

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

FERNANDA IRMA REMUS HAMESTER

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE MUTAÇÕES *SLIPPAGE* EM MARCADORES
MICROSSATÉLITES AUTOSSÔMICOS DE INTERESSE FORENSE
EM UMA POPULAÇÃO DO SUDESTE BRASILEIRO**

Porto Alegre
2017

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR**

Fernanda Irma Remus Hamester

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE MUTAÇÕES *SLIPPAGE* EM MARCADORES
MICROSSATÉLITES AUTOSSÔMICOS DE INTERESSE FORENSE
EM UMA POPULAÇÃO DO SUDESTE BRASILEIRO**

Porto Alegre

2017

FERNANDA IRMA REMUS HAMESTER

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE MUTAÇÕES *SLIPPAGE* EM MARCADORES
MICROSSATÉLITES AUTOSSÔMICOS DE INTERESSE FORENSE
EM UMA POPULAÇÃO DO SUDESTE BRASILEIRO**

Dissertação apresentada ao Programa de Pós-graduação em Biologia Celular e Molecular da PUCRS como requisito para a obtenção do grau de Mestre em Biologia Celular e Molecular.

Orientadora: Profa. Dra. Clarice Sampaio Alho

Porto Alegre

2017

Ficha Catalográfica

H215i Hamester, Fernanda Irma Remus

Identificação e caracterização de mutações slippage em marcadores microssatélites autossômicos de interesse forense em uma população do sudeste brasileiro / Fernanda Irma Remus Hamester . – 2017.

48 f.

Dissertação (Mestrado) – Programa de Pós-Graduação em Biologia Celular e Molecular, PUCRS.

Orientadora: Profa. Dra. Clarice Sampaio Alho.

1. Taxa de Mutação. 2. STR. 3. Teste de Paternidade. 4. Microssatélite. I. Alho, Clarice Sampaio. II. Título.

Fernanda Irma Remus Hamester

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE MUTAÇÕES *SLIPPAGE* EM MARCADORES
MICROSSATÉLITES AUTOSSÔMICOS DE INTERESSE FORENSE
EM UMA POPULAÇÃO DO SUDESTE BRASILEIRO**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em
Biologia Celular e Molecular, da Faculdade de Biociências da
Pontifícia Universidade Católica do Rio Grande do Sul.

Aprovado em 29 de março de 2017

BANCA EXAMINADORA

Prof. Dr. Sandro Luis Bonatto – PUCRS

Prof. Dr. Marcelo Malaghini – UFPR – ICPR

Dra. Ana Carolina Mardini

Porto Alegre

2017

AGRADECIMENTOS

A minha orientadora, Professora Doutora Clarice Sampaio Alho, pela orientação, confiança, e auxílio, essenciais em todas as etapas de desenvolvimento desse trabalho.

A minha colega de profissão e co-orientadora de coração, Doutora Ana Paula Magalhães Lebouté pela prestatividade, paciência, auxílio e por todo o conhecimento compartilhado durante o desenvolvimento desse trabalho.

Agradeço ao Carlos Henrique Ares Motta pela confiança, prestatividade e oportunidade de realizarmos o levantamento de uma pequena parte do vasto banco de dados do Núcleo de Perícias Laboratoriais do Instituto de Medicina Social e de Criminologia de São Paulo (IMESC).

Agradeço a Débora Santos da Silva pelo auxílio e prestatividade na realização da coleta dos dados no Instituto de Medicina Social e de Criminologia de São Paulo (IMESC).

A minha família, por todas as oportunidades que me proporcionaram e pelo apoio dado durante esses anos.

Agradeço aos amigos e colegas de laboratório pela rica troca de conhecimento.

A CAPES e demais órgãos de fomento pelo financiamento que viabilizou a realização deste trabalho.

RESUMO

Estimativas bem definidas de taxas de mutação em locos polimórficos tetra-nucleotídeos do tipo STR (*short tandem repeats*) são um pré-requisito para exames de investigação genética de paternidade nas rotinas de investigações cíveis e criminais. Estudando 15 locos STR autossômicos de interesse forense, nós identificamos 193 mutações do tipo *slippage* (189 *one-step* e 4 *two-steps*) em 148.875 transferências alélicas entre duas gerações a partir de 5.171 casos de paternidade, nos quais havia relação biológica confirmada entre os indivíduos (15.096 indivíduos; 4.754 trios e 417 duos; 9.925 meioses). A taxa global de mutação foi $1,3 \times 10^{-3}$. As taxas mais elevadas foram observadas nos locos vWA ($2,8 \times 10^{-3}$), FGA e D18S51 ($2,7 \times 10^{-3}$ para ambos), enquanto nos locos TH01 e TPOX não foram observadas mutações nessa amostra populacional. A taxa de mutação decorrente de meiose paterna ($1,8 \times 10^{-3}$) foi seis vezes superior à taxa decorrente da origem materna ($0,3 \times 10^{-3}$), exceto para o locus D21S11.

PALAVRAS-CHAVE

Taxa de Mutação, STR, Teste de Paternidade, Microsatélite.

ABSTRACT

Well-defined estimates of mutation rates in highly polymorphic tetra-nucleotide short tandem repeats (STR) loci are a prerequisite for human identification (paternity) in the routines of civil and criminal investigations. Studying 15 autosomal STR loci of forensic interest (CSF1PO, D13S317, D16S539, D18S51, D19S433, D21S11, D2S1338, D3S1358, D5S818, D7S820, D8S1179, FGA, TH01, TPOX and vWA) we detected 193 slippage mutations (189 one-step and 4 two-step mutations) in 148,875 parent–child allelic transfers from 5,171 paternity cases with true biological relationship (15,096 individuals; 4,754 trios and 417 duos; 9,925 meioses). The overall mutation rate was 1.3×10^{-3} . The highest rates were observed in vWA (2.8×10^{-3}), FGA and D18S51 (2.7×10^{-3} for both) loci, while TH01 and TPOX loci presented no mutations in this populational sample. Mean slippage mutation rate of paternal origin (1.8×10^{-3}) was six times higher than that observed from maternal origin (0.3×10^{-3}), except for locus D21S11.

KEYWORDS

Mutation Rate, Short Tandem Repeat, Paternity Testing, Microsatellite.

LISTA DE ABREVIATURAS

- **DNA** – Deoxyribonucleic Acid, Ácido desoxirribonucleico
- **LR** – Likelihood Ratio
- **MMM** – Multistep mutation model, Modelo de mutação de passos múltiplos
- **PB** – Pares de base
- **PCR** – Polymerase chain reaction, Reação em cadeia da polimerase
- **SMM** – Stepwise mutation model, Modelo de mutação de passo a passo
- **SSR** – Single sequence repeat, Repetição de sequência simples
- **STR** – Short tandem repeat, Repetição curta em tandem
- **VNTR** – Variable number of Tandem Repeat, Número Variável de Repetições em Tandem

SUMÁRIO

INTRODUÇÃO	11
REFERENCIAL TEÓRICO	12
POLIMORFISMOS <i>SHORT TANDEM REPEATS</i> (STR)	12
VARIABILIDADE E MUTAÇÃO EM STR	13
<i>STEPWISE MUTATION</i> e MECANISMO DE MUTAÇÃO <i>DNA SLIPPAGE</i>	15
<i>STEPWISE MUTATION MODEL</i> (SMM)	16
FATORES QUE INFLUENCIAM AS MUTAÇÕES DE PASSO DOS STRs	17
TAXAS DE MUTAÇÃO DOS MICROSSATÉLITES	22
JUSTIFICATIVA	24
OBJETIVOS	25
LEVANTAMENTO DOS DADOS E CRITÉRIOS DE INCLUSÃO NO ESTUDO	26
ARTIGO CIENTÍFICO	27
CONSIDERAÇÕES FINAIS	41
REFERÊNCIAS BIBLIOGRÁFICAS	42
ANEXO	46

INTRODUÇÃO

Os testes de investigação de vínculo genético tornaram-se uma ferramenta extremamente útil com o avanço da biologia molecular, em especial o teste de paternidade, devido à necessidade da certeza de sua confirmação em casos cíveis, como disputas legais para fins de pensão alimentícia e de herança, em casos criminais envolvendo estupro, rapto, troca ou abandono de crianças (além de casos clínicos de diagnóstico pré-natal e/ou aconselhamento genético). A realização desses testes é feita pela análise de marcadores moleculares que fornecem informações únicas e exclusivas dos indivíduos, da mesma forma como é uma impressão digital; uma impressão digital do material genético ou o *DNA fingerprint*.

As regiões do genoma, ou lócus, selecionadas para esta finalidade precisam atender alguns critérios para a validação do seu uso na determinação de vínculo genético. Inicialmente é necessário que tais marcadores apresentem alto grau de polimorfismo com variação intrapopulacional. Também é necessário que apresentem um padrão de herança mendeliana codominante, tendo dados de frequências alélicas e genotípicas de acordo com os princípios de Hardy-Weinberg. Os lócus utilizados devem segregar de forma independente (segunda Lei de Mendel), possibilitando o emprego da regra da multiplicação nos cálculos estatísticos (HAMMOND *et al.*, 1994). Assim, tais marcadores apresentarão um elevado poder de discriminação, sendo úteis para os cálculos de probabilidade de paternidade.

