,01, respectivamente. Os valores de G_{ST} obtidos a partir do banco de dados original foram considerados estatisticamente relevantes se o intervalo de confiança não incluiu o valor “zero”.

A diversidade genética é uma medida de biodiversidade que mede a variação genética dentro de cada espécie, tanto intra-populacional quanto inter-populacional. A diversidade inter-populacional entre o Distrito Federal e os estados do Mato Grosso do Sul, Paraná, Rio de Janeiro e São Paulo foi estimada para as cinco populações em conjunto, obtendo um único valor de G_{ST} e para as populações duas a duas, em um total de dez pares, obtendo-se um valor de G_{ST} para cada par analisado.

4.5.7 Diversidade Alélica

A diversidade alélica foi um parâmetro estimado para as cinco populações brasileiras (Distrito Federal, Mato Grosso do Sul, Paraná, Rio de Janeiro e São Paulo). Com o emprego do aplicativo GDA e pelo método da contagem direta, foram estabelecidas a quantidades de alelos e de alelos privados por *locus* por população, bem como a quantidade total de alelos por *locus* contabilizada no conjunto das cinco populações analisadas. Por alelo privado entende-se neste trabalho como sendo um determinado alelo presente em uma população e ausente nas demais populações analisadas.

O número de alelos em uma amostra (*allelic richness* - riqueza alélica) é uma medida fundamental da diversidade genética. No entanto, quando são comparadas populações com diferentes tamanhos amostrais, esta medida não é fácil de usar, porque se espera que grandes amostras contenham mais alelos que pequenas amostras. Nas cinco populações analisadas, o tamanho amostral variou de 165 cromossomos na amostra do Rio de Janeiro a 625 cromossomos na amostra de Brasília. Para compensar esta diferença amostral, a contagem de alelos e de alelos privados por *locus* por população foi submetida a um processo estatístico denominado de “rarefação”, método este utilizado para estimar quantos alelos são esperadas em uma amostra de tamanho especificado, tomada a partir de um taxon. Por taxon, entende-se uma população ou um conjunto de populações que tenham sido colocados na mesma categoria. A estimativa foi realizada com o aplicativo HP-Rare (108), cujas rotinas se baseiam em procedimento desenvolvido para comparar o número de espécies em diferentes amostras ecológicas. Para tanto, arbitra-se um tamanho amostral não maior que a amostra de menor tamanho, número este que será usado pelo aplicativo para a realização da estimativa da diversidade alélica.

4.5.8 Equilíbrio de Ligação Genotípica

Os STRs analisados neste trabalho estão localizados, evidentemente, em um único cromossomo, o X. Dois STRs situam-se no braço curto e os demais no braço

longo (Figura 1). Por esta razão foram realizados testes para verificar eventuais divergências da hipótese nula (H_0) no que diz respeito ao equilíbrio de ligação: os marcadores analisados são independentes entre si, ou seja, os genótipos ocorrem de forma aleatória entre os marcadores, de forma que o genótipo observado em um *locus* não permite inferir qual será observado em outro *locus*. O estudo foi realizado com o auxílio do programa Arlequin ver. 3.0 (100) utilizando apenas a amostra de origem masculina, com o objetivo de verificar a associação preferencial de alelos entre pares de *loci*, com base em um teste exato de desequilíbrio de ligação. Os procedimentos são análogos ao teste exato de Fisher, empregando uma tabela de contingência 2 x 2. As estimativas dos valores exatos de p foram computadas pelo método da Cadeia de Markov, com 10000 dememorizações. Uma probabilidade exata menor que 0,05 indicou um desequilíbrio significativo.

Para uma melhor compreensão do significado do equilíbrio de ligação envolvendo marcadores STR, os resultados obtidos com os CrX STRs foram comparados com estudos levados a termo com 10 STRs do cromossomo Y (DYS456, DYS390, DYS389II, DYS458, DYS19, DYS439, DYS635, DYS392, DYS438 e DYS448) dentre os marcadores que compõem o sistema multiplex de amplificação Y-filer (*AmpFISTR® Yfiler® PCR Amplification Kit - Applied Biosystems, CA, USA*) e com 10 STRs autossômicos (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, vWA e TPOX), dentre os 13 incluídos no sistema CODIS (base de dados de DNA criminal dos Estados Unidos). O tamanho amostral para os três tipos de marcadores foi equivalente ($N = 204$, sendo N o número de haplótipos). Embora 205 homens tenham sido genotipados para os marcadores do cromossomo X, em uma amostra não foi possível amplificar o alelo em um *locus*, sendo esta amostra excluída dos cálculos. Para os marcadores autossômicos e do cromossomo Y foram selecionados 10 STRs com grau de variabilidade o mais próximo possível dos CrX STRs testados, elegendo-se como parâmetros a heterozigose observada e a diversidade alélica. Os 10 AS STRs escolhidos localizavam-se em 10 cromossomos diferentes, para simular a situação de independência completa. Os haplótipos dos AS STRs foram identificados com base na análise de 102 filhos (204 haplótipos) O estabelecimento destes dois critérios eram essenciais para a correta identificação dos haplótipos de origem materna e paterna para cada filho.

4.5.9 Estimativas de contribuição parental na composição atual da população

O Brasil é um país de dimensões continentais, com uma população multi-étnica e de composição variável entre suas regiões. Neste trabalho foram estudados marcadores genéticos polimórficos com frequências alélicas variáveis entre diferentes populações, permitindo inferir a contribuição de diferentes grupos ancestrais na composição de uma amostra populacional.

A estimativa da constituição ancestral da população do Distrito Federal foi feita com base na premissa de que ela é resultante fundamentalmente da mistura de três populações parentais, duas imigrantes e uma nativa: européia, africana subsaariana e ameríndia. Para as análises de mistura ancestral, o grupo europeu foi representado por duas amostras de Portugal (N=936), composta de indivíduos das regiões central e sul do país (11). Da África subsaariana foram usados dados de uma população da região de Karamoja (N=393), situada no Nordeste de Uganda, denominada Karimojong (109). Finalmente a parental ameríndia foi constituída pela reunião de amostras de 13 tribos, cinco delas habitando o norte da Argentina (Guarani Mbya n=25, Toba Formosa n=16, Toba do Chaco n=4, Wichi Chaco n=7 e Wichi Formosa n=4), uma o Paraguai (Pilaga n=14) e sete do Brasil dispersas nas regiões norte, centro-oeste, sudeste e sul do país (Kaingang n=17, Guarani n=11, Wai Wai n=8, Xavante n=7, Gavião n=8, Zoro n=6, Suruí n=4), (78). O tamanho amostral da população ameríndia foi 131 alelos.

Apenas seis CrX STRs puderam ser utilizados na estimativa da composição ancestral. O fator limitante foi a amostra ameríndia, para a qual haviam dados relativos a 11 microssatélites do cromossomo X mas apenas seis deles compartilhados com o sistema de amplificação X-Decaplex usado no presente trabalho (DXS9898, DXS7133, GATA31E08, DXS7423, DXS7132 e DXS6789). As frequências alélicas das populações parentais foram obtidas das publicações referidas. Para a população européia as frequências foram calculadas pela média ponderada das duas amostras, uma vez que o tamanho amostral diferia entre elas. As frequências alélicas de cada um dos STRs na população do Distrito Federal foram estimadas por contagem direta, utilizando-se o programa Arlequin ver. 3.0 (100). Foram estimadas também as proporções de contribuições parentais para as populações do Mato Grosso do Sul, Paraná, Rio de Janeiro e São Paulo, cujos

genótipos foram obtidos do material suplementar relacionado ao estudo colaborativo conduzido pelo Grupo Espanhol-Português da ISFG com 15 populações ibero-americanas (11), gentilmente disponibilizados pela autora mediante comunicação pessoal. A análise de mistura foi realizada com o aplicativo ADMIX95 (<http://www.genetica.fmed.edu.uy/software.htm> - Acesso em 08 abr. 2010), tendo como fundamento o método de Identidade Gênica entre as populações ancestrais e as miscigenadas. As proporções de mistura foram feitas com base na estimativa não ponderada.

A estimativa de mistura ancestral da população do Distrito Federal utilizando marcadores STRs do cromossomo X foi comparada com dados publicados para a mesma população mediante a análise de 14 marcadores autossômicos (APO, AT3, D1, DRD2-A, ECA, FXIIB, FyNull, GC, LPL, OCA2, PV92, Rb2300, Sb19.3 e TPA25) do tipo AIM (Marcador Informativo de Ancestralidade), assumindo igualmente a européia, africana e ameríndia como populações parentais (110).

4.5.10 Distâncias Genéticas e Dendrograma

Os padrões de variação genética dentro e entre populações contemporâneas podem ser explicadas por mutações e deriva genética ocorrida após a separação das mesmas ou então pela existência de fluxo gênico e podem ser abordadas a partir de cálculos de distância genética e sua representação em dendrogramas (111). Diversas metodologias de estimativas de distância genética estão disponíveis. A distância de Reynolds (112) baseia-se em coancestralidade. Ela estima a distância entre populações com base nos valores de F_{ST} , medida utilizada para quantificar os desvios das frequências genotípicas em relação ao esperado de acordo com os postulados de Hardy-Weinberg, tendo como origem a divisão da população em unidades, sob efeito de deriva genética (113). Ao contrário da distância de Nei, esta medida estima uma quantia proporcional ao período decorrido desde a separação das populações, tempo necessário para originar o nível de variância entre as frequências alélicas das populações x e y comparadas, e pode ser calculada pela seguinte expressão (113):

$$D_R^2 = \sum_i (x_i - y_i)^2 / 2(1 - \sum_i x_i y_i)$$

Dentre os métodos existentes para se obter árvores evolutivas os mais utilizados são UPGMA (*Unweighted Pair Group Method with Arithmetic Mean* – (114) e *Neighbor-joining* (115), inferidos a partir de uma matriz de distâncias.

Taxas evolutivas podem variar entre linhagens diferentes, o que é admitido pelo método *Neighbor-joining* e o torna mais confiável que o UPGMA quando se estuda populações naturais e isoladas. O método UPGMA tem como principal condição a ocorrência de taxas constantes de evolução, o que resulta em ramos com mesmos comprimentos a partir de um determinado nó. Tal método é mais aconselhável em análises de similaridade, em situações onde as populações comparadas não apresentam parentesco direto e padrão de ramificação. Considerando-se as populações aqui analisadas, o método UPGMA foi considerado mais adequado para a inferência de similaridade entre as populações.

O cálculo de distâncias genéticas (112) e a construção de dendrogramas UPGMA foram feitos utilizando-se quatro programas que fazem parte do pacote PHYLIP versão 3.5c: Seqboot, Gendist, Neighbor e Consense (116).

O programa Seqboot foi utilizado para construir um banco de dados contendo 1.000 matrizes distintas obtidas a partir do banco de dados original (contendo frequências alélicas), pelo procedimento de *bootstrap* (117). Com a utilização do programa Gendist, foram calculadas as distâncias genéticas entre as populações, para cada uma das 1.000 matrizes distintas (devido à grande quantidade de matrizes, estes dados não são apresentados). Para cada uma destas matrizes, foram construídos dendrogramas distintos pelo método UPGMA, com o uso do programa Neighbor. Em seguida, os 1.000 dendrogramas obtidos foram analisados pelo programa Consense, obtendo-se um dendrograma “ótimo”, com maior suporte estatístico. Por fim este dendrograma consenso foi visualizado utilizando-se o programa TreeView, versão Win32 (118).

Para que os dendrogramas apresentassem os comprimentos dos ramos proporcionais às distâncias que separam as populações, apenas os programas Gendist e Neighbor foram empregados. Posteriormente, os valores de *bootstrap* de cada ramo, obtidos pela análise mencionada no parágrafo acima, foram adicionados ao dendrograma.

4.5.11 Uso dos CrX STRs na resolução de casos reais de paternidade deficientes

Com a finalidade de testar o poder dos dez marcadores STRs do cromossomo X, os mesmos foram usados para integrar os cálculos estatísticos em quatro casos reais de investigação de paternidade nos quais o suposto pai era falecido e foi substituído por parentes consanguíneos. Os quatro casos foram previamente analisados com uma bateria de 21 ou 22 marcadores STRs autossômicos, porém os resultados obtidos deixaram dúvidas quanto à paternidade pretendida. Os procedimentos laboratoriais para a obtenção dos perfis genéticos das pessoas testadas seguiram os parâmetros previamente descritos. A valoração estatística foi realizada com o emprego de planilhas do aplicativo Excel do pacote Microsoft Office 2007 (*Microsoft Corporation*), usadas rotineiramente no IPDNA.

Em dois dos quatro casos o perfil genético do suposto pai foi estimado com base na genotipagem da suposta avó (mãe do suposto pai). Nestes dois casos, as fórmulas empregadas para a realização dos cálculos do Índice de Paternidade foram as mesmas utilizadas para as análises de paternidade com o trio mãe, filha e suposto pai, substituindo na fórmula o suposto pai pela suposta avó. Isto porque o homem herda da sua mãe um dos dois anéis do cromossomo X, com uma probabilidade de 0,5 para cada um e o transmite às suas descendentes do sexo feminino com a probabilidade de 1,0. Sendo assim, a transmissão do pai para a filha pode ser desconsiderada nos cálculos. Nos outros dois casos, o perfil alélico do suposto pai pode ser reconstituído pela análise de uma filha biológica e a respectiva mãe. Uma vez que o homem é hemizigoto para o cromossomo X, a probabilidade dele transmitir para um descendente do sexo feminino o único alelo de um determinado marcador é sempre 1,0. Nestes dois casos, os cálculos do Índice de Paternidade foram realizados conforme proposto por Ayres e Powley (119), modificadas por não ter sido levado em consideração sub-estruturação populacional.

