

**Universidade Federal do Rio Grande do Sul  
Faculdade de Medicina  
Programa de Pós-Graduação em Medicina: Ciências Médicas**

**FREQÜÊNCIAS E DISTRIBUIÇÃO HAPLOTÍPICA DE STRs DO  
CROMOSSOMO Y EM INDIVÍDUOS DO RIO GRANDE DO SUL**

Rodrigo Rodenbusch

Orientador: Profa. Dra. Maria Luiza Saraiva Pereira

Dissertação de Mestrado

2008

**Universidade Federal do Rio Grande do Sul  
Faculdade de Medicina  
Programa de Pós-Graduação em Medicina: Ciências Médicas**

**FREQÜÊNCIAS E DISTRIBUIÇÃO HAPLOTÍPICA DE STRs DO  
CROMOSSOMO Y EM INDIVÍDUOS DO RIO GRANDE DO SUL**

Rodrigo Rodenbusch

Profa. Dra. Maria Luiza Saraiva Pereira

Dissertação apresentada ao Programa de Pós-Graduação em Medicina: Ciências Médicas, UFRGS, como requisito parcial para obtenção do título de Mestre.

Dissertação de Mestrado

2008

## **Agradecimentos**

- A minha orientadora Prof. Dra. Maria Luiza Saraiva Pereira, por orientar todos esses anos e me ensinar muitas coisas de ordem profissional e pessoal que levarei sempre comigo;
- Aos meus “ex” colegas e ainda amigos do Laboratório de Identificação Humana, Ana Carolina Mardini, André Gastaldo e Hugo Bock, pelas horas de trabalho e brincadeiras juntos;
- Ao pessoal do Serviço de Genética Médica - HCPA, principalmente as “bolsistas da Luiza” pelo apoio e amizade de todas;
- Aos meus colegas do Laboratório de Investigação de Paternidade - CDCT - FEPPS, que me acolheram de forma extraordinária;
- A Taiane Moura dos Santos, por ter estado ao meu lado durante um curto mais importante período da minha vida;
- Ao meu filho Vinícius Rodenbusch, que apesar de não compreender ainda, é a coisa mais preciosa da minha vida, minha fonte de inspiração para tudo o que faço;
- A minha irmã Rochelle e meus afilhados João Pedro e Vitória, pela palavra amiga e carinho nas horas que eu precisava;
- Aos meus pais Valdir e Sirlei, pelo amor, pelo carinho e por sempre acreditaram em mim;

A todos, do fundo do coração, muito obrigado.

## Sumário

<b>Lista de abreviaturas e siglas</b> .....	5
<b>Resumo</b> .....	6
<b>Abstract</b> .....	7
<b>1- Introdução</b> .....	8
<b>2- Revisão da Literatura</b> .....	10
<b>2.1- Histórico da identificação humana</b> .....	10
2.1.1- Sistema ABO .....	10
2.1.2- Sistema HLA .....	11
2.1.3- DNA - Ácido desoxirribonucléico .....	12
<b>2.2- Identificação humana através da molécula do DNA</b> .....	13
2.2.1- Vantagens do DNA sobre a sorologia tradicional .....	13
2.2.2- Transmissão da informação genética .....	14
2.2.3- Polimorfismos e regiões hipervariáveis .....	14
<b>2.3- O DNA como ferramenta na identificação humana</b> .....	16
2.3.1- Estudos antropológicos .....	16
2.3.2- Perícia forense .....	16
2.3.3- Investigação de paternidade .....	18
a) Marcadores autossômicos .....	19
b) Marcadores do cromossomo Y .....	20
<b>3- Objetivos</b> .....	23
<b>4- Referências</b> .....	24
<b>5- Artigo em português</b> .....	30
<b>6- Artigo em inglês</b> .....	46

## Lista de Abreviaturas e Siglas

CODIS	sistema de DNA índice combinado ( <i>combined DNA index system</i> )
DNA	ácido desoxirribonucléico ( <i>deoxyribonucleic acid</i> )
FBI	escritório federal de investigação dos EUA ( <i>federal bureau of investigation</i> )
GSE	grupos sanguíneos eritrocitários
HLA	antígenos leucocitários humanos ( <i>human leukocyte antigens</i> )
NRY	região não recombinante do cromossomo Y ( <i>non-recombining region of the Y chromosome</i> )
PAR1	região pseudo-autossômica 1 ( <i>pseudoautosomal region 1</i> )
PAR2	região pseudo-autossômica 2 ( <i>pseudoautosomal region 2</i> )
pb	pares de bases
PCR	reação da polimerase em cadeia ( <i>polymerase chain reaction</i> )
Rh	Fator Rhesus
RNA	ácido ribonucléico ( <i>ribonucleic acid</i> )
STR	repetições sucessivas pequenas ( <i>short tandem repeat</i> )
VNTR	repetições sucessivas de número variável ( <i>variable number of tandem repeats</i> )

## Resumo

Os marcadores genéticos denominados *short tandem repeats* (STR) são amplamente utilizados na identificação humana, tanto aqueles localizados nos cromossomos autossômicos como os encontrados no cromossomo Y. Durante a última década, muitas pesquisas demonstraram a existência de vários polimorfismos no cromossomo Y. O objetivo deste estudo foi caracterizar, na população regional, os diferentes haplótipos utilizando 12 *loci* (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 e DYS439) e determinar as taxas de diversidade haplotípica e poder de identificação individual. A população testada foi composta por 162 pares de pais/filhos do sexo masculino. O DNA foi isolado a partir de sangue periférico, utilizando o kit comercial *Wizard<sup>®</sup> Genomic DNA Purification Kit* (Promega). As regiões de interesse foram amplificadas pela PCR multiplex com oligonucleotídeos marcados com fluorescência e os produtos analisados por eletroforese capilar no analisador genético *ABI 3100* (Applied Biosystems). Na amostra analisada, 151 haplótipos distintos foram encontrados, onde os dois haplótipos mais comuns apresentaram frequências de 0,0265 e 0,0199, respectivamente. Outros 6 haplótipos distintos apresentaram frequência de 0,0132 e a frequência dos outros 143 haplótipos foi 0,0066 (1 ocorrência cada). Os resultados obtidos permitiram ainda a determinação de um valor de diversidade haplotípica de 0,9982 e um poder de discriminação individual de 0,9321. Além disso, após a determinação destas frequências e das taxas estudadas foi possível comprovar que a utilização desses marcadores apresenta um alto poder de discriminação na correta identificação de indivíduos, tanto em casos de paternidade como na área forense.

## **Abstract**

The genetic markers named short tandem repeats (STR) are largely employed in human identification, those located on autosomal chromosomes as well as on Y-chromosome. During the last decade, several studies showed the occurrence of various polymorphisms on Y-chromosome. The aim of this study was to identify distinct haplotypes using 12 loci (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439) and to determine haplotype diversity rates and individual identification power. Tested population was composed by 162 pairs father-male sib. DNA was isolated from peripheral blood, using the Wizard<sup>®</sup> Genomic DNA Purification Kit (Promega). Regions of interest were amplified by multiplex PCR with fluorescent primer and products were analyzed by capillary electrophoresis in the genetic analyzer ABI 3100 (Applied Biosystems). In the tested population, 151 distinct haplotypes were found, where frequencies of two common haplotypes were established to be 0.0265 and 0.0199, respectively. Individual frequencies of other 6 haplotypes were 0.0132 and frequency of the remaining 143 were 0.0066 (1 event each). Obtained results also allowed to determine haplotype diversity value of 0.9982 and individual discrimination power of 0.9321. Finally, results of this study indicate that application of these markers is responsible for a high discrimination power in individual identification on both paternity and forensic cases.

## 1- INTRODUÇÃO

A busca pela caracterização individual assim como o estabelecimento de laços familiares proporcionaram diversas descobertas ao longo do tempo. Estas descobertas começaram a ser desvendadas no início do século XX, com a descrição do grupo sanguíneo ABO. Com o avanço da ciência juntamente com a descoberta de outros sistemas sanguíneos, os processos de identificação humana sofreram progressos consideráveis.

Em meados de 1970, o sistema de antígenos leucocitários humanos (HLA) foi reconhecido cientificamente como um método para avaliação em testes de paternidade e utilizado em conjunto com os outros sistemas. Estes testes, usando múltiplos sistemas, forneceram um meio prático para a exclusão de pelo menos 95% dos supostos pais; no entanto, ainda não era totalmente adequado para este fim<sup>1,2</sup>.

Os resultados das análises de DNA, ao contrário das análises convencionais de grupos sanguíneos e sistemas protéicos (que dependem da disponibilidade de sangue ou fluídos corporais relacionados), não dependem da natureza do material ou célula analisada, pois a informação genética está contida integralmente em todas as células somáticas de um indivíduo. Outra vantagem da análise do DNA é que este é fisicamente muito mais resistente do que as proteínas<sup>3</sup>.

A identificação humana através da estrutura do DNA é a metodologia mais moderna aplicada mundialmente e está baseada em diferenças na seqüência de DNA de uma pessoa para outra. Estas diferenças genéticas são conhecidas como polimorfismos.

Os polimorfismos genéticos correspondem a formas alternativas de seqüências de nucleotídeos e, quando analisados ao nível de DNA, estas seqüências diferem entre si por alterações do tipo troca de base ou por alterações no tamanho do fragmento analisado. Dentro deste segundo grupo, podemos identificar os polimorfismos originados pela alteração do número de seqüências de repetições consecutivas (*in tandem*) que pode ser dividido em duas classes: as repetições sucessivas de número variável (VNTR) e as repetições sucessivas pequenas (STR)<sup>4</sup>.

Na identificação humana, as análises de STR localizadas em cromossomos autossômicos são o primeiro foco de uma análise. Durante a última década, muitas pesquisas descreveram um alto grau de polimorfismo nos STR localizados no cromossomo Y (Y-STR), tornando-se estes, em alguns casos, uma poderosa ferramenta na identificação humana, principalmente em casos de violência sexual e em teste de paternidade quando o suposto pai é falecido ou não é encontrado<sup>5</sup>.

A utilização dos Y-STR como ferramenta de identificação está baseada nas diferenças entre os cromossomos X e Y e os autossômicos, onde os marcadores no cromossomo Y se encontram em regiões não-recombinantes (NRY) ao X<sup>6</sup>. Estas regiões facilitam a identificação do indivíduo do sexo masculino que contribui com material genético em uma mistura macho/fêmea (X/Y), uma vez que estes *loci* são geneticamente ligados e herdados inalterados de pai para filho<sup>7</sup>.

## **2- REVISÃO DA LITERATURA**

### **2.1- Histórico da identificação humana**

#### **2.1.1- Sistema ABO**

A utilização de amostras biológicas teve seu início no princípio do século XX, com descoberta de Landsteiner do grupo ABO. Esta descoberta só foi aplicada para a identificação humana em 1910, por Von Dungern e Hirsfeld, e ficou conhecida como herança mendeliana do grupo sanguíneo. O próximo evento significativo ocorreu em 1927, quando o sistema MN foi descrito por Landsteiner e Levine, e este evento foi seguido pela descoberta, em 1940, do sistema Rh por Levine e colaboradores<sup>1</sup>.

A aplicação do sistema Rh fornece uma média de exclusão de paternidade de aproximadamente 55% dos homens supostamente acusados. Por outro lado, a combinação dos sistemas ABO, MN e Rh, tornou possível excluir 70% dos homens suspeitos por uma paternidade. Assim estes testes começaram a ser considerados pelo seu poder de exclusão, mesmo que estes tivessem alcance limitado<sup>8</sup>.

Com o passar dos anos, outros sistemas de grupos sanguíneos, herdados independentemente entre si, foram descritos. Entre eles podemos citar os sistemas Kell, Lewis, Duffy e outros<sup>9,10</sup>.

Atualmente mais de 20 sistemas de grupos sanguíneos com determinação genética são conhecidos e mais de 400 antígenos em hemácias identificados por sorologia. Os antígenos das hemácias são epítomos localizados na superfície celular capazes de induzir a produção de anticorpos. Alguns pertencem a proteínas estruturais da membrana celular, outros a ramificações de glicoproteínas, localizando-se no

glicocálice, e outros são moléculas do plasma que se ligam à superfície das células vermelhas<sup>11</sup>.

