

**UNIVERSIDADE ESTADUAL PAULISTA JULIO DE MESQUITA FILHO**

**Faculdade de Ciências Farmacêuticas**

**JOYCE APARECIDA MARTINS LOPES FERRAZ**

**ESTUDO DE POLIMORFISMOS DO CROMOSSOMO X NA  
POPULAÇÃO DA REGIÃO SUDESTE DO BRASIL**

**ARARAQUARA-SP**

**2011**

**UNIVERSIDADE ESTADUAL PAULISTA JULIO DE MESQUITA FILHO**

**Faculdade de Ciências Farmacêuticas**

**JOYCE APARECIDA MARTINS LOPES FERRAZ**

**ESTUDO DE POLIMORFISMOS DO CROMOSSOMO X NA  
POPULAÇÃO DA REGIÃO SUDESTE DO BRASIL**

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

Orientadora: Profa. Dra. Regina Maria Barretto Cicarelli

Co-orientador: Prof. Dr. Ángel Carracedo Alvarez

**ARARAQUARA-SP**

**2011**

### Ficha Catalográfica

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

F381e Ferraz, Joyce Aparecida Martins Lopes  
Estudo de polimorfismo do cromossomo X na população da região  
sudeste do Brasil / Joyce Aparecida Martins Lopes Ferraz. – Araraquara,  
2011  
109 f.

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

1. X-STR. 2. Cromossomo X. 3. Identificação humana. 4. Sudeste do  
Brasil. 5. Aplicação forense. 6. DNA. I. Cicarelli, Regina Maria Barretto,  
orient. II. Carracedo Alvarez, Ángel, coorient.. III. Título.

CAPES: 40300005

**DEDICATÓRIA**

Aos

Meus pais Gilberto Martins e Zulmira Furlan Fraga Martins pelos ensinamentos que me tornaram uma pessoa diferenciada, pela dedicação ao meu crescimento pessoal e profissional, pela família maravilhosa e única que somos.

Meus Irmãos Jaqueline Martins, Gilberto Martins Filho e Juliana Lara Martins pelos laços de amor e carinho intenso que nos une, pelas palavras verdadeiras e o apoio incondicional.

Ao meu marido José Roberto Lopes Ferraz Filho por estar em minha vida há anos, tornando-a muito mais prazerosa e tranqüila, por acompanhar meu crescimento e me apoiar em minhas escolhas, por ser um marido e pai exemplar. Te amo amor!

Ao meu filho João Pedro Martins Lopes Ferraz, que em tão pouco tempo trouxe grandes mudanças em minha vida, fazendo-me refletir sobre diversos aspectos que me tornam uma pessoa melhor! Você é único amor!

Ao meu primeiro afilhado Miguel Martins Priuli, cujo apelido é “Alegria” e já mostra o quanto está sendo gostoso poder vivenciar seu crescimento nesses cinco meses.

## **AGRADECIMENTOS**

À Professora Regina Maria Barretto Cicarelli pela orientação, confiança e disposição em trabalhar com pesquisa. Por ter contribuído efetivamente no meu crescimento acadêmico profissional.

Ao pesquisador Ángel Carracedo Alvarez pela co-orientação, receptividade e disposição em ajudar sempre.

Às pesquisadoras Leonor Gusmão e Paula Sánches-Diz por aceitarem contribuir nesse projeto, pelo esclarecimento de diversas dúvidas à respeito do tema estudado, pelas valiosas informações fornecidas.

À Bárbara Kawamura, aluna de iniciação científica que co-oriento e que tem se dedicado e demonstrado grande interesse em continuar com essa linha de pesquisa.

À Raquel de Freitas Figueiredo e Jeane Cristina Costa pela grande amizade e pelo apoio técnico no desenvolvimento desse projeto. Vocês foram fundamentais!

À Greiciane pela serenidade e companheirismo em vários momentos, pela troca de idéias e pelo auxílio na coleta das amostras utilizadas neste trabalho.

Ao grupo da paternidade pela amizade e pela ajuda que me prestam, muitas vezes, à distância.

Ao professor Otávio H. Thiemann do Instituto de Física de São Carlos-SP por permitir a utilização do analisador genético em seu laboratório e, em especial, à Susana pelo suporte técnico e confiança.

Às companheiras de república Angélica, Ana, Sônia e Marcela pela amizade, pelos momentos de alegria e incertezas que juntas vivenciamos e por me apoiarem na busca deste sonho.

Às amigas Claudia e Sueli do laboratório de Parasitologia pelos momentos descontraídos e muito agradáveis, especialmente, na hora do almoço.

À Faculdade de Ciências Farmacêuticas da Universidade Estadual Paulista Júlio de Mesquita Filho pelos ensinamentos acadêmicos, pela alma científica e pelo núcleo profissional de excelência!

À CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela bolsa de estudo concedida durante o doutorado.

Ao PADC–UNESP (Programa de Apoio ao Desenvolvimento Científico da UNESP) pelo auxílio financeiro para a realização deste trabalho.

À FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) pelo auxílio financeiro para a realização deste trabalho.

Aos demais, amigos e familiares, que mesmo distante torceram por mim e fazem parte desta conquista!



*“O que vale na vida não é o ponto de partida e sim a caminhada.*

*Caminhando e semeando, no fim, terás o que colher.”*

*Cora Coralina*

## SUMÁRIO

### CAPÍTULO I

Resumo

Abstract

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

Lista de Símbolos

<b>1. Introdução.....</b>	<b>24</b>
<b>2. Revisão Bibliográfica.....</b>	<b>26</b>
2.1. Marcadores biológicos para identificação humana – panorama histórico.....	27
2.2. Genoma humano e polimorfismos genéticos.....	28
2.3. Cromossomo X.....	33
2.3.1. Evolução do cromossomo X.....	33
2.3.2. Características do cromossomo X.....	34
2.3.3. Ligação e desequilíbrio de ligação para o cromossomo X.....	37
2.3.4. Aplicações dos marcadores genéticos do cromossomo X.....	40
2.3.4.1. Análise de relações biológicas.....	40
2.3.4.2. Análise de vestígio biológico.....	45
2.3.4.3. Ancestralidade.....	45
2.3.5. Marcadores X-STR: sistemas <i>multiplex</i> e frequências alélicas.....	46
<b>3. Objetivos.....</b>	<b>49</b>
<b>4. Materiais e Métodos.....</b>	<b>51</b>
4.1. Aspectos éticos.....	52

4.2. Casuística.....	52
4.3. Coleta do material biológico e extração do DNA.....	52
4.4. Reação de amplificação dos X-STRs.....	53
4.5. Preparo da amostra para eletroforese.....	57
4.5.1. Preparo de amostras para ABI377.....	58
4.5.2. Preparo de amostras para ABI3500.....	59
4.6. Determinação dos perfis alélicos e análise estatística.....	59
4.7. Análise estatística nos casos de relações biológicas.....	61
4.8. Elaboração do banco de dados brasileiro para X-STRs.....	61
<b>5. Resultados e Discussão.....</b>	<b>62</b>
5.1. Avaliação da metodologia.....	63
5.2. Variação genética.....	64
5.3. Parâmetros estatísticos de avaliação forense.....	69
5.3.1. Desequilíbrio de ligação entre os marcadores X-STR.....	74
5.4. Comparação entre populações.....	77
5.5. Análise de X-STRs nos casos de relações biológicas.....	83
5.6. Banco de dados brasileiro para X-STRs.....	85
5.6.1. X-STRs no Brasil.....	86
5.6.2. Dados dos X-STRs.....	88
5.6.2.1. Localização.....	89
5.6.2.2. Outras informações.....	89
5.6.2.3. Frequências alélicas.....	91
5.6.2.4. Parâmetros estatísticos.....	91
5.6.2.5. Marcadores em desequilíbrio de ligação e haplótipos.....	92
5.6.3. Submeter dados.....	93

5.6.4. Contribuidores.....	93
<b>6. Conclusões.....</b>	<b>94</b>
<b>7. Referências Bibliográficas.....</b>	<b>96</b>
<b>Anexos.....</b>	<b>104</b>
Anexo A.....	105
Anexo B.....	106
Anexo C.....	107
Anexo D.....	108
Anexo E.....	109

## **CAPÍTULO II**

Artigo publicado

Genetic profile characterization of 10 X-STRs in four populations of the southeastern region of Brazil.

## **CAPÍTULO I**

## RESUMO

Os marcadores polimórficos *short tandem repeats* do cromossomo X (X-STR) são de utilização recente na prática forense, sendo aplicados, principalmente, com a finalidade de complementar os dados obtidos com marcadores genéticos autossômicos. Tendo em vista a necessidade de ampliação dos dados da população brasileira em relação aos X-STRs, este projeto teve o objetivo de determinar as frequências alélicas e os parâmetros estatísticos de interesse forense para 10 X-STRs na população de São Paulo-SP, Rio de Janeiro-RJ, Belo Horizonte-BH e Vitória-ES (Sudeste do Brasil). Posteriormente, os X-STRs foram genotipados em casos deficientes de relação biológica e, com o objetivo de compilar os dados genéticos populacionais brasileiros publicados para X-STRs, este projeto desenvolveu o Banco Genético Brasileiro do Cromossomo X (BGBX).

Para isso, foram analisados 1001 indivíduos não relacionados e residentes nas cidades descritas acima e 3 casos deficientes de relação biológica. Os marcadores X-STRs foram amplificados em sistema *decaplex* e submetidos à eletroforese em analisador genético ABI377 e ABI3500. Os programas GeneScan e GeneMapper ID-X foram utilizados para a determinação alélica e, a planilha Excel e o programa Arlequin, para a determinação das frequências alélicas, parâmetros estatísticos e análise de distância genética entre diferentes populações.

Em relação aos X-STRs analisados, os marcadores DXS6809 e GATA172D05 foram os mais discriminativos, enquanto os marcadores DXS8378, DXS7133 e DXS7423 apresentaram os menores poder de discriminação. Diferenças significativas de frequências alélicas foram obtidas dentre as populações brasileiras e entre estas e demais populações estrangeiras, sendo que uma maior distância genética foi obtida em relação à população africana de Uganda. O sistema *decaplex* se mostrou uma poderosa ferramenta para a prática

forense e demonstrou a necessidade de banco de dados genético específico para X-STRs. O BGBX está acessível em [www.bgbx.com.br](http://www.bgbx.com.br) e, até o momento, compilou informações de frequências alélicas, parâmetros estatísticos, desequilíbrio de ligação e haplótipos de X-STR para 24 populações. Esta proposta tem imediata aplicação forense e, por centralizar os dados brasileiros, facilita a busca destes na web, disseminando as pesquisas conduzidas com X-STRs em nosso país e aumentando o nosso reconhecimento perante a comunidade internacional.

**Palavras-chave:** X-STR, Cromossomo X, Identificação Humana, Sudeste, Brasil.

## ABSTRACT

The short tandem repeat polymorphic markers of X chromosome (X-STR) are of recent use in forensic practice, been applied mainly in order to complement the data obtained with autosomal genetic markers. Given the needs of expansion of the Brazilian population data in relation to X-STRs, this project aimed to determine the allele frequencies and the statistical parameters of forensic interest to 10 X-STRs in the population from São Paulo-SP, Rio de Janeiro-RJ, Belo Horizonte-BH e Vitória-ES (Southeast of Brazil). Subsequently, the X-STRs were genotyped in deficient kinship cases and, with the aim of compiling the Brazilian genetic population data published for X-STRs, this project developed the Brazilian Genetic Database of Chromosome X (BGBX).

For this, we have analyzed 1001 unrelated individuals, residents in the cities previously described, and three deficient kinship cases. The X-STR markers were amplified in decaplex system and submitted to electrophoresis on ABI377 and ABI3500 genetic analyzers. GeneScan and GeneMapper ID-X softwares were used to determine the allele and, Excel spreadsheet and Arlequin software, for the determination of allele frequencies, statistical parameters and genetic distance analysis between different populations.

Regarding the X-STRs analyzed, markers DXS6809 and GATA172D05 were the most discriminative, while the markers DXS8378, DXS7133 and DXS7423 showed the lowest discrimination power. Significant differences in allele frequencies were obtained within Brazilian populations and between these and other foreign populations, and a greater genetic distance was obtained in relation to the African population of Uganda. The decaplex system proved to be a powerful tool for forensic practice and demonstrated the need of specific genetic database for X-STRs. The BGBX is available at [www.bgbx.com.br](http://www.bgbx.com.br) and, to date, compiled information of allele frequencies, statistical parameters, linkage disequilibrium and



X-STR haplotypes to 24 populations. This proposal has immediate use in the forensic practice, and by centralizing the Brazilian data, facilitates the search of these data on the Web, disseminating the research conducted with X-STRs in our country and increasing our recognition before the international community.

**Keywords:** X-STR, X Chromosome, Human Identification, Southeast, Brazil.

## LISTA DE FIGURAS

Figura 1.	Esquema de um marcador minissatélite (D1S80) e um marcador STR (TH01).....	31
Figura 2.	Representação esquemática dos cromossomos sexuais X e Y com as regiões de homologia (coloridas) entre tais cromossomos.....	36
Figura 3.	Distribuição de 30 marcadores X-STRs de uso forense.....	46
Figura 4.	Idiograma do cromossomo X com a localização citogenética e física dos 10 X-STRs estudados neste trabalho.....	54
Figura 5.	Eletroferograma com perfil alélico de amostra feminina obtido em ABI377. Marcação 6-FAM (painel azul: marcadores DXS8378, DXS9898, DXS7133, GATA31E08), VIC (painel verde: marcadores GATA172D05, DXS7423, DXS6809) e NED (painel preto: DXS7132, DXS9902, DXS6789).....	63
Figura 6.	Eletroferograma com perfil alélico de amostra feminina obtido em ABI3500. Marcação 6-FAM (painel azul: marcadores DXS8378, DXS9898, DXS7133, GATA31E08), VIC (painel verde: marcadores GATA172D05, DXS7423, DXS6809) e NED (painel preto: DXS7132, DXS9902, DXS6789).....	64
Figura 7.	Página inicial do banco de dados BGBX.....	86

Figura 8. Estados brasileiros com dados de X-STR cadastrados no BGBX até o momento (marcados com um círculo). Ao clicar no círculo, os estudos correspondentes a este Estado são listados.....

## LISTA DE TABELAS

Tabela 1.	Tabela 1: Estrutura de repetição e nomenclaturas utilizadas para os 10 X-STRs.....	55
Tabela 2.	Sequência dos <i>primers</i> para os 10 X-STRs (adaptada de GUSMÃO et al., 2009).....	57
Tabela 3.	Perfil alélico das amostras de DNA padrão utilizadas como referência (adaptada de GUSMÃO et al., 2009).....	60
Tabela 4.	Teste exato para diferenciação das populações masculina e feminina de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.....	65
Tabela 5.	Frequência alélica e diversidade gênica dos 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.....	65
Tabela 6.	Heterozigose esperada e observada e teste exato para o equilíbrio de Hardy-Weinberg calculados para os 10 X-STRs na população feminina de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.....	69
Tabela 7.	Chance média de exclusão em trios envolvendo filha (MECT) para os 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.....	71
Tabela 8.	Chance média de exclusão em duo pai/filha (MECD) para os 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.....	71

Tabela 9.	Poder de discriminação masculino (PDM) para os 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.....	72
Tabela 10.	Poder de discriminação feminino (PDF) para os 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.....	73
Tabela 11.	Valor de $p$ obtido no teste de desequilíbrio de ligação para todos os pares de marcadores analisados na população de São Paulo, Rio de Janeiro, Vitória, Belo Horizonte e Sudeste.....	75
Tabela 12.	Comparação para locus simples entre populações brasileiras. Distância genética $F_{ST}$ (1ª linha) e correspondente valor de $p$ (2ª linha) e $p$ valor do teste exato para diferenciação das populações (3ª linha).....	78
Tabela 13.	Comparação em nível haplotípico entre populações brasileiras. Distância genética $F_{ST}$ (abaixo da diagonal) e correspondente valor de $p$ (acima da diagonal) entre populações brasileiras.....	80
Tabela 14.	Distância genética $F_{ST}$ para nível haplotípico entre populações brasileiras e internacionais.....	81
Tabela 15.	Perfis observados para 15 AS-STRs e 10 X-STRs em três casos de avaliação do vínculo biológico.....	84

## LISTA DE ABREVIATURAS E SIGLAS

AMOVA	<i>Analysis of molecular variance</i> , análise de variação molecular
AS-STR	<i>Short tandem repeats</i> autossômicos
BGBX	Banco Genético Brasileiro do Cromossomo X
CEP	Comitê de Ética em Pesquisa
CNV	<i>Copy-number variation</i>
DG	Diversidade gênica
DNA	Ácido desoxirribonucléico
FO-USP	Faculdade de Odontologia da Universidade de São Paulo
GHEP-ISFG	Grupo de Línguas Espanhola e Portuguesa da Sociedade Internacional de Genética Forense
He	Heterozigose esperada
HLA	<i>Human Leucocyte Antigen</i>
HEMOES	Centro de Hemoterapia e Hematologia do Espírito Santo
HEMOMINAS	Fundação Centro de Hematologia e Hemoterapia de Minas Gerais
HEMORIO	Instituto Estadual de Hematologia Arthur de Siqueira Cavalcanti
Ho	Heterozigose observada
IP	Índice de Paternidade
ISFG	<i>International Society of Forensic Haemogenetics</i>
LD	<i>Linkage Disequilibrium</i> , desequilíbrio de ligação
LINE	<i>Long Interspersed Nuclear Elements</i>
LR	<i>Likelihood Ratio</i> , razão de verossimilhança
MEC	<i>Mean exclusion chance</i> , chance média de exclusão
MECD	Chance Média de Exclusão em Duo pai/filha

MECT	Chance Média de Exclusão em Trios envolvendo filha
MySQL	<i>Structured Query Language</i>
NCBI	<i>National Center for Biotechnology Information</i>
OMIM	<i>Online Mendelian Inheritance in Man</i>
PAR	<i>Pseudoautosomal Region</i> , região pseudoautossômica
PDF	Poder de discriminação feminino
PDM	Poder de discriminação masculino
PhP	<i>Hypertext preprocessor</i>
PIC	<i>Polymorphism information content</i> , conteúdo de informação polimórfica
SINE	<i>Short interspersed nuclear elements</i>
SNP	<i>Single nucleotide polymorphisms</i>
SSR	<i>Simple sequence repeats</i>
STR	<i>Short tandem repeats</i>
UNESP	Universidade Estadual Paulista Júlio de Mesquita Filho
VNTR	<i>Variant number of tandem repeats</i>
XAR	<i>X-added region</i>
XCR	<i>X-conserved region</i>
X-STR	<i>Short tandem repeats</i> do cromossomo X
XTR	<i>X-transposed region</i>
Y-STR	<i>Short tandem repeats</i> do cromossomo Y

## LISTA DE SÍMBOLOS

g	Gramas
cM	Centi-Morgan
M	Molar
Mb	Mega par de base
mm	Milímetro
mL	Mililitro
ng	Nanograma
pb	Par de base
qsp	Quantidade suficiente para
rpm	Rotações por minuto
$\mu$ L	Micro litro
$\mu$ M	Micromolar
~	aproximadamente
%	por cento
$^{\circ}$ C	Graus Celsius



## **1. INTRODUÇÃO**

A identificação humana por meio da análise do ácido desoxirribonucléico (DNA) constitui, atualmente, uma ferramenta indispensável na Genética Moderna, auxiliando de forma imensurável tanto nas questões de identificação civil como nas de identificação criminal (FRANÇA, 2001).

Os exames forenses de DNA consistem em testes comparativos do perfil genético do periciando com seus familiares – ascendentes ou descendentes – ou com o próprio perfil do indivíduo que se busca identificar, uma vez que o perfil pode ser obtido em qualquer material biológico a partir de objetos de uso pessoal deste. Em ambas as situações têm-se a vantagem de não haver necessidade de registros anteriores para se chegar à identificação individual. No entanto, é importante ressaltar que a análise do DNA constitui-se em uma evidência a mais no caso, não é a sentença. A análise forense do DNA sempre será um elemento a se somar ao processo investigativo.

Todos os indivíduos possuem um patrimônio genético constituído pela combinação de diversos marcadores herdados dos seus progenitores. Estes marcadores correspondem a sequências de DNA que variam de um indivíduo para outro, denominados polimorfismos. Na prática forense, são analisados principalmente marcadores nucleares autossômicos e, em menor escala, do cromossomo Y e DNA mitocondrial, sendo a utilização dos marcadores do cromossomo X algo recente (SZIBOR et al., 2003a).

Em alguns casos, como na exumação, em que pouco material biológico está disponível e, em casos deficientes de vínculo biológico, como os exames de DNA realizados na ausência do suposto pai, a análise dos tradicionais marcadores autossômicos pode não solucionar o caso. Nestas situações, faz-se necessária a análise de outros marcadores como as regiões polimórficas do cromossomo X (SZIBOR et al., 2003a), as quais são capazes de complementar a análise dos primeiros de forma muito eficaz.

**2. REVISÃO  
BIBLIOGRÁFICA**

## 2.1. Marcadores biológicos para identificação humana – panorama histórico

A identificação humana é o processo científico pelo qual se determina a identidade de uma pessoa, através do estabelecimento de um conjunto de caracteres que a individualize, fazendo-a igual apenas a si mesma (FRANÇA, 2001). Dentre os processos de identificação humana, tem-se o campo da Ciência que envolve a análise de marcadores biológicos protéicos e genéticos.

Inicialmente, diversos marcadores protéicos foram tipados para detectar polimorfismos que poderiam ser potencialmente usados para diferenciar indivíduos. No século XX, a descoberta dos grupos sanguíneos ABO, marcou o início dos estudos modernos sobre a variação genética humana (JORDE; WATKINS; BAMSHAD, 2001), sendo tal sistema utilizado por várias décadas em testes de identificação humana e em evidências relacionadas a crimes. Posteriormente, outros sistemas de marcadores protéicos como os componentes grupo-específicos (transferrina, albumina, ceruloplasmina, haptoglobina, fosfoglicomutase-1, fosfatase ácida, esterase D, etc) mostraram-se variáveis entre os grupos populacionais e passaram, também, a ser utilizados.

Marcando uma segunda fase, em 1954, foi demonstrada a ocorrência de um sistema de histocompatibilidade mediado por antígenos na superfície dos leucócitos, conhecido como complexo HLA (*Human leucocyte antigen*), determinado por genes alélicos muito próximos, localizados no braço curto do cromossomo 6, com acentuado poder de discriminação (CALABREZ, 1999).

O uso de sistemas protéicos para a aquisição de altos níveis de diferenciação entre os indivíduos apresenta, no entanto, alguns inconvenientes, tais como: a baixa estabilidade em amostras biológicas expostas ao ambiente; o baixo poder de discriminação e a ausência dos mesmos marcadores protéicos em todos os tecidos de um indivíduo (WEEDN; SWARNEN,

1998). A tipagem de polimorfismos genéticos contornou muitas dessas limitações, dando início à terceira fase do desenvolvimento das ciências forenses voltadas à identificação humana.

A tipagem do material genético surgiu em 1984 quando o geneticista inglês Alec Jeffreys e seu grupo de trabalho descobriram que, em um dos *introns* do gene codificador da proteína mioglobina, havia certas regiões do DNA que variavam de um indivíduo para o outro, produzindo perfis genéticos conhecidos como “impressão digital do DNA” ou “DNA *fingerprinting*” (JEFFREYS; WILSON; THEIN, 1985).

A análise do DNA foi utilizada por Jeffreys, oficialmente pela primeira vez, em 1985, na Inglaterra, para solucionar um problema de imigração no qual um menino de 13 anos, após uma longa estadia com seu pai em Gana, retorna sozinho para a Inglaterra e é detido pela imigração sob alegação de passaporte forjado (JEFFREYS; BROOKFIELD; SEMEONOFF, 1985).

Um ano depois, também na Inglaterra, esta técnica foi utilizada pela primeira vez em um caso criminal, no qual vários homens entre 13 e 30 anos foram tipados para que o verdadeiro estuprador e assassino de 2 jovens fosse identificado. O caso ficou conhecido como *Enderby* (Queen x Pitchfork) e, desde então, a análise do DNA passou a ser potencialmente utilizada na rotina forense (TANDE, 1989).

## **2.2. Genoma humano e polimorfismos genéticos**

O genoma humano é o conjunto de toda a informação genética existente nas células do organismo e, nas células somáticas, está contido em 46 cromossomos no núcleo (22 pares de cromossomos homólogos e 1 par de cromossomos sexuais X/Y), denominado de DNA

nuclear, e em uma dupla fita circular presente no interior da organela celular mitocôndria, denominado DNA mitocondrial (BUTLER, 2005).

O DNA nuclear é biparental (metade dos cromossomos herdada da mãe e a outra metade herdada do pai) e possui apenas uma cópia por célula; enquanto o DNA mitocondrial é uniparental (herdado apenas da mãe em todos os indivíduos) e possui muitas cópias por célula, pois uma única célula pode ter dezenas a centenas de mitocôndrias e em cada mitocôndria podem existir dezenas a centenas de cópias do DNA mitocondrial (ROBIN; WONG, 1998). Nos gametas (óvulo e espermatozóide), há diferença em relação ao DNA nuclear, o qual é haplóide, apresentando 23 cromossomos de origem materna (no caso do óvulo) ou de origem paterna (no caso do espermatozóide).

Uma primeira visão da sequência do genoma humano foi reportada em 2001, cobrindo ~90% do material eucromático e, embora um marco, ainda estava altamente imperfeita, contendo cerca de 250 mil *gaps* e muitos erros na sequência de nucleotídeos (*International Human Genome Sequencing Consortium*, 2001). Em 2004, uma sequência mais completa e com melhor qualidade foi publicada, perfazendo ~99% do genoma eucromático (~93,5% do genoma humano), com apenas 341 *gaps* e um erro de nucleotídeo por 100 mil bases (*International Human Genome Sequencing Consortium*, 2004).