As fórmulas empregadas para realização dos cálculos do Índice de Paternidade para os CrX STRs bem como os heredogramas dos quatro casos analisados encontram-se no Anexo I. A Probabilidade de Paternidade (PP) é uma decorrência direta do Índice de Paternidade (IP) e foi obtida pela fórmula:

$$PP = IP / IP + 1$$

Para que o resultado fosse apresentado em termos percentuais, o valor obtido foi multiplicado por 100. Para a realização dos cálculos foi usado o Banco de Dados de frequência alélica publicado por (11).

5 RESULTADOS E DISCUSSÃO

5.1 DISTRIBUIÇÃO DAS FREQUÊNCIAS ALÉLICAS

A distribuição das frequências alélicas obtidas para as 10 regiões genômicas estudadas na população do Distrito Federal estão apresentadas nas Tabela 5, Tabela 6 e Tabela 7. Devido ao modo de herança diferenciado entre os sexos, as frequências alélicas na amostra de mulheres (Tabela 6) foram calculadas separadamente da amostra de homens (Tabela 5). A comparação dos dados por teste de diferenciação alélica com o emprego do programa Arlequin não evidenciou diferenças estatisticamente relevantes entre as duas amostras em qualquer um dos *loci*, o que permitiu um cálculo único de frequência alélica (Tabela 7). Resultados semelhantes já foram observados por outros autores (11).

Todos os alelos observados nos dez *loci* analisados já haviam sido previamente encontrados em outras populações, de forma que nenhum alelo novo foi identificado. Dez alelos distribuídos em seis STRs apresentaram frequências acima de 30%, sendo o mais frequente o alelo 14 do DXS7423 (0,4224).

Com relação ao número de alelos observados na população do Distrito Federal o *locus* DXS6809 liderou, com 17 alelos, enquanto que nos *loci* DXS8378, DXS7133, GATA172D05, DXS7423 e DXS9902, oito alelos foram observados em cada um (Tabela 7). A quantidade de alelos por *locus* pode ser considerada alta, ao se contrastar com o observado em outras populações como Portugal (11), Coréia (69), Espanha (81) e Estados Unidos (120). Nessas populações foi observada sistematicamente uma menor quantidade de alelos por *locus*, sendo que as amostras apresentam tamanhos comparáveis com a amostra deste trabalho. Esta maior diversidade alélica na população do Distrito Federal deve refletir a origem da população brasileira, constituída por mistura de contingentes europeu, africano e ameríndio, adicionada ao fato da população do Distrito Federal ter sido moldada devido a fluxos migratórios de todo o país, possivelmente representando uma amostragem aproximadamente fiel do que seja a população brasileira do ponto de vista dos marcadores analisados.

Tabela 5 - Distribuição das frequências alélicas para os 10 CrX STRs na amostra masculina da população do Distrito Federal

Alelos	DXS8378	DXS9898	DXS7133	GATA31E08	GATA172D05	DSX7423	DXS7132	DXS9902	DXS6789	Alelos	Alelos	DXS6809
6					0,1716					6	28	0,0098
7	0,0049	0,0049	0,0049	0,0049	0,0147					7	29	0,0244
8			0,0049	0,0146	0,1324	0,0049		0,0049		8	30	0,0683
8.3		0,1902								8.3	30.1	
9	0,0244		0,3024	0,1902	0,1667			0,0244		9	31	0,1073
10	0,3024	0,0732	0,1854	0,0439	0,2108			0,0439		10	31.1	0,0098
11	0,4000	0,1220	0,4341	0,1463	0,2255		0,0146	0,3366		11	32	0,1415
12	0,2341	0,3171	0,0488	0,3317	0,0784	0,0049	0,0878	0,4146		12	32.1	0,0049
12.1								0,0293		12.1	33	0,3122
13	0,0293	0,2146	0,0195	0,1561		0,0780	0,2439	0,1366		13	33.1	0,0098
13.3		0,0098								13.3	34	0,2098
14	0,0049	0,0634		0,1024		0,3024	0,3512	0,0098	0,0146	14	35	0,0537
14.3										14.3	36	0,0293
15		0,0049		0,0098		0,4390	0,1854		0,0829	15	36.1	0,0049
15.3							0,0293			15.3	37	0,0098
16						0,1220	0,0634			16	37.1	0,0049
16.3							0,0146			16.3	38	0,0048
17						0,0488				17		
17.3							0,0049			17.3		
18									0,0146	18		
18.3							0,0049			18.3		
19									0,0634	19		
20									0,3122	20		
21									0,2537	21		
22									0,1366	22		
23									0,0390	23		
24									0,0098	24		

Tabela 6 - Distribuição das frequências alélicas para os 10 CrX STRs na amostra feminina da população do Distrito Federal

Alelos	DXS8378	DXS9898	DXS7133	GATA31E08	GATA172D05	DSX7423	DXS7132	DXS9902	DXS6789	Alelos	Alelos	DXS6809
6					0,1619					6	28	0,0071
7		0,0071		0,0048	0,0190					7	29	0,0143
8			0,0071	0,0167	0,1452	0,0048				8	30	0,0690
8.3		0,1381								8.3	30.1	0,0024
9	0,0071	0,0048	0,3262	0,1833	0,1310			0,0095		9	31	0,1405
10	0,3143	0,0571	0,1929	0,0619	0,2214			0,0429		10	31.1	0,0095
11	0,3643	0,1738	0,3738	0,1381	0,2333	0,0024	0,0143	0,3976		11	32	0,1667
12	0,2762	0,2905	0,0667	0,2500	0,0857	0,0024	0,0810	0,3381		12	32.1	0,0024
12.1								0,0119		12.1	33	0,2786
13	0,0357	0,2429	0,0190	0,2405	0,0024	0,0810	0,1762	0,1905	0,0024	13	33.1	
13.3		0,0095								13.3	34	0,1690
14		0,0690	0,0143	0,0976		0,3571	0,3690	0,0095		14	35	0,0976
14.3							0,0024			14.3	36	0,0286
15	0,0024	0,0071		0,0071		0,4143	0,2548		0,0833	15	36.1	
15.3							0,0095			15.3	37	0,0071
16						0,1048	0,0595		0,0786	16	37.1	
16.3							0,0143			16.3	38	0,0071
17						0,0333	0,0071		0,0048	17		
17.3							0,0071			17.3		
18									0,0048	18		
18.3							0,0048			18.3		
19									0,0429	19		
20									0,3262	20		
21									0,2619	21		
22									0,1405	22		
23									0,0500	23		
24									0,0048	24		

Tabela 7 - Distribuição das frequências alélicas combinadas de homens e mulheres para os 10 CrX STRs na população do Distrito Federal

Alelos	DXS8378	DXS9898	DXS7133	GATA31E08	GATA172D05	DSX7423	DXS7132	DXS9902	DXS6789	Alelos	Alelos	DXS6809
6					0,1651					6	28	0,0080
7	0,0016	0,0064	0,0016	0,0048	0,0176					7	29	0,0176
8			0,0064	0,0160	0,1410	0,0048		0,0016		8	30	0,0688
8.3		0,1552								8.3	30.1	0,0016
9	0,0128	0,0032	0,3184	0,1856	0,1426			0,0144		9	31	0,1296
10	0,3104	0,0624	0,1904	0,0560	0,2180			0,0432		10	31.1	0,0096
11	0,3760	0,1568	0,3936	0,1408	0,2308	0,0016	0,0144	0,3776		11	32	0,1584
12	0,2624	0,2992	0,0608	0,2768	0,0833	0,0032	0,0912	0,3632		12	32.1	0,0032
12.1								0,0176		12.1	33	0,2896
13	0,0336	0,2336	0,0192	0,2128	0,0016	0,0800	0,2368	0,1728	0,0016	13	33.1	0,0032
13.3		0,0096								13.3	34	0,1824
14	0,0016	0,0672	0,0096	0,0992		0,3392	0,3552	0,0096	0,0048	14	35	0,0832
14.3							0,0016			14.3	36	0,0288
15	0,0016	0,0064		0,0080		0,4224	0,1984		0,0832	15	36.1	0,0016
15.3							0,0160			15.3	37	0,0080
16						0,1104	0,0560		0,0752	16	37.1	0,0016
16.3							0,0144			16.3	38	0,0048
17						0,0384	0,0048		0,0048	17		
17.3							0,0064			17.3		
18									0,0080	18		
18.3							0,0048			18.3		
19									0,0496	19		
20									0,3216	20		
21									0,2592	21		
22									0,1392	22		
23									0,0464	23		
24									0,0064	24		
Número de alelos	8	10	8	9	8	8	12	8	12			17

No *locus* GATA172D05 não foi possível identificar o alelo em uma amostra de indivíduo do sexo masculino, apesar de terem sido realizadas cinco amplificações com DNA genômico obtido por dois métodos de extração diferentes (extração rápida com NaOH e orgânica fenol-clorofórmio). Este achado foi caracterizado como um provável alelo nulo devido a mutação na região de pareamento de um dos iniciadores. Como não se dispunha de parentes do doador para serem genotipados nem de iniciadores externos não marcados para sequenciamento, não foi possível confirmar tecnicamente a hipótese do alelo nulo. Não foram identificadas duplicações alélicas na amostra masculina ou padrão tri-alélico na amostra feminina.

A diversidade genética da população do Distrito Federal foi comparada com as das populações dos estados de Mato Grosso do Sul, Paraná, Rio de Janeiro e São Paulo. A Tabela 8 apresenta a quantidade de alelos observados por *locus* por população bem como a quantidade estimada após ser submetida a processo de rarefação. Em nove dos dez *loci* analisados, a população do Distrito Federal apresentou o maior número de alelos, sendo que em cinco *loci* uma ou duas das demais populações apresentaram número semelhante. A correção pelo processo de rarefação diminuiu a diferença nos valores estimados entre as populações. Na média, o Distrito Federal apresentou um número ligeiramente superior às demais.

Tabela 8 – Número de alelos por *locus* observados nas cinco populações e estimados após o processo de rarefação

	DF		MS		PR		RJ		SP	
	+	#	+	#	+	#	+	#	+	#
DXS8378	8	5,6	6	5,8	6	5,2	5	4,9	6	5,0
DXS9898	10	8,6	9	7,9	7	6,8	7	7,0	9	8,1
DXS7133	8	6,7	6	5,5	7	6,2	5	5,0	8	7,2
GATA31E08	9	8,3	9	8,2	7	6,9	9	8,8	9	8,6
GATA172D05	8	7,2	8	7,8	7	6,7	7	7,0	7	6,5
DXS7423	8	6,2	7	6,0	8	6,7	6	6,0	6	5,5
DXS6809	17	12,4	11	9,9	10	9,6	13	12,6	12	11,1
DXS7132	12	9,8	10	9,1	9	8,1	11	10,6	12	9,7
DXS9902	8	6,9	9	7,2	9	8,1	7	7,0	10	8,1
DXS6789	12	9,8	11	9,8	11	9,4	11	9,1	9	8,9
Média	10	8,1	8,6	7,7	8,1	7,4	8,1	8,0	8,8	7,9

+ - número de alelos observados

- número de alelos estimados após a rarefação

Os alelos privados observados em cada população encontram-se na Tabela 9. A população do Distrito Federal liderou a contagem, com onze dos vinte e um alelos observados, equivalendo a aproximadamente metade do total. A correção pelo método da rarefação continuou estimando uma quantidade maior de alelos privados para a população do Distrito Federal, agora perfazendo 30% do número total de alelos estimados (Tabela 10).

Embora seja a menor e mais recentemente constituída das cinco populações estudadas, no Distrito Federal foi observada uma maior quantidade de alelos em média por *locus* e uma maior quantidade de alelos privados, mesmo após a uniformização do tamanho amostral pelo método da rarefação. A razão para estes resultados deve ser encontrada na formação da população do Distrito Federal, resultante de fluxos migratórios de contingentes populacionais de todas as unidades da federação brasileira.