Devido ao baixo poder discriminatório dos grupos sanguíneos eritrocitários (GSE), dentre outras desvantagens, estes foram substituídos na maioria dos centros responsáveis por este tipo de identificação.

### **2.1.2- Sistema HLA**

Em 1954, foi demonstrado a ocorrência de um sistema de histocompatibilidade mediado por antígenos na superfície dos leucócitos, conhecido por complexo HLA (*histocompatibility leucocyte antigen*) determinado por genes alélicos muito próximos localizados no braço curto do cromossomo 6, com acentuado poder de discriminação individual ou determinação da individualidade genética<sup>12</sup>.

Os estudos sorológicos do HLA são realizados por intermédio da análise dos antígenos na superfície dos linfócitos do sangue dos indivíduos, usando anticorpos HLA. Estes anticorpos são de origem humana ou produzidos pela tecnologia monoclonal, mas podem apresentar forte reação cruzada entre os alelos, necessitando de vários reagentes para a classificação correta dos antígenos<sup>13</sup>.

Até a década de 80, a tipificação sanguínea era realizada por métodos sorológicos que detectavam antígenos e não alelos específicos. Os antígenos eram identificados com o prefixo HLA seguido por letras maiúsculas para a denominação do *locus* gênico e um ou dois dígitos para identificação numérica do antígeno<sup>14, 15</sup>.

Os genes do complexo HLA são os mais polimórficos dos mamíferos e é muito provável que estes polimorfismos sejam importantes para garantir a sobrevivência das espécies<sup>16</sup>. A herança de um haplótipo materno e um paterno em co-dominância faz com

que cada pessoa possa expressar em suas células diferentes moléculas do HLA. Quanto mais polimórfico for o complexo HLA numa população, maior será a probabilidade de herdar haplótipos diferentes e, portanto, ser heterozigoto para todos os *loci*<sup>17</sup>. Além da sorologia tradicional, outro método de identificação é a tipagem molecular do sistema HLA, que permite avaliar diretamente todos os genes HLA. Essa tecnologia baseia-se no estudo dos genes dos *loci* HLA – A, B, C, através da amplificação dos mesmos<sup>15</sup>.

### **2.1.3- DNA - Ácido desoxirribonucléico**

Os ácidos nucleicos foram descobertos, em 1869, por Friedrich Miescher, um médico suíço de 22 anos de idade. Miescher isolou uma macromolécula, nunca antes identificada, à qual deu o nome de nucleína, posteriormente denominada ácido nucleico. No início do século XX, o bioquímico Kossel comprovou a existência de dois tipos de ácidos nucleicos: o ácido desoxirribonucléico (ADN ou DNA) e o ácido ribonucléico (ARN ou RNA).

O DNA era apenas estudado do ponto de vista bioquímico até 1953, quando James Watson e Francis Crick sugeriram a estrutura da molécula do DNA. Seu formato é descrito como uma estrutura em dupla hélice, com moléculas que apresentam uma estrutura com átomos de carbono e nitrogênio, denominadas bases nitrogenadas<sup>18</sup>.

A utilização do DNA como ferramenta voltada à identificação humana veio com a publicação de Jeffreys e colaboradores sobre certas regiões do genoma humano que produziam uma espécie de “impressões digitais” de DNA<sup>4</sup>. Esses pesquisadores usaram essa metodologia para solucionar um problema de imigração na Inglaterra. A partir deste caso, a identificação humana, através da técnica de genotipagem do DNA, tem

sido uma potente ferramenta no esclarecimento de diversos crimes e casos de vínculos familiares<sup>19</sup>.

## **2.2- Identificação humana através da molécula do DNA.**

### **2.2.1- Vantagens do DNA sobre a sorologia tradicional**

O objetivo da identificação ou caracterização do material biológico é limitar ou reduzir o número de indivíduos que poderiam ser a fonte deste material. Quando temos uma população limitada ou restrita, esta identificação pode ser feita por marcadores com baixo poder discriminatório. Porém, nos casos em que a população não é limitada, os métodos de maior poder discriminatório tornam-se indispensáveis. O DNA, ao contrário da sorologia tradicional (ABO, Rh, Mn, HLA, entre outros), pode atingir um alto poder discriminatório, necessário para inferir a identificação<sup>20</sup>.

Os exames envolvendo o grupo sanguíneo ABO podem discriminar aproximadamente um em cada três indivíduos na população, e, mesmo com marcadores sorológicos adicionais, os valores em média são de um em alguns milhares, enquanto que, com o DNA, os resultados podem chegar a um em alguns bilhões ou mais.

Uma das principais vantagens das análises de DNA em relação à sorologia tradicional é que estas não dependem da natureza do material ou célula analisada, pois a informação genética está contida integralmente em todas as células somáticas de um indivíduo.

Outra vantagem da análise do DNA é que este é fisicamente muito mais resistente à degradação do que as proteínas. As proteínas podem ser desnaturadas de

forma relativamente mais fácil e sua estrutura terciária, que é importante na tipagem, é altamente vulnerável e esta desnaturação. A informação genética, por sua vez, é encontrada na seqüência nucleotídica, que independe da conformação da molécula<sup>3</sup>.

### **2.2.2- Transmissão da informação genética**

O material genético de um indivíduo é encontrado dentro do núcleo das células somáticas, em estruturas denominadas cromossomos. Os cromossomos ocorrem aos pares nas células chamadas diplóides, isto é, possuem dois conjuntos de cromossomos, um de origem paterna outro de origem materna.

Cada espécie possui um número constante de cromossomos. Na espécie humana, as células diplóides apresentam 46 cromossomos ou 22 pares de cromossomos homólogos e um par de cromossomos sexuais.

A transferência do material genético dos pais para os filhos ocorre através de um processo de divisão celular conhecido como meiose, onde uma célula dá origem a quatro novas células com a metade do número de cromossomos da célula inicial. Esse tipo de divisão ocorre no processo de formação de gametas e está de acordo com a 1ª lei de Mendel<sup>21</sup>.

### **2.2.3- Polimorfismos e regiões hipervariáveis**

A variabilidade genética entre indivíduos da raça humana é extremamente alta. Dois genomas humanos escolhidos ao acaso e sem vínculo familiar entre eles diferem aproximadamente em uma de cada 500 bases do DNA. Levando em conta que o genoma humano tem cerca de 3 bilhões de bases, podemos estimar que existam 6 milhões de diferenças entre duas pessoas<sup>22</sup>.

Essas diferenças na seqüência de DNA de uma pessoa para outra são conhecidas geneticamente como polimorfismos e são as bases para a identificação humana através do DNA.

Os polimorfismos genéticos correspondem a formas alternativas de seqüências de nucleotídeos e, quando analisados ao nível do DNA, estas seqüências diferem entre si por alterações do tipo troca de base ou por alterações no tamanho do gene. Dentro deste segundo grupo, podemos identificar os polimorfismos originados pela alteração do número de seqüências de repetições consecutivas (*in tandem*) que pode ser dividido em duas classes: as repetições sucessivas de número variável (VNTR) e as repetições sucessivas pequenas (STR)<sup>4,23</sup>.

Os marcadores polimórficos conhecidos como VNTR (ou minissatélites) compreendem várias unidades de repetições sucessivas, cada uma com 8 a 80 pares de base (pb). O número exato de repetições assim como o comprimento da região de VNTR varia de um alelo para outro e diferentes alelos podem ser identificados pelo tamanho do seu fragmento<sup>23</sup>.

Os microssatélites ou *short tandem repeats* (STR) são seqüências de 1 a 6 pb que estão repetidas várias vezes ao longo do genoma e se caracterizam por serem altamente polimórficas<sup>23</sup>. Os STR se tornaram uma poderosa ferramenta para identificação humana, uma vez que, para sua detecção, podemos utilizar a técnica da reação da polimerase em cadeia (PCR), amplificando-se o DNA a partir de oligonucleotídeos situados nas regiões adjacentes não repetitivas<sup>25</sup>.

## **2.3- O DNA como ferramenta na identificação humana**

### **2.3.1- Estudos antropológicos**

Os marcadores genéticos são candidatos ideais para diversas aplicações, incluindo estudos antropológicos com reconstrução de grupos filogenéticos, aplicações na ciência forense e na investigação de paternidade/maternidade<sup>26</sup>.

As diversas aplicações destes marcadores levam em conta as propriedades dos STR em estudo. Nos estudos forenses, os STR escolhidos devem ser bastante polimórficos, com um elevado número de alelos. Ao contrário, nos estudos antropológicos, os *loci* de STR preferidos são aqueles que apresentam um baixo número de alelos e que carregam alelos típicos de uma população específica<sup>27,28</sup>.

Os estudos antropológicos estão baseados inicialmente na caracterização da distribuição alélica de uma determinada população em determinados *loci* previamente selecionados. Através das frequências alélicas e de cálculos estatísticos, são estimadas as distâncias genéticas, comparando a população em estudo com diferentes populações mundiais<sup>29,30</sup>.

### **2.3.2- Perícia forense**

Na prática forense, um dos objetivos da identificação humana através da molécula de DNA, é testar uma hipótese de que determinada pessoa é a fonte doadora de uma evidência biológica. Dentre os possíveis resultados desta investigação constam a exclusão (as amostras biológicas possuem origens diferentes), os resultados inconclusivos (não é possível determinar se as amostras provêm ou não do mesmo doador) e a inclusão (as amostras são similares e originaram-se da mesma fonte)<sup>31</sup>.

O fato de duas amostras terem o mesmo perfil para um grupo de marcadores genéticos em especial não significa, obrigatoriamente, que elas apresentem a mesma origem. Quando a identificação genética de duas amostras é igual, torna-se necessário expressar numericamente a significância deste evento. O número de marcadores empregados, a presença de sub-estruturas na população e a mistura de amostras podem interferir nos resultados. A expressão estatística dos resultados deve basear-se na presença ou não de misturas de material biológico, como é freqüentemente observado em casos de abuso sexual<sup>32</sup>.

Outra característica muito importante para o sucesso da identificação forense está relacionada à qualidade das amostras ou rastros biológicos encontrados na cena do crime ou no corpo da vítima e sua conservação. A exposição do DNA a fatores como luz solar, microorganismos e componentes químicos pode provocar a degradação da molécula. Logo, quanto melhor for a coleta e a preservação do material coletado, melhor será a análise do material genético extraído.

Nas tragédias em massa e em guerras, o foco principal das investigações é determinar a identidade das vítimas e enviar os restos mortais para seus familiares, isto é praticado de acordo com as convenções de Genebra. Para que isso ocorra o mais breve possível, os peritos comparam dados anteriores das vítimas, como por exemplo, fraturas e deformações ósseas, histórico dentário, assim como as roupas que as vítimas utilizavam no momento da morte<sup>33</sup>. Quando esses dados não estão presentes ou são insuficientes, as técnicas de genotipagem do DNA são ferramentas para uma correta identificação destas vítimas. Esta tecnologia de identificação humana através do DNA, incluindo análises de STR e análises de DNA mitocondrial, foi o método escolhido na

identificação de pessoas desaparecidas entre 1991-1995 durante a guerra da Croácia; por exemplo<sup>34</sup>.

### **2.3.3- Investigação de paternidade**

A investigação de vínculo genético consiste na comparação de regiões presentes no DNA, que apresentam variação entre as pessoas da população.

O compartilhamento de alelos entre o filho e o suposto pai, quando diversos marcadores são analisados, permite confirmar a paternidade com uma probabilidade igual ou superior a 99,99%. Da mesma forma, quando os alelos não são compartilhados entre o filho e o suposto pai em três ou mais marcadores, este suposto pai é excluído da possibilidade de ser o pai biológico deste filho com uma probabilidade de 100%, ou seja, a probabilidade de paternidade é igual a zero. Atualmente, a possibilidade de atribuímos ou excluirmos a paternidade biológica com segurança se dá através das análises das regiões hipervariáveis espalhadas pelo genoma<sup>35</sup>.