O genoma humano apresenta sequências codificantes, contidas em genes que codificam proteínas (*exons* codificadores representam ~1,2% do genoma eucromático) (LANDER, 2011) e RNAs e sequências não codificantes (*International Human Genome Sequencing Consortium*, 2001). Dentre estas últimas há as sequências de DNA repetitivas que correspondem à ~45% do genoma humano e englobam uma classe denominada *Simple Sequence Repeats* (SSR), a qual é de grande utilidade na genética forense, como veremos a seguir.

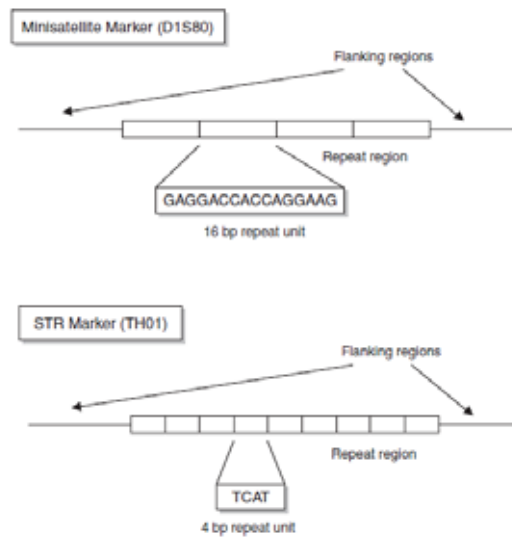
A grande aplicação forense do DNA decorre do seu alto poder de discriminação ocasionado por polimorfismos genéticos existentes no DNA. Inclusive para gêmeos monozigóticos, pesquisas recentes com *copy-number variation* (CNV) têm mostrado diferenças genéticas entre tais indivíduos (BRUDER et al., 2008), o que reforça a aplicação do DNA na identificação humana.

Os polimorfismos genéticos do DNA são classificados em dois tipos:

- Polimorfismos de sequência, originados por substituição de um ou mais nucleotídeos numa sequência de DNA.
- Polimorfismos de comprimento, originados por inserções ou deleções de um ou mais nucleotídeos.

Dentre os polimorfismos de sequência, os mais frequentes são os *Single Nucleotide Polymorphisms* (SNPs). Dentre os polimorfismos de comprimento, destacamos os marcadores genéticos VNTRs (*Variant Number of Tandem Repeats*) ou minissatélites, repetições consecutivas de unidades constituídas por aproximadamente 8 a 100 pares de base (pb), utilizados inicialmente nos exames de DNA, e os marcadores microssatélites, SSR ou mais usualmente conhecidos por *Short Tandem Repeats* (STRs), repetições consecutivas de unidades constituídas por 2 a 7 pb (Figura 1) (BUTLER, 2011), utilizados atualmente na rotina forense.

Figura 1. Esquema de um marcador minissatélite (D1S80) e um marcador STR (TH01).



Fonte: BUTLER (2005).

Os marcadores STR são classificados de acordo com o número de nucleotídeos que formam o motivo (dinucleotídeos, trinucleotídeos, tetranucleotídeos, etc.), correspondendo a aproximadamente 3% do genoma humano (*International Human Genome Sequencing Consortium*, 2001).

Segundo Butler (2005), os STRs de maior valor para a identificação humana são aqueles que apresentam:

- alto poder de discriminação, normalmente maior que 90%, com heterozigose observada maior que 70%;
- robustez e reprodutibilidade de resultados quando em *multiplex* com outros marcadores;
- taxa de mutação não muito elevada a fim de se estabelecer o vínculo biológico entre indivíduos relacionados e nem muito baixa a fim de se diferenciar as pessoas entre si;



- tamanho aparente<sup>1</sup> dos alelos na faixa de 90-500 pb. Tamanhos menores melhor para análise de amostras de DNA degradado;
- estreita faixa de tamanho de alelos, reduzindo a perda de alelos na genotipagem pela amplificação preferencial de alelos menores;
- baixa formação de *stutters*, fragmentos de DNA tipicamente com uma ou mais unidade de repetição a menos que o alelo verdadeiro, beneficiando a análise de amostras com mistura de material biológico proveniente de diferentes indivíduos.

Atualmente, as análises forenses empregam marcadores genéticos presentes tanto no DNA nuclear dos cromossomos autossômicos, bem como dos cromossomos sexuais (X e Y), e no DNA mitocondrial, sendo que a seleção de um ou outro tipo de marcador irá depender das particularidades de cada caso.

Marcadores do cromossomo Y são ditos de linhagem paterna, pois são passados de pai para filho; enquanto o DNA mitocondrial é de linhagem materna, sendo transferido da mãe para seus filhos. No caso do cromossomo X, a mulher herda um cromossomo de cada um dos pais e o homem herda apenas o cromossomo materno. Diferentemente dos marcadores autossômicos, em que homens e mulheres herdam dois cromossomos: um de herança materna e, o outro, de herança paterna. Em virtude do mecanismo de transmissão (somente materna, paterna ou ambas) e do tipo de polimorfismo (sequência ou comprimento), um maior poder de discriminação é obtido com os marcadores STR autossômicos (AS-STR) e, portanto, estes são os de escolha nos exames de DNA.

No entanto, como veremos a seguir, outros marcadores podem auxiliar a análise dos tradicionais AS-STR. Assim, os marcadores STR do cromossomo X (X-STRs) tem o seu

---

<sup>1</sup>Tamanho obtido na eletroforese de fragmentos, o qual considera a região de flaqueamento da estrutura de repetição do STR e fatores que influenciam no tamanho obtido como o fluorocromo utilizado na marcação.

valor na prática forense, constituindo-se em uma importante ferramenta em casos específicos de investigação de relação biológica.

## **2.3. Cromossomo X**

### **2.3.1. Evolução do cromossomo X**

As propriedades originais deste cromossomo são uma consequência da evolução dos cromossomos sexuais nos mamíferos. Tais cromossomos evoluíram de um par de autossomos homólogos (denominados proto-cromossomos sexuais) dentre os últimos 300 milhões de anos (OHNO, 1967), quando o proto-cromossomo Y adquiriu um gene determinante do sexo masculino e, posteriormente, outros genes benéficos para este, mas não para o sexo feminino. Assim, a fim de se manter a heterogeneidade entre os cromossomos X e Y, por mecanismos de seleção, uma barreira para a recombinação foi desenvolvida entre eles, isolando regiões determinantes do sexo ao longo dos dois cromossomos (GUSMÃO; ALVES, 2010).

Na ausência de recombinação, o acúmulo de eventos de mutação levou à degeneração do cromossomo Y, tornando os cromossomos sexuais heteromórficos. Neste processo, o cromossomo Y perdeu praticamente todos os traços do ancestral autossômico e tem desenvolvido um único padrão de sequência repetitiva (SKALETSKY et al., 2003; JOBLING; TYLER-SMITH, 2003). Por contraste, os elementos funcionais foram conservados no cromossomo X e este é, fisicamente, o mais estável cromossomo nuclear entre os mamíferos placentários (SCHAFFNER, 2004).

No trabalho realizado por Ross et al. (2005), a origem autossômica dos cromossomos sexuais foi ilustrada através do alinhamento entre as sequências do cromossomo X humano e as sequências de dois cromossomos autossomos de galinha. Esse estudo mostrou que o braço

longo do cromossomo X apresenta grande homologia com o braço curto do cromossomo 4 de galinha e é formado por uma região conservada (*X-conserved region* ou XCR), descendente de um proto-cromossomo X; já seu braço curto, apresenta grande homologia com o braço longo do cromossomo 1 de galinha, sendo constituído por uma pequena região (Xcen -p11.3) conservada e pela adição de um grande fragmento de DNA (*X-added region* ou XAR), decorrente da translocação com um segundo cromossomo autossomo há, aproximadamente, 105 milhões de anos.

### 2.3.2. Características do cromossomo X

O cromossomo X possui muitas características que são únicas no genoma, como o fato da mulher herdá-lo de cada um dos pais, mas o homem herdar apenas um, o materno.

Este é uma molécula de DNA linear com ~155 mega pares de base (Mb), uma baixa densidade gênica (4% do total de genes humanos), genes de tamanho pequeno, um baixo índice de C-G (39% comparado com a média de 41% do genoma) e é altamente rico em regiões repetitivas, tais como *interspersed repeats* (~56% da euromatina comparados à média de 45% do genoma), duplicações de segmentos (~2,6%) e extensas regiões de DNA satélite. Foram anotados 1.098 genes, 700 pseudogenes (644 processados e 56 não processados) e 153.146 SNPs (ROSS et al. 2005).

Dentre as regiões repetitivas, os elementos intercalares curtos (SINEs, *short interspersed nuclear elements*) foram encontrados abaixo da média e, os longos (LINEs, *long interspersed nuclear elements*) da família 1, apareceram acima da média, compreendendo 29% da sequência do cromossomo X (média do genoma é de 17%) (ROSS et al. 2005).

Apesar de apenas 4% dos genes do genoma humano estarem presentes no cromossomo X, estes são de grande interesse na genética médica, uma vez que cerca de 10% das doenças

com herança mendeliana têm sido atribuídas a este cromossomo. Isto ocorre pelo fato do homem herdar apenas um cromossomo X, revelando de forma direta o fenótipo decorrente de uma mutação recessiva. Muitas desordens genéticas, incluindo às ligadas ao cromossomo X, podem ser acessadas pela base de dados OMIM (*Online Mendelian Inheritance in Man*) (HAMOSH, 2005).

A baixa densidade gênica do cromossomo X pode, simplesmente, refletir seu ancestral autossômico. No entanto, acredita-se que a seleção pode ter favorecido a transposição de certa classe de genes do cromossomo X para os autossomos durante a evolução dos mamíferos. Tal classe deveria abranger genes cujo produto protéico seria requerido em dose dupla em ambos os sexos ou genes cuja mutação em tecidos somáticos masculinos seria letal.

Com a finalidade de tornar proporcional os produtos gênicos do cromossomo X entre homens e mulheres, um dos cromossomos X da mulher é inativado no início do desenvolvimento (CHOW et al., 2005) e sofre uma reativação nas células germinativas para recombinar com o outro cromossomo X feminino. No homem, o cromossomo X não se recombina ao longo de todo o seu comprimento, sendo tal processo restrito às regiões pseudoautossômicas (PAR, *pseudoautosomal region*) 1 e 2, localizadas nas extremidades dos braços dos cromossomos sexuais.

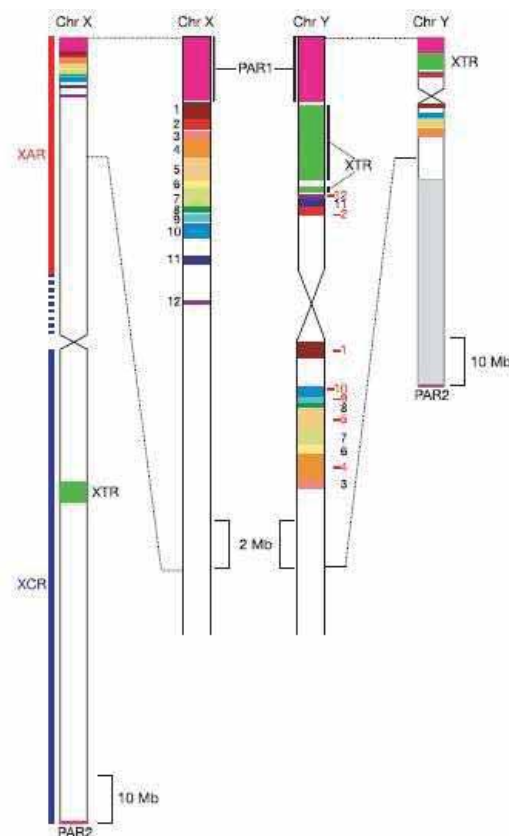
A PAR1, localizada na extremidade do braço curto dos cromossomos X e Y, compreende 2,7 Mb e os genes desta região estão presentes em duas cópias em homens e mulheres, não estando sujeitos ao processo de inativação do cromossomo X. Esta região apresenta uma frequência de recombinação elevada e é necessária para a normal segregação dos cromossomas sexuais durante a meiose (ELLIS; GOODFELLOW, 1989).

Em relação à PAR2, localizada na extremidade do braço longo dos cromossomos sexuais, esta possui 0,33 Mb e foi criada pela duplicação de uma região do cromossomo X para o Y, desde a divergência das linhagens de humanos e chimpanzés (FREIJE et al., 1992).

Nesta região, a recombinação não ocorre em todas as meioses e alguns genes estão sujeitos à inativação do cromossomo X, refletindo a condição deste cromossomo antes da duplicação (FREIJE et al., 1992; ROSS et al., 2005).

Além das PAR, os cromossomos sexuais apresentam homologia em regiões não recombinantes (Figura 2), localizadas na XAR e em uma região de transposição do cromossomo X (XTR, *X-transposed region*), situada na XCR, formada pela duplicação de material do X para o Y há 4,7 milhões de anos (SKALETSKY et al., 2003; PAGE et al., 1984).

Figura 2: Representação esquemática dos cromossomos sexuais X e Y com as regiões de homologia (coloridas) entre tais cromossomos.



Fonte: ROSS et al., 2005.

Os marcadores X-STR, aplicados em testes de identificação humana, estão localizados em regiões não recombinantes entre os cromossomos sexuais; mas, a recombinação ocorre entre os cromossomos X da mulher, garantido a variabilidade desses marcadores. Assim, da mesma forma que os autossomos, o cromossomo X se recombina nas mulheres e, similarmente ao cromossomo Y e DNA mitocondrial, em homens tem um modelo que permite determinar diretamente o haplótipo.

Distintas características genéticas são, portanto, esperadas para o cromossomo X em relação aos autossomos, principalmente: 1 – uma menor diversidade genética devido à menor taxa de mutação em mulheres do que em homens e um menor tamanho populacional efetivo (*International Human Genome Sequencing Consortium*, 2001; LI; YI; MAKOVA, 2002); 2 – uma deriva genética mais acentuada devido ao menor tamanho populacional efetivo e, conseqüentemente, uma estrutura genética mais pronunciada; 3 – um maior desequilíbrio de ligação, uma vez que o cromossomo X recombina apenas na mulher (SCHAFFNER, 2004).

### **2.3.3. Ligação e desequilíbrio de ligação para o cromossomo X**

Pelo fato dos marcadores X-STR estarem localizados no mesmo cromossomo, por apenas 2/3 dos cromossomos X se recombinarem à cada geração e por serem mais jovens e apresentarem menor taxa de mutação que os marcadores AS-STRs, ligação e desequilíbrio de ligação (LD) podem ocorrer entre os X-STRs, sendo necessário que tais parâmetros sejam avaliados para a posterior utilização destes marcadores na prática forense (SZIBOR et al., 2003a; SZIBOR, 2007).

Ligação ou *linkage* é um parâmetro estatístico que está relacionado à dependência física entre *loci*, ou seja, se refere à co-segregação de *loci* localizados próximos (TILLMAR et al., 2008). Enquanto que o desequilíbrio de ligação ou *linkage disequilibrium* é um parâmetro

estatístico que se refere à associação não randômica de alelos de diferentes *loci*, não necessariamente no mesmo cromossomo, medindo o desvio da frequência haplotípica do produto das frequências dos respectivos alelos na população (SZIBOR et al., 2003a; SZIBOR, 2007).

Para um conjunto de marcadores, a ligação pode ser avaliada pela determinação da frequência de recombinação entre estes dentre os indivíduos em uma família (*pedigree*). Já o LD pode ser estimado pelas frequências alélicas e haplotípicas obtidas em uma população (TILLMAR et al., 2008).

A taxa de recombinação meiótica é estimada para pares de marcadores e transformada em distância genética (expressa em centi-Morgan, cM), sendo que 1 cM equivale a 1 esperada recombinação por 100 meioses. A distância física e genética entre marcadores não estão estritamente correlacionadas, ou seja, a distância física de 1 Mb não corresponde a uma distância genética de 1 cM. Portanto, a suposição de que marcadores estejam ligados, baseando-se na distância física entre eles, deve ser confirmada pela análise da taxa de recombinação (SZIBOR et al., 2003a). Comparações entre a fração de recombinação obtida do *pedigree* e do *Rutgers combined linkage-physical map* (mapa onde a distância genética é inferida a partir da distância física) confirmou que *hotspots* de recombinação estão presentes ao longo do cromossomo X, não sendo esta influenciada apenas pela distância física entre marcadores (INTURRI et al., 2011).

Marcadores distanciados entre si por 50 cM são ditos não ligados e espera-se que haja ausência de LD entre eles. Quando a distância é inferior a 50 cM, os *loci* são ditos ligados, e à medida que esta distância diminui, aumentam-se as chances de ocorrer LD. Segundo Szibor (2007), marcadores com distância genética entre si inferiores a 0,5 cM (*loci* intimamente ligados) representam estáveis haplótipos e, em exames para avaliação de parentesco, devem ser analisados como um todo e não através dos seus constituintes alélicos; fato este que é de

grande utilidade em casos de reconstrução do perfil do suposto pai através da análise de seus ascendentes e/ou descendentes.

Portanto, quando os marcadores estão intimamente ligados, LD pode ocorrer e tal fato tem sido relatado na literatura: DXS101 - DXS7424 (ELDEMAN et al., 2002a); DXS6801 – DXS6809 – DXS6789 (SZIBOR et al., 2005) e DXS10079 – DXS10074 – DXS10075 (HERING et al., 2006). Segundo Hering et al. (2004), quando os marcadores estão separados por 5 cM, o LD entre eles já é improvável.

No entanto, o LD não é determinado apenas pela distância genética entre os STRs, podendo ser ocasionado por vários outros fatores como: seleção, mutação, deriva aleatória, efeito fundador, miscigenação ou subdivisão populacional (CHAKRAVARTI, 1999; ZARRABEITIA et al., 2006). Assim, os dados de LD em uma população não podem ser extrapolados para outra. Para marcadores com forte LD, a frequência dos haplótipos não pode ser estimada pelas frequências alélicas, sendo necessário obter tal dado diretamente da população (SZIBOR et al., 2003a; SZIBOR, 2007).

Szibor et al. (2003a) estabeleceram quatro grupos de ligação para o cromossomo X, localizados em Xp22.2, Xq12, Xq26 and Xq28 e *kits* foram desenvolvidos com a inclusão de marcadores pertencentes a estes grupos:

- Mentype Argus X-UL, Biotype: analisava um marcador por grupo de ligação (DXS7132, DXS7423, DXS8378 e HPRTB) mais a amelogenina.

- Mentype Argus X-8, Biotype: analisa dois marcadores por grupo de ligação (DXS7132, DXS7423, DXS8378, DXS10074, DXS10101, DXS10134, DXS10135, HPRTB) mais a amelogenina.

- Investigator Argus X-12, QIAGEN: analisa três marcadores por grupo de ligação (DXS7132, DXS7423, DXS8378, DXS10074, DXS10079, DXS10101, DXS10103, DXS10134, DXS10135, DXS10146, DXS10148 e HPRTB) mais a amelogenina.



### 2.3.4. Aplicações dos marcadores genéticos do cromossomo X

Em virtude do alto poder de individualização e praticabilidade, a análise dos STRs tornou-se uma rotina amplamente utilizada na prática forense. A grande maioria dos exames de vínculo biológico, aborda os STRs localizados em autossomos e cromossomo Y, sendo a análise de marcadores do cromossomo X algo ainda recente.

A vantagem da aplicação dos X-STRs em certas situações se dá pela diferença no número de cromossomos existente entre os tradicionais marcadores autossômicos e os do cromossomo X para os indivíduos do sexo masculino. Estes apresentam sempre dois alelos para os marcadores autossômicos e, apenas um, para os do cromossomo X. Assim, no mecanismo de herança genética, o pai transfere 100% do seu perfil genético (haplótipo) de cromossomo X para a filha, o que não ocorre com os autossômicos, nos quais apenas 50% da informação genética paterna é herdada.

#### 2.3.4.1. Análise de relações biológicas.

a) casos com incompatibilidade(s) entre filha e suposto pai (duo ou trio)

Segundo normas de padronização dos exames periciais no Brasil (MINISTÉRIO DA JUSTIÇA-SENASP):

Em casos de identificação de vínculo genético em primeiro grau, deve ser considerada exclusão a não coincidência em no mínimo dois loci, nos casos onde os genótipos forem heterozigotos e, de no mínimo três loci, quando pelo menos um dos genótipos forem homozigotos;

Onde não ocorrer coincidência em no mínimo dois loci, outros marcadores deverão ser analisados para ser certificada a exclusão ou confirmada a ocorrência de mutação.

Em casos duo ou trio, uma ou poucas incompatibilidade(s) sugerem mutação ou a possibilidade de relação biológica entre o suposto pai testado e o verdadeiro pai biológico.

Os X-STRs são capazes de complementar a análise dos marcadores autossômicos de forma muito eficiente nestas situações, pois como o pai transfere 100% do seu perfil para a filha, os X-STRs apresentam maior chance média de exclusão (MEC, *mean exclusion chance*) que os autossômicos, com semelhante conteúdo de informação polimórfica (PIC, *polymorphism information content*), ou seja, possuem uma maior capacidade de excluir de uma situação de investigação de paternidade, indivíduos não relacionados à criança testada.

Se os supostos pais são pai e filho, no caso de STRs autossômicos, estes obrigatoriamente compartilham 50% dos alelos e, conseqüentemente, o exame com um deles pode aparecer com apenas uma inconsistência genética, o que dificulta a diferenciação entre mutação paterna e relativa proximidade do suposto pai com o verdadeiro pai biológico da criança. No caso dos X-STRs, nenhum alelo entre pai e filho é idêntico por descendência, uma vez que estes herdam o cromossomo X de suas mães e ambas são diferentes, facilitando a resolução desses casos (SZIBOR et al., 2003a). Já no caso dos supostos pais serem dois irmãos, ambos receberam o cromossomo X da mesma mãe e por isso há uma chance de 50% destes possuírem alelos idênticos por descendência, da mesma forma como ocorre com os STRs autossômicos, sendo ambos os sistemas equivalentes nesta situação (SZIBOR et al., 2003a). No trabalho de Toscanini et al. (2009), uma inconsistência foi obtida entre criança e suposto pai após a análise de 17 marcadores autossômicos (kits PowerPlex 16, Promega e Identifiler, Applied Biosystems) e os cálculos estatísticos realizados sugeriam que o suposto

pai testado deveria ser o irmão do verdadeiro pai biológico. A análise de 10 X-STRs foi capaz de excluir em 6 locus o indivíduo testado como pai biológico da criança.

Em um estudo realizado por Silveira et al. (2007), quatro casos complicados de paternidade, compreendendo dois trios (mãe/filha/suposto pai) e duas duplas (filha/suposto pai), obtiveram baixo índice de paternidade devido à presença de uma suposta mutação paterna, permanecendo inconclusivos. Ao se analisar 8 X-STRs, em todos os casos, o haplótipo paterno foi diferente do apresentado pela filha e o suposto pai foi excluído como pai biológico desta. Em outro estudo, 5 X-STRs confirmaram a relação paterna (caso 1) e materna (caso 2) após presença de uma mutação em 15 STRs autossômicos (kit Identifiler, Applied Biosystems) (CHEN et al., 2009). Assim, a aplicação dos X-STRs mostrou ser uma poderosa ferramenta na resolução de casos que apresentaram poucas inconsistências após análise de autossômicos.

b) casos com dificuldade de se analisar o material biológico.

Em situações de parentesco com dificuldade para se analisar o material biológico, tais como em exumação, embora o material a ser genotipado é do suposto pai, o número de STRs amplificados com sucesso pode ser menor do que geralmente se obtém com amostras frescas (como sangue) e o laudo pode permanecer inconclusivo. Como os X-STRs possuem maiores valores de MEC que aos autossômicos, suficiente poder estatístico pode ser obtido quando um pequeno número destes marcadores são analisados (SZIBOR et al., 2003a).

c) casos deficientes de paternidade

Nestes casos, o material biológico do suposto pai não está disponível e o DNA de seus parentes é analisado para a reconstrução do seu perfil genético. Com a análise dos tradicionais STRs autossômicos, estes casos podem permanecer inconclusivos, pois para se ter uma boa reconstrução do perfil do suposto pai, vários parentes deste precisam ser genotipados. No entanto, este fato pode encarecer o custo do exame e/ou algumas pessoas podem se negar a doar o material biológico e/ou podem haver poucos parentes disponíveis, obtendo-se um baixo poder estatístico para o vínculo biológico investigado.

Na reconstrução com emprego de X-STRs, a mãe do suposto pai é bastante importante, pois os alelos paternos podem ser facilmente determinados a partir dela, já que o pai herda apenas um cromossomo X, o materno (ROSS et al., 2005). Assim, quando a suposta avó paterna está presente, tais casos podem ser facilmente solucionados; o que não ocorre com os autossômicos, sendo necessária a tipagem de mais parentes do suposto pai. Além disso, se a suposta avó não está presente, seu perfil pode ser facilmente determinado através de suas filhas e, muito mais facilmente, a partir dos irmãos do suposto pai, o que aumenta as circunstâncias em que bons resultados podem ser obtidos com os X-STRs (SZIBOR et al., 2003a). Em 4 casos de investigação de vínculo biológico, nos quais estavam presentes a mãe, a criança e a suposta avó paterna, ao adicionar 6 X-STRs à análise anteriormente feita com autossômicos (kit Identifiler, Applied Biosystems by Life Technologies), o valor da razão de verossimilhanças (LR) foi 118 a 6.500 vezes maior do que o obtido apenas com autossômicos (AQUINO et al., 2009).

No caso do cromossomo X, por exemplo, quando dois indivíduos do sexo feminino possuem o mesmo pai, estas também apresentam o mesmo cromossomo X paterno. Assim, em uma investigação em que duas irmãs ou meias-irmãs são tipadas, a presença de quatro

diferentes alelos em um marcador do cromossomo X permite a exclusão da paternidade comum entre elas, até mesmo se nenhum dos pais biológicos ou seus parentes forem analisados (SZIBOR et al., 2003a). Através de marcadores autossômicos, ambas poderiam aparentar serem irmãs, já que o pai pode ter doado um de seus alelos para uma filha e o seu outro alelo para a outra. Turrina; Atzei; De Leo (2007) utilizaram apenas 4 X-STRs para auxiliar na resolução de um caso de meia-irmãs, no qual a análise de 15 autossômicos não foi suficiente.