Tabela 9 – Alelos privados, frequência e população onde foram observados

Marcador	Alelo	Frequência	População
DXS8378	8	0,003	PR
DXS8378	7	0,002	DF
DXS8378	15	0,002	DF
DXS7423	18	0,006	PR
DXS7423	9	0,003	SP
DXS7423	11	0,002	DF
DXS6809	39,1	0,006	RJ
DXS6809	38,1	0,006	RJ
DXS6809	36,1	0,002	DF
DXS6809	37,1	0,002	DF
DXS6809	33,1	0,003	DF
DXS6809	30,1	0,002	DF
DXS6809	31,1	0,010	DF
DXS713	9	0,003	SP
DXS7132	14,3	0,002	DF
DXS7132	18,3	0,005	DF
DXS9902	16	0,007	MS
DXS9902	13,1	0,003	SP
DXS9902	11,1	0,003	SP
DXS6789	25	0,006	RJ
DXS6789	13	0,002	DF

Tabela 10 – Número de alelos privados observados por população e estimados após o processo de rarefação

População	Observados	Estimados após a rarefação
Distrito Federal	11	5,3
Mato Grosso do Sul	1	2,5
Paraná	2	2,2
Rio de Janeiro	3	3,9
São Paulo	4	3,5

5.2 PARÂMETROS POPULACIONAIS E FORENSES

A Tabela 11 apresenta os resultados observados para os parâmetros estudados: Heterozigose observada, Heterozigose esperada, Poder de Discriminação Feminino, Poder de Discriminação Masculino, Probabilidade Média de Exclusão para Trio, Probabilidade Média de Exclusão para Duo. A heterozigose observada em cada um dos *loci* analisados não diferiu significativamente da heterozigose esperada, de forma que não foi detectado excesso de homozigoto em nenhum dos dez marcadores.

O teste de aderência ao equilíbrio de Hardy-Weinberg realizado na amostra de origem feminina demonstrou que a proporção de genótipos observada não desviou significativamente das proporções esperadas exceto nos marcadores DXS7133 e DXS6789, para os quais foram obtidos desvios significativos ($p = 0,0481$ e $0,0172$ respectivamente). Não foram detectados alelos nulos na amostra masculina e não se verificou excesso de homozigoto na amostra feminina, em ambos os *loci*, que pudessem explicar estes resultados. Resultados semelhantes foram observados em populações de Portugal, Argentina e nos estados do Mato Grosso do Sul e Paraná (11). Aplicando a correção de Bonferroni, e assumindo dez testes para a população, o desequilíbrio deixa de ser observado.

O Poder de Discriminação foi calculado separadamente para a amostra feminina (PD_F) e para a masculina (PD_M), por *locus* individualmente (Tabela 11) e para o conjunto dos 10 *loci*. Como era de se esperar, os valores alcançados para a amostra feminina foram sistematicamente superiores aos obtidos para a amostra masculina. O PD_F variou de 0,844 para o *locus* DXS8378 a 0,950 para o *locus* DXS6809. Os extremos para o PD_M foram observados igualmente com estes dois

marcadores (0,692 e 0,828 respectivamente). Juntamente com o DXS7133, GATA172D05, DXS7423 e DXS9902, o DXS8378 apresentou a menor quantidade de alelos em comparação com os demais, no entanto, apresentou uma distribuição de frequências alélicas mais concentrada em poucos alelos, justificando daí o seu menor Poder de Discriminação em comparação com os outros quatro CrX STRs. Os maiores valores obtidos com o DXS6809 são explicados pelo elevado número de alelos, 17 ao todo e pela distribuição das frequências dos referidos alelos.

O PD_F e o PD_M de cada um dos *loci* foram comparados com os dados obtidos para 15 populações íbero-americanas, entre elas quatro brasileiras (11). Na maioria das comparações (75%) o PD_F e o PD_M obtidos na população em estudo mostraram valores mais elevados. Resultados semelhantes foram obtidos quando comparados a populações da Colômbia (121) e Coréia (69).

Com base nos valores individuais de cada marcador, foi calculado o Poder Médio de Discriminação do sistema como um todo para mulheres (PMD_F) e para homens (PMD_M). Os valores alcançados foram 0,99999999996 para o PMD_F e 0,9999995 para o PMD_M .

A Probabilidade Média de Exclusão por *locus* foi calculada considerando o trio mãe, filha e suposto pai (PME_T) e o duo filha e suposto-pai (PME_D). Os resultados encontram-se na Tabela 11. Os valores obtidos para o PME_T foram invariavelmente superiores aos obtidos para o PME_D . Isto porque, na ausência da mãe, o alelo paterno obrigatório não será identificado nos *loci* nos quais a filha é heterozigota, de forma que um homem possuindo qualquer dos dois alelos da filha não será excluído do rol dos possíveis pais. A Probabilidade Média de Exclusão guardou semelhança com o Poder de Discriminação, uma vez que ambos são parâmetros dependentes diretamente da distribuição alélica. Estes resultados foram igualmente comparados com os dados obtidos para 15 populações íbero-americanas (11), com resultados similares. A Probabilidade Média de Exclusão Cumulativa para Trio ($PMEC_T$) e para Duo ($PMEC_D$), considerando os *loci* agrupados, foi de 0,99999795 e 0,99988927, respectivamente.

Tabela 11 - Parâmetros estatísticos populacionais e forenses na população do Distrito Federal analisada com os 10 CrX STRs

	Ho	He	HW	PD_F	PD_M	PME_T	PME_D
DXS8378	0,6571	0,6927	0,6150	0,8444	0,6921	0,6313	0,4854
DXS9898	0,8238	0,7993	0,6447	0,9311	0,7986	0,7703	0,6450
DXS7133	0,7048	0,7038	0,0481	0,8595	0,7032	0,6508	0,5069
GATA31E08	0,8000	0,8112	0,7028	0,9378	0,8105	0,7842	0,6622
GATA172D05	0,7952	0,8252	0,7488	0,9451	0,8245	0,8004	0,6824
DSX7423	0,7143	0,6870	0,8628	0,8486	0,6864	0,6334	0,4892
DXS6809	0,8238	0,8286	0,4206	0,9499	0,8279	0,8074	0,6933
DXS7132	0,7000	0,7668	0,5475	0,9113	0,7662	0,7322	0,5992
DXS9902	0,7286	0,6937	0,8675	0,8503	0,6932	0,6376	0,4933
DXS6789	0,7857	0,7933	0,0172	0,9297	0,7927	0,7654	0,6399

Ho – Heterozigose observada

He – Heterozigose esperada

HW – Equilíbrio de Hardy-Weinberg

PD_F - Poder de Discriminação Feminino

PD_M - Poder de Discriminação Masculino

PME_T – Probabilidade Média de Exclusão para Trio

PME_D – Probabilidade Média de Exclusão para Duo

Tomando por base os parâmetros Heterozigose observada, Poder de Discriminação e Probabilidade Média de Exclusão, o STR mais informativo foi o DXS6809. Não por acaso este é o marcador com maior quantidade de alelos (17) sendo que o alelo mais observado (alelo 33) apresentou frequência abaixo de 30% (28,96%). De acordo com os mesmos parâmetros, o menos informativo foi o DXS8378. Neste STR foram encontrados apenas oito alelos, sendo que três deles apresentam frequências acima de 25% e respondem em conjunto por 94,88% dos alelos observados.

Na comparação dos parâmetros Heterozigose observada, Poder de Discriminação e Probabilidade Média de Exclusão entre o conjunto dos STRs do cromossomo X e o dos STRs autossômicos, as médias para a Heterozigose observada e para o Poder de Discriminação para os AS STRs alcançaram valores um pouco superiores aos obtidos para os CrX STRs (Tabela 12). Já a Probabilidade Média de Exclusão entre as duas classes de marcadores diferiu substancialmente, com os CrX STRs apresentando valores superiores aos AS STRs ($0,721 \pm 0,074$ para CrX STRs e $0,567 \pm 0,116$ para AS STRs). Uma vez que foram obtidos valores muito próximos entre as duas classes de marcadores para Heterozigose observada e para o Poder de Discriminação, pode-se inferir que as frequências alélicas também

eram semelhantes na média. Nestas condições, a Probabilidade Média de Exclusão para marcadores do cromossomo X é sempre maior que para marcadores autossômicos (122), em razão de o homem apresentar apenas um cromossomo X.

Tabela 12 - Média e desvio padrão da Heterozigose observada, Poder de Discriminação e Probabilidade Média de Exclusão obtidas com AS-STR e CrX STR.

	Heterozigose observada	Poder de Discriminação	Probabilidade Média de Exclusão²
	\bar{X}	\bar{X}	\bar{X}
AS STRS	0,780 ± 0,063	0,920 ± 0,038	0,567 ± 0,116
CrX STR	0,753 ± 0,059	0,901 ± 0,044	0,721 ± 0,074

¹Poder de Discriminação Feminino

²Probabilidade Média de Exclusão para Trio

A estimativa do Coeficiente de Endogamia (f) com base na amostra feminina foi de 0,01026. Para um Intervalo de Confiança de 95%, os valores variaram de -0,0015081 a 0,033932, sendo considerado portanto estatisticamente não significativo. Desta forma, admite-se um baixo Coeficiente de Endogamia para a população estudada, sugerindo a ocorrência de acasalamentos aleatórios. Estes resultados são consistentes com os apresentados por Godinho (110), obtidos a partir da análise de marcadores AIMS e STRs autossômicos na mesma população, sendo que a amostra dos STRs autossômicos foi retirada dos casos de paternidade do IPDNA da Polícia Civil de Brasília e, portanto, do mesmo estrato populacional da atual amostra de STRs do cromossomo X. Como a origem de ambas as amostras foi de uma instituição pública de perícia que realiza exames de paternidade para os indivíduos beneficiários da justiça gratuita, é provável que estes resultados digam respeito principalmente ao estrato social menos favorecido economicamente da população.

Considerando que se trata de uma única população e 10 marcadores, a análise de desequilíbrio de ligação entre os CrX STRs com base na comparação de todos os pares de *loci* possíveis realizou 45 testes (número de testes dado pela fórmula " $N!/2$ " onde N é o número de *loci*). Cinco dos 45 testes apresentaram valores de $p < 0,05$, sugerindo associação gamética (DXS8378/DXS7133, DXS7133/GATA172D05, GATA31E08/DXS6789, GATA172D05/DXS6809 e DXS7423/ DXS7132). Aplicando a correção de Bonferroni, com um valor de p

considerado de 0,0011 em razão da quantidade de testes realizados, nenhum desvio significativo foi mais observado, sugerindo independência entre os marcadores testados. Com base nestes resultados, as frequências haplotípicas em homem, considerando os dez marcadores do cromossomo X analisados, podem ser inferidas, na população estudada, pela multiplicação das frequências alélicas.

O resultado da análise de Desequilíbrio de Ligação tem implicação direta na valoração estatística em análises de vínculo genético de paternidade na área forense. Uma vez que as frequências haplotípicas podem ser inferidas pela multiplicação das frequências alélicas, quando o suposto pai pode ser testado ou seu genótipo reconstituído pela genotipagem de suas filhas biológicas, as frequências dos alelos identificados nos dez *loci* do cromossomo X testados podem integrar o cálculo do Índice de Paternidade. Face aos resultados obtidos para os parâmetros Heterozigose observada, Poder de Discriminação, Probabilidade de Exclusão, aderência ao equilíbrio de Hardy-Weinberg e Desequilíbrio de Ligação, pode-se recomendar o emprego das frequências alélicas observadas nos dez STRs compondo o sistema X-Decaplex, como base de dados para cálculos estatísticos no âmbito da genética forense, na população do Distrito Federal.

A análise de Desequilíbrio de Ligação com os AS STRs produziu resultados até certo ponto inesperados. Dos 45 testes realizados, em três observou-se um $p < 0,05$. Mesmo após a correção de Bonferroni, foi observado desequilíbrio de ligação significativo em um par de *loci* (Tabela 14). Os testes com o grupo de Y STRs apresentou resultado bastante diferente dos observados com os autossômicos e os STRs do cromossomo X. Dos 45 pares de *loci* testados, 35 apresentaram resultados significativos ($p < 0,05$). Mesmo após a correção de Bonferroni, este estado de desequilíbrio se manteve em 29 pares (Tabela 15).

Esses resultados devem ser interpretados com base no modo de herança, que é distinto entre os três tipos de marcadores analisados. Os STRs do cromossomo Y localizam-se evidentemente em um único cromossomo, apresentam herança uniparental paterna e não sofrem recombinação na meiose, esperando-se que sejam gerados haplótipos estáveis, cindidos apenas por eventos mutacionais. Conseqüentemente era de se esperar um forte desequilíbrio de ligação entre eles, corroborando dados da literatura (123, 124, dentre outros). Os STRs autossômicos apresentam herança bi-parental, segregam-se mediante recombinação na meiose e os aqui testados localizam-se em cromossomos distintos. Desta forma era de se

esperar um estado de completa independência entre os *loci* testados, podendo-se atribuir o desequilíbrio obtido a provável estratificação populacional. Os CrX STRs apresentam um modo de herança híbrido. Na transmissão materna ocorre recombinação e segregação, enquanto que na paterna não, em razão do homem possuir apenas um cromossomo X. Isto faz com que dois terços dos cromossomos X recombinem a cada geração. O haplótipo presente no homem é transmitido apenas às filhas e é considerado instável, uma vez que são desfeitos na progênie destas filhas. Desta forma, um haplótipo formado será quebrado no máximo duas gerações adiante, explicando assim a ausência de associação entre alelos de diferentes marcadores do cromossomo X.