REFERENCIAL TEÓRICO

POLIMORFISMOS *SHORT TANDEM REPEATS* (STR)

Um tipo de marcador que atende essas expectativas é o *Short Tandem Repeat* (STR), que é um polimorfismo de comprimento caracterizado por unidades de repetição (sequência 2 a 6 nucleotídeos), as quais se repetem adjacentes uma a outra (configuração em tandem). Também são conhecidos pelas nomenclaturas *Single Sequence Repeat* (SSR) ou microssatélite. Centenas de milhares de polimorfismos STR estão dispersos no genoma humano (WEBER, WONG, 1993; VALDES, MONTGOMERY, FREIMER, 1993), correspondendo a 3% de todo o conteúdo genético. A maioria dos STRs está localizada em regiões não codificadoras e apenas 8% deles estão em regiões codificadoras (ELLEGREN, 2000). Para cada locus STR investigado, o perfil genético de um indivíduo conterá dois alelos, que poderão apresentar variação na quantidade de unidades de repetição presentes em cada cromossomo, sendo denominados pelo número de repetições presentes em cada cópia.

O início das análises de polimorfismos STR baseadas na técnica da Reação em Cadeia da Polimerase (PCR), em meados da década de 1990, representou um grande avanço na ciência forense expandindo a utilização dos exames de DNA para fins de investigação de paternidade e identificação humana (EDWARDS *et al.*, 1991). O desenvolvimento da técnica da PCR aumentou a sensibilidade das análises, permitindo que com apenas 0,5-1ng de DNA seja possível obter sucesso na amplificação e obtenção de perfil genético. Os sistemas de amplificação de DNA comercialmente disponíveis, atualmente, permitem a amplificação de vários loci STR simultaneamente em uma única reação (PCR multiplex). O sistema de PCR multiplex é possível devido ao tamanho pequeno dos produtos amplificados (amplicons) e ao emprego de diferentes pares de *primers* marcados com fluoróforos (*dye-labeled*) específicos que possuem temperaturas de pareamento muito próximas. Esses *primers* marcados possibilitam a leitura dos produtos amplificados por eletroforese capilar em sequenciadores automatizados e, posteriormente, interpretação dos perfis genéticos em *softwares* específicos (Buckleton *et al.*, 2005; Butler, 2006).

Os resultados obtidos das análises dos marcadores STR através da técnica de PCR demonstraram rapidez, alta sensibilidade, facilidade de interpretação dos resultados e confiabilidade (WEBER, MAY, 1989), favorecendo sua aplicação em testes de vínculo genético, identificação forense e diagnósticos clínicos (HAMMOND *et al.*, 1994), bem como em estudos de evolução genética e mapeamento de doenças genéticas (WEBER, WONG, 1993).

VARIABILIDADE E MUTAÇÃO EM STR

Os microsatélites utilizados na determinação de vínculo genético apresentam alto grau de polimorfismo intrapopulacional, proporcionando ampla variabilidade genética. Esses polimorfismos são gerados por mutações, que promovem alterações no seu arranjo original de repetições, seja pela substituição de um nucleotídeo, pela inserção ou deleção de um pequeno segmento do DNA ou, mais frequentemente em STRs, pelo aumento ou diminuição do número de repetições. Caso isso ocorra em uma célula da linhagem germinativa, essa alteração poderá ser transmitida aos descendentes (BRINKMANN *et al.*, 1998; ELLEGREN, 2004) e passará a integrar o conjunto genético das gerações subsequentes. Alguns alelos mutantes podem ser identificados por uma observação direta do perfil genético, por exemplo, de um trio (mãe, pai e filho), no qual o filho terá um alelo diferente em comprimento daqueles que estão presentes em seus genitores. Em uma análise de parentesco/paternidade de um trio verdadeiro, por exemplo, essa diferença seria chamada de inconsistência. O QUADRO 1 ilustra a análise de 15 loci autossômicos em trio verdadeiro (mãe, filha, pai), onde se detecta mutação de adição (ganho) de uma unidade de repetição no alelo de D18S51 (em destaque) transmitido do pai para a filha. No locus D18S51, o genótipo da mãe é heterozigoto 18-20, o do pai é homozigoto 12-12 e o genótipo da filha é heterozigoto 13-20; esta herdou o alelo 20 da mãe, logo o alelo 13 tem de origem paterna (no entanto, o pai tem o alelo 12). Isto sugere a ocorrência de mutação de um passo de adição de unidade de repetição. Aceita-se essa hipótese de mutação devido a uma probabilidade mais elevada, pois a outra opção seria haver ocorrido uma mutação em 18 ou 20 (maternos) para 13 e, simultaneamente, outra mutação de 12 (paterno) para 20, o que matematicamente é menos esperado.

QUADRO 1 - ANÁLISE DE 15 LÓCUS AUTOSSÔMICOS EM TRIO VERDADEIRO (MÃE, FILHA, PAI), ONDE SE DETECTA MUTAÇÃO DE ADIÇÃO DE UMA UNIDADE DE REPETIÇÃO NO ALELO DE D18S51 (EM DESTAQUE) TRANSMITIDO DO PAI PARA A FILHA.

LÓCUS	MÃE BIOLÓGICA		FILHA		PAI BIOLÓGICO	
AMEL	X	X	X	X	X	Y
vWA	16	18	14	18	14	17
D16S539	10	12	10	11	9	11
D2S1338	19	25	19	25	19	24
D8S1179	10	15	13	15	13	13
D21S11	30.2	31	30.2	31	28	30.2
D18S51	18	20	13	20	12	12
D19S433	13.2	15	13.2	14	14	15
TH01	6	7	6	6	6	7
FGA	20	22	22	24	24	24
D3S1358	15	15	15	15	16	18
D13S317	13	14	13	13	8	13
D7S820	8	9	9	9	7	9
CSF1PO	10	12	10	12	10	12
D5S818	9	13	13	13	11	13
TPOX	9	10	8	10	8	8

Sem o reconhecimento da possibilidade de haver mutações espontâneas nas células germinativas dos pais, e sem a inclusão da probabilidade de mutação no cálculo de um teste de paternidade, tais inconsistências poderiam induzir a uma falsa exclusão da paternidade biológica.

Para que tenha um elevado poder de discriminação, uma das premissas para a seleção de um locus para ser usado em análises forenses é que ele apresente alto grau de polimorfismo; um locus que é muito polimórfico, evidentemente, apresenta alta taxa de mutação, tendo sido justamente por causa dessa alta taxa de mutação que a variabilidade alélica do locus ficou ampla. Sendo assim, é consenso que nos loci de interesse forense a ampla variabilidade alélica está diretamente proporcional à taxa de mutação do locus (PARDI *et al.*, 2005; Butler, 2006).

STEPWISE MUTATION e MECANISMO DE MUTAÇÃO DNA SLIPPAGE

As *stepwise mutations* ou mutações de passo, assim como as mutações de ponto, são geradas por eventos durante a replicação do DNA na meiose da linhagem germinativa parental e são caracterizadas pelo ganho ou perda de uma ou mais unidades de repetição. Essas alterações promovem variações no número de unidades repetitivas, alterando o comprimento dos microssatélites. O mecanismo que melhor explica esse tipo de mutação é produzido por um pareamento incorreto entre as fitas de DNA ("deslizamento" ou "escorregão" das fitas de DNA) e falha no mecanismo de reparo (LEVINSON, GUTMAN, 1987; ELLEGREN, 2004). Esse fenômeno é conhecido como DNA *slippage* ou ainda *polymerase slippage* e é o principal mecanismo de geração de novas mutações em microssatélites. Esse pareamento errôneo, gera um *loop* (laço) em uma das fitas, levando a adição (ganho) ou deleção (perda) de unidades de repetição, como pode ser observado na FIGURA 1 (ELLEGREN, 2004).

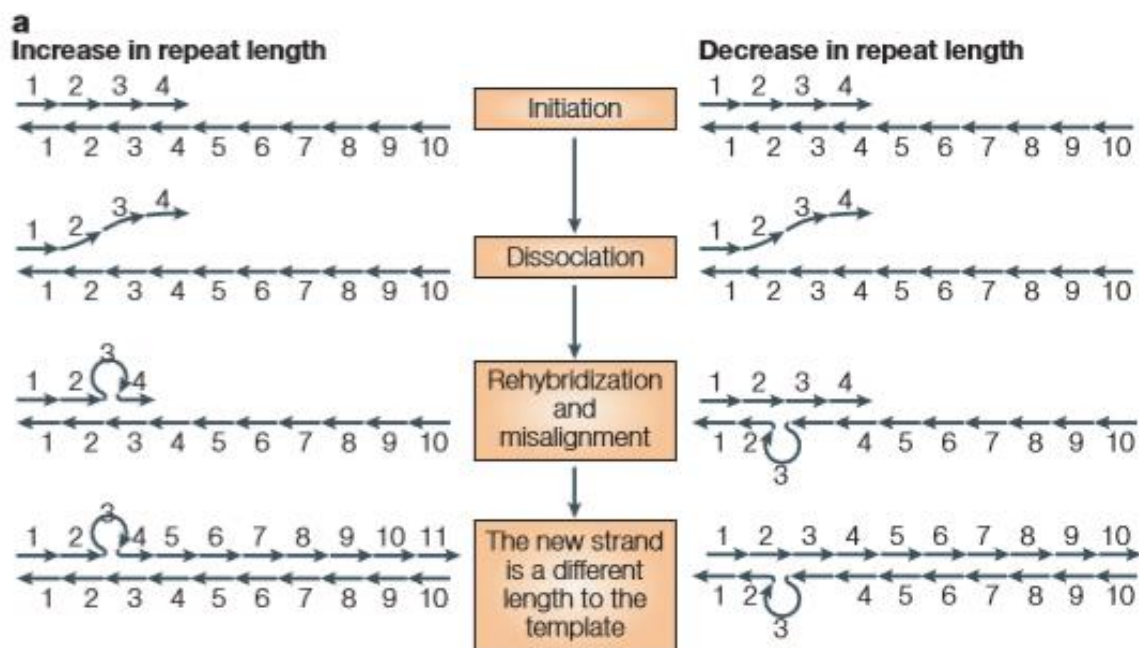


FIGURA 1– DNA *slippage*. Ao início da replicação de um trecho de seqüências de bases repetitivas, as duas fitas se dissociam. Se ao realinharem, a fita recém-sintetizada pairar erroneamente, essa terá um comprimento diferente da fita molde. Se o desalinhamento gerar o *loop* na fita recém-sintetizada, poderá ser gerado um alelo com comprimento maior. Se o desalinhamento gerar o *loop* na fita molde, poderá acarretar diminuição do comprimento do alelo. Fonte: ELLEGREN, 2004.