#### d) casos de maternidade

Em virtude das altas taxas de ilegitimidade paterna vistas na sociedade moderna, a identificação de esqueletos através da análise de mãe-filho é mais requerida que a de pai-filho. Para os testes mães-filhas, os X-STRs são equivalentes aos autossômicos; no entanto, para mães-filhos, tais marcadores são extremamente eficientes, já que o filho possui apenas um alelo para cada X-STR e este deve obrigatoriamente estar presente na mãe. A chance de exclusão nesses casos é idêntica para os testes pai-filha (SZIBOR et al., 2003a). De acordo com Toni; Domenici; Presciuttini (2007), nestes casos, os MECs obtidos são quase duas vezes maiores que os valores conseguidos para STRs autossômicos com moderada heterozigose (0,6-0,75).

#### e) casos de violência sexual com feto do sexo feminino

Após 6 a 8 semanas de gestação, quando ocorre o aborto, é muito difícil obter o material apenas do embrião, recuperando-se, geralmente, uma mistura de material embrionário com sangue materno. Neste caso, apenas autossômicos e X-STRs podem ser

utilizados e a prova de paternidade recai sobre os alelos do embrião que não são compartilhados com a mãe. Como para os X-STRs o homem apresenta apenas um alelo, estes são melhores que os autossômicos, já que a probabilidade de se excluir um indivíduo falsamente acusado é maior. Em casos de incesto (pai-filha), no entanto, a aplicação dos X-STRs fica limitada, pois todos os alelos do embrião estão presentes na mãe, ao contrário dos STRs autossômicos, nos quais pode-se obter alelos do embrião que foram obrigatoriamente herdados do pai (SZIBOR et al., 2003a).

#### 2.3.4.2. Análise de vestígio biológico

Quando precisamos determinar a identidade de uma vestígio biológico, ou seja, identificar quem contribuiu para a formação de tal vestígio, no caso de não haver mistura de material biológico, os marcadores do cromossomo X possuem uma capacidade de discriminação semelhante aos autossômicos quando o perfil é feminino, porém, a discriminação é menor se o perfil for masculino. Assim, os marcadores do cromossomo X não são mais vantajosos que os autossômicos nestas situações (GUSMÃO; ALVES, 2010).

No entanto, em casos de mistura, quando desejamos identificar um componente feminino em vestígio contaminado com material do agressor, os X-STR são de grande valia, pois dificilmente o perfil feminino será idêntico ao masculino (GUSMÃO; ALVES, 2010).

#### 2.3.4.3. Ancestralidade

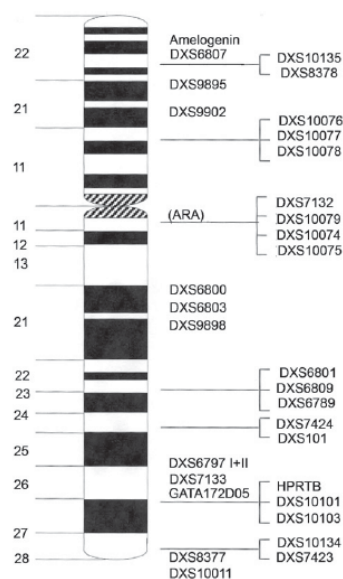
Similarmente ao cromossomo Y e do DNA mitocondrial, os quais permitem analisar apenas a linhagem paterna e materna, respectivamente, a análise de regiões com baixa taxa de recombinação no cromossomo X, também permite estudar a genealogia humana deste

cromossomo. Uma vez que este encontra-se presente em ambos os sexos e apresenta a forma haplotípica no sexo masculino, sua genealogia é interessante, pois fornece informação dos componentes masculino e feminino da população. Assim, a comparação da genealogia do cromossomo X com as do cromossomo Y e DNA mitocondrial deve ser informativa para a história da origem genética das populações (SANTOS-LOPES et al., 2007).

### 2.3.5. Marcadores X-STR: sistemas *multiplex* e frequências alélicas

O potencial dos marcadores genéticos presentes nos autossomos, no cromossomo Y e no DNA mitocondrial tem sido amplamente explorado pela genética populacional e forense, no entanto, falta uma maior contribuição do cromossomo X. Assim, cresce a investigação em tal cromossomo e diversos marcadores têm sido estudados por diferentes grupos. Mais de 50 X-STRs foram descritos e Szibor (2007) em sua revisão reportou as características dos 30 marcadores de uso forense mais conhecidos até aquela data (Figura 3).

Figura 3: Distribuição de 30 marcadores X-STRs de uso forense.



Fonte: Szibor, 2007.

Vários sistemas *multiplex* comercialmente disponíveis tornaram-se amplamente utilizados em genética populacional e forense, principalmente devido à facilidade de utilização e ao seu elevado poder de discriminação, tais como: PowerPlex 16 e PowerPlex 18D (Promega) e AmpFLSTR Identifiler e AmpF $\ell$ STR MiniFiler (Applied Biosystems by Life Technologies) para análise de AS-STRs e MenPlex Argus Y-MH (Biotype), PowerPlex Y (Promega) e AmpFISTR™ Yfiler (Applied Biosystems by Life Technologies) para investigação de Y-STRs. Relativamente aos X-STRs, como já mencionado, três kits comerciais foram desenvolvidos: Mentype Argus X-UL (Biotype), primeiro *kit* produzido, disponível no mercado (mas não no Brasil) de 2003 a 2008; Mentype Argus X-8 (Biotype), lançado em 2005 e disponível comercialmente no momento (mas não no Brasil) e Investigator Argus X-12 (QIAGEN), mais recente *kit*, disponível comercialmente também no Brasil (<http://www.chrx-str.org/>, acesso em 10 nov. 2011).

Alguns trabalhos analisaram marcadores contidos nos *kits* comerciais (TILLMAR et al., 2008; ZALÁN et al., 2007; ZALÁN et al., 2008; VERZELETTI et al., 2007; TIE; UCHIGASAK; OSHIDA, 2010; ZENG et al., 2011; EDELMANN et al., 2011). Porém, muitos grupos desenvolveram sistemas *multiplex* “in house”: 4plex (LEE et al., 2004), 5plex (POETSCH et al., 2005), 6plex (ROBINO et al., 2006), 7plex (BINI et al., 2005), 10plex (GUSMÃO et al., 2009), 12plex (SHIN et al., 2005 e RIBEIRO-RODRIGUES et al., 2010) e 13plex (HWA et al., 2009).

Em 2006, o Grupo de Línguas Espanhola e Portuguesa da Sociedade Internacional de Genética Forense (GHEP-ISFG) elaborou o primeiro trabalho colaborativo com X-STRs, no qual desenvolveu um sistema *decaplex* “in house” que foi caracterizado em 15 populações Ibéricas e Latino Americanas (GUSMÃO et al., 2009). Nesta data, na América Latina não havia licença para comercialização dos kits comerciais da Biotype e o Investigator Argus X-12 ainda não existia, propiciando a utilização de sistemas *multiplex* “in house”.



Alguns autores não têm observado diferenças significativas na distribuição alélica dos X-STRs dentre diferentes regiões da Europa (ZARRABEITIA et al., 2006; ROBINO et al., 2006, ALER et al., 2007). Porém, diferenças significativas foram relatadas entre Norte Americanos classificados como africanos, hispânicos e asiáticos (GOMES et al., 2007a), entre africanos (GOMES et al., 2007b; GOMES et al., 2009a), entre a população Coreana e Germânica, Chinesa e Espanhola (SHIN et al., 2005), entre italianos e chineses (VERZELETTI et al., 2007), entre a população Ibérica e Latino Americana (GUSMÃO et al., 2009), entre 3 populações chinesas e italiana, espanhola, germânica, paquistanesa, dentre outras (LIU et al., 2011), concluindo que um banco de dados comum não pode ser utilizado para os marcadores X-STRs.

Por fim, em 2005, foi anunciado um banco de dados para marcadores do cromossomo X, acessado pelo endereço [www.chrx-str.org](http://www.chrx-str.org) (SZIBOR; HERING; EDELMANN, 2005). No banco pode-se encontrar informações destes marcadores e são cadastrados dados de diversas populações, porém estes não precisam estar publicados desde que atendam a certos padrões de qualidade estabelecidos no banco. O Brasil, com sua população em proporção continental e muito heterogênea, para a qual já se tem relatado diferenças na distribuição alélica de X-STRs (GUSMÃO et al., 2009), não possui um banco de dados para tais marcadores.

### **3. OBJETIVOS**

Os objetivos deste trabalho foram:

- Determinar as frequências alélicas de 10 regiões STR do cromossomo X (DXS8378, DXS9902, DXS7132, DXS9898, DXS6809, DXS6789, DXS7133, GATA172D05, GATA31E08, DXS7423) em indivíduos brasileiros residentes em São Paulo-SP, Rio de Janeiro-RJ, Vitória-ES e Belo Horizonte-MG;
- Avaliar a eficiência forense desses marcadores nas populações estudadas;
- Verificar se há diferença significativa dentre as populações analisadas e entre estas e outras com dados relatados na literatura;
- Comprovar a eficácia do sistema X-STR *decaplex* em casos não solucionados pelos marcadores autossômicos tradicionais;
- Elaborar e implementar um banco de dados para marcadores X-STR caracterizados na população brasileira.

## **4. MATERIAIS E MÉTODOS**

#### **4.1. Aspectos Éticos**

Este projeto possui a aprovação do Comitê de Ética em Pesquisa (CEP) do Instituto Estadual de Hematologia Arthur de Siqueira Cavalcanti – HEMORIO (ANEXO A), da Fundação Centro de Hematologia e Hemoterapia de Minas Gerais - HEMOMINAS (ANEXO B), da Faculdade de Odontologia da Universidade de São Paulo – FO/USP (ANEXO C) e da Faculdade de Ciências Farmacêuticas da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) - Araraquara (ANEXO D).

#### **4.2. Casuística**

Para determinar as frequências alélicas dos marcadores X-STR, foram estudados 1001 indivíduos aparentemente saudáveis, não aparentados e residentes em São Paulo-SP (n = 250; 164 mulheres e 86 homens), Rio de Janeiro-RJ (n = 261; 145 mulheres e 116 homens), Vitória-ES (n = 245; 160 mulheres e 85 homens) e Belo Horizonte-MG (n = 245; 108 mulheres e 137 homens).

As frequências alélicas obtidas neste estudo foram aplicadas em três casos (A, B e C) de determinação do vínculo biológico pela análise do DNA, compreendendo a análise de nove amostras biológicas (Caso A: mãe, criança e mãe do suposto pai; Caso B: mãe, criança, mãe do suposto pai e filha biológica do suposto pai; Caso C: criança e mãe do suposto pai).

#### **4.3. Coleta do material biológico e extração do DNA**

As coletas foram realizadas nos locais em que se obteve as aprovações do CEP e no Centro de Hemoterapia e Hematologia do Espírito Santo – HEMOES. Os casos para análise

do vínculo biológico e a estrutura laboratorial para o desenvolvimento do projeto foram fornecidos pelo Laboratório de Investigação de Paternidade da Faculdade de Ciências Farmacêuticas da UNESP-Araraquara.

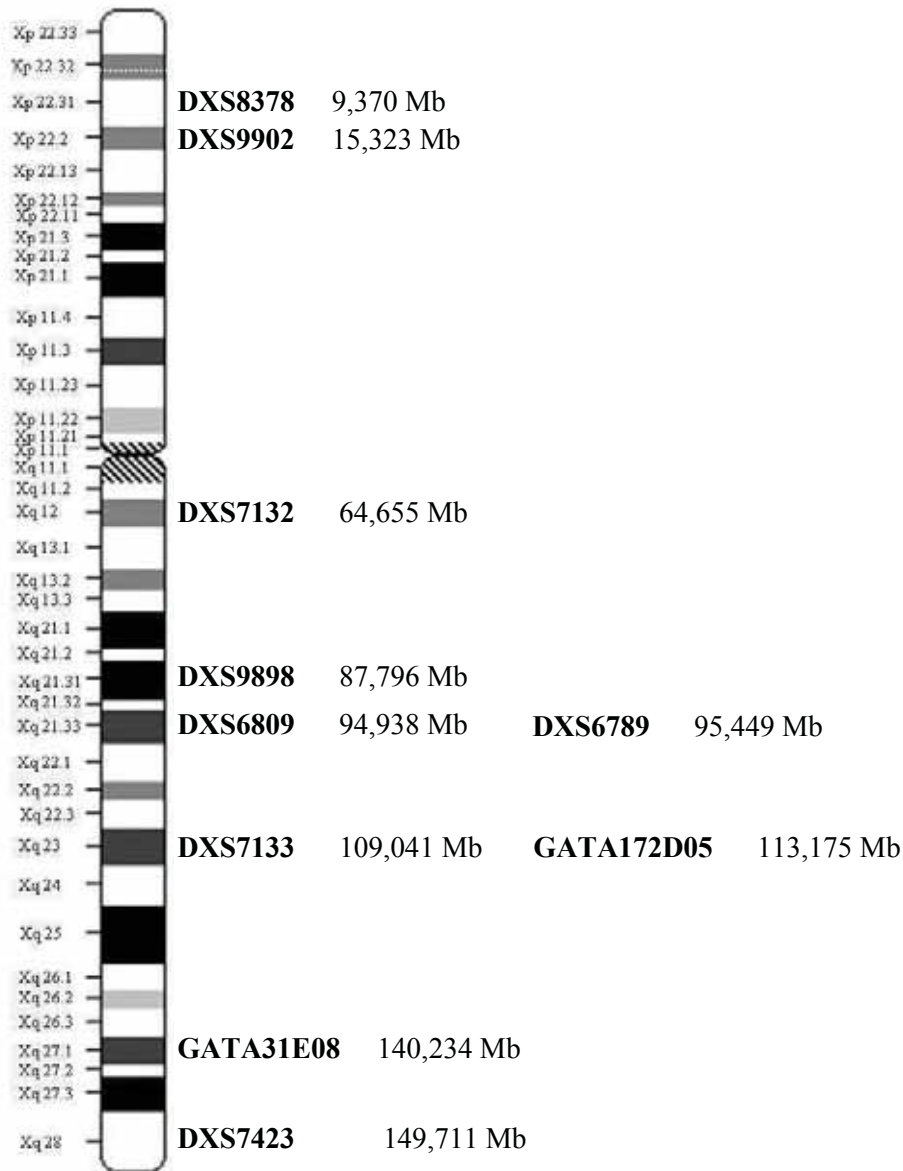
Todas as amostras (sangue) foram coletadas em papel FTA (Whatman). As 1001 amostras, oriundas dos indivíduos da região sudeste, foram extraídas da seguinte forma: 1 disco (1,2 mm) de papel FTA foi cortado com o auxílio do *Harris Micro Punche e Cutting Mat* (Whatman) e depositado em microtubo de 1,5 mL. Procedeu-se à extração com 200 µL de solução FTA *Purification Reagent* (Whatman), seguindo-se de duas lavagens com 100 µL da mesma solução; posteriormente, o disco foi tratado duas vezes com 100 µL de solução DNA *Rehydration* (Promega), incubado à 37 °C por uma a duas horas e armazenado à 4 °C por no máximo uma semana.

As nove amostras incluídas nos casos de investigação de vínculo biológico foram extraídas como segue: 1 disco de papel FTA foi cortado com o auxílio do *Harris Micro Punche e Cutting Mat* (Whatman) e depositado em microtubo de 1,5 mL, no qual adicionou-se 50 µL de água e incubou-se por 30 minutos à temperatura ambiente. Posteriormente, centrifugou-se a 13.000 rpm por 1 min e descartou-se a água. Adicionou-se 100 µL de Chelex 5%, incubando-se a 56 °C por 1 hora. Vortexou-se, incubou-se a 100 °C por 8 minutos, centrifugou-se a 13.000 rpm por 3 minutos e transferiu-se o sobrenadante para novo microtubo, o qual foi armazenado à 4 °C.

#### **4.4. Reação de amplificação dos X-STRs**

Neste trabalho foi realizada a amplificação de 10 X-STRs cuja localização citogenética e física foram obtidas em UCSC *Human Genome Browser* (<http://www.genome.ucsc.edu/cgi-bin/hgGateway>) – construção GRCh37/hg19 e estão apresentadas na Figura 4.

Figura 4: Idiograma do cromossomo X com a localização citogenética e física dos 10 X-STRs estudados neste trabalho.



Fonte: Este trabalho.

A estrutura de repetição dos 10 marcadores X-STR e as nomenclaturas alélicas utilizadas (conforme descrito por Gusmão et al., 2009) estão apresentadas na Tabela 1, a seguir.

Tabela 1: Estrutura de repetição e nomenclaturas utilizadas para os 10 X-STRs.

Marcador	Estrutura de repetição	Referência
DXS8378	(CTAT) <sub>n</sub>	Edelmann et al., 2002b
DXS9898	(TATC) <sub>2</sub> -(ATC)-(TATC) <sub>n</sub>	Hering; Szibor, 2000
DXS7133	(ATAG) <sub>n</sub>	Edelmann et al., 2002b
GATA31E08	(AGGG) <sub>2-3</sub> (AGAT) <sub>n</sub>	Gomes et al., 2009b
GATA172D05	(TAGA) <sub>n</sub>	Edelmann et al., 2002b
DXS7423	(TCCA) <sub>3</sub> -N <sub>8</sub> -(TCCA) <sub>n</sub>	Zarrabeitia et al., 2002a
DXS6809	(CTAT) <sub>n</sub> -(ATCT) <sub>3</sub> -N <sub>9</sub> -(TATC) <sub>n</sub> - (ATCT) <sub>n</sub> -N <sub>10</sub> -(ATCT) <sub>n</sub>	Edelmann et al., 2003
DXS7132	(TCTA) <sub>n</sub>	Edelmann et al., 2002b
DXS9902	(TAGA) <sub>n</sub>	Gusmão et al., 2009
DXS6789	(TATC) <sub>0-1</sub> (TATG) <sub>n</sub> -(TATC) <sub>n</sub>	Hering; Kuhlisch; Szibor, 2001

Obs: Não estão apresentadas as estruturas dos alelos intermediários.

As reações de amplificação foram realizadas em todas as amostras de DNA, através de uma reação *decaplex* que permitiu avaliar os 10 X-STRs propostos em uma única reação. Utilizou-se o termociclador GeneAmp<sup>®</sup> PCR System 9700 (Applied Biosystems by Life Technologies) e a reação foi preparada com pequenas modificações em relação à descrita por Gusmão et al. (2009), como segue.

5,0 µL *Gold STR 2X Buffer* (Promega)

1,0 µL *Primer Mix 10x* (inclui todos os *primers* a uma concentração de 2 µM - Tabela 2)

0,3 µL (1,5 U) de *Platinum<sup>®</sup> Taq DNA Polymerase* (Invitrogen)

Água ultrapura (Milli-Q) autoclavada qsp. 10 µl

1 disco de papel FTA com DNA ou 2,0 µL de solução resultante da extração com Chelex ou

0,5 µL das amostras de DNA padrão (9947A e 9948 a 1 ng/µL, Promega)



Ciclagem:

95 °C – 15 minutos

94 °C – 30 segundos }  
60 °C – 90 segundos } 10x  
72 °C – 60 segundos }

94 °C – 30 segundos }  
58 °C – 90 segundos } 20x  
72 °C – 60 segundos }  
60 °C – 60 minutos

As amostras padrão 9947A (Promega) e 9948 (Promega) foram utilizadas como controles na reação de amplificação.

Tabela 2: Sequência dos *primers* para os 10 X-STRs (adaptada de GUSMÃO et al., 2009).

STR	Orientação	Marcação/Sequência (5'- 3')	Referência
DXS8378	<i>forward</i>	6-FAM TTAGGCAACCCGGTGGTCC	Edelmann et al. (2002b)
	<i>reverse</i>	ACAAGAACGAAACTCCAATC	
DXS9898	<i>forward</i>	6-FAM CGAGCACACCTACAAAAGCTG	Gomes et al. (2007a)
	<i>reverse</i>	TAGGCTCACCTCACTGAGCA	
DXS7133	<i>forward</i>	6-FAM CACTTCCAAAAGGGGAAAAA	Gusmão et al. (2009)
	<i>reverse</i>	ACTTGTACTTGGTGGGAGGAA	
GATA31E08	<i>forward</i>	6-FAM GCAAGGGGAGAAGGCTAGAA	Gusmão et al. (2009)
	<i>reverse</i>	TCAGCTGACAGAGCACAGAGA	
GATA172D05	<i>forward</i>	VIC TAGTGGTGTGTTGCACAG	Edelmann et al. (2002b)
	<i>reverse</i>	ATAATTGAAAGCCCGGATTC	
DXS7423	<i>forward</i>	VIC GTCTTCCTGTCATCTCCCAAC	Edelmann et al. (2002b)
	<i>reverse</i>	TAGCTTAGCGCCTGGCACATA	
DXS6809	<i>forward</i>	VIC TCCATCTTTCTCTGAACCTTCC	Gomes et al. (2007a)
	<i>reverse</i>	TGCTTTAGGCTGATGTGAGG	
DXS7132	<i>forward</i>	NED TCCCCTCTCATCTATCTGACTG	Gomes et al. (2007a)
	<i>reverse</i>	CACTCCTGGTGCCAAACTCT	
DXS9902	<i>forward</i>	NED CTGGGTGAAGAGAAGCAGGA	Gusmão et al. (2009)
	<i>reverse</i>	GGCAATACACATTCATATCAGGA	
DXS6789	<i>forward</i>	NED CTTCATTATGTGCTGGGGTAAA	Gomes et al. (2007a)
	<i>reverse</i>	ACCTCGTGATCATGTAAGTTGG	

#### 4.5. Preparo da amostra para eletroforese

As 1001 amostras provenientes dos indivíduos da região sudeste foram submetidas à corrida eletroforética em analisador genético ABI377 (Applied Biosystems by Life Technologies) e, as amostras oriundas dos casos de determinação do vínculo biológico, à eletroforese capilar em ABI3500 (Applied Biosystems by Life Technologies), preparando-se as amostras de formas diferentes.

#### 4.5.1. Preparo de amostras para ABI377

0,5  $\mu$ L Padrão interno de corrida ROX500 (Applied Biosystems by Life Technologies)

1,5  $\mu$ L *Blue Dextran Loading Solution* (Promega)

1,0  $\mu$ L Produto de amplificação

As amostras foram submetidas à corrida eletroforética no analisador genético ABI377 (Applied Biosystems by Life Technologies) e 1,1  $\mu$ L foram aplicados no gel, o qual foi preparado conforme protocolo abaixo.

Gel poliacrilamida (40 mL):

14,5 g de Uréia (Promega)

0,4 g de Amberlite XAD-4 (Sigma)

4 mL de *Long Ranger Gel Solution 50%* (Cambrex)

20 mL Água ultrapura (Milli-Q) autoclavada

Estes reagentes foram homogeneizados em um agitador magnético por 10 minutos, seguindo-se de adição de 4 mL de Tris Borato EDTA 10X (0,020 M EDTA (Merck); 0,89 M Tris-base (USB); 0,89 M Ácido Bórico (USB); pH 8,0) e deaeração por 10 minutos. Posteriormente, ao gel deaerado, foram adicionados 200  $\mu$ L de Persulfato de Amônio 10% (Sigma-Aldrich) e 28  $\mu$ L de TEMED (Amresco). Com uma seringa de 100 mL aplicou-se o gel sobre o conjunto de placas, previamente limpas e montadas, que foram colocadas no analisador para a realização da corrida eletroforética.

#### 4.5.2. Preparo de amostras para ABI3500

0,3  $\mu$ L Padrão de corrida GeneScan 600 LIZ (Applied Biosystems by Life Technologies)

8,7  $\mu$ L Hi-Di *Formamide* (Applied Biosystems by Life Technologies)

1,0  $\mu$ L Produto de amplificação

Na eletroforese, utilizou-se o polímero POP-4 (Applied Biosystems by Life Technologies) e capilar de 36 cm (Applied Biosystems by Life Technologies).

#### 4.6. Determinação dos perfis alélicos e análise estatística

A análise dos *amplicons* para obtenção dos respectivos perfis alélicos foi realizada com o auxílio do programa GeneScan (Applied Biosystems by Life Technologies) no caso do ABI377 e GeneMapper<sup>®</sup> ID-X v1.2 (Applied Biosystems by Life Technologies) no caso do ABI3500.

As amostras padrão feminina 9947A (Promega) e masculina 9948 (Promega) foram utilizadas como referência para a determinação alélica e o perfil destas estão apresentados na Tabela 3 (adaptada de Gusmão et al., 2009).

Tabela 3: Perfil alélico das amostras de DNA padrão utilizadas como referência (adaptada de GUSMÃO et al., 2009).

<b>Marcador</b>	<b>9947A</b>	<b>9948</b>	<b>Referência</b>
DXS8378	10-11	11	Szibor et al. (2003b)
DXS9898	12-15	13	Gomes et al. (2007a)
DXS7133	9-10	11	Szibor et al. (2003b)
GATA31E08	13	12	Gusmão et al. (2009)
GATA172D05	10	6	Szibor et al. (2003b)
DXS7423	14-15	14	Szibor et al. (2003b)
DXS6809	31-34	31	Szibor et al. (2003b)
DXS7132	12	13	Szibor et al. (2003b)
DXS9902	12	13	Gusmão et al. (2009)
DXS6789	21-22	20	Szibor et al. (2003b)

Com o auxílio do programa Arlequin v. 3.11 (EXCOFFIER; LAVAL; SCHNEIDER, 2005), foram determinadas as frequências alélicas das populações estudadas e realizado o teste exato para avaliar o desequilíbrio de ligação e a divergência do equilíbrio de Hardy-Weinberg. Segundo Desmarais et al. (1998), com o programa Excel foram determinados os diversos parâmetros estatísticos de avaliação forense (chance média de exclusão em trios envolvendo filha – MECT, chance média de exclusão em duo pai/filha – MECD, poder de discriminação masculino – PDM e poder de discriminação feminino - PDF).