Tabela 13- Teste exato para análise de desequilíbrio de ligação entre os marcadores CrX STR analisados aos pares

Marcadores		Valor de P	Desvio Padrão (\pm)
DXS8378	DXS9898	0,148	0,002
DXS8378	DXS7133	0,032*	0,001
DXS9898	DXS7133	0,090	0,002
DXS8378	GATA31E08	0,291	0,002
DXS9898	GATA31E08	0,500	0,002
DXS7133	GATA31E08	0,473	0,002
DXS8378	GATA172D05	0,932	0,002
DXS9898	GATA172D05	0,415	0,003
DXS7133	GATA172D05	0,003*	0,001
GATA31E08	GATA172D05	0,624	0,002
DXS8378	DSX7423	0,533	0,003
DXS9898	DSX7423	0,536	0,004
DXS7133	DSX7423	0,471	0,005
GATA31E08	DSX7423	0,458	0,003
GATA172D05	DSX7423	0,497	0,003
DXS8378	DXS6809	0,643	0,002
DXS9898	DXS6809	0,496	0,001
DXS7133	DXS6809	0,554	0,002
GATA31E08	DXS6809	0,838	0,002
GATA172D05	DXS6809	0,029*	0,001
DSX7423	DXS6809	1,000	0,000
DXS8378	DXS7132	0,838	0,002
DXS9898	DXS7132	0,982	0,002
DXS7133	DXS7132	0,160	0,002
GATA31E08	DXS7132	0,257	0,002
GATA172D05	DXS7132	0,664	0,002
DSX7423	DXS7132	0,303	0,002
DXS6809	DXS7132	0,207	0,002
DXS8378	DXS9902	0,463	0,002
DXS9898	DXS9902	0,256	0,002
DXS7133	DXS9902	0,706	0,003
GATA31E08	DXS9902	0,884	0,002
GATA172D05	DXS9902	0,458	0,004
DSX7423	DXS9902	0,702	0,002
DXS6809	DXS9902	0,261	0,001
DXS7132	DXS9902	0,918	0,001
DXS8378	DXS6789	0,327	0,002
DXS9898	DXS6789	0,473	0,002
DXS7133	DXS6789	0,502	0,002
GATA31E08	DXS6789	0,035*	0,001
GATA172D05	DXS6789	0,205	0,003
DSX7423	DXS6789	0,711	0,002
DXS6809	DXS6789	0,039*	0,001
DXS7132	DXS6789	0,307	0,002
DXS9902	DXS6789	0,542	0,003
Média		0,461	

* par de marcadores em Desequilíbrio de Ligação antes da correção de Bonferroni

Tabela 14 - Teste exato para análise de desequilíbrio de ligação entre os marcadores AS STR analisados aos pares

Marcadores		Valor de P	Desvio Padrão (\pm)
D8S1179	D21S11	0,540	0,002
D8S1179	D7S820	0,316	0,002
D21S11	D7S820	0,258	0,002
D8S1179	CSF1PO	0,741	0,002
D21S11	CSF1PO	0,532	0,003
D7S820	CSF1PO	0,425	0,002
D8S1179	D3S1358	0,193	0,001
D21S11	D3S1358	0,211	0,002
D7S820	D3S1358	0,763	0,002
CSF1PO	D3S1358	0,257	0,003
D8S1179	TH01	0,494	0,003
D21S11	TH01	0,127	0,002
D7S820	TH01	0,017 *	0,001
CSF1PO	TH01	0,676	0,004
D3S1358	TH01	0,753	0,003
D8S1179	D13S317	0,837	0,003
D21S11	D13S317	0,284	0,003
D7S820	D13S317	0,420	0,002
CSF1PO	D13S317	0,701	0,002
D3S1358	D13S317	0,017 *	0,001
TH01	D13S317	0,328	0,003
D8S1179	D16S539	0,418	0,003
D21S11	D16S539	0,383	0,003
D7S820	D16S539	0,266	0,002
CSF1PO	D16S539	0,122	0,002
D3S1358	D16S539	0,389	0,002
TH01	D16S539	0,922	0,002
D13S317	D16S539	0,748	0,003
D8S1179	vWA	0,739	0,003
D21S11	vWA	0,587	0,002
D7S820	vWA	0,908	0,001
CSF1PO	vWA	0,973	0,001
D3S1358	vWA	0,000 +	0,000
TH01	vWA	0,938	0,001
D13S317	vWA	0,559	0,003
D16S539	vWA	0,843	0,003
D8S1179	TPOX	1,000	0,000
D21S11	TPOX	0,988	0,001
D7S820	TPOX	0,707	0,002
CSF1PO	TPOX	0,198	0,002
D3S1358	TPOX	0,571	0,003
TH01	TPOX	0,874	0,003
D13S317	TPOX	0,764	0,002
D16S539	TPOX	0,987	0,001
vWA	TPOX	1,000	0,000
Média		0,551	

* par de marcadores em Desequilíbrio de Ligação

+ par de marcadores em Desequilíbrio de Ligação mesmo após correção a de Bonferroni

Tabela 15 - Teste exato para análise de desequilíbrio de ligação entre os marcadores Y STR analisados aos pares

Marcadores		Valor de P	Desvio Padrão (±)
DYS456	DYS390	0,195	0,002
DYS456	DYS389II	0,003 *	0,000
DYS390	DYS389II	0,000 +	0,000
DYS456	DYS458	0,501	0,002
DYS390	DYS458	0,000 +	0,000
DYS389II	DYS458	0,064	0,001
DYS456	DYS19	0,011 *	0,001
DYS390	DYS19	0,000 +	0,000
DYS389II	DYS19	0,060	0,001
DYS458	DYS19	0,016 *	0,001
DYS456	DYS439	0,000 +	0,000
DYS390	DYS439	0,206	0,004
DYS389II	DYS439	0,441	0,002
DYS458	DYS439	0,031 *	0,001
DYS19	DYS439	0,000 +	0,000
DYS456	DYS635	0,070	0,001
DYS390	DYS635	0,000 +	0,000
DYS389II	DYS635	0,000 +	0,000
DYS458	DYS635	0,000 +	0,000
DYS19	DYS635	0,000 +	0,000
DYS439	DYS635	0,000 +	0,000
DYS456	DYS392	0,062	0,002
DYS390	DYS392	0,000 +	0,000
DYS389II	DYS392	0,006 *	0,001
DYS458	DYS392	0,000 +	0,000
DYS19	DYS392	0,000 +	0,000
DYS439	DYS392	0,000 +	0,000
DYS635	DYS392	0,000 +	0,000
DYS456	DYS438	0,048 *	0,001
DYS390	DYS438	0,000 +	0,000
DYS389II	DYS438	0,001 +	0,000
DYS458	DYS438	0,000 +	0,000
DYS19	DYS438	0,000 +	0,000
DYS439	DYS438	0,000 +	0,000
DYS635	DYS438	0,000 +	0,000
DYS392	DYS438	0,000 +	0,000
DYS456	DYS448	0,059	0,002
DYS390	DYS448	0,000 +	0,000
DYS389II	DYS448	0,117	0,002
DYS458	DYS448	0,000 +	0,000
DYS19	DYS448	0,000 +	0,000
DYS439	DYS448	0,000 +	0,000
DYS635	DYS448	0,000 +	0,000
DYS392	DYS448	0,000 +	0,000
DYS438	DYS448	0,000 +	0,000
Média		0,042	

* par de marcadores em Desequilíbrio de Ligação

+ par de marcadores em Desequilíbrio de Ligação mesmo após a correção de Bonferroni

5.3 ANÁLISE DE MISTURA GENÉTICA E DIVERSIDADE INTER-POPULACIONAL

Os resultados das análises de mistura genética com base nos seis CrX STRs estudados na população do Distrito Federal bem como nas populações do Mato Grosso do Sul, Paraná, Rio de Janeiro e São Paulo encontram-se na Tabela 16.

Utilizando a abordagem não ponderada, na população do Distrito Federal a estimativa aponta para um predomínio da contribuição europeia (61,9%), com a parental ameríndia tendo a menor participação (11,1%). O grupo africano ocupou uma posição intermediária, com 27,9%. Quando comparada com as estimativas para as populações dos outros quatro estados brasileiros, a constituição étnica da população do Distrito Federal guardou semelhanças com as do Mato Grosso do Sul e de São Paulo.

Tabela 16 - Estimativa de contribuição genética ancestral na composição atual de cinco populações brasileiras com base na análise dos CrX STRs.

	Europeu	Africano	Ameríndio
Distrito Federal	0,619±0,060	0,270±0,082	0,111±0,040
Mato Grosso do Sul	0,649±0,085	0,206±0,117	0,145±0,057
Paraná	0,865±0,068	0,018±0,093	0,117±0,046
Rio de Janeiro	0,480±0,081	0,477±0,111	0,043±0,054
São Paulo	0,643±0,057	0,250±0,077	0,108±0,038

O r^2 alcançou um valor não muito elevado (0,734), podendo este resultado ser atribuído ao limitado número de marcadores analisados bem como a pouca diferença na distribuição das frequências alélicas entre as populações parentais, como ocorre em geral em marcadores do tipo STRs.

Na população do Paraná as estimativas demonstraram um amplo predomínio da contribuição europeia (86,5%), com um mínimo de presença africana (1,8%), enquanto que a ameríndia ocupou uma posição intermediária (11,7%). Na população do Rio de Janeiro as contribuições parentais europeia e africana foram equivalentes

(48,0% e 47,7% respectivamente). Nesta população o grupo ameríndio apresentou a menor contribuição na comparação com as demais populações, apenas 4,3%.

Adicionalmente as populações das cinco unidades da federação foram analisadas em conjunto, simulando uma única população. Nesta população brasileira hipotética a contribuição da parental ameríndia foi idêntica à do Distrito Federal (11,1%), a europeia foi ligeiramente superior (66,0% a 61,9%), e a africana inferior na mesma quantidade (23,0% a 27,0%). Na comparação com as populações do Mato Grosso do Sul e de São Paulo as contribuições de cada uma das parentais no conjunto das populações apresentou igualmente pequenas variações. Grandes diferenças nas contribuições foram observadas em relação às populações do Paraná e Rio de Janeiro, uma vez que no Paraná a contribuição africana foi negligível enquanto que no Rio de Janeiro sua participação se equiparou à europeia. Estes resultados parecem corroborar a história assim como são coerentes com dados do IBGE sobre autodeclaração em cor ou raça. Dessa forma, analisadas pelos STRs do cromossomo X, a população da região sul teria sido formada a partir de grandes contingentes europeus, com a contribuição africana e ameríndia menos relevante numericamente. Já o Rio de Janeiro, capital do Brasil por aproximadamente 200 anos, foi o porto por onde imigrou e lá se fixou grande parte dos africanos, enquanto que no Distrito Federal a formação da população resultou da mescla de indivíduos de todas as regiões do Brasil, devendo repetir em pequena escala o que seria esperado para a média da população brasileira. A proporção tri-híbrida observada na população de São Paulo parece refletir tanto o histórico do povoamento inicial como as migrações mais recentes, oriundas de todas as regiões do país, mercê da prosperidade econômica daquele estado. O Mato Grosso do Sul apresentou a maior contribuição de ameríndios na comparação com as demais populações, possivelmente por se tratar de um estado da região centro-oeste, assemelhando-se mais à região norte, onde a ancestralidade ameríndia é mais relevante que em outras regiões do país.

A análise comparativa de mistura ancestral da população do Distrito Federal estimada por marcadores do cromossomo X e marcadores autossômicos encontra-se na Tabela 17. As proporções observadas não diferiram significativamente quando estimadas por uma ou outra classe de marcador. No que se refere à parental europeia, a contribuição medida pelos CrX STRs e pelos marcadores autossômicos praticamente se equivaleram (61,9% e 61,4%, respectivamente). Para a africana e

ameríndia observou-se uma variação não superior a 4% entre as duas classes de marcadores.

Estudos com marcadores genéticos de herança uniparental (mitocondrial e cromossomo Y) têm corroborado os registros históricos de que o povo brasileiro foi constituído, principalmente, por ancestralidade masculina européia e ancestralidade feminina ameríndia e africana (91, 125, 126, 127) registrando-se grandes variações regionais. Em decorrência do modo de herança do cromossomo X, espera-se que seus marcadores genéticos sinalizem principalmente a história genética das mulheres, por contribuírem com 2/3 dos cromossomos X em uma determinada geração. Como corolário, era de se supor uma menor proporção de contribuição européia e uma maior contribuição africana e ameríndia na relação marcadores do cromossomo X/marcadores autossômicos na população do Distrito Federal, dada a proporção de cada uma das parentais na estimativa da constituição de sua população. No entanto, para os marcadores STRs do cromossomo X a estimativa da contribuição européia foi até ligeiramente superior (0,5 ponto percentual). Pelos dados aqui apresentados, na população do Distrito Federal os marcadores autossômicos tipo AIMs e os CrX STRs parecem contar a mesma história.

Tabela 17 – Estimativa de mistura ancestral para a população do Distrito Federal: comparação dos resultados obtidos com base em análises por CrX STRs e por AS STRs.

	Europeu	Africano	Ameríndio
STR do cromossomo X	61,9±0,060	27,0±0,082	11,1±0,040
Autossômicos tipo AIMs	61,4±0,008	22,9±0,008	15,7±0,007

A Figura 4 apresenta a árvore genética (dendrograma), com base na matriz de distância calculada pelo método de Reynolds para o conjunto das cinco populações brasileiras e das parentais européia, africana e ameríndia, analisadas com os STRs do cromossomo X. Aqui serão apresentadas análises referentes às relações da população do Distrito Federal com as demais populações brasileiras. Uma análise das relações entre as outras populações pode ser encontrada em Gusmão *et al.* (11).