STEPWISE MUTATION MODEL (SMM)

O *Stepwise Mutation Model*, ou Modelo de Mutação Passo a Passo, foi introduzido na teoria genética de populações por Kimura & Ohta (1978). Esse modelo partiu de algumas hipóteses que serviram de orientação para estudos posteriores: (i) as alterações mutacionais no comprimento dos arranjos repetitivos se movem um passo na direção positiva (ganho de unidade de repetição) ou um passo na direção negativa (perda de unidade de repetição); (ii) há probabilidades iguais de expansão (ganho) e de contração (perda) do número de repetições; (iii) os alelos crescem infinitamente; (iv) os loci apresentam taxas fixas de mutação, independentemente do motivo do microssatélite e do comprimento dos alelos. Logo após, houve um desenvolvimento teórico matemático na tentativa de aperfeiçoar o modelo proposto, por exemplo por Brinkmann *et al.* (1998).

Estudos posteriores apresentaram algumas adaptações ao modelo de Ohta & Kimura, ao notarem que este modelo não era aplicável para os cálculos de taxa de mutação com mais de um passo (perda ou ganho de mais de uma unidade de repetição). Também foi verificada a ocorrência de diferentes taxas de mutação para diferentes loci analisados (VALDES, MONTGOMERY, FREIMER, 1993; MARDINI *et al.*, 2013). Atualmente, sabe-se que mutações de passo único (perda ou ganho de uma única unidade de repetição) correspondem a aproximadamente 90% dos eventos mutacionais, sendo seguidos das mutações de passo duplo e mais raramente, mutações de passo múltiplo (BRINKAMNN *et al.*, 1998; HUANG *et al.*, 2002; DAUBER *et al.*, 2012; JIA *et al.*, 2015). O mecanismo de mutação que envolve dois ou mais passos (perda ou ganho de duas ou mais unidades de repetição) é o modelo de mutação em passos múltiplos ou MMM (do inglês *multistep mutation model*), o qual é aceitável considerando a chance de re-pareamento após um deslizamento mais amplo das fitas. Ao observar a presença de mutações, pode-se afirmar que há uma predileção em aceitar a ocorrência de uma mutação de perda ou ganho de uma única unidade de repetição, porém isso é apenas resultado de uma análise tendenciosa, e não um fato verdadeiramente confirmado. O QUADRO 2 mostra o detalhe da detecção de uma mutação em D16S539, no qual o Alelo Paterno Obrigatório (APO; presente na filha), é o 12 de uma análise de 15 loci autossômicos em trio verdadeiro (mãe, filha, pai). Com base na probabilidade mais aceitável, pode-se considerar que na meiose do gameta paterno (o qual originou a filha) houve a perda de uma unidade de repetição no alelo paterno original 13 para 12. Mas o fato não é confirmado. Assim, há de se considerar a chance de ter havido uma mutação de múltiplas unidades de repetição, a

qual alterou o alelo paterno original 15 para o 12 (MMM com perda de seis unidades de repetição).

QUADRO 2- DETALHE DO LÓCUS D16S539 EM ANÁLISE DE 15 LÓCUS AUTOSSÔMICOS EM TRIO VERDADEIRO (MÃE, FILHA, PAI), NA QUAL SE DETECTA MUTAÇÃO NO ALELO PATERNO OBRIGATÓRIO TRANSMITIDO PARA A FILHA.

LÓCUS	MÃE BIOLÓGICA		FILHA		PAI BIOLÓGICO	
D16S539	9	9	9	12	13	15

FATORES QUE INFLUENCIAM AS MUTAÇÕES DE PASSO DOS STRs

O DNA *slippage* é o mecanismo que melhor explica a origem do modelo de mutação de passo. Contudo, há fatores também associados com as mutações dos microssatélites, que explicam o aumento da variação, destacando-se: (i) as características da unidade repetitiva (composição de bases e comprimento da unidade de repetição); (ii) o comprimento da sequência (número de repetições); (iii) a composição da estrutura dos microssatélites; (iv) a instabilidade dos STRs nos heterozigotos; (v) o sexo e a idade do indivíduo (relativo à geração de gametas mutados); (vi) as regiões flangeadoras dos microssatélites.

Segundo Kimura e Ohta (1978), independentemente do número de repetições presente no alelo, as probabilidades de expansão ou de contração ocorreriam de forma constante e igual para todos os loci. No entanto, foi demonstrado ser diferente. Em relação ao comprimento da unidade de repetição, dinucleotídeos apresentam uma taxa de mutação 1,5-2 vezes maior que tetranucleotídeos (CHAKRABORTY *et al.*, 1997). Isso pode ser em razão da estrutura, uma vez que dinucleotídeos formarem um *loop* na fita simples de DNA (o que geraria um *slippage*) mais facilmente do que os tetranucleotídeos. Em relação ao tipo de variação, Ellegren H (2000) em seus achados, observou que as expansões e as contrações ocorrem em igual quantidade em STR tetranucleotídeos, enquanto ocorrem mais expansões do que contrações em STR dinucleotídeos.

Ao longo dos anos, notou-se que os loci contendo microssatélites com três, quatro (EDWARDS *et al.*, 1991) e cinco repetições de base (tri-, tetra- e pentanucleotídeos, respectivamente) fornecem produtos amplificados mais confiáveis e de interpretação mais fácil

se comparados aos lócus STR dinucleotídeos (HAMMOND *et al.*, 1994). Tri-, tetra- e pentanucleotídeos são menos propensos a ocorrência de *slippage* (BUCKLETON *et al.*, 2005) e, por esse motivo, esses microssatélites (especialmente os tetranucleotídeos) são os motivos de escolha nos sistemas de PCR multiplex comercialmente disponíveis. Naturalmente as taxas de mutação dos dinucleotídeos, quando comparados com as de tetranucleotídeos, são bem mais elevadas (ELLEGREN, 2000; XU, PENG, FANG, 2000; HUANG *et al.*, 2002) e, ainda, STRs dinucleotídeos apresentam excessiva ocorrência de *slippage* (deslizamento) durante a amplificação artificial (técnica PCR), produzindo grande número de artefatos (denominados *stutter*) que dificultam e podem confundir a interpretação do resultado.

A ocorrência de mutação é também dependente da estrutura interna do microssatélite, ou seja, dos nucleotídeos constituintes da unidade repetitiva. Por exemplo, os motivos de STR que possuem regiões ricas em A-T apresentaram maiores taxas de mutação, estando mais suscetíveis aos eventos *slippage*; enquanto as regiões ricas em G-C mostraram-se mais estáveis, devido à ponte de hidrogênio tripla entre as bases, apresentando menor ocorrência de eventos mutacionais (PARDI *et al.*, 2005).

Da mesma forma, a ocorrência de mutação é também dependente do comprimento do segmento STR; alelos longos apresentaram maiores taxas de mutação (ELLEGREN, 2000; HUANG *et al.*, 2002; LAI, SUN, 2003), ou seja, quanto maior o número de repetições do motivo STR, mais propensos estarão a eventos *slippage*. Também foi observado que, dentro de um lócus, os alelos longos apresentam uma tendência a mutarem para alelos de tamanho menor (WEBER, WONG, 1993; ELLEGREN, 2000), e alelos mais curtos apresentam tendência à expansão (XU, PENG, FANG, 2000; HUANG *et al.*, 2002). Notadamente, há um viés mutacional dependente do tamanho, com uma relação negativa entre a magnitude do microssatélite (tamanho e direção) e a padronização do comprimento do alelo (HUANG *et al.*, 2002). Poder-se-ia esperar, portanto, que a ocorrência de mutações de expansão e de contração de repetições ocorresse em frequências iguais (BRENNER, 2004); mas alguns estudos mostram as expansões de repetições superiores às contrações (WEBER, WONG, 1993; ELLEGREN, 2000).

A contração do tamanho dos alelos longos promove a distribuição equilibrada do comprimento dos mesmos e parece resultar de um balanço entre eventos de DNA *slippage* e mutações de ponto, que interrompem as sequências repetitivas longas, quebrando-as em sequências menores (KRUGLYAK *et al.*, 1998) – alterando assim a mutabilidade do microssatélite; esse mecanismo aleatório pode funcionar evitando o crescimento infinito dos alelos. As mutações de ponto podem ser do tipo substituição de base (transição e/ou transversão), ou do tipo indel (inserção e/ou deleção), e ocorrem com uma frequência menor em relação às taxas de mutações geradas por DNA *slippage* (FAN, CHU, 2007).

