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

JULIANE BENTES PICANÇO

# ESTUDO DE TRIALELIAS NO MARCADOR FORENSE *TPOX* E CARACTERIZAÇÃO DO TERCEIRO ALELO

Porto Alegre
2014

JULIANE BENTES PICANÇO

# ESTUDO DE TRIALELIAS NO MARCADOR FORENSE *TPOX* E CARACTERIZAÇÃO DO TERCEIRO ALELO

Tese apresentada como requisito para a obtenção do grau de Doutor pelo Programa de Pós-Graduação em Biologia Celular e Molecular da Faculdade de Biociências da Pontifícia Universidade Católica do Rio Grande do Sul

Orientador

Profa. Dra. Clarice Sampaio Alho

Porto Alegre

2014

**AGRADECIMENTOS**

A Profa. Dra. Clarice Sampaio Alho, agradeço pela oportunidade de fazer parte de seu grupo de pesquisa por todos esses anos. Além de todo o conhecimento compartilhado e por toda sua paciência em me aturar.

Aos queridos amigos, Déborah Soares, Fernanda Sawitzki, Rodrigo Rodenbusch, Bruna Schroeder, agradeço por toda a ajuda, troca de experiências e pelo momentos de alegrias e descontração que foram essenciais para a execução desse trabalho.

Ao amigo, Paulo Eduardo Raimann, que me ajudou nas análises dos perfis, valeu pelas conversas, dicas e por você também acreditar e confiar que eu conseguiria fazer um bom trabalho.

Aos demais professores e funcionários da Faculdade de Biociências da PUCRS, que me acompanharam nesta caminhada de estudos.

Ao meu marido, Estevam Luis Cruz da Silva, agradeço pela troca de experiências e ensinamentos, carinho, compreensão, dedicação e companheirismo do início ao fim dessa etapa.

Em especial, aos meus pais, Francisco de Aguiar Picanço e Sílvia Mara Bentes Picanço, agradeço pelo amor, carinho e dedicação que recebi, estando presente ou não, durante todos esses anos; pelos momentos em que a saudade teve que ser guardada e pelo exemplo de caráter e valores que levarei para o resto da vida.

*“The only way to do great work*

*is to love what you do.”*

*Steve Jobs*

**RESUMO**

A genotipagem de *short tanden repeats* ( STRs ) é amplamente utilizada na análise de DNA forense, contudo eventuais ocorrências de trialelias podem reduzir o poder da análise. Alelos do lócus de STR *TPOX* tem de 6-14 números diferentes do motivo de repetição em tandem de quatro nucleotídeos (AATG). Apesar dos genótipos tri-alélicos sejam geralmente raro, o padrão tri-alélico do *TPOX* tem uma alta freqüência, variando amplamente entre as populações. Apesar disso, há poucos relatos precisos para divulgar a natureza do terceiro alelo do *TPOX*. Neste trabalho nós realizamos dois estudos: No primeiro, apresentamos os dados obtidos a partir de 45 indivíduos pertencentes à mesma linhagem, em que houve casos de genótipos tri-alélicos do *TPOX*. Os indivíduos foram aparentemente saudável, com um desenvolvimento biológico normal. Percebemos seis casos tri-alélicos nesta família, e todos eles foram em mulheres. A análise do cariótipo mostrou nenhuma ocorrência de trissomia parcial 2p. Todos os casos tri-alélicos tinham o genótipo 8-10-11, provavelmente devido a três cópias da seqüência do STR *TPOX* em todas as células (Padrão tri-alélico tipo 2). Com base nos dados anteriores, assumimos o alelo 10 como sendo o terceiro alelo do *TPOX*. As análises do pedigree mostraram evidências de que o alelo extra do *TPOX* foi o alelo 10, é localizado distante do lócus principal do *TPOX*, e que existe um potencial de ligação entre o alelo- extra-10 do *TPOX* com um marcador em Xq. Este foi o primeiro estudo que incluiu uma grande análise do pedigree, a fim de compreender a natureza do padrão tri-alélico do *TPOX*. No segundo estudo, investigamos se há um único terceiro extra-alelo no padrão tri-alélico do *TPOX*, o que é ele, e onde está. Nós pesquisamos por indivíduos tri-alélicos no lócus *TPOX* em 75.113 famílias brasileiras. Considerando apenas a geração parental (mãe + pai) tivemos 150.226 indivíduos não relacionados avaliados. Desse total, encontramos 88 indivíduos não relacionados com o padrão tri-alélico no lócus *TPOX* (0,06%, 88/150, 226). Dezessete por cento destes indivíduos (15/88) apresentaram heterozigoses com desbalanço de picos, que descrevemos como uma categoria derivada do padrão tri-alélico de Clayton Tipo 2 (um pico mais alto de homozigoto dose dupla, mais um pico de tamanho regular). Neste trabalho, apresentamos dados detalhados de 66 trios (mãe + pai + filho; com verdadeiras relações biológicas) onde o padrão tri-alélico foi observado na mãe ou no pai. Em 39 destas famílias (39/66; 59%) o terceiro alelo-extra do *TPOX* foi transmitido tanto a partir da mãe ou do pai para a criança. Nossas evidências indicaram o alelo 10 como sendo o terceiro alelo-extra do *TPOX*, e ele está no cromossoma X. Os dados apresentados, os quais suportam as hipóteses anteriores, melhoram o entedimento sobre o padrão tri-alélico do lócus *TPOX* do CODIS permitindo o uso do perfil *TPOX* em análises forense, mesmo quando com o padrão tri-alélico. Essa avaliação está agora disponível para diferentes aplicações forenses.