O Distrito Federal mostrou maior afinidade com São Paulo. A origem da população da capital do Brasil, uma mescla de fluxos de todo o país embora com predomínio de algumas regiões ou estados e os vigorosos fluxos migratórios para o estado de São Paulo no século passado atraídos pela sua pujança econômica devem explicar esta proximidade genética. A população do Rio de Janeiro foi a que apresentou a maior distância em relação à do Distrito Federal, devendo refletir o fato do Rio de Janeiro, capital do Brasil desde 1763, ter recebido grandes levas de escravos oriundos da África pelos 100 anos seguintes, até a proibição do tráfico em meados do século 18. Isto fez com que a proporção da parental africana na constituição da população da antiga capital seja bem superior à da população do Distrito Federal (47,7% e 27,0%, respectivamente), quando estimada pelos STRs do cromossomo X.

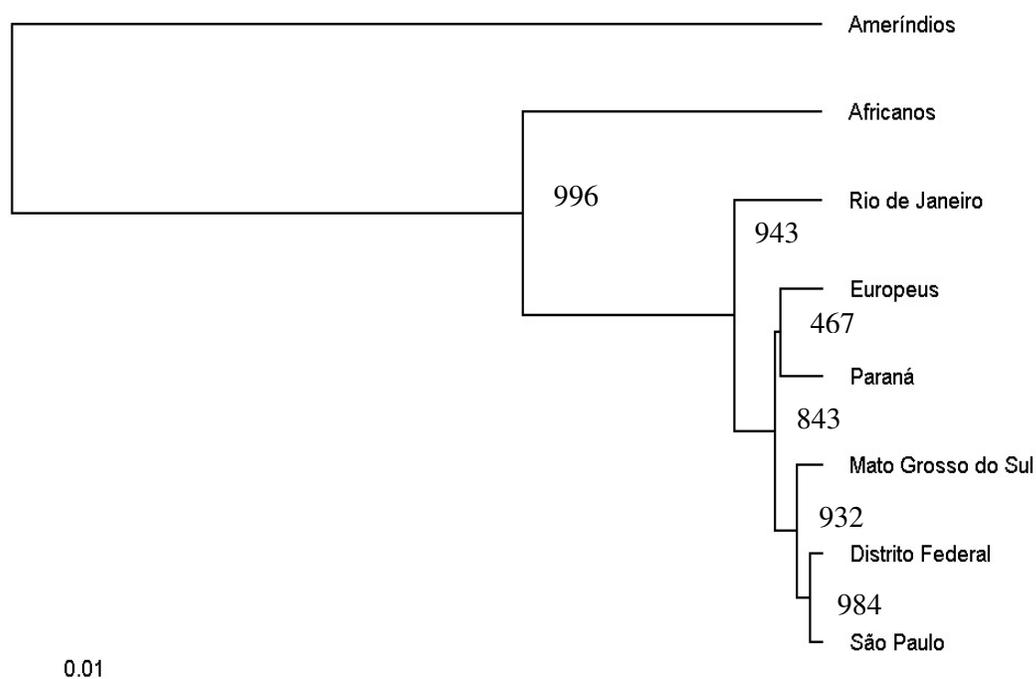


Figura 4 – Dendrograma do tipo UPGMA representativo do relacionamento genético das cinco populações brasileiras e das três parentais avaliadas com base nos dez CrX STRs, com os respectivos valores de *bootstrap*

Os valores de G_{ST} obtidos a partir da análise dos dez STRs do cromossomo X com o aplicativo GDA v1.1 nas amostras masculina e feminina

considerando a população do Distrito Federal e as populações dos outros quatro estados, analisados aos pares, estão apresentados na Tabela 18. Embora os valores de G_{ST} tenham sido baixos, foram observadas distâncias significativas entre o Distrito Federal e os estados do Mato Grosso do Sul, Paraná e Rio de Janeiro, considerando intervalos de confiança de 95% e 99%. Entre o Distrito Federal e São Paulo não foi observada distância significativa considerando os mesmos intervalos. Análise comparativa entre as populações dos outros quatro estados pode ser encontrada em Gusmão *et al.* (11).

Tabela 18 – Matriz de distâncias genéticas com base nos dez CrX STRs entre os pares de populações medidas pelos valores de G_{ST}

População	Valores de G_{ST}			
Distrito Federal				
Mato Grosso do Sul	0,001922			
Paraná	0,004481	0,001902		
Rio de Janeiro	0,003484	0,005532	0,011450	
São Paulo	-0,000292	0,000102	0,003640	0,002158

As distâncias genéticas entre as populações com base no valor de G_{ST} estimado para as cinco populações em conjunto (Distrito Federal, Mato Grosso do Sul, Paraná, Rio de Janeiro e São Paulo) foi de 0,002739. Para um Intervalo de Confiança de 99%, os valores variaram de 0,005219 a 0,000878, sendo considerado portanto estatisticamente significativo ($p < 0,01$). Estes dados demonstram que as populações analisadas diferem razoavelmente no seu conjunto.

5.4 IMPACTO DA INTRODUÇÃO DE DADOS DO X- DECAPLEX EM QUATRO CASOS DE PATERNIDADE DEFICIENTES

Os valores dos Índices de Paternidade (IP) e Probabilidades de Paternidade (PP) para os quatro casos analisados encontram-se na Tabela 19. Os

IPs obtidos com os STRs autossômicos foram multiplicados pelos IPs obtidos com os CrX STRs, gerando um IP conjunto.

Tabela 19 – Índices de Paternidade (IP) e Probabilidades de Paternidade (PP) obtidos com STRs autossômicos, STRs do cromossomo X e para o conjunto das duas classes de marcadores.

	AS STR		CrX STR		Conjunto	
	IP	PP	IP	PP	IP	PP
Caso 01	0,0054	0,54	2,11	67,85	0,01	1,127
Caso 02	10,01	90,92	2,11	67,85	21,12	95,48
Caso 03	55,83	98,24	164.063,94	99,9994	9.159.128,01	99,999990
Caso 04	54	98,18	1.525.584,28	99,99993	82.381.551,27	99,999990

PP apresentada em percentual

Nos quatro casos aqui descritos, a análise estatística com a bateria de AS STRs resultou em IPs bem abaixo de 10.000, valor considerado como o índice mínimo a ser tentado alcançar, conforme protocolo do IPDNA.

Nos casos 1 e 2, os IPs obtidos com os CrX STRs dobraram os índices obtidos com os marcadores autossômicos, mas não alteraram substancialmente os resultados no que tange à resolução dos casos. No caso 1 o resultado com marcadores autossômicos apontou para a não paternidade ($IP < 1$). Adicionando o resultado alcançado com os CrX STRs, o IP continuou abaixo de um. No caso 2, o IP elevou-se de 10 para 20 após agregar o resultado da análise com os CrX STRs. No entanto, o IP não atingiu valor necessário para a definição do caso. O impacto maior da análise com os STRs do cromossomo X ocorreu nos casos 3 e 4, nos quais o perfil alélico do suposto pai foi reconstituído pela genotipagem de uma filha biológica e a respectiva mãe. No caso 3, o IP obtido com autossômicos elevou-se de 55,83 para 9.159.128,01 pela incorporação dos dados dos CrX STRs, o que dá uma Probabilidade de Paternidade Conjunta de 99,999990%. No caso 4 a elevação foi de 54 para 82.381.551,27, com uma Probabilidade de Paternidade Conjunta de 99,999990.

Como explicar as diferenças entre os casos 1 e 2 e os casos 3 e 4 no que tange ao impacto da genotipagem pelos marcados CrX STRs? Nos casos 1 e 2, a suposta avó substituiu o suposto pai na aquisição da informação genética relativa a este último. Nesta situação, a regra da multiplicação dos IPs individuais só pode ser

aplicada para marcadores não ligados. Os dez CrX STRs considerados pertencem a três dos quatro grupos de ligação do cromossomo X (Szibor *et al.*, 2003). Desta forma, três dos dez marcadores testados foram incorporados aos cálculos estatísticos nestes casos, escolhidos entre os mais informativos em cada grupo de ligação (GATA31E08, DXS6809 e DXS9902). Vale salientar que não foram detectadas inconsistências nos outros sete *loci* analisados e não incorporados aos cálculos. Nos casos 3 e 4 o haplótipo do suposto pai pode ser reconstituído por genotipagem de uma filha biológica. Como já discutido anteriormente, não foi observado Desequilíbrio de Ligação entre os dez CrX STRs na população do Distrito Federal, nem nas populações que integram o Banco de Dados utilizado como referência para a realização dos cálculos estatísticos. Desta forma, podem-se prever as frequências haplotípicas com base nas frequências gênicas.

Estes resultados demonstram a validade do sistema multiplex de amplificação X-Decaplex composto de dez STRs do cromossomo X, em casos selecionados de paternidade com o suposto pai ausente, especialmente se seu perfil puder ser reconstituído pela genotipagem de filhas biológicas.

6. CONCLUSÃO

Com base nos dados analisados neste trabalho, pode-se concluir que:

1. O marcador CrX STR mais informativo foi o DXS6809 e o menos informativo, o DXS8378;
 2. Foi observado um provável alelo nulo no marcador GATA172D05 em uma amostra testada, evento que não pode ser confirmado por ausência de parentes do doador e de iniciadores externos não marcados para sequenciamento. Não foram observadas duplicações na amostra masculina ou padrão tri-alélico na feminina;
 3. Os CrX STRs utilizados neste trabalho demonstraram em geral alto Poder de Discriminação e elevada Probabilidade de Exclusão. A análise dos dois parâmetros para o conjunto dos marcadores estudados torna o sistema multiplex de amplificação X-Decaplex bastante adequado para análises de identificação humana;
 4. Os resultados dos parâmetros forenses analisados permitem o uso do Banco de Dados das frequências alélicas dos dez CrX STRs obtidas para a população do Distrito Federal como base de dados para cálculos em análises forenses em casos envolvendo indivíduos desta população;
 5. Com base no estudo dos CrX STRs, a população do Distrito Federal parece ser a que apresenta a maior diversidade genética, quando comparada com as populações dos estados do Mato Grosso do Sul, Paraná, Rio de Janeiro e São Paulo;
 6. As estimativas de distância genética pela análise dos CrX STRs demonstraram que a população do Distrito Federal difere significativamente das populações dos estados do Mato Grosso do Sul, Paraná e Rio de Janeiro;
-

-
7. A contribuição parental européia predominou na estimativa feita para a constituição da população do Distrito Federal, seguida da africana e, em menor proporção, da ameríndia;
 8. A análise de mistura na população do Distrito Federal com base nos CrX STRs considerando as parentais européia, africana e ameríndia demonstrou proporções semelhantes às apresentadas para marcadores autossômicos estimados na mesma população;
 9. Não foi observado Desequilíbrio de Ligação entre os marcadores genéticos na amostra de origem masculina, permitindo desta forma o emprego da regra da multiplicação no cálculo do Índice de Paternidade quando o haplótipo do suposto pai pode ser identificado, considerando para tal os dez marcadores CrX STRs;
 10. A taxa de desequilíbrio de ligação entre os CrX STRs se assemelha à dos AS STRs e difere do observado com os Y STRs;
 11. A incorporação dos CrX STRs na valoração estatística nos quatro casos deficientes de paternidade teve um grande impacto nos casos nos quais o perfil genético do suposto pai pode ser reconstituído pela genotipagem de uma filha, porém não alterou substancialmente o resultado quando o parente testado foi a suposta avó.
-

7 REFERÊNCIAS BIBLIOGRÁFICAS

1. Saferstein R. An introduction to forensic science. In: *Criminalistics*. 5ª edição. Practice Hall. 2001.
 2. Filho GH, Figini AL, Reis ABR. *et al.* Identificação humana. Porto Alegre: Editora Sagra Luzatto. 1ª edição. 1999.
 3. Jeffreys AJ, Brookfield JFY, Semeonoff R. Positive identification of an immigration test-case using human DNA fingerprints. *Nature*. 1985, 317:818–819.
 4. Mullis K, Faloona F, Scharf S. *et al.* Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* 1986, 51:263-273.
 5. Edwards A, Civitello A, Hammond HA. *et al.* DNA typing and genetic mapping at five trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* 1991, 49:746–756.
 6. França GV. Perícia Médico-Legal. In: *Medicina Legal*. 5ª edição, Rio de Janeiro: Guanabara Koogan. 1998-A.
 7. Gusmão L, Butler JM, Carracedo A. *et al.* DNA Commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic. Sci. Int.* 2006, 157:187–197.
 8. Sibille I, Duverneuilb C, De-La-Grandmaison GL. *et al.* Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa. *Forensic. Sci. Int.* 2002, 125:212–216.
 9. Szibor R, Krawczak M, Hering S. *et al.* Use of X-linked markers for forensic purposes. *Int. J. Legal Med.* 2003, 117:67–74.
 10. Trindade-Filho A, Ferreira STG, Oliveira SF. Impact of a Chromosome X STR Multiplex in deficiency paternity cases. *Manuscrito*.
 11. Gusmão L, Sánchez-Díaz P, Alves C. *et al.* A GEP-ISFG collaborative study on the optimization of an X-STR Decaplex: data on 15 Iberian and Latin American populations. *Int. J. Legal Med.* 2009, 123:227-234. doi:10.1007/s00414-008-0309-4.
-