Outro fator associado aos eventos mutacionais dos STRs é a composição da estrutura do alelo. Alelos com STR interrompidos ou heterogêneos (aqueles que contêm unidades irregulares intercaladas em sua composição) apresentam taxas de mutação menores em relação aos alelos não interrompidos ou homogêneos (BRINKMANN *et al.*, 1998). Vários estudos reportaram uma correlação positiva de aumento de mutação nos arranjos repetitivos não interrompidos, ou homogêneos, quanto maior o comprimento das sequências repetitivas (LEOPOLDINO, PENA, 2003; DAUBER *et al.*, 2003), nos quais alelos acima de 10 repetições homogêneas (BRINKMANN *et al.*, 1998; SUN *et al.*, 2015) e acima de 11 repetições homogêneas (KAYSER & SAJANTILA, 2001) apresentaram aumento na taxa de mutação. Conforme sugere Brinkmann *et al.*, uma sequência de repetições interrompida com uma unidade imperfeita promove a estabilização do microssatélite, pois diminui o comprimento da sequência homogênea, conseqüentemente, diminuindo os eventos *slippage*.

A heteroziguidade parece também exercer influência sobre a ocorrência de mutações dos microssatélites, demonstrando maior instabilidade, se comparada à herança em homozigose no mesmo locus. Em 1996, Amos, Sawcer & Feakes revelaram que locus STR em heterozigose apresentaram maior taxa de mutação do que locus em homozigose; eles observaram que a mutabilidade é ainda maior nos locus onde há grande diferença no comprimento dos alelos heterozigotos (conhecido como efeito *allele span*). Um estudo mais recente corroborou com essa hipótese (MASTERS *et al.*, 2010).

A instabilidade dos microssatélites está associada também ao gênero e à idade do indivíduo. Calcula-se que as taxas de mutação de origem paterna são maiores em relação às taxas de mutação de origem materna (LEOPOLDINO, PENA, 2002; DAUBER *et al.*, 2012; QIAN *et al.*, 2015; SUN *et al.*, 2015). Essa diferença está associada ao tipo e ao número de divisões celulares na gametogênese (WEBER, WONG, 1993); a oogônia sofre aproximadamente 22 divisões antes da meiose iniciar e gerar o oócito. A espermatogônia se renova continuamente por mitose, posteriormente algumas passam por meiose para formação do espermatozoide – ou seja, as células espermáticas de um homem velho sofrem mais divisões que as células espermáticas de um jovem, o que justifica a correlação entre a idade do pai com o aumento da ocorrência das mutações (BRINKMANN *et al.*, 1998; ELLEGREN, 2000). Mardini *et al.*, (2013) observaram em seu estudo que a taxa de mutação de origem paterna ($1,8 \times 10^{-3}$) foi significativamente superior à taxa de mutação de origem materna ($0,36 \times 10^{-3}$), mas um dado inesperado foi encontrado, para o locus D21S11 a taxa de mutação paterna ($1,2 \times 10^{-3}$) foi inferior à materna ($1,6 \times 10^{-3}$).

Por fim, cita-se aqui mutações que geram alterações em regiões flanqueantes aos STR, as quais poderão ser importantes no caso do registro do resultado do perfil genético. As

mutações ocorrem com maior frequência em regiões não codificantes do DNA e, principalmente, nas sequências com unidades de repetição em tandem. O tipo de mutação e a sua localização na estrutura do microssatélite podem gerar efeitos diferentes no momento da amplificação daquele marcador genético (CLAYTON *et al.*, 2004; BUCKLETON *et al.*, 2005). Alguns desses efeitos são facilmente detectáveis por genotipagem capilar, pois promovem alteração no comprimento do alelo, enquanto outros possuem efeito mais sutil.

As mutações de ponto promovem alteração de um nucleotídeo. As principais ocorrem 1) pela substituição de um nucleotídeo e são conhecidas como transição ou transversão ou 2) pela inserção ou deleção de um ou mais pares de bases.

Essas variações podem ocorrer na região de pareamento do *primer*, na região flanqueadora ou na região das unidades de repetição dos microssatélites. As mutações de ponto que eventualmente surgirem próximas à extremidade 5', ou no centro da região de anelamento do *primer*, podem reduzir a eficiência do pareamento, fazendo com que ocorra um desbalanço na amplificação do alelo mutante, em relação ao alelo não alterado (FIGURA 2, a). Algumas mutações de substituição podem passar despercebidas em vários sistemas de amplificação, pois, dependendo da localização, não geram alteração no número de repetições e o alelo mutado terá o mesmo comprimento da sequência sem mutação. Por outro lado, se as mutações ocorrerem próximas à extremidade 3' da região de anelamento do primer, podem provocar um efeito mais drástico, como a falha no pareamento do primer e, conseqüentemente, a não amplificação do alelo (FIGURA 2, b). Esse efeito é conhecido como alelo "drop-out" e resulta no aparecimento de alelo nulo (BUTLER, 2006; LI *et al.*, 2014). Suspeita-se da ocorrência de alelo nulo, quando um dos genitores e o filho apresentam uma inconsistência nos genótipos e ambos são homozigotos [Mãe 18-20, Suposto pai 16-16, Filho 18-18]. Nesse caso, indivíduos com genótipo heterozigoto apresentarão um falso genótipo homozigoto. O alelo está presente no DNA, porém devido à falha no anelamento do primer, o mesmo não é amplificado na reação da PCR e não aparecerá na genotipagem dos indivíduos.

As mutações de ponto que porventura surgirem na região flanqueadora ou na região central do microssatélite (onde estão as unidades de repetição) promovem o aparecimento de alelos variantes ou microvariantes (FIGURA 2, c) [Mãe 7-9, Suposto pai 6-10, Filho 6-9.3]. Esses alelos variantes, conforme vão sendo identificados na população, são incorporados aos *ladders* dos kits comerciais. Por exemplo, no marcador HUMTH01 uma mutação no alelo 10 deletou um nucleotídeo de uma unidade repetitiva do tetranucleotídeos e gerou uma variação de trinucleotídeo (X.3) na sequência de unidades repetitivas, fazendo surgir o alelo 9.3 (PUERS *et al.*, 1993). Outros loci, como HUMD21S11 e o HUMvWA, também sofreram mutações nas

estruturas dos STRs tetranucleotídeos, gerando alelos variantes X.2 (MOLLER, MEYER, BRINKMANN, 1994; URQUHART *et al.*, 1994).

Ao se usar painéis com *primers* específicos para a amplificação de lócus, há de se ter garantia de que a região de anelamento do primer (região flanqueante do STR) não está alterada (mutada) a ponto de o primer não conseguir realizar o pareamento. No caso de uma investigação de paternidade, para evitar que se conclua uma exclusão com inconsistência envolvendo lócus em homozigose, esse critério deve ser averiguado e deve ser feita uma amplificação com *primers* que pareiem em uma região diferente, na tentativa de detecção de um possível alelo nulo (CLAYTON *et al.*, 2004; BUTLER, 2006).

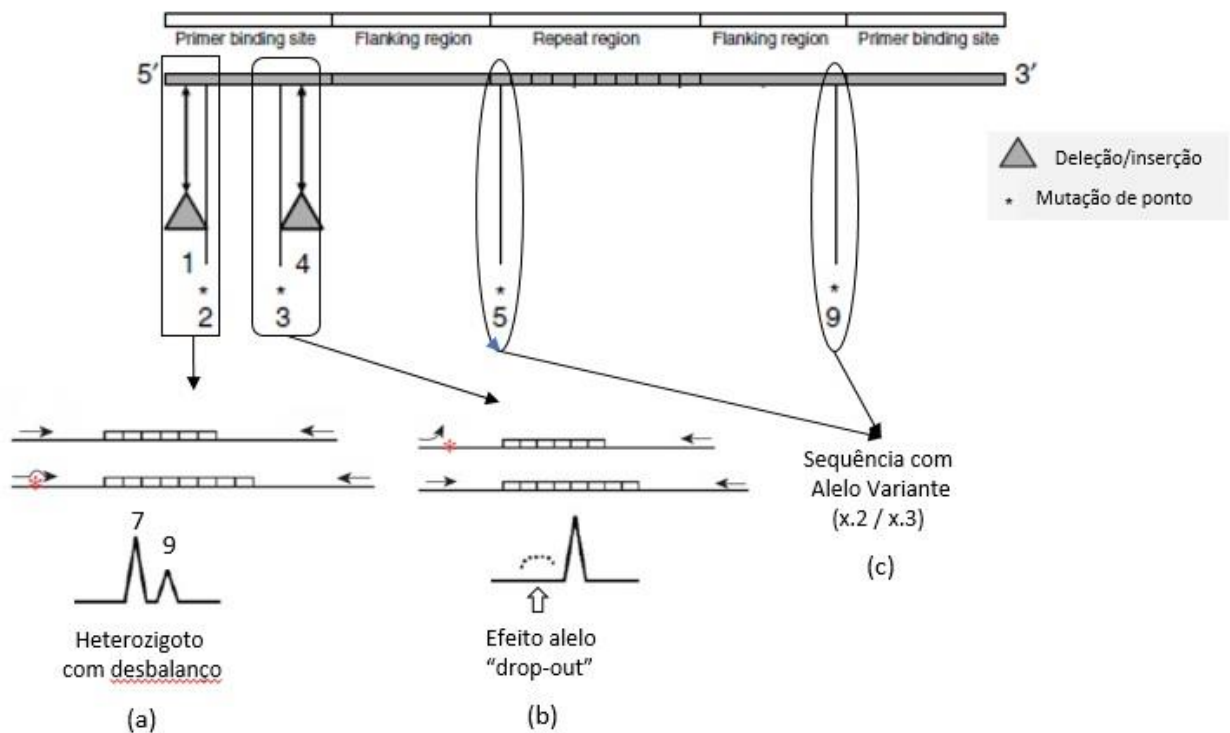


FIGURA 2– Possíveis efeitos das mutações em diferentes localizações da estrutura do microsatélite. Fonte: Buckleton *et al.*, 2005, com adaptação.