Além disso, para avaliar a existência ou não de diferença significativa dentre as populações estudadas e entre estas e outras relatadas na literatura, também com o programa Arlequin, foram realizados o teste exato de diferenciação de populações e o cálculo de distância genética ( $F_{ST}$ ).

#### 4.7. Análise estatística nos casos de relações biológicas

Os casos A, B e C, primeiramente foram analisados com marcadores AS-STRs e, posteriormente, com marcadores X-STRs.

A tipagem e a análise estatística para os AS-STRs foram efetuadas pelo Laboratório de Investigação de Paternidade da UNESP, utilizando-se o Kit PowerPlex 16 (Promega) e o programa Familias v.1.81 (EGELAND et al., 2000), respectivamente.

A tipagem de X-STRs foi feita como parte deste projeto de doutorado, conforme protocolo descrito anteriormente. O cálculo de *Likelihood Ratio* (LR) ou Razão de Verossimilhança foi realizado em planilha excel, segundo Pinto; Gusmão; Amorim (2011), utilizando-se os dados de frequência alélica obtidos neste trabalho para a população de São Paulo.

#### 4.8. Elaboração do banco de dados brasileiro para X-STRs

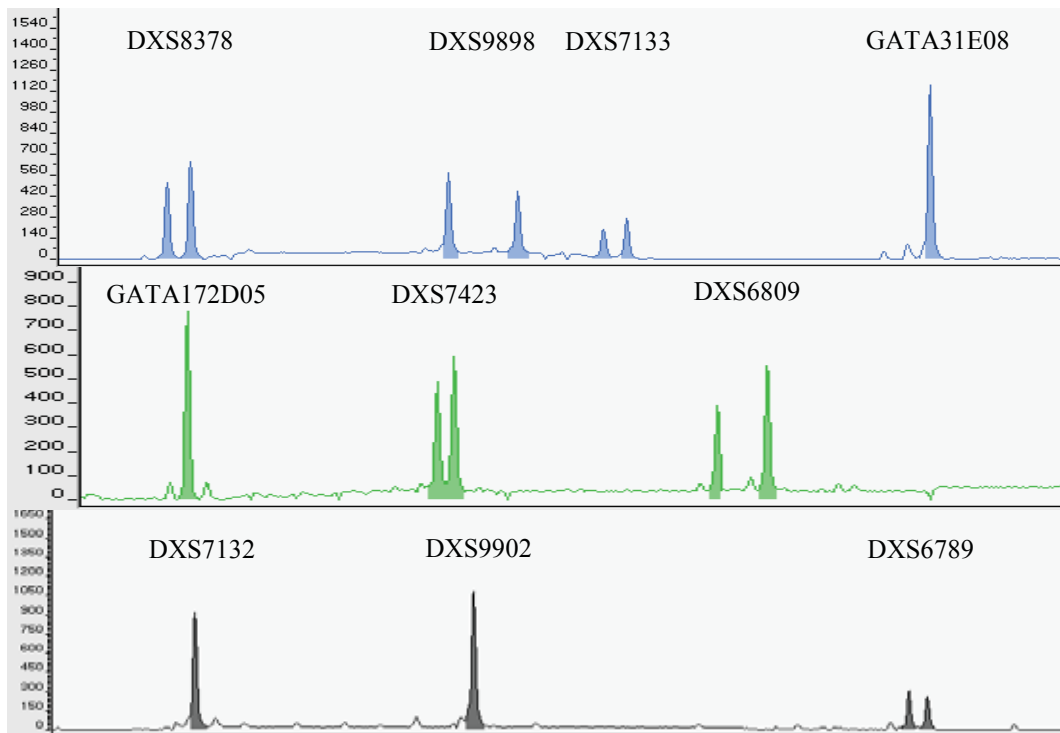
O Banco foi elaborado em parceria com Alexandre E. Cardoso, Programador e *Webdesigner*, autônomo, residente em Araraquara-SP. O sistema foi desenvolvido na linguagem PHP (*Hypertext Preprocessor*) com gerenciador MySQL (*Structured Query Language*), ambos gratuitos. Uma das características marcantes do PHP é a operacionalidade do script PHP no lado do servidor, sendo enviado para o cliente apenas HTML. O gerenciador MySQL é um dos mais poderosos sistemas gerenciadores de banco de dados, caracterizado por ser rápido, *multi-thread* (capacidade de um processo efetuar várias operações simultaneamente) e multi-usuário. Com a finalidade de garantir uma maior interatividade e facilidade para o usuário, utilizou-se o *Framework JavaScript* (Jquery).

## **5. RESULTADOS E DISCUSSÃO**

### 5.1. Avaliação da metodologia

Os eletroferogramas dos perfis alélicos obtidos em analisadores genético ABI377 e ABI3500 estão apresentados nas Figuras 5 e 6, respectivamente. Pode-se notar que a metodologia foi eficaz, uma vez que todos os marcadores foram bem amplificados, obtendo-se uma precisa determinação alélica.

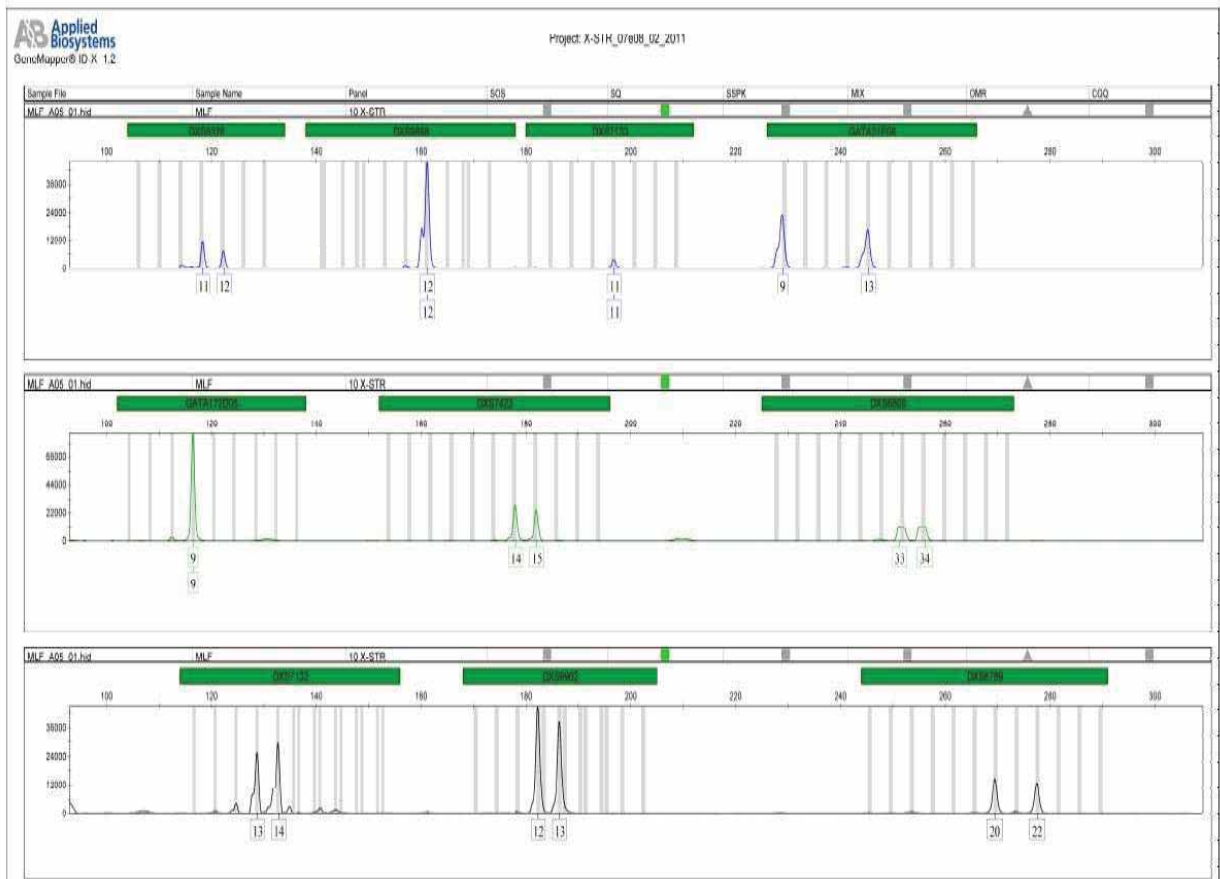
Figura 5: Eletroferograma com perfil alélico de amostra feminina obtido em ABI377. Marcação 6-FAM (painel azul: marcadores DXS8378, DXS9898, DXS7133, GATA31E08), VIC (painel verde: marcadores GATA172D05, DXS7423, DXS6809) e NED (painel preto: DXS7132, DXS9902, DXS6789).



Fonte: Este trabalho.



Figura 6: Eletroferograma com perfil alélico de amostra feminina obtido em ABI3500. Marcação 6-FAM (painel azul: marcadores DXS8378, DXS9898, DXS7133, GATA31E08), VIC (painel verde: marcadores GATA172D05, DXS7423, DXS6809) e NED (painel preto: DXS7132, DXS9902, DXS6789).



Fonte: Este trabalho.

## 5.2 Variação genética

A frequência alélica para os 10 X-STRs foi avaliada separadamente por gênero. Os valores de  $p$  do teste exato para diferenciação das populações variou entre 0,0178 para o DXS7132 no Rio de Janeiro e 0,9752 para DXS9902 em São Paulo, conforme dados da Tabela 4. Após correção de Bonferroni para múltiplas análises (significante nível de 0,005), não foi observada diferença significativa entre a frequência alélica de homens e mulheres para

qualquer um dos *loci* estudados. Assim, a frequência alélica e diversidade gênica dos 10 X-STRs em São Paulo, Rio de Janeiro, Vitória e Belo Horizonte foram calculadas considerando-se a amostra total (Tabela 5).

Tabela 4: Teste exato para diferenciação das populações masculina e feminina de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.

	São Paulo	Rio de Janeiro	Vitória	Belo Horizonte
DXS8378	0,3940	0,9705	0,5880	0,6159
DXS9898	0,7725	0,0998	0,1057	0,9397
DXS7133	0,5686	0,9467	0,1416	0,0387
GATA31E08	0,9545	0,5028	0,8588	0,5730
GATA172D05	0,3020	0,4276	0,5163	0,9404
DXS7423	0,7987	0,4251	0,4844	0,6559
DXS6809	0,4831	0,5041	0,2644	0,5834
DXS7132	0,5077	0,0178	0,4718	0,1266
DXS9902	0,9752	0,9317	0,1602	0,6638
DXS6789	0,5175	0,4081	0,2807	0,2428

Tabela 5: Frequência alélica e diversidade gênica dos 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.

	São Paulo	Rio de Janeiro	Vitória	Belo Horizonte
<b>DXS8378</b>				
N	414	406	405	353
9	0,0121	0,0099	0,0049	0,0027
10	0,3986	0,3128	0,3062	0,3514
11	0,3333	0,3941	0,3704	0,3286
12	0,2367	0,2488	0,2889	0,2919
13	0,0193	0,0320	0,0272	0,0198
14		0,0025	0,0025	0,0057
DG	0,6751	0,6855	0,6866	0,6849
<b>DXS9898</b>				
N	414	406	405	353

7	0,0024	0,0025	0,0049	0,0085
8.3	0,1305	0,1798	0,2049	0,1726
9			0,0025	
10	0,0266	0,0394	0,0346	0,0454
11	0,1908	0,1970	0,1852	0,1871
12	0,3430	0,2611	0,2790	0,3115
13	0,2077	0,2315	0,1975	0,2125
13.3	0,0024	0,0074	0,0074	0,0055
14	0,0845	0,0764	0,0617	0,0453
15	0,0121	0,0049	0,0222	0,0115
DG	0,7796	0,8016	0,8032	0,7909
<b>DXS7133</b>				
N	412	406	401	351
7	0,0072	0,0049	0,0025	
8	0,0097	0,0025	0,0050	0,0057
9	0,4999	0,3103	0,3666	0,3250
10	0,1408	0,2069	0,1746	0,1679
11	0,2913	0,3892	0,3766	0,4243
12	0,0291	0,0591	0,0474	0,0456
13	0,0097	0,0172	0,0150	0,0171
14	0,0122	0,0099	0,0075	0,0144
15			0,0050	
DG	0,6457	0,7073	0,6925	0,6855
<b>GATA31E08</b>				
N	414	406	405	353
7	0,0024	0,0074	0,0074	0,0115
8	0,0048	0,0123	0,0049	0,0084
9	0,1353	0,1626	0,1827	0,2010
10	0,0362	0,0813	0,0543	0,0765
11	0,1715	0,1453	0,1185	0,1162
12	0,2705	0,2488	0,2765	0,2239
13	0,2633	0,2291	0,2296	0,2633
14	0,1039	0,1010	0,1136	0,0823
15	0,0121	0,0123	0,0123	0,0169
DG	0,7994	0,8230	0,8093	0,8158
<b>GATA172D05</b>				
N	414	406	405	353
6	0,1473	0,1897	0,1481	0,1531
7	0,0169	0,0222	0,0272	0,0369
8	0,1546	0,1724	0,1531	0,1246
9	0,1256	0,1576	0,1407	0,1843
10	0,2730	0,2685	0,2963	0,2069
11	0,2126	0,1601	0,1605	0,1953
12	0,0701	0,0271	0,0716	0,0965
13		0,0025	0,0025	0,0028
DG	0,8157	0,8125	0,8174	0,8378
<b>DXS7423</b>				
N	414	406	405	353

8	0,0048	0,0074		
10	0,0024			
12		0,0049	0,0049	0,0057
13	0,0555	0,0714	0,0494	0,0595
14	0,3309	0,3596	0,3531	0,3851
15	0,4855	0,3990	0,4321	0,3684
16	0,0918	0,1158	0,1259	0,1190
17	0,0290	0,0419	0,0346	0,0623
DG	0,6440	0,6929	0,6708	0,6963
<b>DXS6809</b>				
N	414	406	405	353
27	0,0024			0,0057
28	0,0072	0,0049	0,0198	0,0142
29	0,0145	0,0320	0,0321	0,0508
30	0,0531	0,0517	0,0716	0,0313
30.1		0,0025		
31	0,1570	0,1502	0,1506	0,1274
31.1	0,0048	0,0049	0,0074	
32	0,1377	0,1429	0,1259	0,1642
32.1	0,0024			0,0057
33	0,3043	0,2586	0,2914	0,2661
33.1		0,0025		
34	0,1932	0,1970	0,1605	0,2154
35	0,0725	0,0985	0,1012	0,0766
35.1		0,0025		
36	0,0362	0,0320	0,0370	0,0227
37	0,0072	0,0123	0,0025	0,0199
38	0,0072	0,0025		
39		0,0049		
DG	0,8187	0,8387	0,8346	0,8313
<b>DXS7132</b>				
N	412	406	405	353
10	0,0048			
11	0,0194	0,0074	0,0222	0,0169
12	0,1044	0,0887	0,1012	0,0850
13	0,2233	0,2512	0,2642	0,2406
14	0,3277	0,3768	0,3012	0,3854
14.3		0,0025		
15	0,2597	0,1946	0,2198	0,2125
15.3	0,0024	0,0025		
16	0,0388	0,0517	0,0519	0,0426
16.3	0,0072	0,0197	0,0173	0,0058
17	0,0097		0,0173	0,0112
17.3	0,0024	0,0049		
18			0,0025	
18.3			0,0025	
DG	0,7642	0,7478	0,7790	0,7410
<b>DXS9902</b>				

N	414	406	405	353
8	0,0024		0,0025	
9	0,0072	0,0246	0,0222	0,0170
10	0,0314	0,0443	0,0222	0,0395
11	0,3840	0,3251	0,3457	0,3230
11.1		0,0025		
12	0,3285	0,3719	0,3679	0,3400
12.1	0,0072	0,0074	0,0148	0,0055
13	0,2295	0,2044	0,2198	0,2605
13.1	0,0048	0,0049		
14	0,0048	0,0148	0,0049	0,0087
15				0,0028
16				0,0028
DG	0,6925	0,7131	0,6973	0,7122
<b>DXS6789</b>				
N	414	406	405	353
14	0,0121	0,0074	0,0074	0,0028
15	0,0821	0,1034	0,1012	0,1217
16	0,1376	0,0616	0,0568	0,0680
17	0,0097	0,0025	0,0025	0,0058
18	0,0048	0,0049	0,0049	0,0084
19	0,0362	0,0493	0,0593	0,0426
20	0,3019	0,2906	0,3407	0,3114
21	0,2005	0,2438	0,2494	0,2209
22	0,1667	0,1773	0,1309	0,1302
23	0,0362	0,0493	0,0420	0,0766
24	0,0121	0,0074	0,0049	0,0112
25		0,0025		
DG	0,8141	0,8071	0,7877	0,8122

N, número de cromossomos analisados (mulheres somam 2 cromossomos e, homens , 1 cromossomo);  
DG, diversidade gênica.

Considerando-se as quatro populações, 6 a 18 alelos foram observados em cada *locus* com um total de 106 alelos para os 10 X-STRs. Os valores de diversidade gênica foram acima de 64% com o valor médio entre as populações variando de 68% para DXS7423, DXS7133 e DXS8378, a 83% para DXS6809. Rio de Janeiro apresentou a maior média de diversidade gênica (76,2%).

Em relação ao sexo masculino, considerando-se os 10 X-STRs, nenhum perfil foi idêntico dentre ou entre os homens de São Paulo-SP, Rio de Janeiro-RJ, Vitória-ES e Belo Horizonte-MG.

### 5.3. Parâmetros estatísticos de avaliação forense

Conforme citado por Szibor et al. (2003a), segundo as recomendações da ISFH - *International Society of Forensic Haemogenetics* (BÄR et al., 1997) - para a aplicação forense de marcadores microssatélites, há a necessidade destes possuírem propriedades genético-populacionais adequadas: equilíbrio de Hardy-Weinberg, alto nível de polimorfismo, desequilíbrio de ligação conhecido, entre outros.

Assim, diversos parâmetros estatísticos para avaliação forense dos 10 X-STRs foram determinados e estão apresentados nas Tabelas de 6 a 10.

Tabela 6: Heterozigose esperada e observada e teste exato para o equilíbrio de Hardy-Weinberg calculados para os 10 X-STRs na população feminina de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.

	São Paulo			Rio de Janeiro		
	He	Ho	P-HWE	He	Ho	P-HWE
DXS8378	0,6773	0,7012	0,9636	0,6889	0,7034	0,9730
DXS9898	0,7789	0,7927	0,4624	0,7990	0,7103	0,1143
DXS7133	0,6566	0,5644	0,0057	0,7162	0,6690	0,0096
GATA31E08	0,8023	0,7561	0,0062	0,8212	0,8138	0,1755
GATA172D05	0,8232	0,8232	0,2342	0,8043	0,7931	0,8085
DXS7423	0,6481	0,6890	0,4784	0,6760	0,7103	0,4441
DXS6809	0,8170	0,7988	0,0085	0,8414	0,8276	0,7654
DXS7132	0,7624	0,7669	0,1530	0,7460	0,6965	0,3787
DXS9902	0,6962	0,6829	0,0719	0,7203	0,6207	0,0256
DXS6789	0,8254	0,8049	0,3450	0,8031	0,8069	0,9098
	Vitória			Belo Horizonte		
	He	Ho	P-HWE	He	Ho	P-HWE
DXS8378	0,6859	0,6875	0,3611	0,6826	0,6296	0,0242
DXS9898	0,7989	0,8062	0,9734	0,7941	0,7870	0,8182
DXS7133	0,7039	0,6478	0,1217	0,6886	0,6667	0,4038
GATA31E08	0,8094	0,8000	0,2870	0,8124	0,8148	0,9898

GATA172D05	0,8131	0,7687	0,1047	0,8394	0,8796	0,8630
DXS7423	0,6605	0,6812	0,0264	0,6982	0,6574	0,2393
DXS6809	0,8260	0,8062	0,5918	0,8428	0,8611	0,9116
DXS7132	0,7760	0,7250	0,0774	0,7276	0,7222	0,6497
DXS9902	0,6971	0,6437	0,3349	0,7060	0,6852	0,5685
DXS6789	0,7861	0,7750	0,4480	0,8198	0,8241	0,3875

He, heterozigose esperada; Ho, heterozigose observada; P-HWE, valor de  $p$  para o equilíbrio de Hardy-Weinberg.

Se uma população está em equilíbrio de Hardy-Weinberg, presume-se que esta é infinita, que não há eventos de mutação e seleção, que os cruzamentos são aleatórios e que o fluxo gênico existente não é capaz de alterar a composição alélica desta população. Assim, as frequências de seus alelos permaneceriam inalteradas ao longo das gerações e as proporções genóticas atingiriam um equilíbrio estável (BEIGUELMAN, 1996). Embora as condições estabelecidas para a obtenção do equilíbrio de Hardy-Weinberg não sejam satisfeitas completamente por nenhuma população real, na prática, a maioria dos estudos realizados demonstra uma distribuição genotípica de acordo com o Equilíbrio de Hardy-Weinberg.

Em nosso estudo, o desequilíbrio ( $p \leq 0,05$ ) foi observado para os marcadores DXS7133, GATA31E08 e DXS6809 em São Paulo; DXS7133 e DXS9902 no Rio de Janeiro; DXS7423 em Vitória e DXS8378 em Belo Horizonte. A presença de alelo nulo já foi reportada para o marcador DXS7133 na população de São Paulo e Norte de Portugal (GUSMÃO et al., 2009) e pode ter ocasionado o aumento de homozigotos neste marcador. No caso dos marcadores DXS6809 e DXS9902 o desequilíbrio pode ser decorrente da presença de alelos raros nestes *loci*. O DXS8378 e DXS7423 foram os menos polimórficos nas respectivas populações em que o desequilíbrio foi observado, porém este não pode ser justificado pela presença de alelo nulo, uma vez que não há excesso de homozigotos e não tem sido descrito alelo nulo na literatura para tais *loci*. Assim, um estudo mais aprofundado

pode ser realizado para uma melhor avaliação. Após correção de Bonferroni (nível de significância de 0,005), no entanto, nenhum desequilíbrio foi obtido.

Tabela 7: Chance média de exclusão em trios envolvendo filha (MECT) para os 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.

	São Paulo	Rio Janeiro	Vitória	Belo Horizonte
<b>DXS8378</b>	0,6076	<b>0,6214</b>	0,6201	<b>0,6166</b>
DXS9898	0,7458	0,7696	0,7723	0,7575
DXS7133	0,5876	0,6528	0,6342	0,6278
GATA31E08	0,7680	0,7967	0,7803	0,7880
<b>GATA172D05</b>	0,7879	0,7833	0,7911	<b>0,8137</b>
<b>DXS7423</b>	<b>0,5822</b>	0,6380	<b>0,6103</b>	0,6416
<b>DXS6809</b>	<b>0,7940</b>	<b>0,8169</b>	<b>0,8133</b>	0,8079
DXS7132	0,7246	0,7071	0,7430	0,6983
DXS9902	0,6314	0,6600	0,6379	0,6551
DXS6789	0,7880	0,7790	0,7576	0,7861

Verde: maiores valores; Vermelho: menores valores.

Tabela 8: Chance média de exclusão em duo pai/filha (MECD) para os 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.

	São Paulo	Rio Janeiro	Vitória	Belo Horizonte
<b>DXS8378</b>	0,4606	<b>0,4749</b>	0,4734	<b>0,4696</b>
DXS9898	0,6146	0,6434	0,6472	0,6290
DXS7133	0,4408	0,5089	0,4896	0,4827
GATA31E08	0,6417	0,6782	0,6572	0,6672
<b>GATA172D05</b>	0,6666	0,6604	0,6709	<b>0,6999</b>
<b>DXS7423</b>	<b>0,4365</b>	0,4943	<b>0,4651</b>	0,4980
<b>DXS6809</b>	<b>0,6758</b>	<b>0,7057</b>	<b>0,7009</b>	0,6939
DXS7132	0,5898	0,5696	0,6118	0,5593
DXS9902	0,4858	0,5171	0,4930	0,5111
DXS6789	0,6678	0,6565	0,6303	0,6657

Verde: maiores valores; Vermelho: menores valores.



Considerando-se uma situação de paternidade com o trio (mãe, filha e suposto pai), o MECT indica a probabilidade média de um indivíduo qualquer da população e não relacionado à criança testada ser excluído da paternidade. O mesmo é válido para o MECD, porém nesta situação tem-se apenas pai e filha. Para o sistema *decaplex* estudado, os valores combinados de MECT foram entre 0,999996 e 0,999998 e para o MECD ficou entre 0,99983 e 0,99989.

Considerando uma criança com genótipo hipotético AB, como no MECD trabalha-se na ausência da mãe, não é possível determinar quais alelos esta doou para a filha, portanto, o pai da criança poderá ser qualquer homem da população que tenha o alelo A ou B presentes nesta. Assim, para os *loci* em que a filha é heterozigota, aumentam-se os pais possíveis e, portanto, diminui-se o MEC.

Tabela 9: Poder de discriminação masculino (PDM) para os 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.

	São Paulo	Rio Janeiro	Vitória	Belo Horizonte
<b>DXS8378</b>	0,6735	<b>0,6838</b>	0,6849	<b>0,6829</b>
DXS9898	0,7778	0,7996	0,8013	0,7886
DXS7133	0,6442	0,7055	0,6908	0,6836
GATA31E08	0,7975	0,8209	0,8073	0,8135
<b>GATA172D05</b>	0,8137	0,8105	0,8154	<b>0,8355</b>
<b>DXS7423</b>	<b>0,6424</b>	0,6911	<b>0,6691</b>	0,6944
<b>DXS6809</b>	<b>0,8167</b>	<b>0,8366</b>	<b>0,8326</b>	0,8290
DXS7132	0,7624	0,7460	0,7771	0,7389
DXS9902	0,6908	0,7113	0,6956	0,7102
DXS6789	0,8122	0,8051	0,7857	0,8099

Verde: maiores valores; Vermelho: menores valores.

Tabela 10: Poder de discriminação feminino (PDF) para os 10 X-STRs na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.