-
12. Cavalli-Sforza LL, Menozzi P, Piazza A. The History and Geography of Human Genes. 1st ed. Abridged Paperback Edition. New Jersey: Princeton University Press. 1996.
 13. Myres JL. Geographical History in Greek Lands. Oxford: Claredon Press, 1953. *Apud* CAVALLI-SFORZA, L.L, MENOZZI, P, PIAZZA, A. The History and Geography of Human Genes. 1st ed. Abridged Paperback Edition. New Jersey: Princeton University Press. 1996.
 14. Rackham H. Books I e II In: Pliny: Natural History. Cambridge, Mass.: Harvard University. 1979. *Apud* CAVALLI-SFORZA, L.L, MENOZZI, P, PIAZZA, A. The History and Geography of Human Genes. 1st ed. Abridged Paperback Edition. New Jersey: Princeton University Press. 1996.
 15. Buckleton J, Triggs CM, Walsh SJ. Forensic X Chromosome Analysis. In: Forensic DNA Evidence Interpretation. CRC Press, Boca Raton FL, 2005.
 16. Colb S. King Solomon in The 21st Century. The U.S. Supreme Court. Disponível em: <<http://writ.news.findlaw.com/colb/20000628.html>>. Acesso em 25 mar. 2010.
 17. França GV. Antropologia Médico-Legal. In: Medicina Legal. 5^a edição, Rio de Janeiro: Guanabara Koogan. 1998-B.
 18. Alcântara HR. Exame para Investigação do Vínculo Genético. *In*: Perícia Médica Judicial. 2^a ed. Guanabara Koogan. Rio de Janeiro. 2006.
 19. Harper PS. Landmarks in Medical Genetics. Editora Oxford University Press. New York – USA. 2004.
 20. Goodwin W, Linacre A, Hadi S. An Introduction to Forensic Genetic. West Sussex, Inglaterra. Editora John Wiley & Sons Ltda. 2007.
 21. Chakraborty R, Kidd KK. The utility of DNA typing in forensic work. *Science*. 1991, 254:1735-1739.
 22. Wyman AR, White R. A highly polymorphic locus in human DNA. *Proceedings of the National Academy of Sciences*. 1980, 77:6754-6758.
 23. Tautz D, Trick M, Dover GA. Cryptic simplicity in DNA is a major source of genetic variation. *Nature*. 1986, 322:652-656.
 24. Estoup A, Cornuet JM. Microsatellite evolution: interferences from population data. In: *Microsatellites - Evolutions and Applications*. Oxford University Press. 1^a edição, 2001.
-

-
25. Ellegren H. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics*. 2000, 16:551-558.
 26. Nakamura Y, Leppert M, O`Connell P. *et al.* Variable number of tandem repeat markers for human gene mapping. *Science*. 1987, 235:1616-1622.
 27. Pena SDJ, Santos PC, Campos MCBN. *et al.* Paternity testing in Brazil by DNA fingerprint with the multilocus probes F10. *Ciência e Cultura*. 1994, 45:236-240.
 28. Ellegren H. Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*. 2004, 5:435-445.
 29. Subramanian S, Mishra RK, Singh L. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biology*. 4: R13, 2003. Disponível em: <<http://genomebiology.com/2003/4/2/R13>>. Acesso em: 15 abr 2010.
 30. Hagelberg E, Gray IC, Jeffreys AJ. Identification of the skeletal remains of a murder victim by DNA analysis. *Nature*. 1991, 352:427 – 429.
 31. Jeffreys A, Allen MJ, Hagelberg E. *et al.* Identification of the skeletal remains of Josef Mengele by DNA analysis. *Forensic. Sci. Int.* 1992, 56:65-76.
 32. Mullis K, Faloona F, Scharf, S. *et al.* Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* 1986, 51:263-73.
 33. Findlay I, Taylor A, Quirke P. *et al.* DNA fingerprinting from single cells. *Nature*. 1997, 389:555-556.
 34. Roewer L, Arnemann J, Spurr NK. *et al.* Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Human Genetics*. 1992, 89:389–394.
 35. Mahtani MM, Willard HF. A polymorphic X-linked tetranucleotide repeat locus displaying a high rate of new mutation: implications for mechanisms of mutation at short tandem repeat loci. *Hum. Mol. Genet.* 1993, 2:431–437.
 36. Tun Z, Honda K, Nakatome, M. *et al.* Simultaneous detection of multiple STR loci on sex chromosomes for forensic testing of sex and identity. *Journal of Forensic Science*. 1999, 44:772–777.
 37. Schaffner SF. The X Chromosome in Population Genetics. *Nature Reviews Genetics*, 2004, 5:43-51.
-

-
38. Edelmann J, Szibor R. Validation of the HumDXS6807 short tandem repeat polymorphism for forensic application. *Electrophoresis*. 1999, 20:2844–2846.
 39. Edelmann J, Hering S, Michael M. *et al.* 16 X-chromosome STR loci frequency data from a German population. *Forensic Sci. Int.* 2001, 124:215–218.
 40. Hering S, Kuhlisch E, Szibor R. Development of the X linked tetrameric microsatellite marker HumDXS6789 for forensic purposes. *Forensic. Sci. Int.* 2001, 119:42–46.
 41. Hering S, Brundirs N, Kuhlisch E. *et al.* DXS10011 - studies on structure, allele distribution in three populations and genetic linkage to further q-telomeric chromosome X markers. *Int. J. Legal Med.* 2004, 118:313–319.
 42. Zarrabeitia T, Amigo T, Sanudo C. *et al.* Sequence structure and population data of two X-linked markers: DXS7423 and DXS8377. *Int. J. Legal Med.* 2002, 116:368–371.
 43. Gomes I, Pereira R, Mayr WR. *et al.* Evaluation of DXS9902, DXS7132, DXS6809, DXS7133, and DXS7423 in humans and chimpanzees: sequence variation, repeat structure, and nomenclature. 10.1007/s00414-009-0357-4. *Int. J. Legal Med.* 2009, 10.1007/s00414-009-0357-4.
 44. Griffiths AJF, Miller JH, Suzuki DT. *et al.* *Introdução à Genética*. Rio de Janeiro. Editora Guanabara Koogan. 7^a edição. 2009.
 45. Fridolfsson AK, Cheng H, Copeland NG. *et al.* Evolution of the avian sex chromosomes from an ancestral pair of autosomes. *Proc. Natl. Acad. Sci. USA.* 1998, 95:8147–8152.
 46. Crow EW, Crow JF. 100 Years Ago: Walter Sutton and the Chromosome Theory of Heredity. *Genetics*. 2002, 160:1-4.
 47. Szibor R. X-chromosomal markers: Past, present and future. *Forensic. Sci. Int.: Genetics*. 2007, 1:93–99.
 48. Lyon MF. Gene action in the X-chromosome of the mouse (*Mus musculus* L.), *Nature*. 1961, 190:372–373.
 49. Rice WR. Evolution of the Y Sex Chromosome in Animals. *Bioscience*. 1996, 46:331-343.
 50. Lahn BT, Page DC. Four evolutionary strata on the human Y chromosome. *Science*. 1999, 286:964–967.
-

-
51. Petit C, Levilliers J, Weissenbach J. Physical mapping of the human pseudo-autosomal region; comparison with genetic linkage map. *Nature*. 1988, 328 (6127):273–275.
 52. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ. *et al*. The male-specific region of the human Y chromosome: a mosaic of discrete sequence classes. *Nature*. 2003, 423:825–837.
 53. Jobling MA, Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. *Nature Rev. Genet*. 2003, 4:598–612.
 54. Waterson RH, Lindblad-Toh K, Birney E. Initial sequencing and comparative analysis of the mouse genome. *Nature*. 2002, 420:520–562.
 55. Lander ES. Initial sequencing and analysis of the human genome. *Nature*. 2001, 409:860–921.
 56. Bohossian HB, Skaletsky H, Page DC. Unexpectedly similar rates of nucleotide substitution found in male and female hominids. *Nature*. 2000, 406:622–625.
 57. Li WH, Yi S, Makova K. Male-driven evolution. *Curr. Opin. Genet. Dev*. 2002, 12:650–656.
 58. Makova KD, Li WH. Strong male-driven evolution of DNA sequences in humans and apes. *Nature*. 2002, 416:624–626.
 59. Chang BH, Li WH. Estimating the intensity of male-driven evolution in rodents by using X-linked and Y-linked Ube 1 genes and pseudogenes. *J. Mol. Evol*. 1995, 40:70–77.
 60. Huang W, Chang BHJ, GU X. *et al*. Sex differences in mutation rate in higher primates estimated from AMG intron sequences. *J. Mol. Evol*. 1997, 44:463–465.
 61. Anagnostopoulos T, Green PM, Rowley G. *et al*. DNA variation in a 5-Mb region of the X chromosome and estimates of sex-specific/type-specific mutation rates. *Am. J. Hum. Genet*. 1999, 64:508–517.
 62. Lawson LJ, Hewitt GM. Comparison of substitution rates in ZFX and ZFY introns of sheep and goat related species supports the hypothesis of male-biased mutation rates. *J. Mol. Evol*. 2002, 54:54–61.
 63. Ebersberger I, Metzler D, Schwarz C. *et al*. Genomewide comparison of DNA sequences between humans and chimpanzees. *Am. J. Hum. Genet*. 2002, 70:1490–1497.
-

-
64. Kong A, Gudbjartsson DF, Sainz J. *et al.* A high-resolution recombination map of the human genome. *Nature Genet.* 2002, 31:241–247.
 65. Mann JD, Cahan A, Gelb AG. *et al.* A sex-linked blood group. *The Lancet.* 1962, 279 (7219):8-10.
 66. Hearne CM, Todd JA. Tetranucleotide repeat polymorphism at the HPRT locus, *Nucleic Acids Res.* 1991, 19:5450.
 67. Edwards A, Hammond HA, Jin L. *et al.* Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups, *Genomics.* 1992, 12:241–253.
 68. Desmarais D, Zhong Y, Chakraborty R. *et al.* Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA). *J. Forensic Sci.* 1998, 43:1046–1049.
 69. Shin SH, Yua JS, Park SH. *et al.* Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. *Forensic. Sci. Int.* 2005, 147:35–41.
 70. Ross MT, Grafham DV, Coffey AJ. *et al.* The DNA sequence of the human X chromosome. *Nature.* 2005, 434:325–337.
 71. King RC, Stansfield WD, Mulligan PK. *A Dictionary of Genetics.* Oxford University Press. 7^a edição, 2006.
 72. Szibor R, Hering S, Kuhlisch E. *et al.* Haplotyping of STR cluster DXS6801-DXS6809-DXS6789 on Xq21 provides a powerful tool for kinship testing, *Int. J. Legal Med.* 2005, 119:363–369.
 73. Nagaraja R, Macmillan S, Jones C. *et al.* Integrated YAC/STS physical and genetic map of 22.5 Mb of human Xq24–q26 at 56-kb inter-STS resolution. *Genomics.* 1998, 52:247–266.
 74. Hubert R, Macdonald M, Gusella J. *et al.* High resolution localization of recombination hot spots using sperm typing. *Nature Genet.* 1994, 7:420–424.
 75. May CA, Shone AC, Kalaydjieva L. *et al.* Crossover clustering and rapid decay of linkage disequilibrium in the Xp/Yp pseudoautosomal gene SHOX. *Nature Genet.* 2002, 31:272–275.
 76. Jeffreys AJ, Neumann R, Panayi M. *et al.* Human recombination hot spots hidden within regions of strong marker association. *Nature Genetics.* 2005, 37:601-606.
-

-
77. Zarrabeitia MT, Alonso A, Martin J. *et al.* Study of six X-linked tetranucleotide microsatellites: population data from five Spanish regions. *Int. J. Legal Med.* 2006, 120:147-150.
 78. Leite FPN, Santos SEB, Rodríguez EMR. *et al.* Linkage Disequilibrium Patterns and Genetic Structure of Amerindian and Non-Amerindian Brazilian Populations Revealed by Long-Range X-STR Markers. *Am. J. Phys. Anthropol.* 2009, 139:404–412.
 79. Edelmann J, Lessig R, Klitschar M. *et al.* Advantages of X-chromosomal microsatellites in deficiency paternity testing: presentation of cases. *International Congress Series.* 2004, 1261:257– 259.
 80. Lee H, Park MJ, Jeong CK. *et al.* Genetic 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.
 81. Zarrabeitia MT, Pinheiro F, Pancorbo MM. *et al.* Analysis of 10 X-linked tetranucleotide markers in mixed and isolated populations. *Forensic. Sci. Int.: Genetics* 2009, 3:63–66.
 82. Martins JA. Araraquara-SP, 2008. Dissertação [Mestrado em Biologia Molecular] – Universidade Estadual Paulista Julio de Mesquita Filho.
 83. Rodrigues EMR, Leite FPN, Hutz MH. *et. al.* 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.
 84. Asamura H, Sakai H, Kobayashi K. *et al.* MiniX-STR multiplex system population study in Japan and application to degraded DNA analysis. *Int. J. Legal Med.* 2006, 120:174–181.
 85. Bein G, Driller B, Schürmann M. *et al.* Pseudo-exclusion from paternity due to maternal uniparental disomy 16. *Int. J. Legal Med.* 1998,111:328–330.
 86. Crow JF. Visão Geral *In: A Avaliação do DNA como Prova Forense.* Editora FUNPEC, Ribeirão Preto, 2001.
 87. Owens K, King MC. Genomic Views of Human History. *Science.* 1999, 286:451-453.
 88. Neves W, Piló LB. Quando e como os Humanos chegaram à América. *In: O povo de Luzia.* São Paulo: Editora Globo. 2008.
-