Se todos os polimorfismos de uma pessoa fossem testados, poderíamos afirmar que aquele perfil é único. Mas isso não acontece, conseqüentemente existe uma certa probabilidade estatística de que dois indivíduos, escolhidos ao acaso, venham a ter um perfil idêntico em certa quantidade de marcadores. Para casos como este, existem cálculos para mensuração da razão de verossimilhança. Este grau de semelhança é medido pela freqüência de incidência de um determinado perfil (perfil encontrado no suposto pai) em relação a qualquer outra pessoa da população<sup>36</sup>.

Os resultados dos cálculos da freqüência da incidência de um determinado perfil genético na população são normalmente expressos com notações que indicam quantas

vezes o perfil se repete na população, por exemplo, 1 em 5 bilhões de pessoas ou 1 em 20 bilhões de pessoas<sup>37</sup>.

Na investigação de paternidade, para cada marcador, deverá haver coincidência entre um dos alelos do filho com um dos presentes em sua mãe (alelo obrigatório). O outro alelo do filho, deverá estar presente no pai, e recebe o nome de alelo paterno-obrigatório<sup>38</sup>.

Em casos de paternidade onde a mãe e o filho são heterozigotos e apresentam o mesmo genótipo ou quando a mãe é falecida, o alelo paterno freqüentemente não pode ser determinado. Em consequência, a probabilidade da exclusão ou inclusão seria reduzida consideravelmente<sup>39</sup>. Para que isso não ocorra e esta probabilidade não seja reduzida, um número maior de marcadores deve ser utilizado para que se possa determinar a origem dos alelos paternos<sup>40</sup>.

#### **a) Marcadores autossômicos**

Diversos sistemas de STR são utilizados e reconhecidos na área de investigação de paternidade devido ao alto grau de evidências, permitindo o esclarecimento de diversas questões relacionadas aos casos de vínculo genético. Desta maneira, os sistemas STR substituíram de vez os sistemas baseados em grupos sanguíneos<sup>41</sup>.

Para os exames interlaboratoriais e Banco de Dados, os marcadores moleculares utilizados devem seguir os treze marcadores definidos pelo *combined DNA index system* (CODIS) que incluem os seguintes STRs: TPOX, D3S1358, D5S818, FGA, CSF1PO, D7S820, D8S1179, TH01, vWA, D13S317, D16S539, D18S51 e D21S11 e a amelogenina. A escolha desses marcadores moleculares pelo FBI, tornou-se

praticamente padrão nos bancos de perfis de criminosos e em laboratórios forenses em todo o mundo. Estes marcadores estão disponíveis em kit comerciais<sup>42</sup>.

Na identificação humana, as análises de STR localizadas em cromossomos autossômicos são o primeiro foco de uma investigação. Porém, em alguns casos, a utilização de marcadores uniparentais em conjunto com marcadores autossômicos se faz necessária. Os marcadores localizados cromossomo Y são transmitidos através de linhagens patrilineas e são bastante empregados na resolução de casos de agressões sexuais e em casos de paternidade, onde o suposto pai não se encontra disponível<sup>43</sup>.

#### **b) Marcadores do cromossomo Y**

O cromossomo Y representa apenas 2% do genoma humano e possui um tamanho aproximado de 60 Mb. Este cromossomo apresenta duas regiões pseudo-autossômicas (PAR1 e PAR2) localizadas nas extremidades do braço curto e longo do cromossomo. Estas regiões são homologas a regiões do cromossomo X e são responsáveis pelo correto pareamento dos dois cromossomos sexuais durante a meiose<sup>6</sup>.

A utilização dos Y-STRs como ferramenta de identificação está baseada nas diferenças entre os cromossomos X e Y e os autossômicos, onde os marcadores no cromossomo Y se encontram em regiões não-recombinantes (NRY) ao X<sup>6</sup>. Estas regiões facilitam a identificação do indivíduo do sexo masculino que contribui com material genético em uma mistura macho/fêmea (X/Y), uma vez que estes *loci* são geneticamente ligados e herdados inalterados de pai para filho<sup>7</sup>.

De acordo com a comunidade forense europeia, os marcadores DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS385a/b constituem o chamado

haplótipo mínimo dos Y-STRs. Quando acrescido do marcador YCAII a/b, este conjunto de *loci* também pode ser chamado de haplótipo estendido dos Y-STR<sup>44</sup>.

Diversos Y-STR estão sendo utilizados como marcadores em análises de identificação humana. Entre eles podemos citar os marcadores denominados DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439 entre os mais frequentemente utilizados<sup>45</sup>.

Vários grupos de pesquisas espalhados pelo mundo trabalham no intuito de determinar as frequências dos Y-STR em suas populações, como por exemplo a publicação que determinou e comparou a frequência de 27 Y-STR na população dos Estados Unidos, em três subgrupos: caucasianos, afro-americanos e hispânicos<sup>46</sup>.

Um estudo realizado no Brasil, comparando as cinco regiões geopolíticas brasileiras, indicou que 99,97% das variações haplotípicas são encontradas dentro das regiões e apenas 0,03% são variações encontradas entre as regiões. Este estudo sugeriu que um único banco nacional de Y-STRs poderia ser utilizado como base em estudos forenses e de identificação humana<sup>47</sup>.

Mesmo havendo estudos que apontem não existir diferenças significativas entre as regiões brasileiras, alguns estudos são realizados para o estabelecimento de frequências locais. Entre estes podemos citar os estudos realizados com 9 Y-STR na população da região amazônica<sup>48</sup> e o trabalho de Cainé e colaboradores realizado em Santa Catarina com 8 Y-STRs<sup>49</sup>.

Outros estudos, além de tentarem estabelecer as frequências locais, tentam estabelecer as frequências e as diferenças haplotípicas de sub-populações que são frequentemente encontradas no Brasil. O estudo de Domingues e colaboradores, realizado com uma sub-população de afro-descendentes residentes no Rio de Janeiro,

indicou haver uma diferença altamente significativa entre os afro-descendentes e o restante da população do Rio de Janeiro<sup>50</sup>.

Comparações com amostras de outras regiões do Brasil mostraram que heterogeneidade existe, indicando que banco nacional de Y-STRs para todo o país deve levar em consideração a sub-estruturação da população<sup>50</sup>.

Devido a estes fatos, achamos importante estabelecer as frequências destes marcadores na população do Rio Grande do Sul, uma vez que o estado foi colonizado por diversas etnias e ainda apresenta uma forte sub-estruturação populacional, principalmente em regiões de colonização alemã e italiana. Além disso, o estabelecimento da estimativa das taxas de mutações nestes marcadores se faz necessário para que se possa evitar uma falsa exclusão em casos onde possa existir uma discrepância alélica entre o pai biológico e o filho devido a uma mutação<sup>5</sup>.

### **3- OBJETIVOS**

#### **3.1- Objetivo geral:**

Estabelecer, na população regional, a frequência de 12 marcadores localizados na região não-recombinante (NRY) do cromossomo Y, assim como determinar as taxas de mutação de cada um deles.

#### **3.2- Objetivos específicos:**

- Estabelecer frequência de 12 marcadores (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439), em indivíduos do sexo masculino do Rio Grande do Sul;
- Comparar as frequências encontradas na população estudada com as frequências já estabelecidas em outras populações;
- Identificar os diferentes haplótipos encontrados na população e determinar suas frequências;
- Calcular a diversidade gênica que cada *locus* apresenta.

#### 4- REFERÊNCIAS

1. Silver H. Paternity testing. *Crit Rev Clin Lab Sci* 1989; 27(5): 391-408.
2. Walker RH, Meyers MA, Phillips LM. The probability of exclusion of the HLA-A, B system in North American whites and blacks in parentage tests. *Transfusion* 1987; 27(1): 75-79.
3. Schneider PM. Basic issues in forensic DNA typing. *Forensic Sci Int* 1997; 88(1): 17-22.
4. Jeffreys AJ, Wilson V, Thein SL. Hypervariable 'minisatellite' regions in human DNA. *Nature* 1985; 314: 67-73.
5. Kurihara R, Yamamoto T, Shi-Lin Li RU, Yoshimoto T, Kamitama HOK, Katsumata Y. Mutations in 14 Y-STR loci among Japanese father-son haplotypes. *Int J Legal Med* 2004; 118: 125-131.
6. Quintana-Murci L, Krausz C, McElreavey K. The human Y chromosome: function, evolution and disease. *Forensic Sci Int* 2001; 118: 169-181.
7. Krenke BE, Viculis L, Richard ML, Prinz M, Milne SC, Ladd C, Gross AM, Gornall T, Frappier JRH, Eisenberg AJ, Barna C, Aranda XG, Adamowicz MS, Budowle B. Validation of a male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex. *Forensic Sci Int* 2005; 148: 1-14.
8. Davey, FR., Hubbell CA, Lauenstein KJ, Tinnesz C, Henry JB. Analysis of paternity. The use of HLA and red cell antigens. *Transfusion* 1984; 24(4): 340-342.
9. Singh G, Johns MM, Paul G. Paternity testing: analysis of six blood groups and HLA markers, with particular reference to comparison of races. *Am J Clin Pathol* 1982; 78(5): 748-752.

10. Sensabaugh GF. Uses of polymorphic red cell enzymes in forensic science. *Clin Haematol* 1981; 10(1): 185-207.
11. Hurley CK. HLA diversity: detection and impact on unrelated hematopoietic stem cell donor characterization and selection. *Int J Hematol* 2002; 76(2): 152-154.
12. Lopez de Castro JA, Barbosa JA, Krangel MS, Biro PA, Strominger JL. Structural analysis of the functional sites of class I HLA antigens. *Immunol Rev* 1985; 85: 149-168.
13. Singh N, Agrawal S, Rastogi AK. Infectious diseases and immunity: special reference to major histocompatibility complex. *Emerg Infect Dis* 1997; 3: 41-49.
14. Marsh SGE, Ekkehard AD, Bodmer WF, Bontrop RE, Dupont B, Erlich HA. Nomenclature for factors of the HLA system. *Tissue Antigens* 2002; 60: 407-464.
15. Schreuder GM, Hurley CK, Marsh SG, Lau M, Maiers M, Kollman C, Noreen HJ. The HLA dictionary 2001: a summary of HLA-A, -B, -C, -DRB1/3/4/5, -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR, and -DQ antigens. *Hum Immunol* 2001; 62(8): 826-849.
16. Van Rood JJ. The impact of the HLA-system in clinical medicine. *Shweiz Med Wschr* 1993; 123: 85-92.
17. Turner D. The human leucocyte antigen (HLA) system. *Vox Sanguinis* 2004; 87: 87-90.
18. Rich A. The double helix: a tale of two puckers. *Nat Struct Biol* 2003; 10(4): 247-249.

19. Jeffreys AJ, Brookfield JFY, Semeonoff R. Positive identification of an immigration test-case using human DNA fingerprints. *Nature* 1985; 317: 818-819.
20. Kirby LT. Case applications. In: *DNA fingerprinting: an introduction*. New York: Oxford University Press. 1992: 217-260.
21. Griffiths AJF *Chromosome Theory of Inheritance*. In: *An Introduction to Genetic Analysis*, 6<sup>a</sup> Ed. New York: W. H. Freeman Company. 1996: 55-90.
22. Lee HC, Gaensslen RE. DNA and other polymorphisms in forensic science. *Mosby Yearbook* 1990; 114-134.
23. Jeffreys AJ, Wilson V, Thein SL. Individual-specific 'fingerprints' human DNA. *Nature* 1985; 316: 76-79.
24. Tautz D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 1989; 17: 6463-6467.
25. Weir BS. Independence of VNTR alleles defined as fixed bins. *Genetics* 1992 130: 873-887.
26. Agrawal S, Khan F. Reconstructing recent human phylogenies with forensic STR loci: a statistical approach. *BMC Genetics* 2005; 6: 47-53.
27. Rowold DJ, Herrera RJ. Inferring recent human phylogenies using forensic STR technology. *Forensic Sci Int* 2003; 133(3): 260-265.
28. Shriver MD, Jin L, Ferrell RE, Deka R. Microsatellite data support an early population expansion in Africa. *Genome Res* 1997; 7(6): 586-591.
29. Khan F, Pandey AK, Tripathi M, Talwar S, Bisen PS, Borkar M, Agrawal S. Genetic affinities between endogamous and inbreeding populations of Uttar Pradesh. *BMC Genetics* 2007; 8(1): 12-22.