TAXAS DE MUTAÇÃO DOS MICROSSATÉLITES

Como apresentado, diferentes estudos têm mostrado que há uma dinâmica variada da ocorrência de mutações, bem como identificado marcadores com maiores taxas de mutação e passos mutacionais que são mais frequentes. Há taxas de mutação específicas dependentes do locus (KAYSER, SAJANTILA, 2001; MARDINI *et al.*, 2013), com notório aumento na taxa de mutação nos locos mais polimórficos (aqueles com escadas alélicas mais amplas e/ou com microvariantes) (BUTLER, 2006; DAUBER *et al.*, 2012). Diversas abordagens foram desenvolvidas para estudar as taxas de mutação, como abordagem familiar, abordagem com modelos biológicos, estudos populacionais e de linhagem celular. A abordagem familiar, amplamente utilizada, consiste em uma análise direta, na qual tanto a taxa de mutação, quanto o tipo de mutação, podem ser diretamente analisados através da transmissão dos STRs dos pais aos filhos (HUANG *et al.*, 2002).

As taxas de mutação para os microsatélites autossômicos humanos foram estimadas entre 10^{-2} a 10^{-4} eventos por locus por geração (WEBER, WONG, 1993; BRINKMANN *et al.*, 1998; SAJANTILA, LUKKA, SYVÄNEN, 1999; XU *et al.*, 2000; HUANG *et al.*, 2002; MARDINI *et al.*, 2013). Índices semelhantes foram observados para os microsatélites do cromossomo Y, de 2×10^{-3} a 4×10^{-3} (SAJANTILA, LUKKA, SYVÄNEN, 1999; KAYSER *et al.*, 2000), o que levou os pesquisadores a concluir que as mutações dos STRs, tanto autossômicos, quanto do cromossomo Y, apresentam comportamento semelhante e podem ser analisados através dos mesmos critérios.

A probabilidade de ocorrência de mutações em dois locos simultâneos, em uma mesma linhagem germinativa, é muito menor. Um estudo que comparou trios (Mãe-Pai-Criança) verdadeiros e falsos, ambos com duas inconsistências, observou a proporção de haver dois trios falsos para cada trio verdadeiro. Na avaliação de uma paternidade, o recomendado é avaliar o Índice de Paternidade (IP) de todos os locos, e considerar que possíveis inconsistências sejam mutações e não resultados de exclusão de paternidade. Nas análises de investigação de vínculo genético, é calculado o Índice de Paternidade Combinado (IPC) e o valor mínimo geralmente empregado para considerar o vínculo como legítimo é de 10.000 (dez mil), em estudos com 15 locos, pelo menos. Para casos em que se observe uma inconsistência no genótipo do trio, deve-se incluir nesse cálculo a taxa de mutação estimada para o locus que apresentar a mutação (BRENNER, 2004). Também é chamada a atenção para que o analista não se detenha apenas no número rígido de locos excludentes ou inconsistentes. A avaliação deve ser baseada na análise estatística incluindo a taxa de mutação específica para o locus (ou para o alelo específico, se houver) e outros parâmetros como a estrutura repetitiva do locus, o

número de passos mutacionais, o modelo de mutação dos microssatélites (BRINKMANN *et al.*, 2001).

JUSTIFICATIVA

Conhecer as taxas de mutação locus-específicas e os tipos de mutação mais frequentes para os marcadores microsatélites utilizados em análises forenses é fundamental para a correta interpretação dos perfis genético e para a determinação dos critérios de exclusão dos testes de vínculo genético. Estudos variados e dispersos foram capazes de identificar locus que apresentam as maiores taxas de mutação, alelos mais propensos a certos eventos mutacionais, entre outras características. Porém nota-se uma carência de um estudo mais abrangente e que considere, na mesma amostra populacional, a relação das taxas de mutação com os fatores apresentados que influenciam os eventos mutacionais. A comunidade forense seria muito beneficiada, contudo, se pudesse contar com um levantamento amplo que identifique as taxas de mutações por locus, por categorias de alelos (número de unidades repetitivas), por sexo do indivíduo e que caracterize cada evento mutacional. No Brasil, o centro público que congrega o maior número de casos cíveis já avaliados é o do Estado de São Paulo (Instituto de Medicina Social e de Criminologia; IMESC) e os dados gerados por esse centro nunca foram levantados. Diante da necessidade de se poder contar com dados precisos acerca de mutações que refinem as análises de paternidade e de identificação humana, esse trabalho se propõe a explorar o banco de informações do IMESC e realizar o levantamento da ocorrência de eventos mutacionais em locus interesse forense. Os resultados do presente estudo serão referência para todas as unidades de investigação forense cível ou criminal.

OBJETIVOS

- 1- Identificar e caracterizar eventos mutacionais do tipo *slippage* em marcadores microssatélites autossômicos de interesse forense em um banco público de informações genéticas de indivíduos brasileiros.

- 2- Unir os dados obtidos com os demais dados Brasileiros já disponíveis a fim de fornecer à comunidade forense informações acerca das taxas de mutações por locus.

LEVANTAMENTO DOS DADOS E CRITÉRIOS DE INCLUSÃO NO ESTUDO

O levantamento dos dados arquivados no IMESC foi realizado de forma manual, através da conferência dos casos em suas pastas físicas e, posteriormente, tabulados em uma planilha eletrônica. Ao total, foram conferidos 14.756 casos (73,2% deles foram casos conclusivos e nos quais se confirmou haver vínculo biológico entre os indivíduos), tendo sido detectados 408 eventos mutacionais. Aplicada a primeira triagem de qualidade, de seleção pelo tipo de mutação e pelo tipo e número de locos estudados, permaneceram 7.060 casos no estudo, dos quais 5.171, com vínculo biológico confirmado através do Índice de Paternidade Combinado com o valor de *Likelihood Ratio* (LR) superior a 10.000. Os casos nos quais foram verificadas inconsistências alélicas entre as gerações, que foram a base do estudo das taxas de mutação, tinham como característica, obrigatoriamente:

1- Tratar-se de evento único de inconsistência (mutação entre gerações).

2- O vínculo biológico entre os indivíduos confirmado com $LR > 10.000$, após a inclusão de valores globais de taxa de mutação no cálculo.

3- Os locos com inconsistência submetidos à amplificação com o uso de, pelo menos, dois painéis comerciais fabricados por empresas diferentes, a fim de garantir análises com diferentes sequências de *primers*.

Ao final, 193 eventos mutacionais foram considerados em condições de serem incluídos na presente análise e foram tratados conforme o Modelo de *Stepwise Mutation*. Uma amostra de 215 outros casos com presença de inconsistências alélicas entre as gerações aguardam análises complementares e/ou duplicatas para serem incluídos em uma fase avançada deste estudo.

ARTIGO CIENTÍFICO

Artigo submetido à publicação na revista científica ELECTROPHORESIS, redigido conforme Instruções aos Autores disponível em https://wol-prod-cdn.literatumonline.com/pb-assets/assets/15222683/ELPS_Instructions_to_authors_Oct_2017-1521035045717.pdf.

Short Communication

Slippage mutation rates in 15 autosomal short tandem repeat loci for forensic purposes in a Southeastern Brazilian population

Fernanda Irma Remus Hamester ^{1,3,*}; Débora Santos da Silva ²; Ana Paula Magalhães Leboutte ³; Carlos Henrique Ares Motta ²; Clarice Sampaio Alho ¹

¹ Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Laboratório de Genética Humana e Molecular - Genética Forense, Porto Alegre, RS, Brazil.

² Instituto de Medicina Social e Criminologia (IMESC), Núcleo de Perícias Laboratoriais, São Paulo, SP, Brazil.

³ Instituto-Geral de Perícias do Estado do Rio Grande do Sul (IGP-RS), Divisão de Genética Forense, Porto Alegre, RS, Brazil.

***Corresponding author:** Fernanda Irma Remus Hamester. Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Escola de Ciências: Laboratório de Genética Humana e Molecular - Genética Forense, Av. Ipiranga, 6681 - Prédio 12; Porto Alegre, RS - BRAZIL; ZIP CODE: 90619-900. E-mail: fer.hamester@gmail.com

Abbreviations: **STR**, short tandem repeats; **DNA**, Deoxyribonucleic Acid; **FTA®** cards, Flinders Technology Associates cards; **PCR**, Polymerase Chain Reaction.

Keywords: Mutation Rates, Short Tandem Repeat, Paternity Testing, Microsatellite.

ABSTRACT

Well-defined estimates of mutation rates in highly polymorphic tetra-nucleotide short tandem repeat (STR) loci are a prerequisite for human identification in the routines of civil (paternity) and criminal investigations. Studying 15 autosomal STR loci of forensic interest (CSF1PO, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, FGA, TH01, TPOX and vWA), we detected 193 slippage mutations (189 one-step and four two-step mutations) in 148,875 parent–child allelic transfers from 5,171 paternity cases with true biological relationship (15,096 individuals; 4,754 trios and 417 duos; 9,925 meioses) from São Paulo State, a very representative population of Brazil. Overall mutation rate was 1.3×10^{-3} and the highest rates were observed at loci vWA (2.8×10^{-3}), FGA and D18S51 (2.7×10^{-3} for both) while loci TH01 and TPOX presented no mutations. Mean slippage mutation rate of paternal origin (1.8×10^{-3}) was six times higher than that observed from maternal origin (0.3×10^{-3}).