-
89. Salzano FM, Freire-Maia N. Populações brasileiras, aspectos demográficos, genéticos e antropológicos. 1967. Companhia Editora Nacional, São Paulo, *Apud* ALVES-SILVA, J. *et al.* The Ancestry of Brazilian mtDNA Lineages. *Am. J. Hum. Genet.* 2000, 67:444–461.
 90. Salzano FM, Bortolini MC. The evolution and genetics of Latin American populations. Cambridge: Cambridge University Press. 2002. *Apud* LEITE, F.P.N. *et al.* Linkage Disequilibrium Patterns and Genetic Structure of Amerindian and Non-Amerindian Brazilian Populations Revealed by Long-Range X-STR Markers. *Am. J. Phys. Anthropol.* 2009, 139:404–412.
 91. Alves-Silva J, Santos MS, Guimarães PEM. The Ancestry of Brazilian mtDNA Lineages. *Am. J. Hum. Genet.* 2000, 67:444–461.
 92. Rodrigues EMR, Santos NPC, Santos AKCR. *et al.* Assessing Interethnic Admixture Using an X-Linked Insertion-Deletion Multiplex. *Am. J. of Human Biology.* 2009, 21:707–709.
 93. Salzano FM. Em busca das raízes. *Ciência Hoje.* 1986, 5:48-53.
 94. IBGE. Instituto Brasileiro de Geografia e Estatística. Contagem da População - 2007. Disponível em: <<http://www.ibge.gov.br/home/estatistica/populacao/contagem2007/contagem.pdf>>. Acesso em: 25 abr. 2010.
 95. Farias YS. Migrantes no Distrito Federal: Quem são? De onde vem? Por que vem? 2008. Disponível em: <<http://vsites.unb.br/ceam/webceam/nucleos/omni/observa/downloads/migrantes.PDF>>. Acesso em: 25 abr. 2010.
 96. Richards B, Skoletsky R, Shuber AP. *et al.* Multiplex PCR amplification from the CFTR gene using DNA prepared from buccal brushes/swabs. *Human Molecular Genetics.* 1993, 2:159-163.
 97. Edelmann J, Deichsel D, Hering S. *et al.* Sequence variation and allele nomenclature for the X-linked STRs DXS9895, DXS8378, DXS7132, DXS6800, DXS7133, GATA172D05, DXS7423 and DXS8377. *Forensic Sci. Int.* 2002, 129:99–103.
 98. Hering S, Szibor R. Development of the X-linked tetrameric microsatellite marker DXS9898 for forensic purposes. *J. Forensic Sci.* 2000, 45:929–931.
 99. Edelmann J, Deichsel D, Plate I. Validation of the X-chromosomal STR DXS6809. *Int. J. Legal Med.* 2003, 117:241–244.
-

-
100. Excoffier L, Laval G, Schneider S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online*. 2005, 1:47–50.
 101. Lewis PO, Zaykin D. GDA (Genetic Data Analysis) version 1.1. 2002. Disponível em <<http://lewis.eeb.uconn.edu/lewishome/software.html>>. Acesso em: 12 ago. 2010.
 102. Nei M. *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY, USA. 1987.
 103. Zaykin D, Zhivotovsky L, Weir BS. Exact tests for association between alleles at arbitrary numbers of loci. *Genética*. 1995, 96:169-178.
 104. Krüger J, Fuhrmann W, Lichte KH. *et al.* Zur Verwendung der sauren Erythrocytenphosphatase bei der Vaterschaftsbegutachtung. *Dtsch. Z. Gerichtl. Med.* 64: 127–146, 1968. *Apud* HERING, *et al.*, Development of the X linked tetrameric microsatellite marker HumDXS6789 for forensic purposes. *Forensic. Sci. Int.* 2001, 119:42–46.
 105. Dalton GC, Maia FAS, Oliveira SF. *et al.* Genetic profile of Federal District of Brazil based on 18 STR autosomal loci. *Forensic Sci. Int.: Genetics Supplement Series*. 2009, 2:373-375.
 106. Wright S. The genetical structure of populations. *Ann. Eugen.* 1951, 15:323-54.
 107. Weir B, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution*. 1984, 38:1358-1370.
 108. Kalinowski ST. HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes* 2005, 5:187-189.
 109. Gomes I, Pereira V, Gomes V. The Karimojong from Uganda: Genetic characterization using an X-STR decaplex system. *Forensic. Sci. Int.: Genetics*. 2009, 3:127–128.
 110. Godinho NMO. O Impacto das Migrações na Constituição Genética de Populações Latino-Americanas. Brasília, 2007. 92p, capítulo 4. Tese [Doutorado em Biologia Animal] – Universidade de Brasília.
 111. Relethford JH. Genetics of Modern Human Origins and Diversity. *Annual Review of Anthropology*. 1998, 27:1-23.
 112. Reynolds J, Weir BS, Cockerham CC. Estimation of the coancestry coefficient: basis for a short term genetic distance. *Genetics*. 1983, 105:767-779.
-

-
113. Meyer D. Árvores evolutivas humanas. In: Série Monografias. Vol 3. Sociedade Brasileira de Genética, Ribeirão Preto. 1995.
 114. Sokal RR, Michener CD. A statistical method for evaluating systematic relationship. Univ. Kansas. Sci. Bull. 1958, 38:1409-1438.
 115. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 1987, 4:406-425.
 116. Felsenstein J. Phylip (phylogeny inference package). Version 3.6. Department of Genetics, University of Washington, Seattle. 2004. Disponível em: <<http://evolution.genetics.washington.edu/phylip.html>>. Acesso em: 10 ago 2010.
 117. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 1985, 39:783-791.
 118. Page RDM. TreeView: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences. 1996, 12:357-358.
 119. Ayres KL, Powley WM. Calculating the exclusion probability and paternity index for X-chromosomal loci in the presence of substructure. Forensic Sci. Int. 2005, 149:201–203.
 120. Gomes I, Prinz M, Pereira R. *et al.* Genetic analysis of three US population groups using an X-chromosomal STR decaplex. Int. J. Legal Med. 2007, 121:198–203.
 121. Pico A, Castillo A, Vargas, C. *et al.* Genetic profile characterization and segregation analysis of 10 CrX STRs in a sample from Santander, Colombia. Int. J. Legal Med. 2008, 122:347–351.
 122. Krawczak M. Kinship testing with X-chromosomal markers: mathematical and statistical issues. Forensic Sci. Int. Genet. 2007, 1:111-114.
 123. Park MJ, Lee HY, Yoo JE. *et al.* Forensic evaluation and haplotypes of 19 Y-chromosomal STR loci in Koreans. Forensic. Sci. Int. 2005, 152:133-147.
 124. Soares-Vieira JA, Billerbeck AEC, Iwamura ESM. Population and mutation analysis of Y-STR loci in a sample from the city of São Paulo (Brazil). Genetics and Molecular Biology. 2008, 31:651-656.
 125. Batista-dos-Santos SE, Rodrigues JD, Ribeiro-dos-Santos AKC. *et al.* Differential contribution of indigenous men and women to the formation of an urban population in the Amazon region as revealed by mtDNA and Y- DNA. Am. J. Phys. Anthropol. 1999, 109:175–180.
-

126. Abe-Sandes K, Silva Jr. WA, Zago MA. Heterogeneity of the Y chromosome in Afro-Brazilian populations. *Human Biology*. 2004, 76:77-86.
127. Barcelos R, Ribeiro G, Silva Jr. WA. *et al.* Male contribution in the constitution of the Brazilian Centro-Oeste population estimated by Y-chromosome binary markers. *International Congress Series*. 2006, 1283:228-230.

EXTRAÇÃO DE DNA COM NaOH (SANGUE TOTAL)

DATA _____ Primeiro Responsável _____ Segundo Responsável _____

N	Nº Prot.	Nº Amostra	Identificação da amostra	Tipo de amostra	2º responsável
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					

MATERIAL NECESSÁRIO:

<u>TRIS HCL 1M (pH = 8,0)</u>	<u>NaOH 50 mM</u>
Pesar 47,29g de Tris-HCl e dissolver em 100 mL de água deionizada (ultra-pura) sob agitação. Completar o volume com água deionizada para 300 mL. Ajustar o pH até 8,0 e conservar em geladeira entre 2 – 8°C) (Validade: 6 meses);	Pesar 1g de NaOH e dissolver em 100 mL de água deionizada sob agitação e completar o volume com água deionizada para 500 mL. Conservar em temperatura ambiente 20 a 25°C. (Validade: 6 meses).

PROCEDIMENTOS:

- | | |
|--|---|
| 1. () Aliquotar 50µL de sangue total em microtubo estéril identificado; | 9. () Incubar a 95°C por 5 minutos; |
| 2. () Adicionar 1 mL de água estéril; | 10. () Agitar o tubo em alta rotação no vortex por 5-10 segundos; |
| 3. () Temp/tempo de incubação: TA/6 min (agitar a cada 2 min); | 11. () Neutralizar a solução com 60µL de Tris-HCl pH 8,0; |
| 4. () Centrifugar 1 min/12.000g; | 12. () Agitar o tubo em alta rotação no vortex por 5-10 segundos; |
| 5. () Descartar 950µL do sobrenadante; | 13. () DNA extraído; |
| 6. () Repetir os passos 2 a 5; | 14. () Realizar etapa de amplificação (utilizar a amostra à temperatura ambiente); |
| 7. () Adicionar 600µL de NaOH 50 mM; | 15. Conservar o DNA entre 0° e – 20 °C. |
| 8. () Agitar o tubo em alta rotação no vortex por 5-10 segundos; | |

LOCAL DE ARMAZENAMENTO DAS AMOSTRAS: _____

OBS: _____

PREPARAÇÃO DO MASTER MIX E AMPLIFICAÇÃO DO DNA

Data

Primeiro Responsável

Segundo Responsável

Planilha nº

X-DECAPLEX (10,0 µL)			
Componente do Mix	1 amostra	nº de amostras+1	Volume final
H ₂ O	5,0 µL		µL
Tampão Taq Gold	1,0 µL		µL
MgCl ₂ 25mM (Taq Gold)	1,0 µL		µL
dNTP	0,5 µL		µL
10x Primer mix	1,0 µL		µL
Taq Gold	0,5 µL		µL
Volume total	9,0 µL	distribuir 9,0µL do Mix por microtubo de 0,2mL	
DNA [1,0 ηg/µl]	1,0 µL		

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
Volume de reação:						Início:					
Programa nº:						Término:					
Termociclador nº:						Responsável final:					

OBSERVAÇÕES:

PREPARAÇÃO DE AMOSTRAS PARA APLICAÇÃO NO ABI 3130

Data _____

Primeiro Responsável _____

Segundo Responsável _____

ELETROFORESE

Formamida HiDi:	$(n+2) \times 8,7 \mu\text{L} =$
GS 500 Liz:	$(n+2) \times 0,3 \mu\text{L} =$
Distribuir 9 μL do mix (Formamida + Liz) em cada poço	
Produto de PCR:	1,0 μL (em amostras com pouco DNA este volume pode ser aumentado p/ até 5 μL)

A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23	A24
B13	B14	B15	B16	B17	B18	B19	B20	B21	B22	B23	B24
C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24
D13	D14	D15	D16	D17	D18	D19	D20	D21	D22	D23	D24
E13	E14	E15	E16	E17	E18	E19	E20	E21	E22	E23	E24
F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24
G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24
H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	H23	H24

Nº PLACA = _____**PERITO****PRÉ CORRIDA – CORRENTE MÁXIMA =** _____ **RESPONSÁVEL:****OBSERVAÇÕES:**

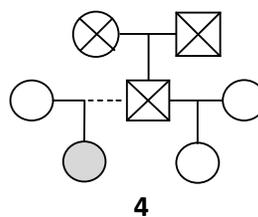
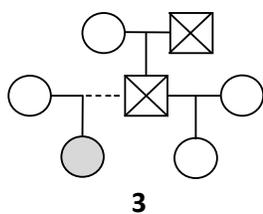
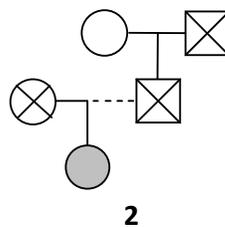
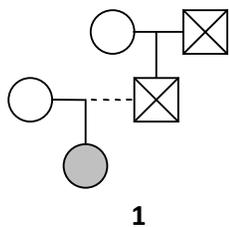
Fórmulas para cálculo do Índice de Paternidade (IP) com marcadores STRs do cromossomo X em análise de vínculo genético, com substituição do genótipo do suposto pai pelo da suposta avó, considerando valor de θ "0"