**PALAVRAS CHAVE:** Padrão Tri-alélico, lócus *TPOX,* Perfil de STR

**ABSTRACT**

Genotyping of polymorphic short tandem repeats (STRs) loci is widely used in forensic *DNA* analysis. *STR* loci eventually present tri-allelic pattern as a genotyping irregularity and, in that situation, the doubt about the tri-allele locus calculation can reduce the analysis strength. Alleles at the *TPOX* STR locus have 6–14 different numbers of a four-nucleotide (AATG) repeat motif arranged in tandem. Although tri-allelic genotypes are generally rare, the *TPOX* tri-allelic pattern has a higher frequency, varying widely among populations. Despite this, there are few accurate reports to disclose the nature of the *TPOX* third allele. In this work we performed two studies: In the first one we present data obtained from 45 individuals belonging to the same pedigree, in which there were cases of tri-allelic *TPOX* genotypes. The subjects were apparently healthy with a normal biological development. We noticed six tri-allelic cases in this family, and all of them were women. Karyotype analysis showed no occurrence of partial 2p trisomy. All the tri-allelic cases had the genotype 8–10–11, probably due to three copies of the *TPOX* STR sequence in all cells (Type 2 tri-allelic pattern). Based on previous data we assumed an allele 10 as the *TPOX* third allele. The pedigree analyses show evidence that the *TPOX* extra allele was an allele 10, it is placed far from the main *TPOX* locus, and that there is a potential linkage of the *TPOX* extra-allele-10 with one marker on Xq. This was the first study that included a large pedigree analysis in order to understand the nature *TPOX* tri-allelic pattern. In the second study, we investigate whether there is a single third-extra allele in the *TPOX* tri-allelic pattern, what it is, and where it is. We looked for *TPOX* tri-allelic subjects in 75,113 Brazilian families. Considering only the parental generation (mother+father) we had 150,226 unrelated subjects evaluated. From this total, we found 88 unrelated subjects with tri-allelic pattern in the *TPOX* locus (0.06%; 88/150,226). Seventeen percent of these subjects (15/88) presented heterozygosis with peak imbalance, which we describe as a derived category to Clayton's Type 2 tri-allelic pattern (a higher peak of double dose homozygote plus a regular sized peak). In this paper we presented detailed data from 66 trios (mother+father+child; with true biological relationships) where the tri-allelic pattern was observed in the mother or in the father. In 39 of these families (39/66; 59%) the third-extra *TPOX* allele was transmitted either from the mother or from the father to the child. Our evidence indicated an allele 10 as the third-extra *TPOX* allele, and it is on the X chromosome. The present data, which support the previous hypothesis, improve the knowledge about tri-allelic pattern of *TPOX* CODIS' locus allowing the use of *TPOX* profile in forensic analyses even when with tri-allelic pattern. This evaluation is now available for different forensic applications.