30. Pérez-Lezaun A, Calafell F, Mateu E, Comas D, Bosch E, Bertranpetit J. Allele frequencies for 20 microsatellites in a worldwide population survey. *Hum Hered* 1997; 47(4): 189-196.
31. Inman K, Rudin N. The origin of evidence. *Forensic Sci Int* 2002; 126(1): 11-16.
32. Melgaço MPS, Figueiredo ALS, Paradela ER. Perícias em DNA: a coisa certa pode ser feita de forma errada? Um estudo de caso hipotético. *Jus Navigandi*, Teresina, n. 1509, 19/08/2007 [internet]; Disponível em: <http://jus2.uol.com.br/doutrina/texto.asp?id=10285>. Acesso em: 04/12/2007.
33. Skinner M, Alempijevic D, Djuric-Srejcic M. Guidelines for international forensic bio-archaeology monitors of mass grave exhumations. *Forensic Sci Int* 2003; 134: 81-92.
34. Primorac D, Andelinovic S, Definis-Gojanovic M, Drmic I, Rezic B, Baden MM, et al. Identification of war victims from mass graves in Croatia, Bosnia, and Herzegovina by use of standard forensic methods and DNA typing. *J Forensic Sci* 1996; 41: 891-894.
35. Cifuentes LO, Martinez EH, Acuna MP, Jonquera HG. Probability of exclusion in paternity testing: time to reassess. *J Forensic Sci* 2006; 5(2): 349-350.
36. Evett IW, Lambert JA, Buckleton JS, Weir BS. Statistical analysis of a large file of data from STR profiles of British Caucasians to support forensic casework. *Int J Legal Med* 1996; 109(4): 173-177.
37. Weir BS. Statistical methods employed in evaluation of single-locus probe results in criminal identity cases. *Methods Mol Biol* 1998; 98: 83-96.

38. Allen RW, Fu J, Reid TM, Baird M. Considerations for the interpretation of STR results in cases of questioned half-sibship. *Transfusion* 2007; 47(3):515-519.
39. Lee HS, Lee JW, Han GR, Hwang JJ. Motherless case in paternity testing. *Forensic Sci Int* 2000; 114: 57–65.
40. Zhao G, Yang Q, Huang D, Yu C, Yang R, Chen H, Mei K. Huang D, YU C. Study on the application of parent-of-origin specific DNA methylation markers to forensic genetics. *Forensic Sci Int* 2005; 154: 122-127.
41. Junge A, Brinkmann B, Fimmers R, Madea B. Mutations or exclusion: an unusual case in paternity testing. *Int J Legal Med* 2005; 19: 1-4.
42. Grubwieser P, Zimmermann B, Niederstätter H, Pavlic M, Steinlechner M, Parson W. Evaluation of an extended set of 15 candidate STR loci for paternity and kinship analysis in an Austrian population sample. *Int J Legal Med* 2007; 121(2): 85-89.
43. Ayadi I, Mahfoudh-Lahiani N, Makni H, Ammar-Keskes L, Rebaï A. Combining autosomal and Y-chromosomal short tandem repeat data in paternity testing with male child: methods and application. *J Forensic Sci* 2007; 52(5): 1068-1072.
44. Butler JM. Recent Developments in Y-Short Tandem Repeat and Y-Single Nucleotide Polymorphism Analysis. *Forensic Sci Rev* 2003; 15: 91-111.
45. Gusmao L, Sanchez-Diz P, Alves C, Quintáns B, García-Poveda E, Geada H, Raimondi E, Fuente SS, Vide MC, Whittle MR, Zarrabeitia MT, Carvalho MG, Negreiros M, Solla LP, Riancho JA, Campos-Sánchez R, Vieira-Silva C, Toscanini U, Amorim A, Carracedo A. Results of the GEP-ISFG collaborative

- study on the Y chromosome STRs GATA A10, GATA C4, GATA H4, DYS437, DYS438, DYS439, DYS460 and DYS461: population data. *Forensic Sci Int* 2003; 135: 150-157.
46. Butler JM, Decker AE, Vallone PM, Kline MC. Allele frequencies for 27 Y-STR loci with U.S. Caucasian, African American, and Hispanic samples. *Forensic Sci Int* 2006; 156: 250–260.
47. Grattapaglia D, Kalupniek S, Guimarães CS, Ribeiro MA, Diener PS, Soares CN. Y-chromosome STR haplotype diversity in Brazilian populations. *Forensic Sci Int* 2005; 149(1): 99-107.
48. Palha Tde J, Rodrigues EM, dos Santos SE. Y-chromosomal STR haplotypes in a population from the Amazon region, Brazil *Forensic Sci Int* 2007; 166(3): 233-239.
49. Cainé L, Corte-Real F, Vieira DN, Carvalho M, Serra A, Lopes V, Vide MC. Allele frequencies and haplotypes of 8 Y-chromosomal STRs in the Santa Catarina population of southern Brazil. *Forensic Sci Int* 2005; 148: 75-79.
50. Domingues PM, Gusmão L, da Silva DA, Amorim A, Pereira RW, de Carvalho EF. Sub-Saharan Africa descendents in Rio de Janeiro (Brazil): population and mutational data for 12 Y-STR loci. *Int J Legal Med* 2007; 121(3): 238-241.

# DETERMINAÇÃO DAS FREQUÊNCIAS DE STRS NO CROMOSSOMO Y EM INDIVÍDUOS DO SUL DO BRASIL

**Título curto: Y-STRs em indivíduos do Rio Grande do Sul-Brasil**

Rodrigo Rodenbusch<sup>1</sup>, Ana Carolina Mardini<sup>1</sup>, Aline Albeche Farias Estivalet<sup>1</sup>, André Zoratto Gastaldo<sup>1</sup>, Simone Schumacher<sup>1</sup>, Maria Helena Albarus<sup>1</sup>, Hugo Bock<sup>1</sup>, Roberto Giugliani<sup>1,2,3</sup>, Maria Luiza Saraiva-Pereira<sup>1,2,4</sup>.

<sup>1</sup>Lab. de Identificação Genética – Centro de Pesquisas – HCPA; <sup>2</sup>Serviço de Genética Médica – HCPA; <sup>3</sup>Departamento de Genética – UFRGS; <sup>4</sup>Departamento de Bioquímica – UFRGS, Porto Alegre, Brasil.

Endereço para correspondência:

Maria Luiza Saraiva-Pereira, PhD  
Laboratório de Identificação Genética – Centro de Pesquisas  
Hospital de Clínicas de Porto Alegre, RS, Brasil  
Rua Ramiro Barcelos, 2350  
CEP 90035-903, Porto Alegre, RS, Brasil  
Tel: + 55 51 21018011  
Tel: + 55 51 21018010  
e-mail: mlpereira@hcpa.ufrgs.br

**Resumo:** Durante a última década, muitas pesquisas demonstraram a existência de polimorfismos em diferentes marcadores distribuídos pelo genoma. Entre eles, os denominados *short tandem repeats* (STR) são largamente utilizados, tanto aqueles localizados nos cromossomos autossômicos como os encontrados no cromossomo Y, para a identificação de indivíduos do sexo masculino. O objetivo deste estudo foi estabelecer, na população regional, a frequência de 12 *loci* (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 e DYS439) e determinar as taxas de mutação de cada *locus*. A população testada foi composta por 162 pares de pais e filhos cuja paternidade tinha sido previamente confirmada pelo uso de STRs autossômicos, atingindo uma probabilidade de paternidade igual ou superior a 99,99%. Na amostra analisada, foram determinados 151 haplótipos distintos e mutações foram encontradas nos *loci* DYS19, DYS390, DYS439 e DYS437. Os resultados obtidos permitiram ainda a determinação da frequência dos alelos desses marcadores na nossa população. Desta forma, foi possível determinar a ocorrência limitada de mutações nessa amostra.

Palavras-Chaves: Cromossomo Y – Frequências Alélicas – Polimorfismos – Y-STRs – Y-Haplótipos

## **Introdução**

A identificação humana através do DNA é a metodologia mais moderna aplicada mundialmente e está baseada em diferenças na seqüência de DNA de uma pessoa para outra; estas diferenças genéticas são conhecidas como polimorfismos.

STRs (ou microssatélites) consistem em seqüências de 1 a 6 pb que estão repetidas várias vezes ao longo do genoma e são altamente polimórficas (Tautz, 1989). Os STRs se tornaram uma poderosa ferramenta para identificação humana, uma vez que, para sua detecção, podemos utilizar a técnica da Reação da Polimerase em Cadeia (PCR), amplificando-se o DNA a partir de oligonucleotídeos situados nas regiões adjacentes não repetitivas (Weir, 1992).

Na identificação humana, as análises de STRs localizadas em cromossomos autossômicos são o primeiro foco de uma investigação. Durante a última década, muitas pesquisas descreveram um alta freqüência de polimorfismos nos STRs do cromossomo Y (Y-STR), tornando-se estes, em alguns casos, uma poderosa ferramenta na identificação humana, principalmente em casos de violência sexual e em teste de paternidade quando o suposto pai é falecido ou não é encontrado (Kurihara *et al*, 2004).

O cromossomo Y representa apenas 2% do genoma humano e possui um tamanho aproximado de 60 Mb. Este cromossomo apresenta duas regiões pseudo-autossômicas (PAR1 e PAR2) localizadas nas extremidades do braço curto e longo do cromossomo. Estas regiões são homologas a regiões do cromossomo X e são responsáveis pelo

correto pareamento dos dois cromossomos sexuais durante a meiose (Quintana-Murci *et al*, 2001).

A utilização dos Y-STRs como ferramenta de identificação está baseada nas diferenças entre os cromossomos X e Y e os autossômicos, onde os marcadores no cromossomo Y se encontram em regiões não-recombinantes (NRY) ao cromossomo X (Quintana-Murci *et al*, 2001). Estas regiões facilitam a identificação do indivíduo do sexo masculino que contribui com material genético em uma mistura macho/fêmea (Y/X), uma vez que estes *loci* são geneticamente ligados e herdados inalterados de pai para filho (Krenke *et al*, 2005).

De acordo com a comunidade forense europeia, os marcadores DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS385a/b constituem o chamado haplótipo mínimo dos Y-STRs (ou *minimal haplotype*). Quando acrescido do marcador YCAII a/b, este conjunto de *loci* também pode ser chamado de haplótipo estendido dos Y-STRs (ou *extended haplotype*) (Butler, 2003).

Apesar de diversos sistemas de multiplex de Y-STRs estarem sendo desenvolvidos em diversas populações, os estudos sobre taxas de mutações desses *loci* ainda são limitados. O estabelecimento da estimativa das taxas de mutações em teste de paternidade é necessário para que se possa evitar uma falsa exclusão em casos onde possa existir uma discrepância alélica entre o pai biológico e o filho devido a uma mutação (Kurihara *et al*, 2004).

## **Materiais e Métodos:**

### *População:*

A população consiste de 162 amostras de filhos do sexo masculino e seus respectivos pais, residentes no Rio Grande do Sul (Brasil) que foram previamente selecionados com um índice de paternidade igual ou superior a 99,99%, confirmadas através de marcadores nos cromossomos autossômicos.

### *Extração de DNA e Amplificação por PCR:*

O DNA foi isolado a partir de sangue total pelo método de precipitação em excesso de sais, utilizando o kit comercial Wizard<sup>®</sup> Genomic DNA Purification (Promega Corporation Cat. #A1120). As regiões de interesse foram amplificadas pela PCR multiplex com oligonucleotídeos marcados com fluorescência através do kit comercial PowerPlex<sup>®</sup> Y System (Promega Corporation Cat.# DC6761) que é composto por 12 *loci*: DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 e DYS439. A amplificação foi mediada pela enzima AmpliTaq Gold<sup>®</sup> DNA polymerase (Applied Biosystems).