	São Paulo	Rio Janeiro	Vitória	Belo Horizonte
<b>DXS8378</b>	0,8275	<b>0,8376</b>	0,8359	<b>0,8331</b>
DXS9898	0,9186	0,9298	0,9315	0,9242
DXS7133	0,8168	0,8606	0,8478	0,8441
GATA31E08	0,9295	0,9437	0,9358	0,9397
<b>GATA172D05</b>	0,9395	0,9368	0,9417	<b>0,9512</b>
<b>DXS7423</b>	<b>0,8119</b>	0,8515	<b>0,8317</b>	0,8538
<b>DXS6809</b>	<b>0,9438</b>	<b>0,9536</b>	<b>0,9527</b>	0,9497
DXS7132	0,9058	0,8966	0,9162	0,8912
DXS9902	0,8450	0,8654	0,8496	0,8609
DXS6789	0,9405	0,9359	0,9259	0,9401

Verde: maiores valores; Vermelho: menores valores.

O Poder de Discriminação corresponde à probabilidade de se escolher dois indivíduos ao acaso na população e estes apresentarem genótipos diferentes entre si para um determinado marcador. Neste trabalho, ao avaliar a eficiência do sistema como um todo, o poder de discriminação feminino combinado variou entre 0,99999999990 e 0,99999999997 e, o poder de discriminação masculino combinado ficou ente 0,99999991 e 0,99999995. Os valores do poder de discriminação são menores em homens pelo fato destes apresentarem apenas um cromossomo X.

Em relação aos parâmetros estatísticos calculados, os valores estimados foram semelhantes para todos os grupos, com o Rio de Janeiro sendo o mais diverso, seguido por Belo Horizonte, Vitória e São Paulo.

O marcador DXS6809 foi o mais discriminativo em São Paulo, Rio de Janeiro e Vitória, enquanto o marcador GATA172D05 foi o mais discriminativo em Belo Horizonte. DXS8378 foi o menos discriminativo no Rio de Janeiro e Belo Horizonte, assim como o

DXS7423 foi o menos discriminativo em São Paulo e Vitória. No geral, o mais diverso *locus* foi o DXS6809 enquanto DXS8378, DXS7133 e DXS7423 foram os menos polimórficos, sendo que a ordem destes últimos variou entre as populações brasileiras comparadas, de acordo com dados obtidos por Gusmão et al. (2009).

Os parâmetros calculados foram comparados àqueles reportados por outras populações na literatura e a maioria dos marcadores foram mais discriminativos nas populações brasileiras aqui estudadas do que em populações Ibéricas de Portugal e Espanha e Latino Americanas de Argentina, Costa Rica e Colômbia (GUSMÃO et al., 2009). Por outro lado, estes parâmetros apresentaram valores mais altos em populações africanas de Angola, Moçambique e Uganda (GOMES et al. 2007b; GOMES et al., 2009a), o que era esperado, uma vez que, pela análise de marcadores autossômicos, do cromossomo Y e do DNA mitocondrial, populações africana subsaarianas são conhecidas por serem o grupo humano mais diversificado (GOMES et al., 2009a). No caso do Brasil, os altos índices obtidos decorrem da sua história etnológica, a qual consiste em cinco séculos de mistura entre populações de origem européia, africana e ameríndia, principalmente.

Assim, os dados obtidos confirmam a poderosa aplicação do sistema *decaplex* para identificação humana na população de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte.

### 5.3.1. Desequilíbrio de Ligação entre os marcadores X-STR

O teste exato para avaliação de LD foi realizado para todos os marcadores estudados, utilizando-se os dados da amostra masculina, e estão apresentados na Tabela 11.

Tabela 11: Valor de  $p$  obtido no teste de desequilíbrio de ligação para todos os pares de marcadores analisados na população de São Paulo, Rio de Janeiro, Vitória, Belo Horizonte e Sudeste.

Par de STR/N	São Paulo	Rio de Janeiro	Vitória	Belo Horizonte	Sudeste
	86	116	85	137	424
DXS8378 - DXS9898	0,9773	0,2208	0,3199	0,6587	0,1320
DXS8378 - DXS7133	0,8943	0,3910	0,3783	0,3802	0,2261
DXS9898 - DXS7133	0,1193	0,5978	0,5261	0,7459	0,1897
DXS8378 – GATA31E08	0,1989	0,7620	0,8250	0,4728	0,7536
DXS9898 – GATA31E08	0,6296	0,6578	0,3386	0,8884	0,9743
DXS7133 – GATA31E08	0,7650	0,0758	0,8844	0,1101	0,0113
DXS8378 – GATA172D05	0,8674	0,3034	0,6814	0,9263	0,6698
DXS9898 - GATA172D05	0,0070	0,0686	0,6821	0,9946	0,6254
DXS7133 - GATA172D05	0,1526	0,0123	0,7871	0,6293	0,0044
GATA31E08 - GATA172D05	0,4119	0,3696	0,8612	0,7975	0,8516
DXS8378 – DXS7423	0,8850	0,5195	0,3260	0,6277	0,1281
DXS9898 – DXS7423	0,1194	0,9227	0,4326	0,5813	0,2152
DXS7133 - DXS7423	0,1049	0,1309	0,3985	0,1272	0,0428
GATA31E08 - DXS7423	0,6450	0,5278	0,4402	0,1719	0,4707
GATA172D05 - DXS7423	0,5151	0,8124	0,7580	0,7471	0,6504
DXS8378 – DXS6809	0,6941	0,1692	0,6024	0,4657	0,0799
DXS9898– DXS6809	0,1648	0,5573	0,5647	0,4322	0,2836
DXS7133– DXS6809	0,1931	0,6644	0,6595	0,8400	0,6280
GATA31E08– DXS6809	0,9964	0,4146	0,8026	0,0611	0,7285
GATA172D05– DXS6809	0,4956	0,0646	0,8373	0,9644	0,7084
DXS7423– DXS6809	0,5455	0,8707	0,9525	0,0548	0,5551
DXS8378 - DXS7132	0,4911	0,8208	0,2227	0,9957	0,3081
DXS9898 - DXS7132	0,6907	0,9841	0,6679	0,6180	0,2675
DXS7133 - DXS7132	0,3126	0,1192	0,7409	0,9970	0,5323
GATA31E08 - DXS7132	0,7860	0,5392	0,1258	0,5869	0,2116
GATA172D05 - DXS7132	0,9737	0,6371	0,6500	0,6004	0,9416
DXS7423 - DXS7132	0,6532	0,9050	0,2089	0,9913	0,9619
DXS6809 - DXS7132	0,7068	0,7324	0,9681	0,9999	0,8350
DXS8378 – DXS9902	0,2270	0,7094	0,9655	0,5016	0,9827
DXS9898 – DXS9902	0,7828	0,5957	0,8207	0,0000*	0,0485
DXS7133 – DXS9902	0,3914	0,7845	0,9850	0,2039	0,3771
GATA31E08 – DXS9902	0,0259	0,8398	0,4804	0,2841	0,6984
GATA172D05 – DXS9902	0,0474	0,4862	0,2725	0,0691	0,2652

DXS7423 – DXS9902	0,4650	0,9149	0,2597	0,4375	0,6371
DXS6809 – DXS9902	0,1746	0,3646	0,3410	0,9188	0,4033
DXS7132 – DXS9902	0,3550	0,0540	0,3142	0,9992	0,9374
DXS8378 – DXS6789	0,0409	0,4414	0,5379	0,7298	0,7511
DXS9898 – DXS6789	0,1608	0,1355	0,0760	0,6864	0,0630
DXS7133 – DXS6789	0,2526	0,0401	0,1105	0,1551	0,0019
GATA31E08 – DXS6789	0,0895	0,8098	0,1989	0,0011	0,1092
GATA172D05 – DXS6789	0,4916	0,5507	0,3914	0,0105	0,7894
DXS7423 – DXS6789	0,0309	0,8411	0,9512	0,0704	0,0885
DXS6809 – DXS6789	0,5792	0,0064	0,3971	0,8903	0,0276
DXS7132 – DXS6789	0,1129	0,0067	0,7251	0,3409	0,2261
DXS9902 – DXS6789	0,7265	0,5903	0,5039	0,1028	0,0441

N, número de homens analisados. \*Valor de  $p$  significante para o nível de 0,0011, após correção de Bonferroni.

Após correção de Bonferroni (valor de  $p$  significativo para nível de 0,0011), apenas um valor de  $p$  significativo foi obtido para o par de marcadores DXS9898-DXS9902 em Belo Horizonte ( $p \leq 0,0000$ ), os quais estão muito distantes no cromossomo X (mais de 72 Mb e 68 cM, [www.bgbx.com.br](http://www.bgbx.com.br)). O teste de LD na amostra total da região Sudeste não confirmou a associação entre qualquer um dos pares de marcadores investigados.

LD tem mostrado ser específico para cada população. Embora DXS6809 e DXS6789 sejam marcadores intimamente ligados (distantes por 511 Kb e 0,35 cM) e um significativo LD tem sido observado entre esses X-STRs na população germânica (SZIBOR et al., 2005), nenhuma associação foi detectada nesse estudo e em várias outras populações (ROBINO et al., 2006; GOMES et al., 2007a; GOMES et al., 2007b; GOMES et al., 2009a; GUSMÃO et al., 2009; BOBILHO et al., 2011). Tal dado pode ser explicado pelo fato de que a pequena taxa de recombinação existente entre estes, ao longo das gerações, pode restabelecer o equilíbrio de ligação. Assim, para tais marcadores, nas quatro populações estudadas, a frequência haplotípica pode ser obtida através do produto das respectivas frequências alélicas.

Como já mencionado, LD não é decorrente apenas da distância genética entre os STRs, podendo ser ocasionado por vários outros fatores como: seleção, mutação, deriva aleatória, efeito fundador, miscigenação ou subdivisão populacional (CHAKRAVARTI, 1999; ZARRABEITIA et al., 2006). A população brasileira é muito heterogênea (formada principalmente por uma mistura de Europeus, Africanos e Ameríndios), porém, o desequilíbrio observado entre DXS9898-DXS9902 em Belo Horizonte parece estar mais relacionado a efeitos de amostragem do que à subestrutura genética, uma vez que esta associação não foi observada nas outras três miscigenadas populações estudadas neste trabalho e em outras populações brasileiras (GUSMÃO et al., 2009).

O trabalho mais recente publicado com dados de X-STRs na população brasileira apresenta cinco pares de marcadores em desequilíbrio de ligação (DXS9898/DXS9895 no Pará, DXS7423/GATA31E08 no Ceará, DXS6789/DXS6800 no Rio Grande do Sul e GATA172D05/DXS10011 e DXS10011/DXS9895 no Maranhão), os quais estão a uma distância de 10 Mb (DXS7423/GATA31E08) ou mais. Ao agrupar as amostras por região geopolítica, também nenhum desequilíbrio foi detectado (RIBEIRO-RODRIGUES et al., 2011).

#### **5.4. Comparação entre populações**

No caso dos X-STRs, como a mulher possui 2 cromossomos X e o homem apenas um, em relação aos marcadores autossômicos, o tamanho populacional efetivo é menor e, assim, uma deriva genética mais rápida e uma estrutura populacional mais forte é esperada para o cromossomo X (SCHAFFNER, 2004). Desta forma, pesquisas são necessárias para avaliar a extensão dos polimorfismos em diferentes populações e para estabelecer bancos de dados úteis para aplicação forense e estudos antropológicos. Como a população do Brasil é

altamente miscigenada, comparações foram realizadas entre os grupos estudados neste trabalho (São Paulo, Rio de Janeiro, Vitória e Belo Horizonte) e outras populações brasileiras (Paraná e Mato Grosso do Sul) (GUSMÃO et al., 2009) que analisaram o mesmo conjunto de X-STRs. As comparações foram feitas aos pares para simples *locus* (teste exato para diferenciação das populações e distância genética  $F_{ST}$ , Tabela 12) e para nível haplotípico (distância genética  $F_{ST}$ , Tabela 13).

Tabela 12: Comparação para simples *locus* entre populações brasileiras. Distância genética  $F_{ST}$  (1ª linha) e correspondente valor de  $p$  (2ª linha) e  $p$  valor do teste exato para diferenciação das populações (3ª linha).

	SP - RJ	SP - V	SP - BH	SP - MS	SP - PR	RJ - V	RJ - BH	RJ - MS
DXS8378	0,00588	0,00687	0,00135	-0,00050	-0,00011	-0,00082	0,00305	0,00887
	0,03418	0,01660	0,23535	0,42773	0,35840	0,53711	0,11035	0,00977
	0,09366	0,04871	0,18943	0,28669	0,41420	0,78661	0,19595	0,06246
DXS9898	0,00391	0,00418	0,00041	0,00091	0,01116	-0,00072	-0,00005	0,00081
	0,04004	0,02930	0,34277	0,25098	0,00098*	0,59668	0,37793	0,27051
	0,12445	0,07475	0,21299	0,40382	0,00072*	0,47364	0,48802	0,08767
DXS7133	0,03402	0,01718	0,03355	0,01075	-0,00017	0,00077	-0,00036	0,00344
	0,00000*	0,00000*	0,00000*	0,00488*	0,39062	0,25586	0,43652	0,08887
	0,00000*	0,00913	0,00002*	0,00585	0,34421	0,65885	0,64546	0,31913
GATA31E08	0,00076	0,00170	0,00466	0,00149	0,00054	-0,00073	0,00014	0,00182
	0,24219	0,13770	0,02246	0,17773	0,29590	0,62598	0,37793	0,15820
	0,12993	0,20781	0,00589	0,22073	0,45558	0,65586	0,65643	0,24769
GATA172D05	0,00234	-0,00024	0,00349	0,00032	0,00405	0,00070	0,00604	0,00167
	0,10352	0,50195	0,03906	0,32422	0,03516	0,25781	0,00488*	0,15723
	0,01923	0,56507	0,02795	0,23844	0,00569	0,09517	0,00025*	0,01339
DXS7423	0,00452	0,00108	0,01119	0,01129	0,00422	-0,00112	-0,00104	-0,00007
	0,04004	0,19824	0,00488*	0,00293*	0,07227	0,64844	0,57617	0,38281
	0,15128	0,24677	0,00187*	0,00117*	0,04026	0,50261	0,54674	0,35472
DXS6809	-0,00045	-0,00054	0,00091	-0,00016	0,00151	-0,00038	-0,00085	-0,00110
	0,52832	0,57227	0,23047	0,42188	0,19434	0,49023	0,66992	0,69336
	0,70543	0,31205	0,02356	0,73947	0,30107	0,30971	0,25459	0,57519
DXS7132	0,00302	0,00038	0,00154	-0,00131	0,00262	0,00228	-0,00194	-0,00029
	0,07324	0,28613	0,18164	0,67871	0,12109	0,12695	0,89844	0,45508
	0,04033	0,43969	0,64037	0,31642	0,01110	0,02488	0,23026	0,02084
DXS9902	0,00223	0,00008	0,00098	-0,00127	0,00485	-0,00151	0,00041	-0,00073
	0,13867	0,33594	0,24121	0,61914	0,05273	0,75684	0,31055	0,47949
	0,16140	0,33027	0,41107	0,41707	0,00960	0,34468	0,56783	0,47558
DXS6789	0,00299	0,00550	0,00354	0,00380	0,00659	0,00060	0,00007	0,00358

0,04199	0,00684	0,03613	0,03516	0,00879	0,27930	0,37012	0,05078
0,03370	0,00280*	0,00805	0,08582	0,02999	0,85751	0,59307	0,41811

Populações estudadas neste trabalho: SP, São Paulo; RJ, Rio de Janeiro; V, Vitória; BH, Belo Horizonte. Populações estudadas por Gusmão et al. (2009): MS, Mato Grosso do Sul; PR, Paraná. \* Valor de  $p$  significativa para o nível de 0,005, após correção de Bonferroni.

Tabela 12: continuação

	RJ - PR	V - BH	V - MS	V - PR	BH - MS	BH - PR	MS - PR
DXS8378	0,00995	0,00016	0,00631	0,00707	-0,00094	-0,00060	-0,00316
	0,00488*	0,36230	0,03613	0,01855	0,50000	0,44824	0,99219
	0,04025	0,65602	0,11691	0,09087	0,63829	0,51233	0,70971
DXS9898	0,00036	-0,00087	0,00169	0,00036	0,00080	0,00324	0,00395
	0,33301	0,62402	0,17480	0,31348	0,26953	0,08008	0,06738
	0,03576	0,76130	0,11903	0,16107	0,03866	0,01699	0,06783
DXS7133	0,01779	0,00036	-0,00136	0,00597	0,00535	0,01839	0,00133
	0,00098*	0,30371	0,63477	0,03711	0,05273	0,00195*	0,23535
	0,01171	0,78429	0,77225	0,37710	0,14768	0,01680	0,64806
GATA31E08	0,00168	0,00092	0,00232	0,00317	0,00939	0,00542	0,00251
	0,15723	0,24023	0,10938	0,08398	0,00195*	0,02051	0,13379
	0,01905	0,42802	0,19823	0,03360	0,01349	0,00148*	0,10790
GATA172D05	0,01419	0,00497	-0,00242	0,00980	0,00327	0,01649	0,01091
	0,00000*	0,01758	0,97266	0,00488*	0,04980	0,00000*	0,00000*
	0,00000*	0,07298	0,95971	0,00000*	0,20891	0,00000*	0,00000*
DXS7423	0,00020	0,00176	0,00176	0,00033	-0,00035	0,00422	-0,00016
	0,34473	0,15625	0,18359	0,32715	0,44141	0,06836	0,39258
	0,43126	0,28405	0,14033	0,11129	0,65396	0,08757	0,36664
DXS6809	0,00009	0,00269	0,00213	0,00382	-0,00080	-0,00029	0,00096
	0,38281	0,06641	0,10840	0,03906	0,59375	0,46777	0,25879
	0,56753	0,00093*	0,01809	0,02448	0,04919	0,40224	0,17568
DXS7132	-0,00026	0,00274	-0,00022	-0,00021	-0,00099	0,00004	-0,00051
	0,40527	0,10254	0,40527	0,43555	0,59082	0,36914	0,46289
	0,00177*	0,38757	0,46440	0,07526	0,26628	0,13288	0,37619
DXS9902	-0,00159	-0,00022	-0,00145	-0,00002	-0,00244	0,00241	0,00141
	0,74023	0,44629	0,64355	0,34961	0,90430	0,14062	0,21680
	0,02879	0,42508	0,27806	0,02700	0,98275	0,02339	0,01402
DXS6789	0,00567	-0,00030	-0,00066	0,00204	-0,00048	0,00400	0,00032
	0,01660	0,50098	0,56348	0,13770	0,48047	0,04980	0,33105
	0,27186	0,46388	0,82287	0,25002	0,82925	0,14560	0,26855

Populações estudadas neste trabalho: SP, São Paulo; RJ, Rio de Janeiro; V, Vitória; BH, Belo Horizonte. Populações estudadas por Gusmão et al. (2009): MS, Mato Grosso do Sul; PR, Paraná. \* Valor de  $p$  significativa para o nível de 0,005, após correção de Bonferroni.



Tabela 13: Comparação em nível haplotípico entre populações brasileiras. Distância genética  $F_{ST}$  (abaixo da diagonal) e correspondente valor de  $p$  (acima da diagonal) entre populações brasileiras.

	SP	RJ	V	BH	PR	MS
SP		0,00000*	0,00000*	0,00000*	0,00000*	0,00586
RJ	0,00557		0,50586	0,18164	0,00000*	0,01465
V	0,00343	-0,00008		0,03906	0,00000*	0,13086
BH	0,00581	0,00057	0,00128		0,00000*	0,05273
PR	0,00361	0,00474	0,00326	0,00535		0,02734
MS	0,00234	0,00174	0,00078	0,00135	0,0019	

Populações estudadas neste trabalho: SP, São Paulo; RJ, Rio de Janeiro; V, Vitória; BH, Belo Horizonte. Populações estudadas por Gusmão et al. (2009): MS, Mato Grosso do Sul; PR, Paraná. \* Valor de  $p$  significativa para o nível de 0,005, após correção de Bonferroni.

Para simples *locus*, considerando ambas metodologias, os resultados (Tabela 12) mostraram valores de  $p$  significativos após correção de Bonferroni ( $p < 0,005$ ) para DXS9898 entre São Paulo e Paraná; para DXS7133 entre São Paulo e Rio de Janeiro/ Belo Horizonte; para GATA172D05 entre Rio de Janeiro e Belo Horizonte/Paraná e entre Paraná e Vitória/Belo Horizonte/Mato Grosso do Sul; para DXS7423 entre São Paulo e Belo Horizonte/Mato Grosso do Sul. Os marcadores DXS8378, DXS9902, DXS7132, DXS6809, DXS6789 e GATA31E08 não apresentaram diferenças significativas entre as populações testadas. Dos 12 X-STRs estudados por Ribeiro-Rodrigues et al. (2011), a análise de distância genética  $F_{ST}$  também mostrou diferenças significativas na população brasileira para 9 marcadores: DXS7133, DXS10011, HPTRB, DXS6800, DXS7132, DXS9895, GATA31E08, DXS7423 e DXS7130.

Em nível haplotípico, resultados significativos de distância genética ( $p < 0,005$ ) foram encontrados entre Paraná e as quatro populações da região sudeste, bem como entre São Paulo

e as outras três populações estudadas neste trabalho (Tabela 13). O valor global de  $F_{ST}$  na análise de variação molecular (AMOVA) entre as populações brasileiras foi de 0,0028 com valor de  $p$  significativo ( $p$  obtido  $\leq 0,000$ ) e um resultado similar foi obtido quando apenas o grupo sudeste foi analisado. Tal resultado foi reproduzido por Ribeiro-Rodrigues et al. (2011) ( $F_{ST} = 0,0047$ ;  $p = 0,0000$ ) ao comparar a população de 16 Estados brasileiros.

Além disso, análise de distância genética ( $F_{ST}$ ) em nível haplotípico foram realizadas entre populações brasileiras e estrangeiras com dados acessíveis na literatura para o mesmo conjunto de X-STRs (GUSMÃO et al., 2009). Os resultados estão apresentados na Tabela 14.

Tabela 14: Distância genética  $F_{ST}$  para nível haplotípico entre populações brasileiras e internacionais.

	SP		RJ		V		BH	
	$F_{ST}$	p	$F_{ST}$	p	$F_{ST}$	p	$F_{ST}$	p
NP	0,00623	0,00000*	0,00433	0,00000*	0,00232	0,00000*	0,00381	0,00000*
PC	0,00650	0,00000*	0,00365	0,00000*	0,00179	0,02539	0,00443	0,00000*
G	0,00549	0,00000*	0,00586	0,00000*	0,00287	0,00293*	0,00564	0,00000*
Ca	0,01073	0,00000*	0,00979	0,00000*	0,00750	0,00000*	0,01059	0,00000*
RN	0,00672	0,01802	0,01016	0,00000*	0,00704	0,00391*	0,01181	0,00000*
Co	0,00653	0,00000*	0,01061	0,00000*	0,00908	0,00000*	0,01125	0,00000*
BA	0,00551	0,00000*	0,01114	0,00000*	0,00630	0,00000*	0,01125	0,00000*
M	0,01022	0,00000*	0,01288	0,00000*	0,01062	0,00000*	0,01446	0,00000*
ER	0,00955	0,01802	0,01294	0,00195*	0,00938	0,00488*	0,01217	0,00000*
CR	0,00862	0,00000*	0,01490	0,00000*	0,01248	0,00000*	0,01331	0,00000*
A	0,01030	0,00000*	0,01450	0,00000*	0,01065	0,00000*	0,01391	0,00000*
K	0,04461	0,00000*	0,02749	0,00000*	0,03166	0,00000*	0,02671	0,00000*

Populações brasileiras estudadas neste trabalho: SP, São Paulo; RJ, Rio de Janeiro; V, Vitória; BH, Belo Horizonte. Populações internacionais estudadas por Gusmão et al. (2009): NP, Norte de Portugal; CP, Centro de Portugal; G, Galícia (Espanha); Ca, Cantabria (Espanha); RN, Rio Negro (Argentina); Co, Córdoba (Argentina); BA, Buenos Aires (Argentina); M, Misiones (Argentina); ER, Entre Rios (Argentina); CR, Costa Rica; An, Antioquia (Colômbia); K, Karimojong (Uganda). \* Valor de  $p$  significante para o nível de 0,005, após correção de Bonferroni.

Na maioria das comparações (Tabela 14), valores significativos foram obtidos e estes resultados revelaram que o perfil de X-STR proveniente de São Paulo, Rio de Janeiro, Vitória e Belo Horizonte é, em geral, mais similar à população Européia (Portugal e Espanha), seguida pela América Latina (Argentina, Costa Rica e Colômbia), estando mais distante da população africana de Uganda. Posteriormente, similar resultado também foi descrito por Ribeiro-Rodrigues et al. (2011).

Segundo a história de formação da população brasileira, o maior contingente migratório para o Brasil foi de europeus, composto principalmente por portugueses, italianos, espanhóis e germânicos, que chegaram entre os séculos XIX e XX (RIBEIRO 2006). Seguindo os europeus, africanos, trazidos como escravos entre os séculos XVI e XIX, foram o segundo maior grupo de imigrantes. Enfim, em menor escala, também imigraram japoneses, sírios e libaneses. Os dados das populações da região sudeste foram comparados apenas com algumas populações internacionais, não sendo possível a análise de populações asiáticas e de outras de origem africana (apenas Uganda), uma vez que poucos trabalhos estudaram o mesmo conjunto de X-STRs e disponibilizaram os perfis genéticos obtidos.

Segundo Carvalho-Silva et al. (2001), estudos com o cromossomo Y em brasileiros classificados como brancos mostraram uma forte contribuição européia (97%). Por outro lado, a origem do DNA mitocondrial na região sudeste do Brasil, considerando indivíduos brancos, foi de 31% de europeus, 34% de africanos e 33% de ameríndios (ALVES-SILVA et al. 2000). Outro estudo, com marcadores autossômicos em indivíduos brancos da população de São Paulo, também mostrou grande contribuição européia (79%), seguido por africana (14%) e ameríndia (7%) (FERREIRA et al. 2006).