Main text

DNA polymerase slippage generates new alleles during cell division. DNA sequences of repetitive nucleotides are unstable regions of the genome where frequent insertions and deletions of nucleotides can take place in consequence of polymerase slipped-strand mispairing, resulting in new alleles with genome rearrangements (1). During replication, DNA polymerase III slips from the DNA template strand at the repeat region and subsequently reattaches at a more/less-distant site causing a newly created DNA strand with expanded/contracted DNA sequence. Slippage replications have different influence depending on whether they occur in coding or non-coding regions. Variations on coding polynucleotide

sequences result in the formation of abnormal protein structure that can affect the phenotype, but more expanded or contracted new alleles in non-coding DNA regions just increase variability.

Repeated sequences are distributed throughout human genome and different alleles are classified according to the number of contiguous repeat units. Loci with STRs have proven to be especially appropriate for personal identification. The most widely used DNA markers for human identification in civil (paternity) and criminal investigations are the high polymorphic tetra-nucleotide STRs. For more than a decade, commercial multiplex panels to genotype STR with very high accuracy have been available allowing a trouble-free laboratory routine for scientists (2). Despite of this, it is not unusual to observe STR mutations disrupting the work day in a DNA analysis laboratory. Important mutation rates in these markers are noticed in all populations, precisely due to polymerase slippage on their tandem repeats DNA structure.

In general, new alleles originated from polymerase slippage events have the same pattern of mutations with gain or loss of just one, or more rarely two or three, steps of the unit of repeat. However, the rate of STR mutations varies considerably according to the locus or even the population. In the present study, we estimated the mutation rates at the 15 STR loci (including the 13 markers of the Combined DNA Index System of Federal Bureau of Investigation - CODIS, FBI) mostly used commercially in a Southeastern Brazilian population.

A total of 7,060 paternity cases from Paternity Laboratory of Institute of Social Medicine and Criminology (IMESC), São Paulo, Brazil were included in the study. DNA samples were obtained from blood on FTA® cards (Whatman Bioscience, Cambridge, UK) and they were genotyped using AmpFISTR®

Identifiler™ PCR Amplification Kit (Applied Biosystems; Life Technologies, USA). When an allelic inconsistency was found, each DNA sample was re-genotyped using PowerPlex® 18D System PCR kit (Promega Corporation, Madison, WI) in order to confirm the sequence. These panels have different primer design and both analyze 15 human autosomal DNA markers (CSF1PO, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, FGA, TH01, TPOX and vWA) by capillary electrophoresis. Sequencing was performed in ABI PRISM® 3130xl Genetic Analyzer and Data Collection v.3.0 and GeneMapper ID v.3.2.1 software (Applied Biosystems; Life Technologies, USA) were used for genotype identification. In order to follow previous systems to describe mutations, we assumed the shortest path of steps from the paternal/maternal allele occurred to generate the child's discrepant mutated allele (3–6). We used the nomenclature recommended by the DNA Commission of the International Society of Forensic Genetics, as well as the recommendations for population data publication for forensic purposes (7,8). All of the participants consented with the investigation, and this study was carried out in accordance with approved ethical standards. For quality control, the Laboratory Internal Control standards and commercial kits' controls were followed throughout the study.

True biological relationship between mother-child and/or father-child was found in 5,171 (73.2%) out of 7,060 paternity cases (15,096 individuals; 4,754 trios and 417 duos; 9,925 meioses per locus) with a high paternity index value ($LR > 10,000$). We noticed 193 inconsistencies derived from slippage mutations in 193 families (each family presented only a single allelic discordancy) for the 15 STR loci analyzed from these biologically related persons, totalling 148,875

meiosis (Table 1). We confirmed 189 (97.9%) mutations as one-step and four (2.1%) as two-step, the latest ones at locus FGA (22>24 or 26>24; 27>25), CSF1PO (12>14) and D13S317 (11>9). Higher numbers of slippage mutations were observed in the most polymorphic loci, agreeing with data presented in other studies (9,10).

Based on the number of meiosis per locus, the overall mutation rate for the 15 forensic STR loci was 1.3×10^{-3} (95% CI: $1.1 - 1.5 \times 10^{-3}$), which is similar to previous studies with Brazilian populations in the Southern (5) and Southeastern regions (11), Poland and Southern Chinese Han (12) that ranged from $1,18 \times 10^{-3}$ to $1,39 \times 10^{-3}$. It is slightly different from an earlier study including a Southern, Southeastern and Central Brazil populations (4) (0.94×10^{-3}) and data published from the Northeastern region of the country (0.75×10^{-3}) (3) and much higher than the rate reported in a recent study analyzing a population from Central Brazil (0.21×10^{-3}) (13).

The highest STR mutation rates were observed at loci vWA (2.8×10^{-3}), FGA and D18S51 (both with 2.7×10^{-3}), while loci TH01 and TPOX had no mutations. In Southern and Central Brazil studies, these were also the loci that presented the highest mutation rates (5,13) but they reported very low mutation rates for TH01 and TPOX. A study in São Paulo and another one in Poland reported the highest rates for vWA and FGA but the third highest were D12S391 and D5S818, respectively. Both had no mutations in TH01 and TPOX (11,14). The earlier study including a Southern, Southeastern and Central Brazil populations (4) reported D10S1237 and D19S253 with the highest mutation rates and no mutations in TH01 and TPOX, but the first loci were not analyzed

in the present study. The most polymorphic loci are expected to present higher mutation rates (10,15,16).

Our results are also similar to two other Brazilian populations (5,11) in paternal:maternal proportion of mutation origin. The mutation rate generated by paternal meiosis (1.8×10^{-3}) was six times higher than the maternal one (0.3×10^{-3}), even though we have detected 27 discrepant alleles with ambiguous origin.

Differently from all other loci, locus D21S11 presented a paternal mutation rate (1.4×10^{-3}) quite similar to the maternal one (1.9×10^{-3}). The relationship between mutation rate and lineage origin at locus D21S11 should be better elucidated since two previous Brazilian reports (3,5) observed no differences between paternal:maternal proportion. Vieira et al. (13) observed a higher paternal rate but still quite many mutations from maternal origin as well as Sun et al. (12) in a study analyzing a Chinese and Brazilian populations, while Whittle et al. (4) and Martinez et al. (11) found no mutations in female lineage in this locus. This type of results exposes the importance of distinguishing paternal and maternal origin of mutated alleles.

Direction of mutational steps was verified and the same proportion of gain to loss of repeats (1:1; 77 for both) was observed, although the information about expansion or contraction of the new allele was ambiguous in 40 (20.7%). The results reported from other Brazilian groups were 1.2:1 for Rio Grande do Sul State population (5), 1.5:1 for São Paulo State (11) but they have very few mutations compared to the present study and to others in the literature. For the Chinese population, the proportions were $\sim 1.1:1$ and $\sim 1.7:1$ (17,18). It should be noted that the ratio found here might be overestimated (or underestimated) because of the 44 undetermined mutations, a similar ratio should remain in the

unassigned events as described by Lu et al. (18). In addition, it should be considered the differences between the number and/or allele types analyzed by those studies.

Lastly, we analyzed the relation between gain/loss of unit repeats according to the number of tetra-nucleotide unit repeats and classified allelic scale per locus as short (25% smaller ones), moderate (50%) and long (25% longest ones) following the criteria published by Ge et al. (19). The proportion of short: moderate: long was 1:23:3 and details of these analyses are presented in Table S1. Considering polymerase slipped-strand mispairing as a stochastic process, the probability of generating new expanded or contracted alleles distribution should remain balanced over time, and it would be expected to detect a trend to a stationary process. In fact, our data showed shorter alleles more inclined to gain repeats (5 gains: 2 losses), moderate alleles with equivalent proportion to both (68 gains: 59 losses) and longer ones more inclined to lose repeats (4 gains: 16 losses), supporting a stationary distribution trend. Similar trends have been reported by other studies (18–20) while they found the rate of repeat contraction to increase with allele size increasing (21,22).

It is accepted that DNA mutations may appear as a hurdle to conclude an investigation. Slippage mutations during meiosis can raise doubts in paternity/maternity tests when allelic inconsistencies are detected between father-child or mother-child match. Incorporation of mutation rates per locus in calculations and in the reports would elucidate each allelic inconsistency and increase the precision of DNA investigations. Additionally, we suggest the use of mutation rates per locus considering whether mutated allele came from paternal or maternal meiosis.

ACKNOWLEDGEMENTS

First author was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil. This study was supported by grant PRO-FORENSES CAPES-BRAZIL.

The authors have no conflict of interest to declare.