### *Análise e Genotipagem dos Fragmentos Amplificados:*

Os fragmentos de PCR amplificados foram separados por eletroforese capilar utilizando o sistema ABI<sup>®</sup> 3100 Genetic Analyzer (Applied Biosystems) e o marcador interno Internal Lane Standard-ILS600 (Promega Corporation). As amostras foram genotipadas utilizando o software ABI Genotyper<sup>®</sup> Software (Applied Biosystems) em conjunto com

o *PowerTyper<sup>TM</sup> Y Macro* (Promega Corporation), utilizando como parâmetro os dados fornecidos pela escala alélica *PowerPlex® Y Allelic Ladder Mix* (Promega Corporation).

*Análises Estatísticas:*

A frequência dos alelos de todos *loci* foram estimadas diretamente pela fórmula  $f_i = f / \sum f$  ou  $f_i = f/n$ . A diversidade gênica e haplotípica juntamente com os índices de diversidade foram estimadas de acordo com os trabalhos de Nei (1973). O poder de discriminação individual foi calculado pela razão entre o número de haplótipos distintos encontrados e o número de indivíduos da amostra.

A taxa de mutação na população foi estimada através razão entre o número de mutações observadas pelo número de meioses. As diferenças entre as populações foram calculadas através Análise Molecular de Variância (AMOVA) utilizando o programa Arlequin v. 2000.

## **Resultados:**

Entre os 162 haplótipos analisados, listados na Tabela 1, encontramos 151 haplótipos distintos, onde os dois haplótipos mais comuns, H058 e H076, apresentaram frequências de 0,0265 e 0,0199, respectivamente. Outros 6 haplótipos distintos (H001, H023, H057, H067, H088 e H122) apresentaram frequência de 0,0132 e a frequência dos outros 143 haplótipos foi 0,0066 (1 ocorrência cada). De acordo com essa distribuição de frequências, foi possível obter um índice de diversidade haplotípica de 0,9321 e um poder de discriminação individual de 0,9934.

A distribuição das frequências alélicas observadas nos 12 *loci* em estudo são listadas na Tabela 2, juntamente com os valores de diversidade de cada marcador.

Também identificamos a ocorrência de mutações nos *loci* DYS19, DYS390, DYS439 e DYS437 (Fig.1). A presença destas mutações nos permitiu calcular uma taxa de mutação para essa população de  $2,06 \times 10^{-3}$ .

Comparando a população estudada com estudos prévios, como o estudo realizado em Santa Catarina (Cainé *et al.*, 2004), 99,25% de todas as variações haplotípicas são encontradas dentro das populações e apenas 0,75% destas diferenças são encontradas entre as populações, não apresentando diferença significativa (AMOVA,  $p > 0,05$ ). Em outro estudo onde participaram indivíduos das cinco regiões geopolíticas do Brasil (Grattapaglia *et al.*, 2005), a comparação dos indivíduos da região sul com os indivíduos

deste estudo os resultados do AMOVA foram semelhantes indicando uma variação intra-populacional de 98,84% e apenas 1,16% entre as populações ( $p > 0,05$ ).

## **Discussão:**

O cromossomo Y é um componente fundamental do genoma humano, participando de funções essenciais na determinação sexual dos mamíferos. Este também possui um papel importante na identificação humana, através de marcadores polimórficos (Y-STRs) distribuídos ao longo do cromossomo.

Neste estudo, nós observamos 151 haplótipos distintos entre os 162 estudados, permitindo determinar um alto poder de discriminação individual nos marcadores estudado, calculado em 0.9934, e um alto índice de diversidade haplotípica, o qual foi estimado em 0.9321.

Nós também observamos a ocorrência limitada de mutações nessa amostra, com a identificação de 4 mutações em marcadores distintos e amostras diferentes, não sendo observada duas mutações na mesma amostra de filho/pai.

Nós confirmamos que a utilização desses marcadores significa um alto poder de discriminação na identificação de indivíduos em casos de paternidade e disputas forenses com a determinação das frequências desses marcadores e de suas taxas de mutação. Entretanto, estudos adicionais em outras populações devem ser realizados e os dados produzidos devem ser organizados em banco de dados com frequências regionais e mundiais para melhorar a confiança do emprego desses marcadores como uma ferramenta poderosa e adicional disponível para a identificação humana.

## Referências:

- BUTLER JM. Recent Developments in Y-Short Tandem Repeat and Y-Single Nucleotide Polymorphism Analysis. **Forensic Sci Rev**, 2003; 15: 91-111.
- CAINÉ L, CORTE-REAL F, VIEIRA DN, CARVALHO M, SERRA A, LOPES V, VIDE MC. Allele frequencies and haplotypes of 8 Y-chromosomal STRs in the Santa Catarina population of southern Brazil. **Forensic Sci Int**, 2005; 148: 75-79.
- GRATTAPAGLIA D, KALUPNIEK S, GUIMARAES CS, RIBEIRO MA, DIENER PS, SOARES CN. Y-chromosome STR haplotype diversity in Brazilian populations. **Forensic Sci Int**, 2005; 149(1): 99-107
- KRENKE BE, VICULIS L, RICHARD ML, PRINZ M, MILNE SC, LADD C, GROSS AM, GORNALL T, FRAPPIER JRH, EISENBERG AJ, BARNA C, ARANDA XG, ADAMOWICZ MS, BUDOWLE B. Validation of a male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex. **Forensic Sci Int**, 2005; 148: 1-14.
- KURIHARA R, YAMAMOTO T, SHI-LIN LI R U, YOSHIMOTO T, KAMIYAMA H O K, KATSUMATA Y. Mutations in 14 Y-STR loci among Japanese father-son haplotypes. **Int. J. Legal Med**, 2004; 118: 125-131.
- NEI M. Molecular Evolutionary Genetics. **Columbia University Press**, 1973 New York.
- PROMEGA CORPORATION *PowerPlex® Y System Technical Manual*, #TMD018, 2003.
- QUINTANA-MURCI L, KRAUSZ C, McELREAVEY K. The human Y chromosome: function, evolution and disease. **Forensic Sci Int**, 2001; 118: 169-181.
- TAUTZ D Hypervariability of simple sequences as a general source for polymorphic DNA markers. **Nucleic Acids Research**, 1989; 17: 6463-6417.
- WEIR BS. Independence of VNTR alleles defined as fixed bins **Genetics**, 1992 130: 873-887.

## **Tabela e Legenda das Figuras**

Tabela 1 – Haplótipos distintos encontrados nesse estudo.

Tabela 2 – Frequências alélicas e valores de diversidade gênica dos 12 *loci* estabelecidos nesse estudo<sup>a</sup>

Figura 1 – Eletroferograma mostrando ocorrência de mutação no *locus* DSY390 de 23 → 24. a) Escada alélica dos *loci* DSY393, DSY390, DSY385; b) Alelos encontrados no filho; c) Alelos encontrados no pai.

**Tabela 1**

Haplótipo	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS385	n
H001	12	14	30	23	11	13	13	16	12	11	11,14	2
H002	13	12	28	24	11	13	13	15	12	11	11,15	1
H003	13	12	30	24	10	16	12	14	11	12	14,14	1
H004	13	13	29	23	10	11	12	15	9	11	13,16	1
H005	13	13	29	24	9	11	13	14	10	10	13,14	1
H006	13	13	29	24	10	11	13	14	10	12	17,18	1
H007	13	13	29	24	10	13	13	14	12	12	14,15	1
H008	13	13	30	22	9	11	13	14	10	12	12,15	1
H009	13	13	30	23	9	11	13	14	11	12	12,15	1
H010	13	13	30	23	10	11	13	14	10	12	16,16	1
H011	13	13	30	24	10	11	12	14	10	11	16,19	1
H012	13	13	30	24	10	11	13	14	10	12	16,18	1
H013	13	13	30	24	10	11	13	14	10	12	17,19	1
H014	13	13	30	24	10	11	13	14	10	13	16,18	1
H015	13	13	30	24	11	11	13	14	10	12	18,21	1
H016	13	13	30	25	9	12	14	14	10	12	18,18	1
H017	13	13	31	23	10	14	13	14	11	11	15,18	1
H018	13	13	31	23	11	14	13	14	11	13	12,16	1
H019	13	13	31	24	11	11	13	14	10	14	14,15	1
H020	13	13	31	25	11	15	13	15	11	12	14,14	1
H021	13	13	32	26	10	11	13	14	10	13	16,17	1
H022	13	14	29	23	10	13	13	14	9	13	13,15	1
H023	13	14	30	24	9	11	13	14	10	10	13,14	2
H024	13	14	31	24	9	11	13	14	10	10	13,14	1
H025	13	14	31	24	10	11	13	14	10	12	17,19	1
H026	13	14	31	24	11	13	14	16	12	13	11,14	1
H027	14	12	27	22	11	14	11	15	10	12	13,18	1
H028	14	12	28	22	10	11	13	16	10	11	13,14	1
H029	14	12	28	22	10	11	13	16	10	12	12,14	1
H030	14	12	28	22	10	11	13	16	11	11	13,14	1
H031	14	12	28	22	10	11	14	16	10	11	13,13	1
H032	14	12	28	23	10	11	13	16	10	11	13,15	1
H033	14	12	28	23	10	11	13	16	10	11	14,15	1
H034	14	12	28	24	11	13	13	15	12	12	11,15	1
H035	14	12	29	22	10	11	13	16	11	11	13,15	1
H036	14	12	29	23	10	11	12	14	10	11	13,19	1
H037	14	13	27	23	11	13	13	15	12	13	11,14	1
H038	14	13	28	24	10	13	13	15	12	12	11,14	1
H039	14	13	28	24	10	14	13	15	12	12	11,14	1
H040	14	13	28	24	11	13	12	15	12	11	11,15	1
H041	14	13	28	24	11	13	13	15	12	12	11,14	1
H042	14	13	29	22	11	13	13	15	12	13	11,14	1
H043	14	13	29	23	10	11	12	14	10	12	13,18	1
H044	14	13	29	23	10	11	12	15	9	12	13,17	1
H045	14	13	29	23	10	11	12	16	9	12	14,17	1
H046	14	13	29	23	11	13	13	14	12	12	11,13	1
H047	14	13	29	23	11	13	13	15	12	12	11,13	1
H048	14	13	29	23	11	13	13	15	12	12	11,14	1
H049	14	13	29	23	11	13	13	15	12	13	11,13	1
H050	14	13	29	23	12	11	12	14	10	11	13,18	1
H051	14	13	29	24	10	13	12	15	12	13	11,14	1
H052	14	13	29	24	10	13	13	15	11	11	11,14	1
H053	14	13	29	24	10	13	13	15	11	12	11,14	1
H054	14	13	29	24	11	13	13	14	12	13	11,14	1
H055	14	13	29	24	11	13	13	15	11	11	13,13	1
H056	14	13	29	24	11	13	13	15	12	10	12,13	1
H057	14	13	29	24	11	13	13	15 (16) <sup>ab</sup>	12	11	11,14	2
H058	14	13	29	24	11	13	13	15	12	12	11,14	4
H059	14	13	29	24	11	13	13	15	12	12	11,15	1
H060	14	13	29	24	11	13	13	15	12	12	12,14	1
H061	14	13	29	24	11	13	13	15	12	12	12,16	1
H062	14	13	29	24	11	13	13	15	12	13	11,14	1
H063	14	13	29	24	11	13	13	15	13	12	11,14	1
H064	14	13	29	24	11	13	13	15	13	13	11,14	1
H065	14	13	29	24	11	14	13	15	12	12	11,14	1
H066	14	13	29	24	12	13	13	15	12	11	11,12	1