De acordo com tais dados e com a história da colonização do Brasil, para os X-STRs também é esperada uma maior contribuição européia, seguida por africana e ameríndia, embora a contribuição africana deva ser maior do que a observada com marcadores

autossômicos e do cromossomo Y, uma vez que a mulher apresenta dois cromossomos X e o homem apenas um.

Embora as populações comparadas não tenham sido muitas, os resultados obtidos estão de acordo com as informações acima, uma vez que menores distâncias genéticas foram encontradas com populações ibéricas de Portugal e Espanha e Latino Americanas da Argentina, Colômbia e Costa Rica, sendo a maior distância obtida com a população africana de Uganda.

A proximidade com populações Latino Americanas pode ser explicada pelo fato de que a América do Sul é caracterizada pela colonização Ibérica e acasalamentos direcionados entre homens de origem europeia e mulheres de origem africana ou ameríndia (CARVAJAL-CARMONA et al., 2003; MARTINEZ-MARIGNAC et al., 2004; BEDOYA et al., 2006; PICO et al., 2008). Embora a população africana tenha sido o segundo maior grupo de imigrantes no Brasil, a distância obtida de Uganda é explicada pelo fato de que tráfico trouxe escravos oriundos principalmente da África Ocidental, Centro-ocidental (Angola) e Sudeste Africano (Moçambique) (<http://www.ibge.gov.br>; GONÇALVES et al., 2007).

Em conclusão, os dados apresentados demonstram que há diferenças significativas dentre as populações brasileiras e entre estas e populações estrangeiras para o conjunto de X-STRs analisados, o qual se mostrou altamente discriminativo na região sudeste do Brasil.

### **5.5. Análise de X-STRs nos casos de relações biológicas.**

Foram analisados 3 casos (A, B e C) que permaneceram inconclusivos após a análise dos tradicionais marcadores AS-STR, apresentados na Tabela 15.

Tabela 15. Perfis observados para 15 AS-STRs e 10 X-STRs em três casos de avaliação do vínculo biológico.

STR locus	Caso A			Caso B				Caso C	
	M	C	MSP	M	C	MSP	FBSP	C	MSP
<b>D3S1358</b>	16	16	16-17	16-18	16-17	17-18	17-18	15-16	15-17
<b>TH01</b>	7	7-8	9	7-9,3	9-9,3	8-9,3	8-9	7-9,3	6-9,3
<b>D21S11</b>	28-31	28-29	28-31,2	28-30,2	28-30	29	29-31	29-34,2	31,2-33,2
<b>D18251</b>	17-21	17-21	15-17	12-13	13-14	12-13	12-14	12-16	13-16
<b>PENTA E</b>	5-14	5-12	NA	9-13	12-13	13-19	11-19	11-16	16-20
<b>D5S818</b>	11-12	12-13	9-10	12	11-12	11-12	11	11-12	11-12
<b>D13S317</b>	11-12	11	9-11	8-11	8-11	9-12	8-9	12	12
<b>D7S820</b>	10-12	11-12	10-11	12-13	8-13	11-12	10-12	10	9-10
<b>D16S539</b>	11	11	9-11	11-12	11-12	12	11-12	9-10	9-12
<b>CSF1PO</b>	7-10	10-12	12	11-12	11-12	10-12	10-12	10	10-11
<b>PENTA D</b>	8-12	12-14	10-14	13	13-15	9-14	9-15	10-13	11-13
<b>vWA</b>	16	16-18	17-18	14-15	15-16	17-18	18-19	17	15-17
<b>D8S1179</b>	13-14	13	13-14	10-14	10-13	10-15	10	10-13	11-16
<b>TPOX</b>	NA	8	NA	8	8-11	10-11	11	8	8
<b>FGA</b>	21	20-21	20	23-24	22-24	22-25	22-25	20-21	21
<b>Amelo.</b>	XX	XX	XX	XX	XX	XX	XX	XX	XX
†			143,13				514,00		35,94
<b>DXS8378</b>	12	10-12	10-12	11-12	11-12	8-12	8-11	10	10
<b>DXS9898</b>	10-12	10-12	8,3-12	8,3-12	11-12	8,3-13	11-13	8,3-13	11-13
<b>DXS7133</b>	10-11	11	10-11	9	9-11	9	9-10	9-11	9-13
<b>31E08*</b>	10-11	10-13	13	11-13	11-14	9-14	14	9-10	10-14
<b>172D05*</b>	8-9	8-9	8-10	6-11	10-11	8-11	10-11	6-8	6-9
<b>DXS7423</b>	14	14-15	15-17	14-15	13-15	15-16	13-15	14-15	14
<b>DXS6809</b>	31-33	31-33	30-33	31-34	30-34	31-35	30-35	34-35	34
<b>DXS7132</b>	14-15	12-14	12-13	13	13-14	13	13-14	13-14	13-14
<b>DXS9902</b>	11	11-12	12-13	10-13	12-13	11-12	12	11-12	12
<b>DXS6789</b>	15-18	15-21	20-21	18-20	20	16-20	16-20	20	15-20
‡			400,570				3.805,195		328,470
§			57.336,4				1.955877,3		11.805,8

M = mãe; C = filha; MSP = mãe do suposto pai, FBSP = filha biológica do suposto pai. NA, Não amplificado; †, LR combinado para 15 AS-STRs; ‡, LR combinado para 10 X-STRs; §, LR combinado considerando-se AS e X STRs. \*GATA foi removido dos nomes destes *loci*.

LR é definida pelo contraste de hipóteses mutuamente exclusivas. Quanto maior é o valor de LR, maior é a evidência em favor da relação do vínculo biológico investigado.

Normalmente, em um exame de determinação do vínculo biológico com suposto pai presente,

são analisados 15 AS-STRs (kits PowerPlex 16 - Promega ou AmpFlSTR Identifier - Applied Biosystems by Life Technologies) e estes são suficientes para determinar a exclusão do vínculo, bem como, no caso de não exclusão, fornecer um LR (neste caso também chamado de Índice de Paternidade – IP) elevado. Não existe um valor de LR mínimo para se determinar a presença ou não do vínculo biológico investigado; tem-se que quanto maior for este valor, maior é a evidência em favor da relação biológica. Muitos laboratórios consideram o valor de 10.000 seguro para emissão de resultado em favor da relação biológica investigada.

Em casos deficientes de paternidade, ou seja, quando o suposto pai está ausente, a análise tradicional pode gerar um LR baixo, como nos casos apresentados na Tabela 14 (Caso A: 143,13; Caso B: 514,00; Caso C: 35,94), sendo necessário o estudo de outros marcadores STR a fim de se buscar incompatibilidades que levem a uma exclusão da relação biológica questionada ou, na ausência destas, elevar o LR para emissão de um laudo mais consistente à favor do vínculo biológico. De acordo com os dados apresentados, pode-se observar que os X-STRs auxiliam de forma muito eficiente a análise em casos complexos de relação biológica, possibilitando que maiores valores de LR sejam obtidos pela análise de menos marcadores. Neste caso, temos que o estudo de apenas 10 X-STRs proporcionou valores de LR 2,7 a 9 vezes superiores aos obtidos com 15 AS-STRs, com LR combinados (AS-STR x X-STR) maiores que 10.000, variando entre 11.805 e 1.955.877.

## **5.6. Banco de dados brasileiro para X-STRs.**

Com a finalidade de compilar os dados genéticos populacionais publicados para marcadores X-STR na população brasileira e auxiliar na prática forense, gerando mais uma ferramenta útil para a polícia técnico-científica, os laboratórios de DNA e a comunidade

acadêmica, foi desenvolvido o Banco Genético Brasileiro do Cromossomo X – BGBX, o qual está disponível para o público nas versões português e inglês, no endereço [www.bgbx.com.br](http://www.bgbx.com.br).

O Banco é composto de uma página administrativa, utilizada pelos controladores do banco e cujas características não estão apresentados neste trabalho por questão de segurança, e por uma página pública acessada pelo endereço descrito acima.

A página pública do BGBX é composta de diversas seções, como visualizado no menu à esquerda da Figura 7.

Figura 7: Página inicial do banco de dados BGBX.

The image shows the homepage of the Brazilian Genetic Bank of the X Chromosome (BGBX). The header is blue with a chromosome icon and the text "BANCO GENÉTICO BRASILEIRO DO CROMOSSOMO X". There is a language selector for "English". A vertical navigation menu on the left lists various sections. The main content area, titled "Página Inicial", contains introductory text in Portuguese explaining the bank's purpose and listing the available sections.

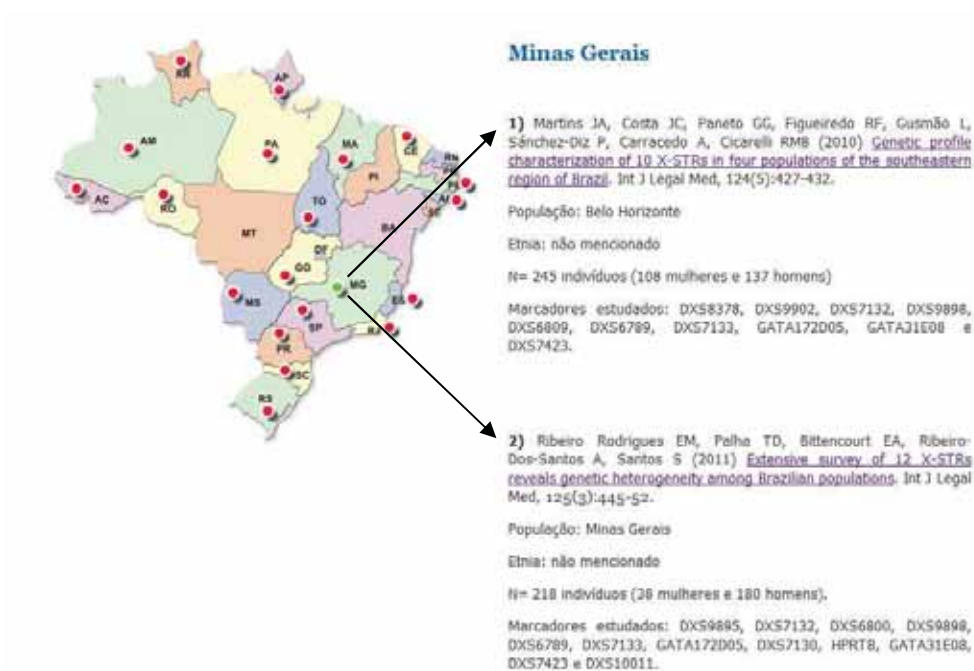
Fonte: Este trabalho ([www.bgbx.com.br](http://www.bgbx.com.br)).

### 5.6.1. X-STRs no Brasil

Esta seção se divide em duas sub-seções: “Mapa” e “Estatística”. Em “Mapa” há a distribuição geográfica dos Estados brasileiros com dados de X-STRs cadastrados no banco. Ao clicar em um Estado, detalhes dos trabalhos registrados serão listados com as seguintes

informações para cada estudo: referência do artigo, população, etnia, número de indivíduos e nomes dos marcadores analisados no trabalho (Figura 8). Assim, o usuário poderá ver todos os trabalhos registrados para uma região particular e selecionar o mais adequado para seus propósitos.

Figure 8: Estados brasileiros com dados de X-STR cadastrados no BGBX até o momento (marcados com um círculo). Ao clicar no círculo, os estudos correspondentes a este Estado são listados.



Fonte: Este trabalho (www.bgbx.com.br).

Em “Estatística”, tem-se a relação das populações e respectivos marcadores X-STRs cadastrados no banco. Até o momento, 10 trabalhos com dados de X-STRs no Brasil estão indexados na base de dados PubMed e foram cadastrados no BGBX, compreendendo 24 populações, a saber:

- 1 de imigrantes japoneses no Brasil (RIBEIRO-RODRIGUES et al., 2010)
- 1 de ameríndios no Brasil (LEITE et al., 2009)



- 22 populações localizadas nas cinco regiões brasileiras: **Norte** - Amazonas, Acre, Amapá, Pará - dados do Pará e Belém, Rondônia, Roraima e Tocantins (RIBEIRO-RODRIGUES et al., 2008 e 2011); **Nordeste** - Alagoas, Ceará, Pernambuco e Maranhão (FERREIRA DA SILVA et al., 2010; RIBEIRO-RODRIGUES et al., 2011); **Centro-Oeste** - Goiás e Mato Grosso do Sul (GUSMÃO et al., 2009; RIBEIRO-RODRIGUES et al., 2011); **Sudeste** - Minas Gerais - dados de Minas Gerais e Belo Horizonte, Rio de Janeiro, São Paulo, Vitória (GUSMÃO et al., 2009; MARTINS et al., 2010; RIBEIRO-RODRIGUES et al., 2011); **Sul** - Paraná, Rio Grande do Sul e Santa Catarina (LEITE et al., 2009; CAINÉ et al., 2007; CAINÉ et al., 2010; RIBEIRO-RODRIGUES et al., 2011).

Nestes trabalhos, um total de 22 marcadores X-STR foram descritos, sendo que os mais analisados estão contidos no sistema 12plex desenvolvido por Ribeiro-Rodrigues et al. (2010), seguido do sistema 10plex (em negrito) desenvolvido por Gusmão et al. (2009) e analisado neste trabalho, os quais totalizam 15 X-STRs diferentes (DXS9895, **DXS8378**, **DXS9902**, **DXS7132**, DXS6800, **DXS9898**, **DXS6809**, **DXS6789**, **DXS7133**, **GATA172D05**, DXS7130, HPRTB, **GATA31E08**, **DXS7423** e DXS10011), uma vez que ambos possuem 7 X-STRs em comum (sublinhados).

#### 5.6.2. Dados dos X-STRs

Dados dos X-STRs inclui as seguintes sub-seções: localização, outras informações, frequências alélicas, parâmetros estatísticos, marcadores em desequilíbrio de ligação e haplótipos.

### 5.6.2.1. Localização

Esta sub-seção apresenta a distribuição dos marcadores X-STRs ao longo do cromossomo X (Idiograma) bem como uma tabela com dados de localização citogenética, genética e física.

A localização citogenética foi obtida de UCSC *Human Genome Browser* criado por *Genome Bioinformatics Group of UC Santa Cruz* (<http://www.genome.ucsc.edu/cgi-bin/hgGateway>). A localização física dos X-STRs foi apresentada como a distância Mb do telômero do braço curto do cromossomo X (Xp) de acordo com UCSC *Human Genome Browser (Assembly Feb. 2009 GRCh37/hg19)* e *National Center for Biotechnology Information – NCBI (Build 37.3)* (<http://www.ncbi.nlm.nih.gov/>).

A localização genética foi descrita como a distância em cM do telômero do Xp e foi obtida dos mapas genéticos Marshfield (<http://research.marshfieldclinic.org/genetics>) e deCODE (<http://www.decode.com/>). Também, foi reportada a distância genética utilizando *Rutgers Combined Linkage-Physical Map* (<http://compgen.rutgers.edu/>).

### 5.6.2.2. Outras informações

Esta sub-seção apresenta informações relevantes sobre os marcadores X-STRs estudados na população brasileira registrada no BGBX. Ao selecionar um marcador de interesse, as seguintes informações são listadas: outros nomes conhecidos do marcador, *links* para *websites* que mencionam o marcador (tais como NCBI, UCSC *Human Genome Browser* e *ChrX database*), dados de localização citogenética, genética e física, diferentes sequências de *primers* descritas para a amplificação do X-STR. Quando acessível, o tamanho dos alelos

ou produto de amplificação obtido com os respectivos pares de *primers* são também apresentados.

É importante ressaltar que diferentes autores podem descrever alelos com tamanhos levemente diferentes, utilizando o mesmo par de *primers* para detectar uma estrutura específica, porque os artigos podem reportar tanto o real tamanho do alelo derivado do sequenciamento como o tamanho aparente do alelo, o qual é influenciado pelos fluorocromos utilizados na marcação dos *primers*, pelo padrão de tamanho (*size standard*) e condições utilizados na eletroforese.

Poucos estudos têm analisado a sequência dos alelos de marcadores X-STRs, a qual é de fundamental importância para se estabelecer uma nomenclatura mais consistente. Além disso, alguns estudos tem reportado diferentes nomenclaturas para um mesmo marcador X-STR, a saber: DXS10011 (WATANABE et al., 2000; HERING et al., 2004), DXS7423 (EDELMANN et al., 2001; EDELMANN et al., 2002b; ZARRABEITIA et al., 2002a; ZARRABEITIA et al., 2002b), GATA31E08 (SHIN et al., 2005; GOMES et al., 2009b), entre outros. Em consequência, cartas ao editor (SZIBOR et al., 2003c; SZIBOR et al., 2009) e estudos mais detalhados, incluindo a análise de grupos populacionais humanos e chimpanzés (GOMES et al., 2009b; GOMES et al., 2009c) foram publicados para explicar tais discrepâncias e fornecer uma nomenclatura mais precisa e segura para a genotipagem. Portanto, esta sub-seção também apresenta a(s) nomenclatura(s) e respectiva(s) estrutura(s) alélica(s) dos marcadores X-STR. Desta forma, se diferentes nomenclaturas para um particular X-STR resultam em diferente número do motivo de repetição, os pesquisadores poderão saber mais facilmente que a distribuição alélica em certos artigos necessitam ser revisada para posterior troca de dados, comparações entre laboratórios e uso em análises estatísticas para avaliar estrutura populacional, por exemplo.

Em adição, o genótipo de algumas amostras referência de DNA controle estão disponíveis (K562, NA9947A, N9948 e NA3657). Para alguns marcadores como GATA31E08, diferentes genótipos estão descritos para a mesma amostra referência em virtude das diferentes nomenclaturas alélicas adotadas pelos autores.

#### 5.6.2.3. Frequências alélicas

Muitos estudos com marcadores X-STRs tem focado na caracterização genética de grupos populacionais humanos, uma vez que estes dados são necessários para estabelecer bases de dados referências para aplicação forense. Como já apresentado, estudos com X-STRs têm mostrado diferenças genéticas entre diversas populações, inclusive no Brasil. Assim, esta sub-seção apresenta as frequências alélicas de vários marcadores X-STRs estudados na população brasileira. Ao escolher o marcador de interesse, o usuário do banco de dados verá uma lista de frequências por população, bem como dados de etnia, número de indivíduos analisados e a referência dos respectivos trabalhos.

#### 5.6.2.4. Parâmetros estatísticos

Nesta sub-seção, ao selecionar o marcador de interesse, o usuário terá acesso aos dados de heterozigose observada, heterozigose esperada, valor de  $p$  para equilíbrio de Hardy-Weinberg e, em acordo com Desmarais et al. (1998), chance média de exclusão em trios envolvendo filha e em duos pai-filha, poder de discriminação feminino e masculino. Os valores reportados são originais dos estudos, não sendo, portanto, calculados pelos organizadores do BGBX.

#### 5.6.2.5. Marcadores em desequilíbrio de ligação e haplótipos

A sub-seção Marcadores - Desequilíbrio de Ligação lista os pares de marcadores X-STR que foram encontrados em LD em certas populações brasileiras, apresentado dados de etnia e número de indivíduos analisados na população para a qual tal desequilíbrio foi observado, bem como a referência do estudo, o valor de  $p$  obtido e o nível de significância adotado (baseado na aplicação da correção de Bonferroni para múltiplas análises em muitos trabalhos) no teste utilizado para avaliar o LD.

Em relação aos 22 X-STRs analisados nas populações brasileiras cadastradas no BGBX, 14 pares de marcadores foram relatados por estarem em LD. Destes, 4 correspondem a grupos de ligação já descritos na literatura DXS10079-DXS10074-DXS10075 (HERING et al., 2006) e DXS7424-DXS101 (EDELMANN et al., 2002a). Para os demais 10 pares, a distância física entre os X-STRs, variou de 9,477 Mb (GATA31E08-DXS7423) a 143,811 Mb (DXS9895-DXS10011) e, a distância genética, obtida do mapa de Rutgers, variou entre 23,6527 cM (GATA31E08-DXS7423) e 171,6096 cM (DXS9895-DXS10011).

Na sub-seção “Haplótipos”, o usuário terá acesso aos haplótipos descritos para conjuntos de X-STRs em LD e suas respectivas frequências na população brasileira. Estes dados são apresentados em formato Excel, o qual é mais flexível para elaboração dos arquivos de dados que serão inseridos em programas como o Arlequin, utilizado para análises comparativas entre populações, por exemplo. Recentemente, em alguns periódicos, este tipo de dado tem sido aceito como material suplementar (eletrônico). Poucos trabalhos têm publicado tais haplótipos; no Brasil, há apenas um estudo para os grupos de ligação DXS10079-DXS10074-DXS10075 e DXS7424-DXS101 nas populações de Alagoas e Rio de Janeiro (FERREIRA DA SILVA et al., 2010).

### 5.6.3. Submeter Dados

Os pesquisadores estão convidados à registrarem seus dados de frequências alélicas e parâmetros estatísticos obtidos para diferentes marcadores X-STR, bem como os dados de frequências de haplótipos X-STR, acessando o menu “Submeter Dados”. Inicialmente, apenas dados publicados em artigos pesquisáveis via PubMed serão aceitos para cadastro no BGBX.

### 5.6.4. Colaboradores

Esta seção apresenta a relação de todos os trabalhos cujos dados genéticos populacionais para X-STR na população brasileira estão cadastrados no BGBX.

Os trabalhos científicos citados no BGBX apresentam *hiperlink* com direcionamento para o respectivo *abstract* na base de dados PubMed. Toda a comunidade forense e científica está convidada a enviar suas questões e sugestões através do formulário de contato no web site.

Por fim, a centralização dos dados brasileiros facilita a busca e acesso destes por internet, divulgando os trabalhos que têm sido realizados e aumentando o reconhecimento internacional da pesquisa brasileira.

O pedido de registro do programa de computador BGBX foi depositado em 18 de agosto de 2011 (nº 12166-1), ANEXO E.

## **6. CONCLUSÕES**

- Foram estabelecidas as tabelas de frequências alélicas para 10 X-STRs na população brasileira de São Paulo-SP, Rio de Janeiro-RJ, Vitória-ES e Belo Horizonte-MG;
  
- A análise estatística para avaliar a eficiência forense dos marcadores X-STR mostrou que o sistema decaplex utilizado foi altamente discriminativo nas populações brasileiras estudadas;
  
- Entre as populações estudadas, bem como entre estas e outras brasileiras e estrangeiras com dados publicados na literatura, observou-se diferenças significativas na distribuição alélica dos marcadores X-STR analisados, não sendo possível a utilização de um banco único ou nacional de frequências alélicas para tais marcadores. Além disso, na análise de distância genética observou-se que a população brasileira está mais próxima de populações Ibéricas de Portugal e Espanha e Latino Americanas da Argentina, Colômbia e Costa Rica, estando mais distante da população africana de Uganda;
  
- O sistema decaplex auxiliou de forma eficiente a análise autossômica em casos complexos de investigação do vínculo biológico;
  
- Desenvolveu-se o Banco de Dados Genético Brasileiro do Cromossomo X – BGBX, acessado pelo endereço [www.bgbx.com.br](http://www.bgbx.com.br).



**7. REFERÊNCIAS  
BIBLIOGRÁFICAS**

ALER, M. et al. Genetic data of 10 X-STRs in a Spanish population sample. **Forensic Sci. Int.**, v. 173, p. 193-196, 2007.

ALVES-SILVA, J. et al. The ancestry of brazilian mtDNA lineages. **Am. J. Hum. Genet.**, v. 67, p. 444-461, 2000.

AQUINO, J. et al. A X-chromosome STR hexaplex as a powerful tool in deficiency paternity cases. **Forensic Sci. Int. Genet.: Suppl. Ser.**, v. 2, p. 45–46, 2009.

BÄR, W. et al. DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. International Society for Forensic Haemogenetics. **Int. J. Legal Med**, v.110, p.175–176, 1997.

BEDOYA, G. et al. Admixture dynamics in Hispanics: a shift in the nuclear genetic ancestry of a South American population isolate. **Proc. Natl. Acad. Sci. USA**, v. 103, p. 7234–7239, 2006.

BEIGUELMAN, B. **Dinâmica dos genes nas famílias e nas populações**. 2. ed. Ribeirão Preto: Sociedade Brasileira de Genética, 1996.

BINI, C. et al. Development of a heptaplex PCR system to analyse X-chromosome STR loci from five Italian population samples: a collaborative study. **Forensic Sci. Int.**, v. 153, p. 231-236, 2005.

BOBILLO, C. et al. Genetic analysis of 10 X-STRs in Argentinian population. **Forensic Sci. Int. Genet.**, v. 5, p. e14-6, 2011.

BRUDER, C.E. et al. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. **Am. J. Hum. Genet.**, v. 82, p. 763-71, 2008.

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

BUTLER, J.M. Short Tandem Repeat (STR) Loci and Kits. In: \_\_\_\_\_. **Advanced Topics in Forensic DNA typing: Methodology**. 1 ed. London: Elsevier Academic Press, 2011. Cap. 5, p. 99-139.

CAINÉ, L.M. et al. Genetic Data of four X-Chromosomal STRs in a population sample of Santa Catarina, Brazil. **J. Forensic Sci.**, v. 52, p. 502-503, 2007.

CAINÉ, L.M. et al. Genetic data of a Brazilian population sample (Santa Catarina) using an X-STR decaplex. **J. Forensic Leg. Med.**, v. 17, p. 272-274, 2010.

CALABREZ, M.C.T. **Influência do calor na análise de DNA extraído de sangue e tecidos humanos**: importância para a identificação de corpos carbonizados. 1999. 92f. Tese (Doutorado) – Faculdade de Medicina da Universidade de São Paulo, 1999.

CARVAJAL-CARMONA, L.G. et al. Genetic demography of Antioquia (Colombia) and the Central Valley of Costa Rica. **Hum. Genet.**, v. 112, p. 534–541, 2003.

CARVALHO-SILVA, D. et al. The phylogeography of Brazilian Y-chromosome lineages. **Am. J. Hum. Genet.**, v. 68, p. 281-286, 2001.