**KEYWORDS:** Tri-allelic pattern, Third-extra allele, *TPOX* STR profile

**LISTA DE ABREVIATURAS**

| | |
|---|---|
| °C | graus Celsius |
| *TPOX* | Thyroid Peroxidase |
| CODIS | Combined DNA Index System |
| DNA | ácido desoxirribonucléico |
| FBI | Federal Bureau of Investigation |
| pb | par de bases |
| PCR | Polymerase Chain Reaction |
| SNP | Single Nucleotide Polymorphism |
| STR | Short Tandem Repeat |
| STRbase | Short Tandem Repeat Internet DataBase |
| PHR | Peak Height Ratio |

**SUMÁRIO**

# CAPITULO 1

## 1 REFERENCIAL TEÓRICO

### ALELOS E TRIALELIA: DEFINIÇÃO E CATEGORIAS

De acordo com a nomenclatura oriunda da genética clássica, a definição de alelo é "uma forma alternativa de um gene" (Lewis, 2004). No entanto, a genética molecular também considera como alelo "uma de várias formas alternativas de um gene", ou "sequência de DNA em uma posição específica do cromossomo (lócus)" (Strachan & Read, 2002). Assim, no intuito de evitar interpretações paralelas, definimos alelo nesse trabalho como sendo "uma sequência de DNA em uma posição específica do cromossomo (lócus)".

A ocorrência de dois alelos em um indivíduo diplóide é baseada na premissa de que um deles tenha sido herdado da mãe e o outro do pai. Através da tecnologia molecular da Reação em Cadeia da Polimerase (PCR) é possível que sejam identificados e amplificados ambos os alelos herdados por um indivíduo. Após a PCR, seguida pela leitura em sistema de eletroforese de gel convencional, os alelos com variação de tamanho são observados como duas 'bandas' (terminologia usual originada das leituras de géis com fragmentos de DNA) em amostras de heterozigotos, ou como uma única banda em amostras de indivíduos homozigotos (Butler, 2001). Na nomenclatura derivada da leitura automatizada de fragmentos de DNA, através da eletroforese do tipo capilar, os alelos com variação de tamanho são visualizados como dois picos em heterozigotos e como um único pico nos indivíduos homozigotos. Excluídos artefatos, as bandas ou os picos extras mostrariam a presença de alelos extras (no caso, um terceiro alelo, como mostrado na Figura 1).

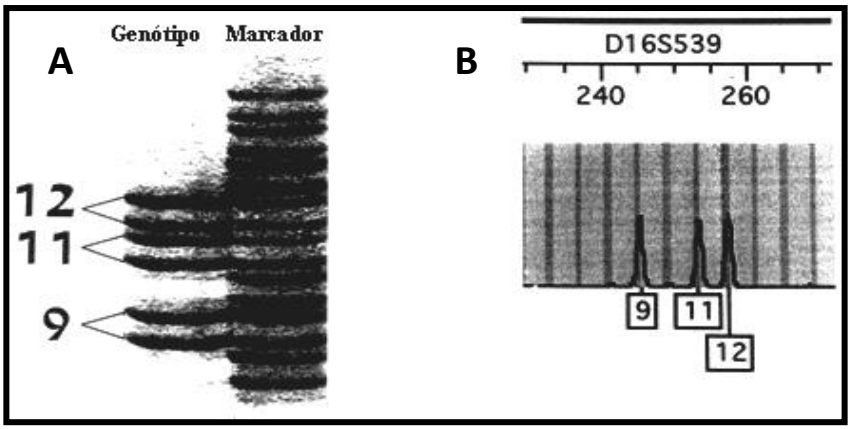



**Figura 1:** Visualização de como se apresenta o padrão tri-alélico do lócus D16S539, visualizado em eletroforese em gel de poliacrilamida (A) e eletroforese capilar (B). Adaptado de Zamir et al. (2002).