H067	14	13	29	25	11	13	13	15	12	12	11,14	2
H068	14	13	29	26	11	13	13	15	12	12	11,14	1
H069	14	13	30	22	10	11	13	15	9	12	13,15	1
H070	14	13	30	23	10	11	12	14	9	11	13,16	1
H071	14	13	30	23	10	11	12	14	10	11	13,15	1
H072	14	13	30	23	11	11	12	15	9	10	13,15	1
H073	14	13	30	24	10	11	12	14	10	12	13,19	1
H074	14	13	30	24	11	13	13	14	12	11	11,14	1
H075	14	13	30	24	11	13	13	15	10	13	12,14	1
H076	14	13	30	24	11	13	13	15	12	12	11,14	3
H077	14	13	30	24	11	13	13	15	13	12	11,13	1
H078	14	13	30	25	10	11	13	14	11	12	11,14	1
H079	14	13	31	24	10	11	13	14	10	12	17,18	1
H080	14	13	31	24	11	13	14	15	12	11	11,14	1
H081	14	14	29	23	11	13	13	14	9	11	14,17	1
H082	14	14	29	24	11	13	13	16	11	12	11,14	1
H083	14	14	30	22	11	13	13	15	12	12	12,14	1
H084	14	14	30	23	10	13	12	16	10	11	17,18	1
H085	14	14	30	23	10	13	13	15	12	13	11,11	1
H086	14	14	30	23	11	13	13	15	12	12	11,14	1
H087	14	14	30	23	11	13	14	14	12	14	11,14	1
H088	14	14	30	24	10	13	13	15	12	12	11,13	2
H089	14	14	30	24	10	13	13	15	12	12	11,14	1
H090	14	14	30	24	10	13	13	15	12	12	11,16	1
H091	14	14	30	24	11	10	12	15	12	11	11,14	1
H092	14	14	30	24	11	13	13	15	12	12	11,14	1
H093	14	14	30	24	11	14	13	15	12	12	11,15	1
H094	14	14	30	25	11	13	13	14	12	13	11,14	1
H095	14	14	31	24	10	13	13	15	12	12	11,15	1
H096	14	15	31	24	10	13	13	15	12	12	11,14	1
H097	14	15	31	24	11	13	12	14	12	11	12,14	1
H098	14	14	33	24	10	11	13	14	10	13	11,11	1
H099	15	12	27	24	10	14	13	14	9	11	14,16	1
H100	15	12	28	23	10	11	13	16	10	11	13,17	1
H101	15	12	28	24	10	11	13	16	9	12	13,16	1
H102	15	12	28	24	10	12	12	16	10	11	14,16	1
H103	15	12	29	21	10	11	14	16	10	11	13,15	1
H104	15	12	29	21	10	11	15	16	10	11	13,16	1
H105	15	12	29	22	11	11	14	15	10	11	14,14	1
H106	15	12	29	23	10	11	14	16	10	11	14,14	1
H107	15	12	29	23	10	11	14	16	10	12	14,14	1
H108	15	12	29	23	11	13	13	14	12	11	11,15	1
H109	15 (16) <sup>a</sup>	12	30	23	10	11	15	16	10	11	13,14	1
H110	15	13	28	24	11	13	13	15	12	13	11,14	1
H111	15	13	29	22	10	11	13	14	10	11	12,19	1
H112	15	13	29	22	11	13	12	15	12	12	11,14	1
H113	15	13	29	22	11	13	13	15	12	13 (12) <sup>a</sup>	11,14	1
H114	15	13	29	23	9	11	12	14	9	12	13,16	1
H115	15	13	29	23	10	11	12	15	9	12	13,18	1
H116	15	13	29	23	10	13	13	14	9	11	14,16	1
H117	15	13	29	23	11	14	13	15	11	12	11,14	1
H118	15	13	29	24	10	14	13	15	12	11	11,14	1
H119	15	13	29	24	11	13	12	15	12	13	11,14	1
H120	15	13	29	24	11	13	13	15	12	11	11,15	1
H121	15	13	29	24	11	13	13	15	12	12	11,14	1
H122	15	13	29	24	11	13	13	15	12	13	11,14	2
H123	15	13	29	24	11	14	13	14	12	11	12,14	1
H124	15	13	29	25	10	11	13	14	11	12	11,15	1
H125	15	13	30	22	10	12	14	15	10	11	16,17	1
H126	15	13	30	23	10	14	13	14	9	11	14,18	1
H127	15	13	30	24	11	11	13	14	10	11	9,17	1
H128	15	13	30	24	11	11	13	14	10	11	12,17	1
H129	15	13	30	25	9	11	13	15	10	12	14,15	1
H130	15	13	31	21	10	11	14	14	11	15	16,17	1
H131	15	13	31	21	11	11	13	14	11	10	16,17	1
H132	15	13	32	21	10	11	13	14	13	15	17,17	1
H133	15	14	30	23	9	11	12	14	9	11	13,16	1
H134	15	14	30	23	10	11	12	15	9	12	14,17	1
H135	15	14	30	23	10	12	15	14	10	11	15,16	1
H136	15	14	31	23	10	12	16	14	10	11	14,16	1
H137	15	14	31	23	10	13	12	14	9	11	14,14	1
H138	15	14	31	23	10	13	13	14	9	10	14,15	1

<b>H139</b>	15	14	31	24	11	13	13	15	12	12	11,15	<b>1</b>
<b>H140</b>	15	14	31	25	10	11	13	14	11	11	11,14	<b>1</b>
<b>H141</b>	16	12	29	23	10	11	15	16	10	10	13,14	<b>1</b>
<b>H142</b>	16	13	28	23	10	11	13	14	10	13	12,12	<b>1</b>
<b>H143</b>	16	13	29	21	9	11	13	16	10	11	11,12	<b>1</b>
<b>H144</b>	16	13	29	23 (24) <sup>a</sup>	11	12	14	15	10	11	15,16	<b>1</b>
<b>H145</b>	16	13	30	21	10	11	12	15	10	11	12,14	<b>1</b>
<b>H146</b>	16	13	30	23	10	11	12	15	9	10	13,16	<b>1</b>
<b>H147</b>	16	13	30	25	11	11	13	14	11	11	11,14	<b>1</b>
<b>H148</b>	16	13	31	21	10	11	14	14	11	12	17,17	<b>1</b>
<b>H149</b>	16	13	31	23	10	12	14	14	10	11	13,16	<b>1</b>
<b>H150</b>	16	13	31	25	10	11	13	14	11	11	12,13	<b>1</b>
<b>H151</b>	17	14	30	24	11	12	13	15	8	12	15,15	<b>1</b>

a: Alelo mutado encontrado no filho

b: Mutação encontrada em 1 do 2 casos

Poder de discriminação individual: 0,9934

Diversidade haplotípica: 0,9321

Tabela 2

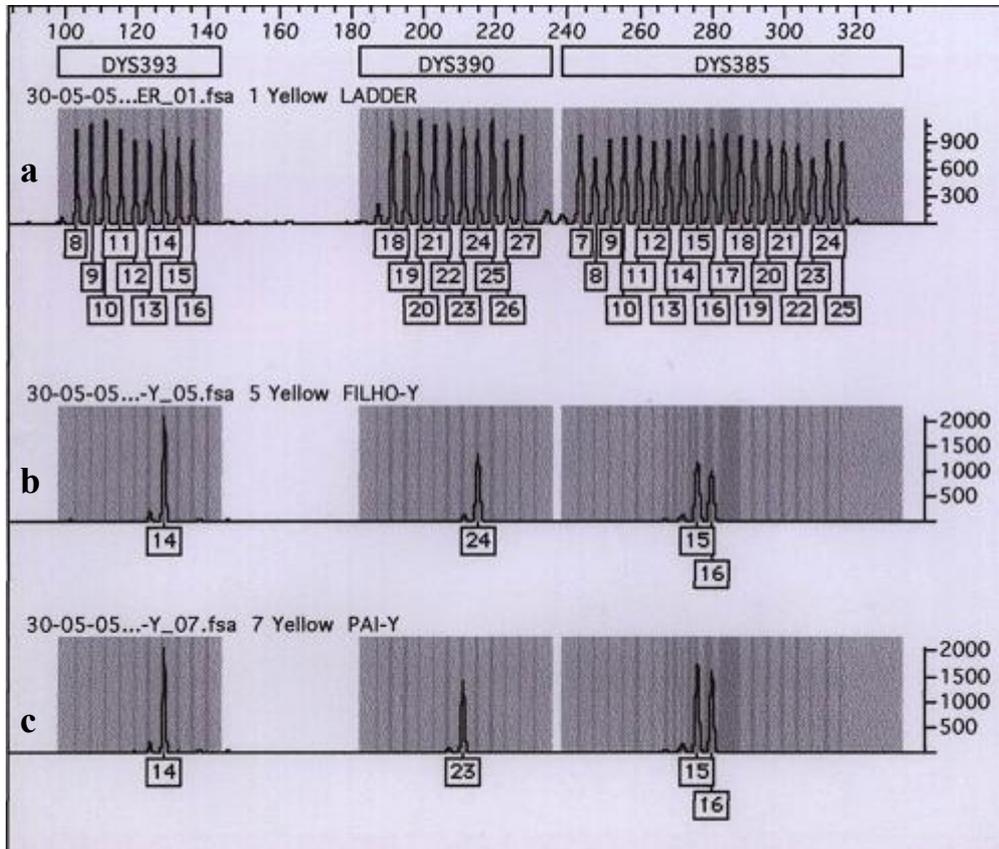
Alelo	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	Haplótipo	DYS395
8					0,0679	0,0123			0,0062		9,17	0,0062
9					0,4630				0,1172		11,11	0,0123
10					0,4568	0,4074			0,3149	0,0618	11,12	0,0123
11					0,0123	0,0432	0,0062		0,1296	0,3519	11,13	0,0371
12	0,0123	0,1481				0,4568	0,1667		0,4074	0,4259	11,14	0,3149
13	0,1605	0,6296				0,4568	0,7160		0,0247	0,1358	11,15	0,0618
14	0,4938	0,2100			0,0679	0,0679	0,0864	0,3765		0,0123	11,16	0,0062
15	0,2531	0,0123			0,0062	0,0062	0,0247	0,4753		0,0123	12,12	0,0062
16	0,0741				0,0062			0,1482			12,13	0,0123
17	0,0062										12,14	0,0432
18											12,15	0,0123
19											12,16	0,0123
20											12,17	0,0062
21				0,0556							12,19	0,0062
22				0,0926							13,13	0,0123
23				0,2963							13,14	0,0494
24				0,4691							13,15	0,0432
25				0,0741							13,16	0,0494
26				0,0123							13,17	0,0123
27											13,18	0,0247
28			0,0185								13,19	0,0123
29			0,1049								14,14	0,0432
30			0,3210								14,15	0,0247
31			0,1358								14,16	0,0185
32			0,0123								14,17	0,0185
33			0,0062								14,18	0,0062
											15,15	0,0062
											15,16	0,0123
											15,18	0,0062
											16,16	0,0062
											16,17	0,0309
											16,18	0,0123
											16,19	0,0062
											17,17	0,0123
											17,18	0,0185
											17,19	0,0123
											18,18	0,0062
											18,21	0,0062
$H^b$	0,6648	0,5408	0,7103	0,6709	0,5758	0,6225	0,4542	0,6142	0,7081	0,6764	$H^c$	0,7660

<sup>a</sup> n= 162

<sup>b</sup> Valores de Diversidade Gênica

<sup>c</sup> Valor de Diversidade Haplótipica

Figura 1



# **FREQUENCIES AND HAPLOTYPE DISTRIBUTION OF Y-CHROMOSOME STR IN INDIVIDUALS FROM SOUTH BRAZIL**

**Short title: Y-STRs in individuals of Rio Grande do Sul-Brazil**

Rodrigo Rodenbusch<sup>1</sup>, Ana Carolina Mardini<sup>1</sup>, Aline Albeche Farias Estivalet<sup>1</sup>, André Zoratto Gastaldo<sup>1</sup>, Simone Schumacher<sup>1</sup>, Maria Helena Albarus<sup>1</sup>, Hugo Bock<sup>1</sup>, Roberto Giugliani<sup>1,2,3</sup>, Maria Luiza Saraiva-Pereira<sup>1,2,4</sup>.