REFERENCES

1. Viguera E. Replication slippage involves DNA polymerase pausing and dissociation. *EMBO J* [Internet]. 2001 May 15;20(10):2587–95. Available from: <http://emboj.embopress.org/cgi/doi/10.1093/emboj/20.10.2587>
2. Buckleton JS, Krawczak M, Weir BS. The interpretation of lineage markers in forensic DNA testing. *Forensic Sci Int Genet*. 2011;
3. de Andrade ES, Gomes AV, Raposo G, Mauricio-da-Silva L, Silva R dos S. Mutation rates at 14 STR loci in the population from Pernambuco, Northeast Brazil. *Forensic Sci Int Genet* [Internet]. 2009 Sep;3(4):e141–3. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S1872497308001841>
4. Whittle MR, Romano NL, Negreiros VAC. Updated Brazilian genetic data, together with mutation rates, on 19 STR loci, including D10S1237. *Forensic Sci Int* [Internet]. 2004 Jan;139(2–3):207–10. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0379073803004894>
5. Mardini AC, Rodenbusch R, Schumacher S, Chula FGL, Michelon CT, Gastaldo AZ, et al. Mutation rate estimates for 13 STR loci in a large population from Rio Grande do Sul, Southern Brazil. *Int J Legal Med* [Internet]. 2013 Jan 10;127(1):45–7. Available from: <http://link.springer.com/10.1007/s00414-011-0642-x>
6. Qian X-Q, Yin C-Y, Ji Q, Li K, Fan H-T, Yu Y-F, et al. Mutation rate analysis at 19 autosomal microsatellites. *Electrophoresis* [Internet]. 2015 Jul;36(14):1633–9. Available from: <http://doi.wiley.com/10.1002/elps.201400558>
7. Carracedo Á, Butler JM, Gusmão L, Parson W, Roewer L, Schneider PM. Publication of population data for forensic purposes. *Forensic Sci Int Genet* [Internet]. 2010 Apr;4(3):145–7. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S1872497310000293>
8. American Association of Blood Banks. AABB Relationship Testing Program Unit. Annual Report Summary for Testing in 2013. [Internet]. Vol. 675. Bethesda, MD; 2013. Available from: <http://www.aabb.org/sa/facilities/Documents/2013-relationship-testing->

summary-report.pdf

9. Butler JM. Genetics and Genomics of Core Short Tandem Repeat Loci Used in Human Identity Testing. *J Forensic Sci* [Internet]. 2006 Mar;51(2):253–65. Available from: <http://doi.wiley.com/10.1111/j.1556-4029.2006.00046.x>
10. García-Aceves ME, Romero Rentería O, Díaz-Navarro XX, Rangel-Villalobos H. Paternity tests in Mexico: Results obtained in 3005 cases. *J Forensic Leg Med* [Internet]. 2018;55(January):1–7. Available from: <https://doi.org/10.1016/j.jflm.2018.02.003>
11. Martinez J, Braganholi DF, Ambrósio IB, Polverari FS, Cicarelli RMB. Mutation rates for 20 STR loci in a population from São Paulo state, Southeast, Brazil. *Ann Hum Biol* [Internet]. 2017 Oct 3;44(7):659–62. Available from: <https://www.tandfonline.com/doi/full/10.1080/03014460.2017.1371222>
12. Sun H, Liu S, Zhang Y, Whittle MR. Comparison of southern Chinese Han and Brazilian Caucasian mutation rates at autosomal short tandem repeat loci used in human forensic genetics. *Int J Legal Med* [Internet]. 2014 Jan 3;128(1):1–9. Available from: <http://link.springer.com/10.1007/s00414-013-0847-2>
13. Vieira TC, Duarte Gigonzac MA, Goulart Rodovalho R, Morais Cavalcanti L, Bernardes Minasi L, Melo Rodrigues F, et al. Mutation rates in 21 autosomal short tandem repeat loci in a population from Goiás, Brazil. *Electrophoresis* [Internet]. 2017 Nov;38(21):2791–4. Available from: <http://doi.wiley.com/10.1002/elps.201700192>
14. Wojtas M, Piniewska D, Polańska N, Stawowiak A, Sanak M. Mutations of microsatellite autosomal loci in paternity investigations of the Southern Poland population. *Forensic Sci Int Genet* [Internet]. 2013 May;7(3):389–91. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1872497312002785>
15. Butler JM. *Advanced topics in forensic DNA typing: interpretation*. Oxford: Academic Press; 2014. 608 p.
16. Zhang Q, Yang M, Pan Y, Zhao J, Qu B, Cheng F, et al. Development of a massively parallel sequencing assay for investigating sequence

- polymorphisms of 15 short tandem repeats in a Chinese Northern Han population. *Electrophoresis* [Internet]. 2018 Nov;39(21):2725–31. Available from: <http://doi.wiley.com/10.1002/elps.201800071>
17. Zhu WQ, Jin B, Su Q, Luo HB, Li YB, Wu J, et al. Mutation Study of 28 Autosomal STR Loci in Southwest Chinese Han Population. *Forensic Sci Int Genet Suppl Ser* [Internet]. 2015 Dec;5:e298–9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1875176815302675>
 18. Lu D, Liu Q, Wu W, Zhao H. Mutation analysis of 24 short tandem repeats in Chinese Han population. *Int J Legal Med* [Internet]. 2012 Mar 18;126(2):331–5. Available from: <http://link.springer.com/10.1007/s00414-011-0630-1>
 19. Ge J, Budowle B, Aranda XG, Planz J V., Eisenberg AJ, Chakraborty R. Mutation rates at Y chromosome short tandem repeats in Texas populations. *Forensic Sci Int Genet*. 2009;3(3):179–84.
 20. Dupuy BM, Stenersen M, Egeland T, Olaisen B. Y-chromosomal microsatellite mutation rates: Differences in mutation rate between and within loci. *Hum Mutat* [Internet]. 2004 Feb;23(2):117–24. Available from: <http://doi.wiley.com/10.1002/humu.10294>
 21. Xu X, Peng M, Fang Z, Xu X. The direction of microsatellite mutations is dependent upon allele length. *Nat Genet* [Internet]. 2000 Apr 1;24(4):396–9. Available from: http://www.nature.com/articles/ng0400_396
 22. Shao C, Lin M, Zhou Z, Zhou Y, Shen Y, Xue A, et al. Mutation analysis of 19 autosomal short tandem repeats in Chinese Han population from Shanghai. *Int J Legal Med* [Internet]. 2016 Nov 28;130(6):1439–44. Available from: <http://link.springer.com/10.1007/s00414-016-1427-z>

Table 1. Slippage mutations observed in 5,171 paternity cases from individuals with confirmed biologic relationship (LR>10,00).

MARKER	PATERNAL ORIGIN				MATERNAL ORIGIN				UNCERTAIN ORIGIN	TOTAL			
	No. of meioses	No. of mutations	Mutation rate (*10 ⁻³)	95% CI ^a (*10 ⁻³)	No. of meioses	No. of mutations	Mutation rate (*10 ⁻³)	95% CI ^a (*10 ⁻³)		No. of meioses	No. of mutations	Mutation rate (*10 ⁻³)	95% CI ^a (*10 ⁻³)
CSF1P0	5131	12	2.3	1.2-4.1	4794	2	0.4	0.0-1.5	1	9925	15	1.5	0.8-2.5
D2S1338	5131	8	1.6	0.7-3.1	4794	1	0.2	0.0-1.2	1	9925	10	1.1	0.5-1.8
D3S1358	5131	8	1.6	0.7-3.1	4794	1	0.2	0.0-1.2	3	9925	12	1.2	0.6-2.1
D5S818	5131	4	0.8	0.2-2.0	4794	0	0	0.0	2	9925	6	0.6	0.2-1.3
D7S820	5131	9	1.8	0.1-3.3	4794	0	0	0	1	9925	10	1.0	0.5-1.8
D8S1179	5131	7	1.4	0.5-2.8	4794	0	0	0	4	9925	11	1.1	0.5-2.0
D13S317	5131	15	2.9	1.6-4.8	4794	3	0.6	0.1-1.8	2	9925	20	2.0	1.2-3.1
D16S539	5131	4	0.8	0.2-2.0	4794	0	0.0	0.0	1	9925	5	0.5	0.2-1.2
D18S51	5131	21	4.1	2.5-6.2	4794	2	0.4	0.0-1.5	4	9925	27	2.7	1.8-4.0
D19S433	5131	3	0.6	0.1-1.7	4794	1	0.2	0.0-1.2	0	9925	4	0.4	0.1-1.0
D21S11	5131	7	1.4	0.5-2.8	4794	9	1.9	0.9-3.6	2	9925	18	1.8	1.1-2.9
FGA	5131	22	4.3	2.7-6.5	4794	4	0.8	0.2-2.1	1	9925	27	2.7	1.8-4.0
TH01	5131	0	0	0	4794	0	0	0	0	9925	0	0	0
TPOX	5131	0	0	0	4794	0	0	0	0	9925	0	0	0
VWA	5131	20	3.9	2.4-6.0	4794	1	0.2	0.0-1.2	7	9925	28	2.8	1.9-4.1
TOTAL		140	1.8	1.5-2.1		24	0.3	0.2-0.5	29		193	1.3	1.1-1.5

a) CI: Confidence interval

Table S1. Distribution of gain/loss of repeats according to allele sizes.