CHAKRAVARTI, A. Population genetics – making sense out of sequence. **Nat. Genet.**, v. 21, p. 56-60, 1999.

CHEN, D.P. et al. Use of X-linked short tandem repeats loci to confirm mutations in parentage caseworks. **Clin. Chim. Acta.**, v. 408, p. 29-33, 2009.

CHOW, J.C. et al. Silencing of the mammalian X chromosome. **Annu. Rev. Genomics Hum. Genet.**, v. 6, p. 69–92, 2005.

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

EDELMANN, J. et al. Sixteen X-chromosome STR loci frequency data from a German population. **Forensic Sci. Int.**, v. 124, p. 215–218, 2001.

EDELMANN, J. et al. Validation of the STR DXS7424 and the linkage situation on the X-chromosome, **Forensic Sci. Int.**, v. 125, p. 217–222, 2002a.

EDELMANN, J. et al. Sequence variation and allele nomenclature for the X-linked STRs DXS9895, DXS8378, DXS7132, DXS6800, DXS7133, GATA172D05, DXS7423 and DXS8377. **Forensic Sci. Int.**, v. 129, p. 99-103, 2002b.

EDELMANN, J. et al. Validation of the X-chromosomal STR DXS6809. **Int. J. Legal Med.**, v.117, p.241–244, 2003.

EDELMANN, J. et al. X-chromosomal haplotype frequencies of four linkage groups using the Investigator Argus X-12 Kit. **Forensic Sci. Int. Genet.** 2011, *in press*.

EGELAND, T. et al. Beyond traditional paternity and identification cases. Selecting the most probable pedigree. **Forensic Sci. Int.**, v. 110, p. 47-59, 2000.

ELLIS, N.; GOODFELLOW, P.N. The mammalian pseudoautosomal region. **Trends Genet.**, v.5, p.406-410, 1989.

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

FERREIRA, L.B. et al. Genomic ancestry of a sample population from the State of São Paulo, Brazil. **Am. J. Human Biol.**, v. 18, p. 702-705, 2006.

FERREIRA DA SILVA, I.H. et al. An X-chromosome pentaplex in two linkage groups: haplotype data in Alagoas and Rio de Janeiro populations from Brazil. **Forensic Sci. Int. Genet.**, v. 4, p. e95-100, 2010.

FRANÇA, G.V. **Medicina Legal**. 6. ed., Rio de Janeiro, Guanabara-Kogan, 2001. cap. 3, p.32-63.

FREIJE, D. et al. Identification of a second pseudoautosomal region near the Xq and Yq telomeres. **Science**, v. 258, p. 1784–1787, 1992.

GOMES, I. et al. Genetic analysis of three US population groups using a X-chromosomal STR decaplex. **Int. J. Legal Med.**, v. 121, p. 198-203, 2007a.

GOMES, I. et al. Analysis of 10 X-STRs in three African populations. **Forensic Sci. Int. Genet.**, v. 1, p. 208-211, 2007b.

GOMES, I. et al. The Karimojong from Uganda: Genetic characterization using an X-STR decaplex system. **Forensic Sci. Int. Genet.**, v. 3, p. e127-128, 2009a.

GOMES, I. et al. X chromosome STRs sequence variation, repeat structure and nomenclature in humans and chimpanzees. **Int. J. Legal Med.**, v. 123, p. 143–149, 2009b.

GOMES, I. et al. Evaluation of DXS9902, DXS7132, DXS6809, DXS7133, and DXS7423 in humans and chimpanzees: sequence variation, repeat structure, and nomenclature. **Int. J. Legal Med.**, v. 123, p. 403-12, 2009c.

GONÇALVES, V.F. et al. The Phylogeography of African Brazilians. **Hum. Hered.**, v. 65, p. 23-32, 2007.

GUSMÃO, L. 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.**, v. 123, p. 227–234, 2009.

GUSMAO, L.; ALVES, C. Marcadores dos cromossomas sexuais em identificação. In: Maria de Fátima Terra Pinheiro. (Org.). In: **Genética Forense. Perspectivas da Identificação Genética**. Porto: Edições Universidade Fernando Pessoa, 2010, p. 105-139.

HAMOSH, A. et al. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. **Nucleic Acids Res.**, v.33, p.514-517, 2005.

HERING S; SZIBOR R. Development of the X-linked tetrameric microsatellite marker DXS9898 for forensic purposes. **J. Forensic Sci.**, v.45, p.929–931, 2000.

HERING S; KUHLISCH E; SZIBOR R. Development of the X-linked tetrameric microsatellite marker HumDXS6789 for forensic purposes. **Forensic Sci. Int.**, v.119, p.42–46, 2001.

HERING, S. 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.**, v. 118, p. 313-319, 2004.

HERING, S. et al. DXS10079, DXS10074 and DXS10075 are STRs located within a 280-kb region of Xq12 and provide stable haplotypes useful for complex kinship cases. **Int. J. Legal Med.** v.120, p.337-345, 2006.

HWA, H.L. et al. Thirteen X-chromosomal short tandem repeat loci multiplex data from Taiwanese. **Int. J. Legal Med.**, v. 123, p. 263-269, 2009.

International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. **Nature**, v. 409, p. 860-921, 2001.

International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. **Nature**, v. 431, p. 931-945, 2004.

INTURRI, S. et al. Linkage and linkage disequilibrium analysis of X-STRs in Italian families. **Forensic Sci. Int. Genet.**, v. 5, p. 152-154, 2011.

JEFFREYS, A.J.; WILSON, V.; THEIN, S.L. Hypervariable "minisatellite" regions in human DNA. **Nature**, v.314, p.67-73, 1985.

JEFFREYS, A.J.; BROOKFIELD, J.F.Y.; SEMEONOFF, R. Positive identification of an immigration test-case using human DNA fingerprints. **Nature**, v.317, p.818-819, 1985.

JOBLING, M.A.; TYLER-SMITH, C. The human Y chromosome: an evolutionary marker comes of age. **Nat. Rev. Genet.**, v. 4, p. 598-612, 2003.

JORDE, L.B.; WATKINS, W.S.; BAMSHAD, M.J. Population genomics: a bridge from evolutionary history to genetic medicine. **Human Mol. Genetics**, v.10, n.20, p.2199-2207, 2001.

LANDER, E.S. Initial impact of the sequencing of the human genome. **Nature**, v. 470, p. 187-197, 2011.

LEE, H.Y. 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.**, v. 118, p. 355-360, 2004.

LEITE, F.P. 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.**, v. 139, p. 404-412, 2009.

LI, W.H.; YI, S.; MAKOVA, K. Male-driven evolution. **Curr. Opin. Genet. Dev.**, v. 12, p. 650-656, 2002.

LIU, Q.L. et al. Development of the nine X-STR loci typing system and genetic analysis in three nationality populations from China. **Int. J. Legal Med.**, v. 125, p. 51-58, 2011.

MARTINEZ-MARIGNAC, V.L. et al. Characterization of admixture in an urban sample from Buenos Aires, Argentina, using uniparentally and biparentally inherited genetic markers. **Hum. Biol.**, v. 76, p. 543-557, 2004.

MARTINS, J.A. et al. Genetic profile characterization of 10 X-STRs in four populations of the southeastern region of Brazil. **Int. J. Legal Med.**, v. 124, p. 427-432, 2010.

MINISTÉRIO DA JUSTIÇA. **Padronização de Exames de DNA em Perícias Criminais**. <http://portal.mj.gov.br/data/Pages/MJ3F6F0588ITEMID4784422F29914B3BB84EE085F6AA9537PTBRNN.htm>. Acessado em 11 nov. 2011.

- OHNO, S. **Sex chromosomes and Sex-linked genes**. Berlin: Springer, 1967.
- PAGE, D.C. et al. Occurrence of a transposition from the X-chromosome long arm to the Y-chromosome short arm during human evolution. **Nature**, v.311, p.119–123, 1984.
- PICO, A. et al. Genetic profile characterization and segregation analysis of 10 X-STRs in a sample from Santander, Colombia. **Int. J. Legal Med.**, v. 122, p. 347–351, 2008.
- PINTO, N.; GUSMÃO, L.; AMORIM, A. X-chromosome markers in kinship testing: a generalisation of the IBD approach identifying situations where their contribution is crucial. **Forensic Sci Int Genet.**, v.5, p.27-32, 2011.
- POETSCH, M. et al. Development of two pentaplex systems with X-chromosomal STR loci and their allele frequencies in a northeast German population. **Forensic Sci. Int.**, v. 155, p. 71-76, 2005.
- RIBEIRO D. **O povo brasileiro: a formação e o sentido do Brasil**. São Paulo: Companhia das Letras, 2006. 435p.
- RIBEIRO-RODRIGUES, E.M. et al. A multiplex PCR for 11 X chromosome STR markers and population data from a Brazilian Amazon Region. **Forensic Sci. Int. Genet.**, v. 2, p. 154-158, 2008.
- RIBEIRO-RODRIGUES, E.M. et al. Genetic data of twelve X-STRs in a Japanese immigrant population resident in Brazil. **Forensic Sci. Int. Genet.**, v. 4, p. e57-58, 2010.
- RIBEIRO-RODRIGUES, E.M. et al. Extensive survey of 12 X-STRs reveals genetic heterogeneity among Brazilian populations. **Int. J. Legal Med.**, v. 125, p. 445-452, 2011.
- ROBIN, E.D.; WONG, R. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. **J. Cell. Physiol.**, v. 136, p. 507–513, 1988.
- ROBINO, C. et al. Development of two multiplex PCR systems for the analysis of 12 X-chromosomal STR loci in a northwestern Italian population sample. **Int. J. Legal Med.**, v. 120, p. 315-318, 2006.
- ROSS, M.T. et al. The DNA sequence of the human X chromosome. **Nature**, v. 434, p. 325-327, 2005.
- SANTOS-LOPES, S.S. et al. A worldwide phylogeography for the human X chromosome. **PLoS ONE**, v. 2, p. e557, 2007.
- SCHAFFNER, S. The X chromosome in population genetics. **Nature Rev.**, v. 5, p. 43-51, 2004.
- SHIN, S.H. et al. Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. **Forensic Sci. Int.**, v. 147, p. 35-41, 2005.

SILVEIRA, D. et al. Use of X-linked short tandem repeat loci in routine parentage casework. **Transfusion**, v. 47, p. 1050-1053, 2007.

SKALETISKY, H. et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. **Nature**, v. 423, p. 825–837, 2003.

SZIBOR, R. et al. Use of X-linked markers for forensic purposes. **Int. J. Legal Med.**, v. 117, p. 67–74, 2003a.

SZIBOR, R. et al. Cell line DNA typing in forensic genetics--the necessity of reliable standards. **Forensic Sci. Int.**, v. 138, p. 37-43, 2003b.

SZIBOR, R. et al. Letter to the Editor: Sequence structure and population data of the X-linked markers DXS7423 and DXS8377—clarification of conflicting statements published by two working groups. **Forensic Sci. Int.**, v. 134, p. 72–73, 2003c.

SZIBOR, R. et al. Haplotyping of STR cluster DXS6801– DXS6809–DXS6789 on Xq21 provides a powerful tool for kinship testing. **Int. J. Legal Med.**, v. 119, p. 363–369, 2005.

SZIBOR, R.; HERING, S.; EDELMANN, J. A new Web site compiling forensic chromosome X research is now online. **Int. J. Legal Med.**, v. 120, p. 252-254, 2005.

SZIBOR, R. X-chromosomal markers: Past, present and future. **Forensic Sci. Int. Gen.**, v. 1, p. 93–99, 2007.

SZIBOR, R. et al. Nomenclature Discrepancies in the HPRTB short tandem repeat. **Int. J. Legal Med.**, v. 123, p. 185–186, 2009.

TANDE, C.M. DNA Typing: A New Investigatory Tool. **Duke Law Journal**, v.1989, n. 2, p.474-494, 1989.

TIE, J.; UCHIGASAK, S.; OSHIDA, S. Genetic polymorphisms of eight X-chromosomal STR loci in the population of Japanese. **Forensic Sci. Int. Genet.**, v. 4, p. e105-108, 2010.

TILLMAR, A.O. et al. Analysis of linkage and linkage disequilibrium for eight X-STR markers. **Forensic Sci. Int. Genet.**, v. 3, p. 37-41, 2008.

TONI, C.; DOMENICI, R.; PRESCIUTTINI, S. Genotype probabilities of pairs of individuals for X-chromosome markers. **Transfusion**, v. 47, p. 1276-1280, 2007.

TOSCANINI U. et al. X-STRs analysis in paternity testing when the alleged father is related to the biological father. **Forensic Sci. Int. Genet.: Suppl. Ser.**, v. 2, p. 234–235, 2009.

TURRINA, S.; ATZEI, R.; DE LEO, D. Polymorphism of four X-chromosomal STRs: DXS7423, DXS7424, DXS8378 and DXS6809 in a North Italian population sample and their use in kinship testing. **Forensic Sci. Int.**, v. 168, p. 241-243, 2007.

VERZELETTI, A. et al. X-chromosome polymorphism on DXS8378, DXS7132, HPRTB and DXS7423 loci in 130 individuals from Brescia (northern Italy). **Leg. Med. (Tokyo)**, v. 9, p. 158-160, 2007.

WATANABE, G. et al. DXS10011: a hypervariable tetranucleotide STR polymorphism on the X chromosome. **Int. J. Legal Med.**, v. 113, p. 249–250, 2000.

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

ZALÁN, A. et al. Hungarian population data of four X-linked markers: DXS8378, DXS7132, HPRTB, and DXS7423. **Int. J. Legal Med.**, v. 121, p. 74-77, 2007.

ZALÁN, A. et al. Hungarian population data of eight X-linked markers in four linkage groups. **Forensic Sci. Int.**, v. 175, p. 73-78, 2008.

ZARRABEITIA, M.T. et al. A new pentaplex system to study short tandem repeat markers of forensic interest on X chromosome. **Forensic Sci. Int.**, v. 129, p. 85–89, 2002a.

ZARRABEITIA, M.T. et al. Sequence structure and population data of two X-linked markers: DXS7423 and DXS8377. **Int. J. Legal Med**, v. 116, p.368–371, 2002b.

ZARRABEITIA, M.T. et al. Study of six X-linked tetranucleotide microsatellites: population data from five Spanish regions. **Int. J. Legal Med.**, v. 120, p. 147-150, 2006.

ZENG, X.P. et al. Genetic polymorphisms of Investigator Argus X-12 amplification system in Guangdong Han population. **Zhonghua Yi Xue Yi Chuan Xue Za Zhi**, v. 28, p. 230-234, 2011.



**ANEXOS**

## ANEXO A

Aprovação pelo Comitê de Ética em Pesquisa do Instituto Estadual HEMORIO.



GOVERNO DO ESTADO DO RIO DE JANEIRO  
SECRETARIA DE ESTADO DE SAÚDE

HEMORIO  
INSTITUTO ESTADUAL DE HEMATOLOGIA  
ARTHUR DE BONDENACAGLIANI

*COMITÊ DE ÉTICA EM PESQUISA – CEP HEMORIO*

Rio de Janeiro, 13 de junho de 2008.

**ASSUNTO:** Parecer consubstanciado de projeto de pesquisa avaliado pelo CEP HEMORIO

Prezada Pesquisadora,

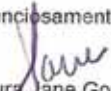
O projeto, "*Estudo de Polimorfismos do cromossomo X na população no Estado do Rio de Janeiro.*", cadastrado no CEP HEMORIO sob o nº 128/08, foi **aprovado** pelo Comitê desta Instituição, conforme a Resolução CNS 196, de 10/outubro de 1996, após resposta às considerações apontadas por este Comitê

Ressaltamos abaixo, algumas orientações fundamentais, as quais o pesquisador deve estar muito atento:

- O sujeito da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado e deve receber uma cópia do Termo de Consentimento Livre e Esclarecido, na íntegra, por ele assinado;
- O pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado e descontinuar o estudo somente após análise das razões da descontinuidade pelo CEP que o aprovou, aguardando seu parecer, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de regime oferecido a um dos grupos da pesquisa que requeira ação imediata;
- O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo. É papel do pesquisador assegurar medidas imediatas adequadas frente a evento adverso grave ocorrido (mesmo que tenha sido em outro centro) e enviar notificações ao CEP e à ANVISA, junto com seu posicionamento;
- Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas. Em caso de projetos do Grupo I ou II apresentados anteriormente à ANVISA, o pesquisador ou patrocinador deve enviá-las também à mesma, junto com o parecer aprovatório do CEP, para serem juntadas ao protocolo inicial.
- Relatórios parciais e final devem ser apresentados ao CEP, inicialmente até 13/01/2009 e ao término do estudo.

Sendo assim, por favor, contate a Coordenação do CEP HEMORIO (Sra. Laura Jane ou Thaís Oliveira) pelo telefone 2299-9442, ramal 2141, a fim de estabelecermos o fluxo de sua pesquisa e tomarmos outras providências pertinentes.

Atenciosamente,

  
Laura Jane Gonçalves Neumann  
Coordenadora do CEP HEMORIO

**ANEXO B**

Aprovação pelo Comitê de Ética em Pesquisa da Fundação Centro HEMOMINAS.



Belo Horizonte, 09 de outubro de 2008

OFÍCIO Nº 89 / 2008  
COMITÊ DE ÉTICA EM PESQUISA  
FUNDAÇÃO HEMOMINAS

Prezado Senhora Pesquisadora, Joyce Aparecida Martins

Encaminhamos o parecer substanciado referente ao seu projeto de pesquisa "Estudo de polimorfismos do cromossomo X na população do Estado de Minas Gerais", nº de registro 194. Seu projeto foi aprovado por nosso Comitê, e então, a partir deste momento, sua pesquisa pode ser desenvolvida no âmbito da Fundação Hemominas.

Desejamos boa sorte e sucesso em seu estudo.

Atenciosamente,

  
Marina Lobato Martins

Secretária do Comitê de Ética em Pesquisa  
Fundação Hemominas

## ANEXO C

Aprovação pelo Comitê de Ética em Pesquisa da Faculdade de Odontologia da USP.




**UNIVERSIDADE DE SÃO PAULO**  
**FACULDADE DE ODONTOLOGIA**

**PARECER DE APROVAÇÃO**  
**Protocolo 24/06**

Com base em parecer de relator, o Comitê de Ética em Pesquisa, **APROVOU** o protocolo de pesquisa "Estudo de frequências alélicas de múltiplos **LOCI STR** do cromossomo X na população do estado de São Paulo e sua contribuição na identificação humana", de responsabilidade do Pesquisador Ricardo Henrique Alves da Silva, sob orientação do Professor Doutor Rogério Nogueira de Oliveira.

Tendo em vista a legislação vigente, devem ser encaminhados a este Comitê relatórios anuais referentes ao andamento da pesquisa e ao término cópia do trabalho em "cd". Qualquer emenda do projeto original deve ser apresentada a este CEP para apreciação, de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas.

São Paulo, 07 de julho de 2006

  
Prof. Dr. **Rogério Nogueira de Oliveira**  
Coordenador do CEP-FOUSP

**ANEXO D**

Aprovação pelo Comitê de Ética em Pesquisa da Faculdade de Ciências Farmacêuticas,  
UNESP.



Protocolo CEP/FCF/CAr nº 02/2009

Interessado: JOYCE APARECIDA MARTINS

Supervisor: Profa. Dra. Regina Maria Barretto Cicarelli

Projeto: Estudo de Polimorfismos do Cromossomo X na População do Estado do Espírito Santo


**Parecer nº 14/2010 – Comitê de Ética em Pesquisa**

O projeto "Estudo de Polimorfismos do Cromossomo X na População do Estado do Espírito Santo", encontra-se adequado em conformidade com as orientações constantes da Resolução 196/96 do Conselho Nacional de Saúde/MS, após passar por adequações pelo pesquisador.

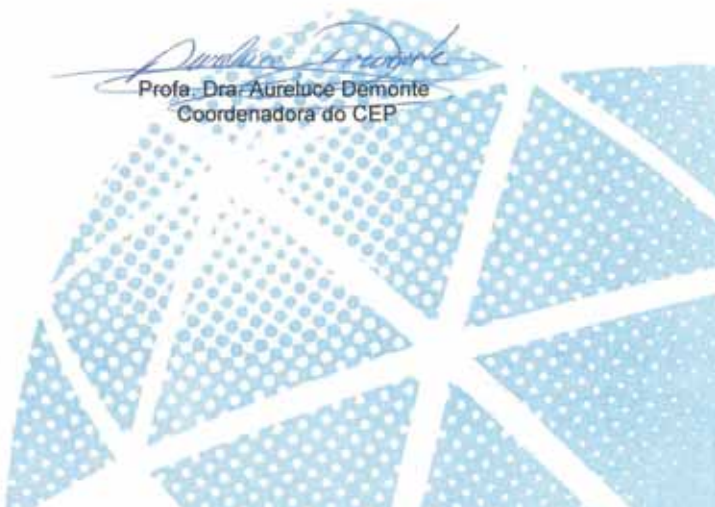
Por essa razão, o Comitê de Ética em Pesquisa desta Faculdade, considerou o referido projeto estruturado dentro de padrões éticos manifestando-se FAVORAVELMENTE à sua execução.

O relatório final do projeto de pesquisa deverá ser entregue em maio de 2011, no qual deverá constar o Termo de Consentimento Livre Esclarecido dos sujeitos da pesquisa.

Araraquara, 15 de abril de 2010.



Prof. Dra. Aureluce Demonte  
Coordenadora do CEP



## ANEXO E

Pedido de registro do programa de computador BGBX (Página 1 do total de 3 páginas).

CÓPIA



**PEDIDO DE REGISTRO DE PROGRAMA DE COMPUTADOR**

Protocolo

**INPI** INSTITUTO NACIONAL DE PROPRIEDADE INTELECTUAL  
BRASILIA 70600-900

18/08/2011 11:37 DCSP

018110031983



0000271106934670

**IDENTIFICAÇÃO DO PEDIDO** (Para uso do INPI)

Número do Pedido \_\_\_\_\_ Protocolo, Data e Hora \_\_\_\_\_

**DADOS DO AUTOR DO PROGRAMA**

Nº de Autores **3** Se mais de um, preencha a "Continuação", com todos os dados solicitados neste Quadro. Date e assine.

CPF\* **008.525.649-88**

Nome **ALEXANDRE EDUARDO CARDOSO**

Nome Abreviado, pseudônimo ou sinal convencional (se houver) \_\_\_\_\_

Data de Nascimento **26/02/1982** Nacionalidade **BRASILEIRO**

Endereço **RUA PRESIDENTE JOÃO BELQUIOR MARQUES GOULART, Nº 64, BAIRRO SANTA CLARA**

Cidade **ARARAQUARA** UF **SP** País **BRASIL**

CEP **14.811-260** Telefone **1133937903** FAX \_\_\_\_\_

E-mail **auin@unesp.br**

**DADOS DO TITULAR DOS DIREITOS PATRIMONIAIS**

Nº de Titulares **1** Se mais de um, preencha a "Continuação", com todos os dados solicitados neste Quadro. Date e assine.

CPF/CNPJ\* **48031918000124**

Nome/Razão Social **UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"**

Nome abreviado, pseudônimo ou sinal convencional (se houver) **UNESP**

Data de Nascimento \_\_\_\_\_ Nacionalidade/Origem \_\_\_\_\_

Endereço **RUA QUIRINO DE ANDRADE, 215**

**CENTRO**

Cidade **SÃO PAULO** UF **SP** País **BRASIL**

CEP **01.049-010** Telefone **11 33937903** FAX \_\_\_\_\_

E-mail **auin@unesp.br**

**SIM**, este Titular é Pessoa Jurídica. Caso afirmativo, assinale a melhor classificação:

Órgão Público     Sociedade com Intuito não Econômico     Microempresa     Software House

Instituição Pública de Ensino ou Pesquisa     Instituição Privada de Ensino ou Pesquisa     Outras

**ENDEREÇO PARA CORRESPONDÊNCIA E CONTATO** (Preencha apenas o necessário)

Toda correspondência será enviada para:  O Procurador ou  O Titular acima ou

Escaninho nº \_\_\_\_\_  Representação INPI em: \_\_\_\_\_  O Endereço abaixo:

Nome **LEOPOLDO CAMPOS ZUANETI**

Endereço **RUA DR. BENTO TEOBALDO FERAZ, 271 - BLOCO 2, BARRA FUNDA**

Cidade **SÃO PAULO** UF **SP** País **BRASIL**

CEP **01.140-070** Telefone **1133937903** FAX \_\_\_\_\_

E-mail **zuaneti@reitoria.unesp.br**

Modelo 1 (folha 12) E

CÓPIA

## **CAPÍTULO II**

## Genetic profile characterization of 10 X-STRs in four populations of the southeastern region of Brazil

Joyce A. Martins · Jeane C. Costa · Greiciane G. Paneto · Raquel F. Figueiredo · Leonor Gusmão · Paula Sánchez-Diz · Ángel Carracedo · Regina M. B. Cicarelli

Received: 31 October 2009 / Accepted: 2 June 2010 / Published online: 16 June 2010  
© Springer-Verlag 2010

**Abstract** Ten X-chromosomal short tandem repeats (DXS8378, DXS9902, DXS7132, DXS9898, DXS6809, DXS6789, DXS7133, GATA172D05, GATA31E08 and DXS7423) were analyzed in four populations of the southeastern region of Brazil (São Paulo, Rio de Janeiro, Vitória and Belo Horizonte). No deviations from the Hardy–Weinberg equilibrium were observed for any of the analyzed loci in the four populations. The average diversity per locus varied between 68% for DXS8378, DXS7133, and DXS7423 and 83%, for DXS6809, with Rio de Janeiro being the most diverse population. Overall power of discrimination values in females varied between 0.99999999990 and 0.99999999997 and between 0.9999991 and 0.9999995 in males. These high values show the potential of this system for forensic application

**Electronic supplementary material** The online version of this article (doi:10.1007/s00414-010-0478-9) contains supplementary material, which is available to authorized users.