<sup>1</sup>Lab. de Identificação Genética – Centro de Pesquisas – HCPA; <sup>2</sup>Serviço de Genética Médica – HCPA; <sup>3</sup>Departamento de Genética – UFRGS; <sup>4</sup>Departamento de Bioquímica – UFRGS, Porto Alegre, Brazil.

Corresponding author:

Maria Luiza Saraiva-Pereira, PhD  
Laboratório de Identificação Genética – Centro de Pesquisas  
Hospital de Clínicas de Porto Alegre, RS, Brasil  
Rua Ramiro Barcelos, 2350  
CEP 90035-903, Porto Alegre, RS, Brasil  
Tel: + 55 51 21018011  
Tel: + 55 51 21018010  
e-mail: mlpereira@hcpa.ufrgs.br

**Abstract:** During the last decade, several studies showed the occurrence of polymorphisms in different loci along the genome. Among them, those called short tandem repeats (STR) are largely employed, on both autosomal chromosomes and on Y-chromosome, for the identification of male individuals. The aim of this study was to establish, in the regional population, frequency of 12 loci (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439) and determine mutation rate in each locus. Sample tested was composed by 162 pairs father-male sib whose paternity was previously confirmed by the autosomal STR, reaching a paternity probability equal or higher than 99.99%. In this sample, we were able to determine 151 different haplotypes and mutations were found in the following loci DYS19, DYS390, DYS439 and DYS437. Obtained results allowed the establishment of frequencies of these allele markers in our population. Finally, we observed a limited mutation rate, indicating that these markers are suitable to be applied for human identification in our population.

Key words: Allelic frequencies – polymorphisms – Y-chromosome – Y-STR – Y-haplotype

## **Introduction**

Human identification through DNA is the most recent methodology employed worldwide and is based on variations of DNA sequence from a person to the other; these genetic differences are known as polymorphisms.

STRs (or microsatellites) are sequences of 1 to 6 bp that can be repeatedly found along the genome and are highly polymorphic (Tautz, 1989). STRs become a powerful tool for human identification, since we can use the polymerase chain reaction (PCR) for their detection, amplifying the DNA from primers located in non-repeat adjacent regions (Weir, 1992).

In human identification, analyses of STRs located on autosomal chromosomes are firstly investigated. During the last decade, several studies showed a high polymorphic degree of STRs on Y-chromosome (Y-STR), becoming a powerful tool in human identification in some cases, mainly in sexual violence and in paternity tests when the alleged father is deceased or not available (Kurihara *et al*, 2004).

Y-chromosome represents only 2% of human genome and is approximately 60 Mb long. This chromosome has two pseudo-autosomal regions (PAR1 and PAR2) located on both ends of chromosome. These regions are homologue to X-chromosome regions and are responsible for correct pair of sexual chromosomes during the meiose (Quintana-Murci *et al*, 2001).

Application of Y-STRs as an identification tool is based on differences between X- and Y-chromosomes and autosomal chromosomes, where Y-chromosome markers are on

non-recombinant regions (NRY) to X chromosome (Quintana-Murci *et al*, 2001). These regions favour the identification of a male individual that contributes with genetic material in a mixture male/female (Y/X), since these loci are genetically linked and unmodified inherited from father to son (Krenke *et al*, 2005).

According to the European forensic community, markers DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS385a/b form the called Y-STRs minimum haplotype (or minimal haplotype). When YCAII a/b marker is added, this loci set is then called Y-STRs extended haplotype (or extended haplotype) (Butler, 2003).

Although several Y-STRs multiplex systems are being developed in various populations, studies on mutation rates of these loci are still limited. The establishment of mutation rates evaluates in paternity testing is essential to avoid a false exclusion in cases where is observed a single allelic divergence between biologic father and son duo to a mutation (Kurihara *et al*, 2004).

## **Material and Methods:**

### *Population:*

Studied population is composed by 162 samples of male sibs and their respective fathers, residents in Rio Grande do Sul (Brazil) that were previously selected with a paternity index of 99.99% or higher, confirmed by markers of autosomal chromosomes.

### *DNA isolation and PCR amplification:*

DNA was isolated from blood through excess salt precipitation method using Wizard<sup>®</sup> Genomic DNA Purification kit (Promega Corporation Cat. #A1120). Regions of interest were amplified by multiplex PCR with fluorescent labeled primers using PowerPlex<sup>®</sup> Y System (Promega Corporation Cat.# DC6761) that is composed by 12 loci: DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439. AmpliTaq Gold<sup>®</sup> DNA polymerase (Applied Biosystems) mediated synthesis of new strands.

### *Analysis and genotyping of amplified fragments:*

Amplified PCR fragments were resolved by capillary electrophoresis in an ABI<sup>®</sup> 3100 Genetic Analyzer (Applied Biosystems) and the Internal Lane Standard-ILS600 (Promega Corporation). Samples were genotyped by ABI Genotyper<sup>®</sup> Software (Applied Biosystems) together with PowerTyper<sup>™</sup> Y Macro (Promega Corporation), using data from PowerPlex<sup>®</sup> Y Allelic Ladder Mix (Promega Corporation).

*Statistical analyses:*

Alleles frequency of all loci were directly estimated by formulae  $f_i = f / \sum f$  or  $f_i = f/n$ . Haplotype and genomic diversity as well as diversity indexes were calculated according to Nei (1973). Individual discrimination power was evaluated by ration between number of distinct haplotypes found and number of individuals in the sample.

Population mutation rate was estimated through ratio between number of observed mutations by number of meioses. Differences among populations were calculated by AMOVA using Arlequin v. 2000 software.

## **Results:**

We have found 151 distinct haplotypes among 162 haplotypes found (Table 1) where frequencies of two more commonly found, H058 and H076, were estimated to be of 0.0265 and 0.0199, respectively. Frequencies of other 6 distinct haplotypes (H001, H023, H057, H067, H088 and H122) was found to be 0.0132 and frequency of the following 143 haplotypes was estimated to be 0.0066 (1 occurrence each). According to this frequencies distribution, we were able to determine a haplotype diversity index of 0.9321 and an individual discrimination power of 0.9934.

Observed allelic frequencies distribution in the 12 studied loci is listed in Table 2, together with diversity value of each marker.

We have also identified mutations in loci DYS19, DYS390, DYS439 and DYS437 (Figure 1). Occurrence of these mutations allowed estimation of a mutation rate of  $2.06 \times 10^{-3}$  in this population.

When comparing the studied population with previous studies as that conducted in Santa Catarina (Cainé *et al*, 2004), 99.25% of all haplotype variations are found within the population and only 0.75% of these differences are found between populations, with no difference statistically significant (AMOVA,  $p > 0.05$ ). When comparing our data with data from a study with individuals from 5 Brazilian geopolitic regions

(Grattapaglia *et al.*, 2005), results were similar indicating an intra-population variation of 98.84% and only 1.16% between populations (AMOVA,  $p > 0.05$ ).

**Discussion:**

Y-chromosome is a fundamental component of human genome, participating of essential functions in mammals' sexual determination. Y-chromosome has also an important role in human identification through polymorphic markers (Y-STRs) distributed along the chromosome.

In this study, we have observed 151 distinct haplotypes in 162 studied that leads to evidence a high individual discrimination power with studied markers, estimated to be 0.9934, and a high haplotype diversity index, calculated in 0.9321.

We have also observed a limited mutation occurrence in this sample with the identification of 4 mutations in distinct markers as well as in different samples. Two mutations were not identified in any pair son/father.

We have then confirmed that employing these markers represents a high discrimination power in individuals' identification in paternity cases as well as in forensic dispute with the establishment of markers' frequencies and their mutation rates. However, further studies in other populations have to be performed and compilation of generated data in databases with regional and worldwide frequencies is essential in order to improve confidence of these markers real utility as an additional and powerful tool available in human identification.

## References:

- BUTLER JM. Recent Developments in Y-Short Tandem Repeat and Y-Single Nucleotide Polymorphism Analysis. **Forensic Sci Rev**, 2003; 15: 91-111.
- CAINÉ L, CORTE-REAL F, VIEIRA DN, CARVALHO M, SERRA A, LOPES V, VIDE MC. Allele frequencies and haplotypes of 8 Y-chromosomal STRs in the Santa Catarina population of southern Brazil. **Forensic Sci Int**, 2005; 148: 75-79.
- GRATTAPAGLIA D, KALUPNIEK S, GUIMARAES CS, RIBEIRO MA, DIENER PS, SOARES CN. Y-chromosome STR haplotype diversity in Brazilian populations. **Forensic Sci Int**, 2005; 149(1): 99-107
- KRENKE BE, VICULIS L, RICHARD ML, PRINZ M, MILNE SC, LADD C, GROSS AM, GORNALL T, FRAPPIER JRH, EISENBERG AJ, BARNA C, ARANDA XG, ADAMOWICZ MS, BUDOWLE B. Validation of a male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex. **Forensic Sci Int**, 2005; 148: 1-14.
- KURIHARA R, YAMAMOTO T, SHI-LIN LI R U, YOSHIMOTO T, KAMIYAMA H O K, KATSUMATA Y. Mutations in 14 Y-STR loci among Japanese father-son haplotypes. **Int. J. Legal Med**, 2004; 118: 125-131.
- NEI M. Molecular Evolutionary Genetics. **Columbia University Press**, 1973 New York.
- PROMEGA CORPORATION *PowerPlex® Y System Technical Manual*, #TMD018, 2003.
- QUINTANA-MURCI L, KRAUSZ C, McELREAVEY K. The human Y chromosome: function, evolution and disease. **Forensic Sci Int**, 2001; 118: 169-181.
- TAUTZ D Hypervariability of simple sequences as a general source for polymorphic DNA markers. **Nucleic Acids Research**, 1989; 17: 6463-6417.
- WEIR BS. Independence of VNTR alleles defined as fixed bins **Genetics**, 1992 130: 873-887.

## **Table and Figure Legends**

Table 1 – Distinct haplotypes found in this study.

Table 2 – Allelic frequencies and genomic diversity values of 12 loci established in this study<sup>a</sup>

Figure 1 – Eletroferogram showing a mutation in DSY390 (23 → 24). a) Allelic ladder of DSY393, DSY390, DSY385 loci; b) Son's alleles; c) Father's alleles.