Mãe	Filha	Suposta avó	Fórmulas (IP)
AA	AA	AA	$1/a$
AA	AA	AB	$1/2a$
AA	AB	AB	$1/2b$
AB	AB	AB	$1/(a+b)$
AB	AA	AA	$1/a$
AB	AB	AA	$1/(a+b)$
AB	AA	AB	$1/2a$
AA	AB	BB	$1/b$
AB	BC	CC	$1/c$
AA	AC	BC	$1/2c$
AB	BC	AC ou CD	$1/2c$
AC	AC	BC	$1/[2(b+c)]$
AC	BC	BC	$1/2b$
AC	CC	BC	$1/2c$
ausente	AB	BC	$1/4b$
ausente	AB	BB	$1/2b$
ausente	AB	AB	$(a+b)/4ab$
ausente	BB	BC	$1/2b$
ausente	BB	BB	$1/b$

Fórmulas para cálculo do Índice de Paternidade (IP) com marcadores STRs do cromossomo X em análise de vínculo genético, com substituição do genótipo do suposto pai pelo de uma filha biológica, considerando valor de θ "0"

Mãe	Filha	Filha do suposto pai	Fórmula (IP)
AA	AA	A	$1/a$
AA	AB	B	$1/b$
AB	AA	A	$1/a$
AB	AB	A	$1/(a+b)$
AB	AC	C	$1/c$
ausente	AA	A	$1/a$
ausente	AB	A	$1/2a$

Heredogramas dos quatro casos de paternidade deficientes analisados



ANEXO 2 – Artigo Submetido para publicação

Elsevier Editorial System(tm) for Forensic Science International: Genetics

Manuscript Draft

Manuscript Number:

Title: Impact of a Chromosome X STR Decaplex in deficiency paternity cases

Article Type: Case Report

Keywords: CrX STR Decaplex, deficiency paternity cases, paternity Index

Corresponding Author: Dr. Silviene Fabiana Oliveira, PhD

Corresponding Author's Institution: Universidade de Brasília

First Author: Aluisio Trindade-Filho, Medical Doctor

Order of Authors: Aluisio Trindade-Filho, Medical Doctor; Samuel T Ferreira, Medical Doctor; Silviene Fabiana Oliveira, PhD

Abstract: Deficiency paternity cases characterized by the absence of the alleged father are a challenge for forensic genetics. We present four cases in which a battery of 21 - 22 autosomal STRs produced results within a range of doubt, by typing the relatives of the alleged father. An additional set of 10 STRs located on chromosome X substantially moved the results toward the alleged fatherhood in half of the cases.

Impact of a Chromosome X STR Decaplex in deficiency paternity cases

Trindade-Filho A.^{1 3} Ferreira STG.² Oliveira SF.^{3 4}

¹ Laboratório de Antropologia Forense, Instituto de Medicina Legal, Polícia Civil do Distrito Federal, Brasília (DF), Brazil.

² Instituto de Pesquisa de DNA Forense, Polícia Civil do Distrito Federal, Brasília (DF), Brazil.

³ Programa de Pós-Graduação em Ciências da Saúde, Faculdade de Ciências da Saúde, Universidade de Brasília, Brasília (DF), Brazil.

⁴ Laboratório de Genética, Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília (DF), Brazil.

Keywords: CrX STR Decaplex, deficiency paternity cases. Paternity Index

Running head: Chr X STR Multiplex in deficiency paternity cases

ADDRESS FOR CORRESPONDENCE:

Aluisio Trindade Filho

Instituto de Medicina Legal - SAISO – Complexo da Polícia Civil - Bloco B - Brasília, DF, Brazil

70.670-000

Phone: +55 61 3362-5692.

e-mail: aluisiotf@yahoo.com.br

Abstract

Deficiency paternity cases characterized by the absence of the alleged father are a challenge for forensic genetics. We present four cases in which a battery of 21 – 22 autosomal STRs produced results within a range of doubt, by typing the relatives of the alleged father. An additional set of 10 STRs located on chromosome X substantially moved the results toward the alleged fatherhood in half of the cases.

Introduction

In the Civil Police laboratory of the Federal District (capital of Brazil) about 11% of civil paternity investigation cases are characterized as involving “deficiency”, i.e. the alleged father is absent. In such cases, close relatives make up for this absence by providing material for analysis, with the disadvantage that the Combined Paternity Index (PI) obtained is invariably lower than the index obtained if the alleged father is available for analysis. In some cases, depending on the relatives available and the combination of genotypes, the exclusive use of autosomal STR markers (AS STR) may be insufficient to obtain PI values above an acceptable limit.

The use of STRs located on chromosome X (CrX STR) in forensic genetics has a very recent history [1; 2; 3]. With respect to CrX STRs, a woman is dizygous and a man is hemizygous, with the latter receiving his single X from his mother. This particular form of inheritance allows these markers to be particularly useful in deficiency paternity cases. Moreover, CrX STRs can be used in half the cases of paternity dispute as there is no allele inherited by descent in the father-son relationship. They are located on a single chromosome, and investigators should take proper precautions when using these markers for this purpose because of genetic linkage and possible linkage disequilibrium between them [4]. We present here the use of this class of markers in four paternity investigation cases that showed inconclusive results with a battery of autosomal STRs.

Methods

Four deficiency paternity cases with a female child and deceased alleged father were analyzed with the X-Decaplex system, composed of 10 CrX STRs - DXS8378, DXS9898, DXS7133, GATA172D05, GATA31E08, DXS7423, DXS6809, DXS7132, DXS9902, and DXS6789 [5]. These STRs are located in three of the four linkage groups described by Szibor for X chromosome [1]. Pedigrees can be found in

Figure 1. DNA was obtained from blood samples and/or oral swab employing the rapid NaOH extraction method [6]. One ng of genomic DNA was amplified in a 10 ul total volume of PCR reaction mix in the GeneAmp 9600 thermocycler (Perkin-Elmer). The products were loaded on a 3130xl Genetic Analyzer (*Applied Biosystems*, CA, USA). Data were collected by DataCollection v3.0 software and alleles were assigned using GeneMapper v3.2. The statistical parameters analyzed were the Combined Paternity Index (PI) and Probability of Paternity (PP), the later assuming a probability a priori of 0.5. The cases had been previously analyzed for 21 or 22 autosomal markers with the use of amplification systems AmpFLSTR Identifiler® (*Applied Biosystems*, CA, USA), Powerplex® 16 System, GenePrint FFFL Multiplex kit and PowerPlex® ES Monoplex System SE33 (Promega Corp., Madison, WI, USA) and failed to show conclusive results.

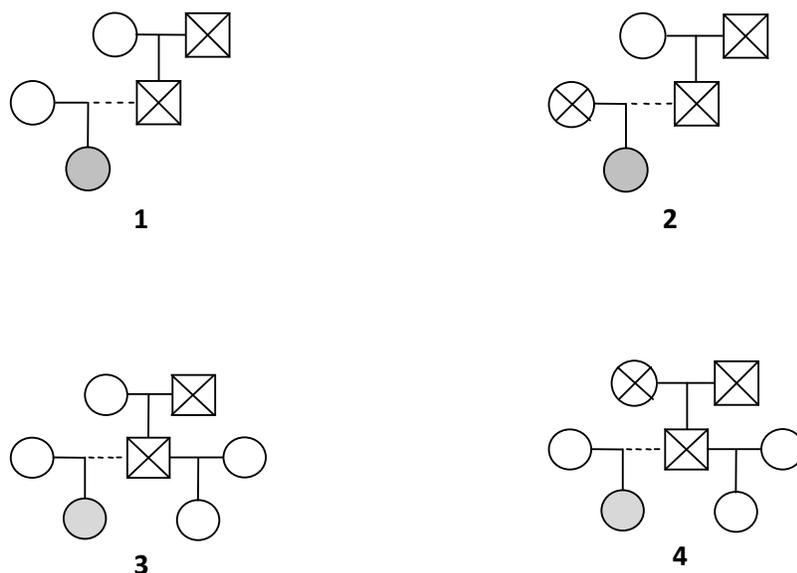


Figure 01 – Pedigrees of the four deficiency paternity cases.

When the alleged father's only available relative was the supposed grandmother (Case 1 and 2), the forensic parameters were estimated using three CrX STRs of the 10 tested: GATA31E08, DXS6809 and DXS9902. The selection criterion was the informativeness of the marker representing each group. In cases 3 and 4 the haplotype of the alleged father could be identified by genotyping biological daughters and their respective mother. The 10 CrX STRs were used in statistical evaluation in both cases. In case 03 the genotype of the alleged paternal

grandmother was used only to identify the paternal allele in a locus where the biological mother and her daughter shared the same heterozygotic genotype.

Results

According to the internal protocol established by our laboratory, after exhausting all available resources in terms of genetic markers, PIs below 10,000 are reported to the justice system as being under the desired threshold. In the four cases described here, statistical analysis with the battery of AS STRs resulted in PIs far above the 10,000 cutoff (Table 1). By typing the four cases with the X-Decaplex system, no inconsistencies were detected.

In cases 1 and 2, the PI obtained with CrX STRs doubled the index obtained with AS markers but did not substantially change the results. In case 1 the result with AS markers pointed to non-paternity (PI <1). Adding the result achieved with CrX STRs, the PI still indicated non-paternity. In case 2 the PI with AS markers moved from 10 to 20 after X-marker typing but, for practical purposes, remained unchanged. Alleles typed in the other seven loci, which are not included in the statistics, were compatible with paternity in both cases. In case 3 the PI with 21 AS markers reached 7,000. In order to raise the PI to 10,000, samples were amplified for AS marker SE33. However, an inconsistency in a repeat unit was found between the putative father's rebuilt genotype and the tested child's genotype. The inconsistency was interpreted as a mutation, knocking the PI to 56. By incorporating additional information from XDecaplex it was possible to achieve a final PI close to 9,000,000, consistent with paternity. The most dramatic change occurred in case 4. A PI of less than 100 achieved by genotyping 21 AS STRs rose to a final PI close to 100,000,000 after joining the PI of the ten CrX markers.

Table 01 – PI and PP for AS STRs, CrX STRs and AA/CrX STRs.

	AS STR		CrX STR		Total	
	PI	PP	PI	PP	PI	PP
Case 01	0.0054	0.54	2.11	67.85	0.01	1.127
Case 02	10.01	90.92	2.11	67.85	21.12	95.48
Case 03	55.83	98.24	164,063.94	99.9994	9,159,128.01	99.999990
Case	54	98.18	1,525,584.28	99.99993	82,381,551.27	99.9999990

Discussion

Due to the mode of inheritance, STRs located on chromosome X are well suited to cases of deficiency paternity. Since inheritance of the X chromosome from mother to child occurs by segregation, if the alleged paternal grandmother replaces the absent alleged father, the use of product rule is restricted to unlinked markers. For this reason, in cases 1 and 2 only three CrX STRs chosen from among the most informative in each linkage group were used for statistical evaluation. It is noteworthy that in case 2 the mother of the tested child was not available, contributing to a PI of only 2 for CrX STRs. In any case, genotyping the mother of the tested child would not substantially change the result since the PI for the AS STRs reached only 10.

Furthermore, the profile of the alleged father can be reconstructed by genotyping his daughters and their respective mothers. In doing so, the product rule depends on the absence of linkage disequilibrium between the loci tested in the population of interest. In a joint study carried out by the Spanish and Portuguese ISFG Working Group (ISFG), no linkage disequilibrium was observed between the loci of the X-Decaplex in 15 Ibero-American populations, four of them from Brazil [5]. Additionally, linkage disequilibrium was not observed in a study conducted by the authors in the population that includes the tested individuals (unpublished data). So, in cases 3 and 4 all CrX markers from the X-Decaplex amplification system were considered in the calculation of PI.

These results are in agreement with the literature and demonstrate the usefulness of CrX STRs in selected cases of deficiency paternity [4; 7], especially if biological daughters of the alleged father and their mothers are available to help reconstruct his haplotype. If the profile of the alleged father can be reconstructed, it should be emphasized that the haplotype from CrX markers is more informative than the genotype from AS markers [1], considering the same allele frequencies and degree of polymorphism among loci tested. Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

- [1] Szibor R, Krawczak, M, Hering S, et al. 2003. Use of X-linked markers for forensic purposes. *Int. J. Legal Med.* 117: 67–74.
 - [2] Szibor R, X-chromosomal markers: past, present and future. 2007. *Forensic Sci. Int. Genet.* 1: 93–99.
 - [3] Shin SH, Yua JS, Park SH, et al. 2005. Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. *Forensic Sci. Int.* 147: 35–41.
 - [4] Edelmann J, Lessiga R, Klintscharb M, Szibor R. Advantages of Xchromosomal microsatellites in deficiency paternity testing: presentation of cases. 2004. *International Congress Series.* 1261: 257– 259.
 - [5] Gusmão L, Sánchez-Diaz P, Alves C, et al. A GEP-ISFG collaborative study on the optimization of an CRX STR Decaplex: data on 15 Iberian and Latin American populations. 2008. *Int. J. Legal Med.* doi:10.1007/s00414-008-0309-4.
 - [6] Richards B, Skoletsky R, Shuber AP. et al. 1993. Multiplex PCR amplification from the CFTR gene using DNA prepared from buccal brushes/swabs. *Human Molecular Genetics.* 2:159-163.
 - [7] Aquino J, Peixe C, Silva D, et al. A X-chromosome STR hexaplex as a powerful tool in deficiency paternity cases. 2009. *Forensic Science International: Genetics Supplement Series.* 2: 45-46.
-