J. A. Martins (✉) · J. C. Costa · G. G. Paneto · R. F. Figueiredo · R. M. B. Cicarelli  
Laboratório de Investigação de Paternidade,  
Faculdade de Ciências Farmacêuticas,  
Universidade Estadual Paulista (UNESP),  
Rodovia Araraquara-Jaú, Km 01,  
14801-902, Araraquara, São Paulo, Brazil  
e-mail: joyce\_apa@hotmail.com

L. Gusmão  
Institute of Molecular Pathology and Immunology  
of the University of Porto (IPATIMUP),  
4200-465 Porto, Portugal

P. Sánchez-Diz · Á. Carracedo  
Genomics Medicine Group, CIBER for Rare Diseases  
(CIBERER), Institute of Legal Medicine,  
University of Santiago de Compostela,  
Santiago de Compostela,  
A Coruña, Spain

and relationships' testing in the studied groups. Genetic comparisons (exact tests of population differentiation and pairwise genetic distances) revealed significant differences between Brazilian and other populations from Europe, Latin America and Africa, as well as among different Brazilian populations.

**Keywords** X chromosome · STRs · Human identification · Genetic population data · Southeastern region · Brazil

### Introduction

The markers most commonly utilized in forensic genetics are autosomal short tandem repeats (STRs), followed by Y-chromosome STRs and mitochondrial DNA. The analysis of X-chromosome STRs (X-STRs) has recently become of importance in some complex cases of biological relationships, when the alleged father is not available, in paternity cases involving close blood relatives, in maternity cases involving a son, and in some other identification cases. In these situations, X-STRs may efficiently complement the autosomal analysis as they present higher mean exclusion chances [1]. In view of the wide application of these markers, several X-STRs multiplex PCR systems have been validated for forensic use, including four to 13 markers (e.g., [2–6]). Although, a great variety of markers have been studied, few are common to different works. Recently, a decaplex system was developed and validated in a collaborative work carried out by the Spanish and Portuguese-Speaking Working Group of the International Society For Forensic Genetics (GHEP-ISFG) working group [7] and data were presented for 15 Iberian and Latin American populations. Until now, this is the system with the greatest amount of data on the Brazilian urban population.



Brazilian genetic data on autosomal [8, 9] and Y-chromosome STRs [10–12] are already available, but few studies have been published regarding the X-STRs in Brazil [5, 7, 13–16]. Thus, the aim of the current study was to analyze the 10 X-STRs standardized by the GHEP-ISFG [7] in capital cities of the southeastern region of Brazil (São Paulo, Rio de Janeiro, Vitória and Belo Horizonte).

## Materials and methods

### DNA samples

After an informed consent, blood samples were collected from 1,001 unrelated individuals living in one of the four capital cities of the southeastern region of Brazil: Rio de Janeiro, Rio de Janeiro State ( $n=261$ , 145 women and 116 men); São Paulo, São Paulo State ( $n=250$ , 164 women and 86 men); Vitória, Espírito Santo State ( $n=245$ , 160 women and 85 men) and Belo Horizonte, Minas Gerais State ( $n=245$ , 108 women and 137 men). Blood samples were collected on flinders technology associates (FTA) cards and genomic DNA was extracted using the FTA method (Whatman, Clifton, NJ, USA).

### Markers genotyping

Amplification of the 10 X-STRs loci was achieved in a single PCR multiplex reaction using a GeneAmp PCR System 9700 thermocycler (Applied Biosystems, Foster City, CA, USA). Multiplex PCR was performed as proposed by Gusmão et al. [7] with minor changes: one disk (1.2 mm) of FTA paper with DNA adsorbed, 5  $\mu$ L of Gold Star Buffer 2X (Promega, Madison, WI, USA), 1  $\mu$ L of 10X primer mix at 2  $\mu$ M each primer and 1.5 U of Platinum Taq DNA Polymerase (Invitrogen, Carlsbad, CA, USA); the solution was made up to a 10  $\mu$ L of final reaction volume with ddH<sub>2</sub>O. Separation and detection were performed in an ABI377 automated sequencer (Applied Biosystems) and genotyping was performed by comparison with DNA control reference samples (9947A and 9948, Promega) and allelic ladders provided by the GHEP-ISFG collaborative study [7] using the GeneScan 2.1 software (Applied Biosystems).

### Statistical analysis

The allele frequencies, gene diversities, exact test of the Hardy-Weinberg equilibrium for female samples, pairwise exact test of linkage disequilibrium for male samples, analysis of molecular variance (AMOVA), population pairwise genetic distances ( $F_{ST}$ ) and pairwise exact test of population differentiation were calculated using the ARLE-

QUIN software version 3.1 [17]. Statistics for forensic efficiency evaluation of each locus, namely mean exclusion chance in trios involving daughters (MECT) as well as in father/daughter duos (MECD), power of discrimination in females (PDF) and in males (PDM) were computed as proposed by Desmarais et al. [18].

## Results and discussion

### Genetic variations

The exact test for population differentiation, after applying the Bonferroni correction (significance level of 0.005), revealed no significant differences between allele distributions of men and women (Supplementary Table S1), so samples were pooled and allele frequencies and gene diversities in each population are presented in Table 1. Gene diversities for all markers were above 64%. The average diversity per loci varied between 68% for DXS8378, DXS7133, and DXS7423 and 83% for DXS6809. Rio de Janeiro showed the highest average gene diversity (above 76.2%).

No identical haplotype-like allelic combinations of the 10 X-STRs markers were found between males within or among the four Brazilian groups studied (Supplementary Table S2). For a significance level of 0.005 (after the Bonferroni correction) no deviations from the Hardy-Weinberg equilibrium were observed for any of the analyzed loci (Table 2).

### Forensic efficiency parameters

Statistical parameters of forensic interest are shown in Table 2. The DXS6809 was the most informative marker in São Paulo, Rio de Janeiro and Vitória, whereas, the GATA172D05 was the most polymorphic in Belo Horizonte. The DXS8378 was the less discriminating locus in Rio de Janeiro and Belo Horizonte, and the DXS7423 was the less discriminating locus in São Paulo and Vitória. Overall, values for the power of discrimination in females varied between 0.9999999990 and 0.9999999997 and between 0.9999991 and 0.9999995 in males. The combined mean exclusion chance values were between 0.999996 and 0.999998 in trios, and between 0.99983 and 0.99989 in duos.

Most markers were more discriminative in the Brazilian samples analyzed here than in Iberian and Latin American populations from Portugal, Spain, Argentina, Costa Rica and Colombia [7]. On the other hand, these parameters were higher in African populations from Angola, Mozambique and Uganda [19, 20] than in the four Brazilian groups.

**Table 1** Allele frequency distribution and gene diversity in capital cities of the southeastern region of Brazil

	São Paulo	Rio de Janeiro	Vitória	Belo Horizonte
<b>DXS8378</b>				
N	414	406	405	353
9	0.0121	0.0099	0.0049	0.0027
10	0.3986	0.3128	0.3062	0.3514
11	0.3333	0.3941	0.3704	0.3286
12	0.2367	0.2488	0.2889	0.2919
13	0.0193	0.0320	0.0272	0.0198
14		0.0025	0.0025	0.0057
GD	0.6751	0.6855	0.6866	0.6849
<b>DXS9898</b>				
N	414	406	405	353
7	0.0024	0.0025	0.0049	0.0085
8.3	0.1305	0.1798	0.2049	0.1726
9			0.0025	
10	0.0266	0.0394	0.0346	0.0454
11	0.1908	0.1970	0.1852	0.1871
12	0.3430	0.2611	0.2790	0.3115
13	0.2077	0.2315	0.1975	0.2125
13.3	0.0024	0.0074	0.0074	0.0055
14	0.0845	0.0764	0.0617	0.0453
15	0.0121	0.0049	0.0222	0.0115
GD	0.7796	0.8016	0.8032	0.7909
<b>DXS7133</b>				
N	412	406	401	351
7	0.0072	0.0049	0.0025	
8	0.0097	0.0025	0.0050	0.0057
9	0.4999	0.3103	0.3666	0.3250
10	0.1408	0.2069	0.1746	0.1679
11	0.2913	0.3892	0.3766	0.4243
12	0.0291	0.0591	0.0474	0.0456
13	0.0097	0.0172	0.0150	0.0171
14	0.0122	0.0099	0.0075	0.0144
15			0.0050	
GD	0.6457	0.7073	0.6925	0.6855
<b>GATA31E08</b>				
N	414	406	405	353
7	0.0024	0.0074	0.0074	0.0115
8	0.0048	0.0123	0.0049	0.0084
9	0.1353	0.1626	0.1827	0.2010
10	0.0362	0.0813	0.0543	0.0765
11	0.1715	0.1453	0.1185	0.1162
12	0.2705	0.2488	0.2765	0.2239
13	0.2633	0.2291	0.2296	0.2633
14	0.1039	0.1010	0.1136	0.0823
15	0.0121	0.0123	0.0123	0.0169
GD	0.7994	0.8230	0.8093	0.8158

**Table 1** (continued)

	São Paulo	Rio de Janeiro	Vitória	Belo Horizonte
<b>GATA172D05</b>				
N	414	406	405	353
6	0.1473	0.1897	0.1481	0.1531
7	0.0169	0.0222	0.0272	0.0369
8	0.1546	0.1724	0.1531	0.1246
9	0.1256	0.1576	0.1407	0.1843
10	0.2730	0.2685	0.2963	0.2069
11	0.2126	0.1601	0.1605	0.1953
12	0.0701	0.0271	0.0716	0.0965
13		0.0025	0.0025	0.0028
GD	0.8157	0.8125	0.8174	0.8378
<b>DXS7423</b>				
N	414	406	405	353
8	0.0048	0.0074		
10	0.0024			
12		0.0049	0.0049	0.0057
13	0.0555	0.0714	0.0494	0.0595
14	0.3309	0.3596	0.3531	0.3851
15	0.4855	0.3990	0.4321	0.3684
16	0.0918	0.1158	0.1259	0.1190
17	0.0290	0.0419	0.0346	0.0623
GD	0.6440	0.6929	0.6708	0.6963
<b>DXS6809</b>				
N	414	406	405	353
27	0.0024			0.0057
28	0.0072	0.0049	0.0198	0.0142
29	0.0145	0.0320	0.0321	0.0508
30	0.0531	0.0517	0.0716	0.0313
30.1		0.0025		
31	0.1570	0.1502	0.1506	0.1274
31.1	0.0048	0.0049	0.0074	
32	0.1377	0.1429	0.1259	0.1642
32.1	0.0024			0.0057
33	0.3043	0.2586	0.2914	0.2661
33.1		0.0025		
34	0.1932	0.1970	0.1605	0.2154
35	0.0725	0.0985	0.1012	0.0766
35.1		0.0025		
36	0.0362	0.0320	0.0370	0.0227
37	0.0072	0.0123	0.0025	0.0199
38	0.0072	0.0025		
39		0.0049		
GD	0.8187	0.8387	0.8346	0.8313
<b>DXS7132</b>				
N	412	406	405	353
10	0.0048			
11	0.0194	0.0074	0.0222	0.0169

Table 1 (continued)

	São Paulo	Rio de Janeiro	Vitória	Belo Horizonte
12	0.1044	0.0887	0.1012	0.0850
13	0.2233	0.2512	0.2642	0.2406
14	0.3277	0.3768	0.3012	0.3854
14.3		0.0025		
15	0.2597	0.1946	0.2198	0.2125
15.3	0.0024	0.0025		
16	0.0388	0.0517	0.0519	0.0426
16.3	0.0072	0.0197	0.0173	0.0058
17	0.0097		0.0173	0.0112
17.3	0.0024	0.0049		
18			0.0025	
18.3			0.0025	
GD	0.7642	0.7478	0.7790	0.7410
DXS9902				
N	414	406	405	353
8	0.0024		0.0025	
9	0.0072	0.0246	0.0222	0.0170
10	0.0314	0.0443	0.0222	0.0395
11	0.3840	0.3251	0.3457	0.3230
11.1		0.0025		
12	0.3285	0.3719	0.3679	0.3400
12.1	0.0072	0.0074	0.0148	0.0055
13	0.2295	0.2044	0.2198	0.2605
13.1	0.0048	0.0049		
14	0.0048	0.0148	0.0049	0.0087
15				0.0028
16				0.0028
GD	0.6925	0.7131	0.6973	0.7122
DXS6789				
N	414	406	405	353
14	0.0121	0.0074	0.0074	0.0028
15	0.0821	0.1034	0.1012	0.1217
16	0.1376	0.0616	0.0568	0.0680
17	0.0097	0.0025	0.0025	0.0058
18	0.0048	0.0049	0.0049	0.0084
19	0.0362	0.0493	0.0593	0.0426
20	0.3019	0.2906	0.3407	0.3114
21	0.2005	0.2438	0.2494	0.2209
22	0.1667	0.1773	0.1309	0.1302
23	0.0362	0.0493	0.0420	0.0766
24	0.0121	0.0074	0.0049	0.0112
25		0.0025		
GD	0.8141	0.8071	0.7877	0.8122

N sample size, GD gene diversity

#### Linkage disequilibrium (LD) analysis

For a significance level of 0.0011 (after the Bonferroni correction for 45 comparisons in each population) only a significant  $p$  value was obtained for the DXS9898–DXS9902 pair of loci in Belo Horizonte ( $p \leq 0.0000$ , Supplementary Table S3), which are quite distant on the chromosome (over 72 Mb). The pairwise LD test for the overall sample from the southeastern region did not confirm associations between any of the maker pairs investigated. LD does not only depend on the distance between marker pairs, but may be associated to a random genetic drift, founder effect, population admixture or stratification, etc. [21]. The Brazilian population is highly heterogeneous; however, the significant DXS9898–9902 LD in Belo Horizonte is thought to be more related to sampling effects than a genetic substructure since this association was not observed in the three other admixed populations studied in this work or in other Brazilian populations [7].

#### Comparisons between Brazilian populations

Pairwise population comparisons between the studied samples and those published for other Brazilian populations that used the same set of markers [7] were performed at a single locus level (exact test of population differentiation and  $F_{ST}$  genetic distance analysis) and for the whole set of markers (AMOVA and  $F_{ST}$  genetic distance analysis).

Considering both methodologies, the results (Supplementary Table S4) showed significant differences after the Bonferroni correction ( $p < 0.005$ ) for DXS9898, DXS7133 GATA172D05 and DXS7423. For the complete set of markers, significant pairwise genetic distances ( $p < 0.005$ ) were found between Paraná and the four populations from the southeastern region, as well as between São Paulo and the other three southeastern populations (Supplementary Table S5). The overall  $F_{ST}$  value found in the AMOVA ( $F_{ST} = 0.0028$ ,  $p \leq 0.000$ ) showed a significant differentiation among Brazilian populations and a similar result was obtained when only the southeastern groups were compared.

#### Comparisons with other populations

The  $F_{ST}$  was performed between Brazilian and other populations from Europe, Latin American and Africa [7], those whose data are available for the same markers studied in this work. In most comparisons, significant genetic distances were obtained (Supplementary Table S6). The results revealed that São Paulo, Rio de Janeiro, Vitória and Belo Horizonte are, in general, are more similar to the European populations (Portugal and

**Table 2** Forensic parameters in the capital cities of the southeastern region of Brazil

	He	Ho	P-HWE	MECT	MECD	PDM	PDF
São Paulo							
DXS8378	0.6773	0.7012	0.9636	0.6076	0.4606	0.6735	0.8275
DXS9898	0.7789	0.7927	0.4624	0.7458	0.6146	0.7778	0.9186
DXS7133	0.6566	0.5644	0.0057	0.5876	0.4408	0.6442	0.8168
GATA31E08	0.8023	0.7561	0.0062	0.7680	0.6417	0.7975	0.9295
GATA172D05	0.8232	0.8232	0.2342	0.7879	0.6666	0.8137	0.9395
DXS7423	0.6481	0.6890	0.4784	0.5822	0.4365	0.6424	0.8119
DXS6809	0.8170	0.7988	0.0085	0.7940	0.6758	0.8167	0.9438
DXS7132	0.7624	0.7669	0.1530	0.7246	0.5898	0.7624	0.9058
DXS9902	0.6962	0.6829	0.0719	0.6314	0.4858	0.6908	0.8450
DXS6789	0.8254	0.8049	0.3450	0.7880	0.6678	0.8122	0.9405
Rio de Janeiro							
DXS8378	0.6889	0.7034	0.9730	0.6214	0.4749	0.6838	0.8376
DXS9898	0.7990	0.7103	0.1143	0.7696	0.6434	0.7996	0.9298
DXS7133	0.7162	0.6690	0.0096	0.6528	0.5089	0.7055	0.8606
GATA31E08	0.8212	0.8138	0.1755	0.7967	0.6782	0.8209	0.9437
GATA172D05	0.8043	0.7931	0.8085	0.7833	0.6604	0.8105	0.9368
DXS7423	0.6760	0.7103	0.4441	0.6380	0.4943	0.6911	0.8515
DXS6809	0.8414	0.8276	0.7654	0.8169	0.7057	0.8366	0.9536
DXS7132	0.7460	0.6965	0.3787	0.7071	0.5696	0.7460	0.8966
DXS9902	0.7203	0.6207	0.0256	0.6600	0.5171	0.7113	0.8654
DXS6789	0.8031	0.8069	0.9098	0.7790	0.6565	0.8051	0.9359
Vitória							
DXS8378	0.6859	0.6875	0.3611	0.6201	0.4734	0.6849	0.8359
DXS9898	0.7989	0.8062	0.9734	0.7723	0.6472	0.8013	0.9315
DXS7133	0.7039	0.6478	0.1217	0.6342	0.4896	0.6908	0.8478
GATA31E08	0.8094	0.8000	0.2870	0.7803	0.6572	0.8073	0.9358
GATA172D05	0.8131	0.7687	0.1047	0.7911	0.6709	0.8154	0.9417
DXS7423	0.6605	0.6812	0.0264	0.6103	0.4651	0.6691	0.8317
DXS6809	0.8260	0.8062	0.5918	0.8133	0.7009	0.8326	0.9527
DXS7132	0.7760	0.7250	0.0774	0.7430	0.6118	0.7771	0.9162
DXS9902	0.6971	0.6437	0.3349	0.6379	0.4930	0.6956	0.8496
DXS6789	0.7861	0.7750	0.4480	0.7576	0.6303	0.7857	0.9259
Belo Horizonte							
DXS8378	0.6826	0.6296	0.0242	0.6166	0.4696	0.6829	0.8331
DXS9898	0.7941	0.7870	0.8182	0.7575	0.6290	0.7886	0.9242
DXS7133	0.6886	0.6667	0.4038	0.6278	0.4827	0.6836	0.8441
GATA31E08	0.8124	0.8148	0.9898	0.7880	0.6672	0.8135	0.9397
GATA172D05	0.8394	0.8796	0.8630	0.8137	0.6999	0.8355	0.9512
DXS7423	0.6982	0.6574	0.2393	0.6416	0.4980	0.6944	0.8538
DXS6809	0.8428	0.8611	0.9116	0.8079	0.6939	0.8290	0.9497
DXS7132	0.7276	0.7222	0.6497	0.6983	0.5593	0.7389	0.8912
DXS9902	0.7060	0.6852	0.5685	0.6551	0.5111	0.7102	0.8609
DXS6789	0.8198	0.8241	0.3875	0.7861	0.6657	0.8099	0.9401

He expected heterozygosity, Ho observed heterozygosity, P-HWE *p* values for Hardy–Weinberg equilibrium test, MECT mean exclusion chance in trios involving daughters, MECD mean exclusion chance in father/daughter duos, PDM power of discrimination in males, PDF power of discrimination in females.

Spain), followed by Latin American (Argentina, Costa Rica and Colombia), and more distant to the African population from Uganda.

In conclusion, the data obtained in this work shows that the present X-STR set is highly informative in the southeastern region of the country and that specific databases for this multiplex system should be used in forensic casework and kinship analysis in Brazilian populations. Until this time, including this article, 21 X-STR markers were reported by 17 Brazilian populations, covering ten urban groups and seven Amerindian groups [5, 7, 13–16]. This decaplex system was studied in six of the ten urban groups and six of these markers (DXS7132, DXS9898, DXS6789, DXS7133, GATA31E08, DXS7423) were the most studied, being reported in more than 15 groups. Therefore, the result of this work will contribute to establish this system as standard for the X-STRs in the Brazilian forensic context.

**Acknowledgements** The authors would like to acknowledge the Instituto Estadual de Hematologia Arthur de Siqueira Cavalcanti, the Fundação Centro de Hematologia e Hemoterapia de Minas Gerais and the Centro de Hematologia e Hemoterapia do Espírito Santo for their assistance in the sample collection. We also acknowledge Otávio H. Thiemann, PhD (Instituto de Física de São Carlos, São Paulo) by his laboratorial assistance. This study was supported by the Programa de Apoio ao Desenvolvimento Científico (PADC) da Faculdade de Ciências Farmacêuticas da UNESP (proc. 2008/7-I) and Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP (proc. 2008/10505-4). JAM was the recipient of a fellowship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the Fundación Carolina; JCC and GGP were the recipient of a fellowship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). PSD is supported by the Isidro Parga Pondal Program (Plan Galego de Investigación, Desenvolvemento e Innovación Tecnolóxica-INCITE (2006–2010) from Xunta de Galicia, Spain). IPATIMUP is partially supported by the Fundação para a Ciência e a Tecnologia, through POCI (Programa Operacional Ciência e Inovação 2010).

## References

- Szibor R (2007) X-chromosomal markers: past, present and future. *Forensic Sci Int Genet* 1:93–99
- Lee HY, Park MJ, Jeong CK, Lee SY, Yoo JE, Chung U, Choi JH, Kim CY, Shin KJ (2004) Genetic characteristics and population study of 4 X-chromosomal STRs in Koreans: evidence for a null allele at DXS9898. *Int J Leg Med* 118:355–360
- Robino C, Giolitti A, Gino S, Torre C (2006) Development of two multiplex PCR systems for the analysis of 12 X-chromosomal STR loci in a northwestern Italian population sample. *Int J Leg Med* 120:315–318
- Gomes I, Prinz M, Pereira R, Meyers C, Mikulasovich RS, Amorim A, Carracedo A, Gusmão L (2007) Genetic analysis of three US population groups using an X-chromosomal STR decaplex. *Int J Leg Med* 121:198–203
- Rodrigues EMR, Leite FPN, Hutz MH, Palha TJBF, Santos AKCR, Santos SEB (2008) A multiplex PCR for 11 X chromosome STR markers and population data from a Brazilian Amazon region. *Forensic Sci Int Genet* 2:154–158
- Hwa HL, Chang YY, Lee JCI, Yin HY, Chen YH, Tseng LH, Su YN, Ko TM (2009) Thirteen X-chromosomal short tandem repeat loci multiplex data from Taiwanese. *Int J Leg Med* 123:263–269
- Gusmão L, Sánchez-Diz P, Alves C et al (2009) A GEP-ISFG collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations. *Int J Leg Med* 123:227–234
- Grattapaglia D, Schmidt AB, Costa e Silva C, Stringher C, Fernandes AP, Ferreira ME (2001) Brazilian population database for the 13 STR loci of the AmpFISTR profiler plus and cofilin multiplex kits. *Forensic Sci Int* 118:91–94
- Whittle MR, Romano NL, Negreiros VA (2004) Updated Brazilian genetic data, together with mutation rates, on 19 STR loci, including D10S1237. *Forensic Sci Int* 139:207–210
- Grattapaglia D, Kalupnick S, Guimarães CS, Ribeiro MA, Diener PS, Soares CN (2005) Y-chromosome STR haplotype diversity in Brazilian populations. *Forensic Sci Int* 149:99–107
- Pereira RW, Monteiro EHG, Hirschfeld GCR, Wang AY, Grattapaglia D (2007) Haplotype diversity of 17 Y-chromosome STRs in Brazilians. *Forensic Sci Int* 171:226–236
- Gois CC, Martins JA, Pereira GA, Freschi A, Paneto GG, Alvarenga VLS, Cicarelli RMB, Hirata MH, Oliveira RN (2008) Genetic population data of 12 STR loci of the PowerPlex Y system in the state of São Paulo population (southeast of Brazil). *Forensic Sci Int* 174:81–86
- Cainé LM, Pontes L, Abrantes D, Lima G, Pinheiro F (2007) Genetic data of four X-chromosomal STRs in a population sample of Santa Catarina, Brazil. *J Forensic Sci* 52:502–503
- Martins JA, Silva RHA, Freschi A, Paneto GG, Oliveira RN, Cicarelli RMB (2008) Population genetic data of five X-chromosomal loci in Bauru (São Paulo, Brazil). *Forensic Sci Int Genet Suppl* 1:155–156
- Leite FPN, Santos SEB, Rodrigues EMR, Callegari-Jacques SM, Demarchi DA, Tsuneto LT, Petzl-Erler ML, Salzano FM, Hutz MH (2009) Linkage disequilibrium patterns and genetic structure of Amerindian and non-Amerindian Brazilian populations revealed by long-range X-STR Markers. *Am J Phys Anthropol* 139:404–412
- Tavares CC, Gusmão L, Domingues C, Domingues P, Silva DA, Aquino J, Peixe C, Amorim A, Carvalho EF (2008) Population data for six X-chromosome STR loci in a Rio de Janeiro (Brazil) sample: usefulness in forensic casework. *Forensic Sci Int Genet Suppl* 1:164–166
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1:47–50
- Desmarais D, Zhong Y, Chakraborty R, Perreault C, Busque L (1998) Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA). *J Forensic Sci* 43:1046–1049
- Gomes I, Alves C, Maxzud K, Pereira R, Prata MJ, Sánchez-Diz P, Carracedo A, Amorim A, Gusmão L (2007) Analysis of 10 X-STRs in three African populations. *Forensic Sci Int Genet* 1:208–211
- Gomes I, Pereira V, Gomes V, Prata MJ, Pinto N, Carracedo A, Amorim A, Gusmão L (2009) The Karimojong from Uganda: genetic characterization using an X-STR decaplex system. *Forensic Sci Int Genet* 3:e127–128
- Chakravarti A (1999) Population genetics—making sense out of sequence. *Nat Genet* 21:56–60