**Table 1**

Haplotype	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS385	n
H001	12	14	30	23	11	13	13	16	12	11	11,14	2
H002	13	12	28	24	11	13	13	15	12	11	11,15	1
H003	13	12	30	24	10	16	12	14	11	12	14,14	1
H004	13	13	29	23	10	11	12	15	9	11	13,16	1
H005	13	13	29	24	9	11	13	14	10	10	13,14	1
H006	13	13	29	24	10	11	13	14	10	12	17,18	1
H007	13	13	29	24	10	13	13	14	12	12	14,15	1
H008	13	13	30	22	9	11	13	14	10	12	12,15	1
H009	13	13	30	23	9	11	13	14	11	12	12,15	1
H010	13	13	30	23	10	11	13	14	10	12	16,16	1
H011	13	13	30	24	10	11	12	14	10	11	16,19	1
H012	13	13	30	24	10	11	13	14	10	12	16,18	1
H013	13	13	30	24	10	11	13	14	10	12	17,19	1
H014	13	13	30	24	10	11	13	14	10	13	16,18	1
H015	13	13	30	24	11	11	13	14	10	12	18,21	1
H016	13	13	30	25	9	12	14	14	10	12	18,18	1
H017	13	13	31	23	10	14	13	14	11	11	15,18	1
H018	13	13	31	23	11	14	13	14	11	13	12,16	1
H019	13	13	31	24	11	11	13	14	10	14	14,15	1
H020	13	13	31	25	11	15	13	15	11	12	14,14	1
H021	13	13	32	26	10	11	13	14	10	13	16,17	1
H022	13	14	29	23	10	13	13	14	9	13	13,15	1
H023	13	14	30	24	9	11	13	14	10	10	13,14	2
H024	13	14	31	24	9	11	13	14	10	10	13,14	1
H025	13	14	31	24	10	11	13	14	10	12	17,19	1
H026	13	14	31	24	11	13	14	16	12	13	11,14	1
H027	14	12	27	22	11	14	11	15	10	12	13,18	1
H028	14	12	28	22	10	11	13	16	10	11	13,14	1
H029	14	12	28	22	10	11	13	16	10	12	12,14	1
H030	14	12	28	22	10	11	13	16	11	11	13,14	1
H031	14	12	28	22	10	11	14	16	10	11	13,13	1
H032	14	12	28	23	10	11	13	16	10	11	13,15	1
H033	14	12	28	23	10	11	13	16	10	11	14,15	1
H034	14	12	28	24	11	13	13	15	12	12	11,15	1
H035	14	12	29	22	10	11	13	16	11	11	13,15	1
H036	14	12	29	23	10	11	12	14	10	11	13,19	1
H037	14	13	27	23	11	13	13	15	12	13	11,14	1
H038	14	13	28	24	10	13	13	15	12	12	11,14	1
H039	14	13	28	24	10	14	13	15	12	12	11,14	1
H040	14	13	28	24	11	13	12	15	12	11	11,15	1
H041	14	13	28	24	11	13	13	15	12	12	11,14	1
H042	14	13	29	22	11	13	13	15	12	13	11,14	1
H043	14	13	29	23	10	11	12	14	10	12	13,18	1
H044	14	13	29	23	10	11	12	15	9	12	13,17	1
H045	14	13	29	23	10	11	12	16	9	12	14,17	1
H046	14	13	29	23	11	13	13	14	12	12	11,13	1
H047	14	13	29	23	11	13	13	15	12	12	11,13	1
H048	14	13	29	23	11	13	13	15	12	12	11,14	1
H049	14	13	29	23	11	13	13	15	12	13	11,13	1
H050	14	13	29	23	12	11	12	14	10	11	13,18	1
H051	14	13	29	24	10	13	12	15	12	13	11,14	1
H052	14	13	29	24	10	13	13	15	11	11	11,14	1
H053	14	13	29	24	10	13	13	15	11	12	11,14	1
H054	14	13	29	24	11	13	13	14	12	13	11,14	1
H055	14	13	29	24	11	13	13	15	11	11	13,13	1
H056	14	13	29	24	11	13	13	15	12	10	12,13	1
H057	14	13	29	24	11	13	13	15 (16) <sup>a,b</sup>	12	11	11,14	2
H058	14	13	29	24	11	13	13	15	12	12	11,14	4
H059	14	13	29	24	11	13	13	15	12	12	11,15	1
H060	14	13	29	24	11	13	13	15	12	12	12,14	1
H061	14	13	29	24	11	13	13	15	12	12	12,16	1
H062	14	13	29	24	11	13	13	15	12	13	11,14	1
H063	14	13	29	24	11	13	13	15	13	12	11,14	1

H064	14	13	29	24	11	13	13	15	13	13	11,14	1
H065	14	13	29	24	11	14	13	15	12	12	11,14	1
H066	14	13	29	24	12	13	13	15	12	11	11,12	1
H067	14	13	29	25	11	13	13	15	12	12	11,14	2
H068	14	13	29	26	11	13	13	15	12	12	11,14	1
H069	14	13	30	22	10	11	13	15	9	12	13,15	1
H070	14	13	30	23	10	11	12	14	9	11	13,16	1
H071	14	13	30	23	10	11	12	14	10	11	13,15	1
H072	14	13	30	23	11	11	12	15	9	10	13,15	1
H073	14	13	30	24	10	11	12	14	10	12	13,19	1
H074	14	13	30	24	11	13	13	14	12	11	11,14	1
H075	14	13	30	24	11	13	13	15	10	13	12,14	1
H076	14	13	30	24	11	13	13	15	12	12	11,14	3
H077	14	13	30	24	11	13	13	15	13	12	11,13	1
H078	14	13	30	25	10	11	13	14	11	12	11,14	1
H079	14	13	31	24	10	11	13	14	10	12	17,18	1
H080	14	13	31	24	11	13	14	15	12	11	11,14	1
H081	14	14	29	23	11	13	13	14	9	11	14,17	1
H082	14	14	29	24	11	13	13	16	11	12	11,14	1
H083	14	14	30	22	11	13	13	15	12	12	12,14	1
H084	14	14	30	23	10	13	12	16	10	11	17,18	1
H085	14	14	30	23	10	13	13	15	12	13	11,11	1
H086	14	14	30	23	11	13	13	15	12	12	11,14	1
H087	14	14	30	23	11	13	14	14	12	14	11,14	1
H088	14	14	30	24	10	13	13	15	12	12	11,13	2
H089	14	14	30	24	10	13	13	15	12	12	11,14	1
H090	14	14	30	24	10	13	13	15	12	12	11,16	1
H091	14	14	30	24	11	10	12	15	12	11	11,14	1
H092	14	14	30	24	11	13	13	15	12	12	11,14	1
H093	14	14	30	24	11	14	13	15	12	12	11,15	1
H094	14	14	30	25	11	13	13	14	12	13	11,14	1
H095	14	14	31	24	10	13	13	15	12	12	11,15	1
H096	14	15	31	24	10	13	13	15	12	12	11,14	1
H097	14	15	31	24	11	13	12	14	12	11	12,14	1
H098	14	14	33	24	10	11	13	14	10	13	11,11	1
H099	15	12	27	24	10	14	13	14	9	11	14,16	1
H100	15	12	28	23	10	11	13	16	10	11	13,17	1
H101	15	12	28	24	10	11	13	16	9	12	13,16	1
H102	15	12	28	24	10	12	12	16	10	11	14,16	1
H103	15	12	29	21	10	11	14	16	10	11	13,15	1
H104	15	12	29	21	10	11	15	16	10	11	13,16	1
H105	15	12	29	22	11	11	14	15	10	11	14,14	1
H106	15	12	29	23	10	11	14	16	10	11	14,14	1
H107	15	12	29	23	10	11	14	16	10	12	14,14	1
H108	15	12	29	23	11	13	13	14	12	11	11,15	1
H109	15 (16) <sup>a</sup>	12	30	23	10	11	15	16	10	11	13,14	1
H110	15	13	28	24	11	13	13	15	12	13	11,14	1
H111	15	13	29	22	10	11	13	14	10	11	12,19	1
H112	15	13	29	22	11	13	12	15	12	12	11,14	1
H113	15	13	29	22	11	13	13	15	12	13 (12) <sup>a</sup>	11,14	1
H114	15	13	29	23	9	11	12	14	9	12	13,16	1
H115	15	13	29	23	10	11	12	15	9	12	13,18	1
H116	15	13	29	23	10	13	13	14	9	11	14,16	1
H117	15	13	29	23	11	14	13	15	11	12	11,14	1
H118	15	13	29	24	10	14	13	15	12	11	11,14	1
H119	15	13	29	24	11	13	12	15	12	13	11,14	1
H120	15	13	29	24	11	13	13	15	12	11	11,15	1
H121	15	13	29	24	11	13	13	15	12	12	11,14	1
H122	15	13	29	24	11	13	13	15	12	13	11,14	2
H123	15	13	29	24	11	14	13	14	12	11	12,14	1
H124	15	13	29	25	10	11	13	14	11	12	11,15	1
H125	15	13	30	22	10	12	14	15	10	11	16,17	1
H126	15	13	30	23	10	14	13	14	9	11	14,18	1
H127	15	13	30	24	11	11	13	14	10	11	9,17	1
H128	15	13	30	24	11	11	13	14	10	11	12,17	1
H129	15	13	30	25	9	11	13	15	10	12	14,15	1
H130	15	13	31	21	10	11	14	14	11	15	16,17	1
H131	15	13	31	21	11	11	13	14	11	10	16,17	1
H132	15	13	32	21	10	11	13	14	13	15	17,17	1
H133	15	14	30	23	9	11	12	14	9	11	13,16	1
H134	15	14	30	23	10	11	12	15	9	12	14,17	1
H135	15	14	30	23	10	12	15	14	10	11	15,16	1

<b>H136</b>	15	14	31	23	10	12	16	14	10	11	14,16	<b>1</b>
<b>H137</b>	15	14	31	23	10	13	12	14	9	11	14,14	<b>1</b>
<b>H138</b>	15	14	31	23	10	13	13	14	9	10	14,15	<b>1</b>
<b>H139</b>	15	14	31	24	11	13	13	15	12	12	11,15	<b>1</b>
<b>H140</b>	15	14	31	25	10	11	13	14	11	11	11,14	<b>1</b>
<b>H141</b>	16	12	29	23	10	11	15	16	10	10	13,14	<b>1</b>
<b>H142</b>	16	13	28	23	10	11	13	14	10	13	12,12	<b>1</b>
<b>H143</b>	16	13	29	21	9	11	13	16	10	11	11,12	<b>1</b>
<b>H144</b>	16	13	29	23 (24) <sup>a</sup>	11	12	14	15	10	11	15,16	<b>1</b>
<b>H145</b>	16	13	30	21	10	11	12	15	10	11	12,14	<b>1</b>
<b>H146</b>	16	13	30	23	10	11	12	15	9	10	13,16	<b>1</b>
<b>H147</b>	16	13	30	25	11	11	13	14	11	11	11,14	<b>1</b>
<b>H148</b>	16	13	31	21	10	11	14	14	11	12	17,17	<b>1</b>
<b>H149</b>	16	13	31	23	10	12	14	14	10	11	13,16	<b>1</b>
<b>H150</b>	16	13	31	25	10	11	13	14	11	11	12,13	<b>1</b>
<b>H151</b>	17	14	30	24	11	12	13	15	8	12	15,15	<b>1</b>

a: Mutant allele found in the son;

b: Mutation identified in 1 out of 2 cases

Individual discrimination power: 0.9934

Haplotype diversity: 0.9321



Table 2

Allele	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	Haplotype	DYS395
8									0.0062		9,17	0.0062
9					0.0679				0.1172		11,11	0.0123
10					0.4630	0.0123			0.3149	0.0618	11,12	0.0123
11					0.4568	0.4074			0.1296	0.3519	11,13	0.0371
12	0.0123	0.1481			0.0123	0.0432	0.0062		0.4074	0.4259	11,14	0.3149
13	0.1605	0.6296				0.4568	0.1667		0.0247	0.1358	11,15	0.0618
14	0.4938	0.2100				0.0679	0.7160			0.0123	11,16	0.0062
15	0.2531	0.0123				0.0062	0.0864	0.3765		0.0123	12,12	0.0062
16	0.0741					0.0062	0.0247	0.4753		0.0123	12,13	0.0123
17	0.0062							0.1482			12,14	0.0432
18											12,15	0.0123
19											12,16	0.0123
20											12,17	0.0062
21				0.0556							12,19	0.0062
22				0.0926							13,13	0.0123
23				0.2963							13,14	0.0494
24				0.4691							13,15	0.0432
25				0.0741							13,16	0.0494
26				0.0123							13,17	0.0123
27											13,18	0.0247
28			0.0185								13,19	0.0123
29			0.1049								14,14	0.0432
30			0.14013								14,15	0.0247
31			0.3210								14,16	0.0185
32			0.1358								14,17	0.0185
33			0.0123								14,18	0.0062
			0.0062								15,15	0.0062
											15,16	0.0123
											15,18	0.0062
											16,16	0.0062
											16,17	0.0309
											16,18	0.0123
											16,19	0.0062
											17,17	0.0123
											17,18	0.0185
											17,19	0.0123
											18,18	0.0062
											18,21	0.0062
$h^b$	0.6648	0.5408	0.7103	0.6709	0.5758	0.6225	0.4542	0.6142	0.7081	0.6764	$h^c$	0.7660

<sup>a</sup>  $n=162$

<sup>b</sup> Genomic diversity values

<sup>c</sup> Haplotype diversity value

Figure 1