Marker	Short					Moderate					Long				
	Alleles	Gains	Loss	G/L	Total	Alleles	Gains	Loss	G/L	Total	Alleles	Gains	Loss	G/L	Total
CSF1PO	6-9	0	0	0	0	10-12.2	6	2	1	9	13-16.3	0	6	0	6
D13S317	7-9	1	0	0	1	10-13	6	5	4	15	14-16	1	3	0	4
D16S539	5-8	0	0	0	0	9-12	2	0	2	4	13-15	0	1	0	1
D18S51	7-11	0	0	0	0	12-20	12	10	2	24	21-27	2	1	0	3
D19S433	5.2-9	0	0	0	0	10-15.2	3	1	0	4	16-17.2	0	0	0	0
D21S11	24-28.2	1	0	0	1	29-34	5	7	5	17	34.2-36	0	0	0	0
D2S1338	13-18	2	0	0	2	19-24	2	3	1	6	25-28	0	2	0	2
D3S1358	9-12	0	0	0	0	13-17	2	1	2	5	18-20	1	3	3	7
D5S818	7-8	0	0	0	0	9-14	2	1	3	6	15-16	0	0	0	0
D7S820	6-8	0	1	0	1	9-13	5	2	2	9	14-16	0	0	0	0
D8S1179	7-10	0	0	0	0	11-16	5	5	1	11	17-19	0	0	0	0
FGA	17-21	1	1	0	2	21.2-31.2	6	15	4	25	32.2-50.2	0	0	0	0
TH01	5-6	0	0	0	0	7-9.3	0	0	0	0	10-12	0	0	0	0
TPOX	5-7	0	0	0	0	8-11	0	0	0	0	12-15	0	0	0	0
vWA	10-13	0	0	0	0	14-20	12	7	9	28	21-24	0	0	0	0
Total		5	2	0	7		68	59	36	163		4	16	3	23

CONSIDERAÇÕES FINAIS

Os resultados obtidos, até o momento, estão em concordância com os demais estudos publicados na comunidade científica. Os pesquisadores têm como perspectivas futuras ampliar a análise, na mesma direção, incluindo outros eventos mutacionais que não selecionados para o presente levantamento. O próximo passo incluirá análises laboratoriais extras e, nos casos em que se julgar necessário, sequenciamento das amostras. Além disso, pretendemos sequenciar os possíveis alelos presentes nos casos nos quais não se pode determinar se a inconsistência teve origem materna ou paterna, a fim de buscar, na sequência de pares de bases do DNA, detalhes que permitam fazer essa classificação. Ao final, pretende-se caracterizar todas as 408 mutações presentes nos 14.756 casos do IMESC avaliados aqui.

REFERÊNCIAS BIBLIOGRÁFICAS

AMOS, W. Microsatellites show mutational bias and heterozygote instability. *Nature genetics*, 1996. v. 13, p. 390–391.

BRENNER, C. . Multiple mutations, covert mutations and false exclusions in paternity casework. *International congress series*, 2004. v. 1261, p. 112–114. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0531513103018430>>.

Brenner CH (2004) Mutations in paternity. <http://www.dna-view.com>. Acessado em 20/09/2015.

Buckleton *et al.* (2005) *Forensic DNA Evidence Interpretation*. CRC Press

BRINKMANN, B. *et al.* Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *American journal of human genetics*, 1998. v. 62, p. 1408–1415.

BUTLER, J. M. Genetics and genomics of core short tandem repeat loci used in human identity testing. *Journal of forensic sciences*, 2006. v. 51, n. 2, p. 253–265.

CHAKRABORTY, R. *et al.* Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proceedings of the national academy of sciences of the united states of america*, 1997. v. 94, n. 3, p. 1041–6. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=19636&tool=pmcentrez&rendertype=abstract>>.

CLAYTON, T. M. *et al.* Primer binding site mutations affecting the typing of str loci contained within the ampflstr SGM Plus kit. *Forensic science international*, 2004. v. 139, n. 2–3, p. 255–259.

DAUBER, E. M. *et al.* Germline mutations of str-alleles include multi-step mutations as defined by sequencing of repeat and flanking regions. *Forensic science international: genetics*, 2012. v. 6, n. 3, p. 381–386.

DAUBER BAR, W., KLINTSCHAR, M., NEUHUBER, F., PARSON, W., GLOCK, B. AND MAYR, W. R., E. M. Mutation rates at 23 different short tandem repeat loci. *International congress series*, 2003. v. 1239, p. 565–567. Disponível em: <<http://www.sciencedirect.com/science/article/B7581-47W664D->

1K0/2/762d7ea04e559861579fc599772c9a99>.

EDWARDS, A. *et al.* Dna typing and genetic mapping with trimeric and tetrameric tandem repeats. *American journal of human genetics*, 1991. v. 49, n. 4, p. 746–756.

ELLEGREN, H. Heterogeneous mutation processes in human microsatellite dna sequences. *Nature genetics*, 2000. v. 24, n. 4, p. 400–2. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/10742106>>.

ELLEGREN, H. Microsatellites: simple sequences with complex evolution. *Nature reviews*, 2004. v. 5, n. June, p. 435–445.

FAN, H.; CHU, J.-Y. A brief review of short tandem repeat mutation. *Genomics, proteomics & bioinformatics*, 2007. v. 5, n. 1, p. 7–14.

HAMMOND, H. A. *et al.* Evaluation of 13 short tandem repeat loci for use in personal identification applications. *Am. j. hum. genet*, 1994. v. 55, p. 175–189.

HUANG, Q.-Y. *et al.* Mutation patterns at dinucleotide microsatellite loci in humans. *American journal of human genetics*, 2002. v. 70, n. 3, p. 625–634.

JIA, Y. S. *et al.* Multistep microsatellite mutation leading to father-child mismatch of fga locus in a case of non-exclusion parentage. *Legal medicine*, 2015. v. 17, n. 5, p. 364–365.

KAYSER, M.; SAJANTILA, A. Mutations at y-str loci: implications for paternity testing and forensic analysis. *Forensic science international*, 2001. v. 118, n. 2–3, p. 116–121.

KIMURA, M.; OHTA, T. Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proc. Nati. Acad. Scd. USA Genetics*, 1978. v. 75, n. 6, p. 2868–2872.

KRUGLYAK, S. *et al.* Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proceedings of the national academy of sciences of the United States of America*, 1998. v. 95, n. 18, p. 10774–10778.

LAI, Y.; SUN, F. The relationship between microsatellite slippage mutation rate and the number of repeat units. *Molecular biology and evolution*, 2003. v. 20, n. 12, p. 2123–2131.

LEOPOLDINO, A. M.; PENA, S. D. J. The mutational spectrum of human autosomal tetranucleotide microsatellites. *Human mutation*, 2003. v. 21, n. 1, p. 71–79.

LEVINSON, G.; GUTMAN, G. A. High frequencies of short frameshifts in poly-ca/tg tandem repeats borne by bacteriophage m13 in escherichia coli k-12. *Nucleic acids research*, 1987. v. 15, n. 13, p. 5323–5338.

LI, F. *et al.* Identification of new primer binding site mutations at th01 and d13s317 loci and determination of their corresponding str alleles by allele-specific pcr. *Forensic science international: genetics*, 2014. v. 8, n. 1, p. 143–146. Disponível em: <<http://dx.doi.org/10.1016/j.fsigen.2013.08.013>>.

MARDINI, A. C. *et al.* Mutation rate estimates for 13 str loci in a large population from rio grande do sul, southern brazil. *International journal of legal medicine*, 2013. v. 127, n. 1, p. 45–47.

MASTERS, B. S. *et al.* Evidence for heterozygote instability in microsatellite loci in house wrens. *Biology letters*, 2010. v. 7, n. 1, p. 127–130.

MEYER, E.; BRINKMANN, B. *Legal medicine* ©. 1994. p. 319–323.

PARDI, F. *et al.* On the structural differences between markers and genomic ac microsatellites. *Journal of molecular evolution*, 2005. v. 60, n. 5, p. 688–693.

PUERS, C. *et al.* Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus humth01[aatg]_n and reassignment of alleles in population analysis by using a locus-specific allelic ladder. *American journal of human genetics*, 1993. v. 53, n. 4, p. 953–958.

QIAN, X. Q. *et al.* Mutation rate analysis at 19 autosomal microsatellites. *Electrophoresis*, 2015. v. 36, n. 14, p. 1633–1639.

SAJANTILA, A.; LUKKA, M.; SYVÄNEN, A.-C. Experimentally observed germline mutations at human micro- and minisatellite loci. *European journal of human genetics*, 1999. v. 7, n. 2, p. 263–266.

SUN, M. *et al.* Mutations of short tandem repeat loci in cases of paternity testing in chinese. *International journal of legal medicine*, 2015. v. 130, n. 5, p. 1203–1204.

URQUHART, A. *et al.* Variation in short tandem repeat sequences--a survey of twelve microsatellite loci for use as forensic identification markers. *International journal of legal medicine*, 1994. v. 107, p. 13–20.

VALDES, A M.; MONTGOMERY, S.; FREIMER, N. B. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics*, 1993. v. 133, n. 3, p. 737–749.

WEBER, J. L.; MAY, P. E. Abundant class of human dna polymorphisms which can be typed using the polymerase chain reaction. *American journal of human genetics*, 1989. v. 44, p. 388–396.

WEBER, J. L.; WONG, C. Mutation of human short tandem repeats. 1993. v. 2, n. 8, p. 1123–1128.

XU, X.; PENG, M.; FANG, Z. The direction of microsatellite mutations is dependent upon allele length. *Nature genetics*, 2000. v. 24, n. 4, p. 396–9. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/10742105>>.

ANEXO

11/22/2018

ScholarOne Manuscripts

**ELECTROPHORESIS**[Home](#)[Author](#)[Review](#)**Submission Confirmation**[Print](#)

Thank you for your submission

Submitted to**ELECTROPHORESIS****Manuscript ID****elps.201800491****Title****Slippage mutation rates in 15 autosomal short tandem repeat loci for forensic purposes in a Southeastern Brazilian population****Authors****Hamester, Fernanda****Silva, Débora****Leboute, Ana Paula****Silveira da Motta, Carlos Henrique****Alho, Clarice****Date Submitted****22- Nov -2018**[Author Dashboard](#)<https://mc.manuscriptcentral.com/elpho>

1/2 11/22/2018

ScholarOne Manuscripts

[© Clarivate Analytics](#) | © ScholarOne, Inc., 2018. All Rights Reserved.

ScholarOne Manuscripts and ScholarOne are registered trademarks of ScholarOne, Inc. ScholarOne Manuscripts Patents [#7,257,767](#) and [#7,263,655](#).

[@ScholarOneNews](#) | [System Requirements](#) | [Privacy Statement](#) | [Terms of Use](#)