Para casos onde ocorra trialelia, foi sugerido o uso da expressão "*three-banded*" para se referir aos indivíduos que apresentam uma terceira banda (ou *3rd-band*) em determinado lócus (Crouse *et al*., 1999; Zamir *et al.*, 2002). Mutações causadoras de um terceiro alelo podem ocorrer de diferentes maneiras e passam, portanto, a serem identificadas como uma terceira banda, ou um terceiro pico nas análises moleculares.

São apresentadas duas explicações para a ocorrência da trialelia na formação do individuo: 1- duplicações genéticas em tandem ou dispersas de uma pequena região de cromossomo (trissomias parciais); ou 2- uma segregação incorreta resultado de uma não disjunção cromossômica mitótica ou meiótica que conduz a uma trissomia completa. Pelos mesmos dois motivos citados, pode ocorrer trialelia em apenas alguns grupos celulares (e não em todas as células do organismo) decorrente de mosaicismos e/ou quimerismos (Brinkmann *et al*., 1998; Zamir *et al*., 2002).

Para a leitura automatizada de fragmentos de DNA, através da eletroforese do tipo capilar, Clayton *et al.* (2004) caracterizam o padrão tri-alélico em duas diferentes categorias (Figura 2): A- Padrão Tipo 1: consistindo em três alelos que apresentam sinais de intensidades desiguais; e B- Padrão Tipo 2: quando os três alelos apresentam a mesma intensidade de sinal. Ele relacionou o padrão Tipo 1 como sendo decorrente de mutações presentes em apenas algumas células somáticas, e o do Tipo 2 como sendo decorrente de rearranjos cromossômicos presentes em todas as células. O padrão tri-alélico Tipo 1, que seja decorrente de uma duplicação cromossômica, pode ter efeito fenotípico, estando relacionado com síndromes clínicas graves (Lukka *et al.,* 2006).

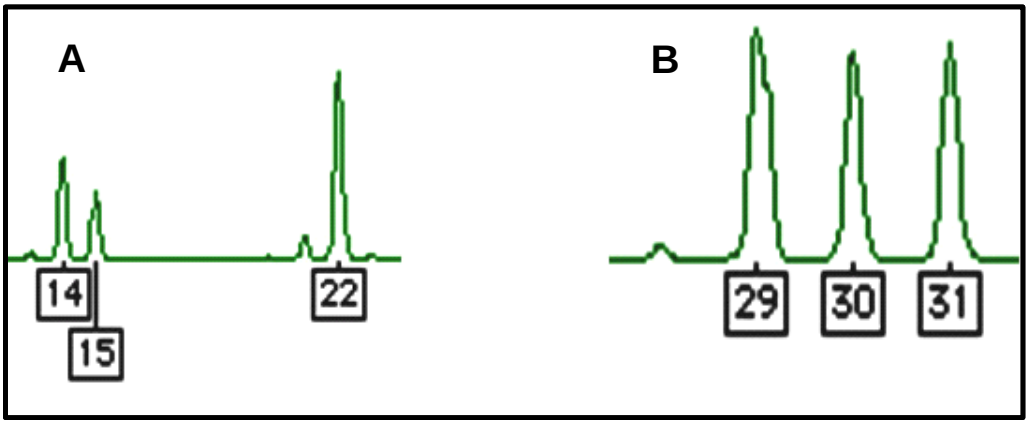



**Figura 2:** Representação esquemática do perfil dos dois tipos de padrão tri-alélico quando aparecem na análise de eletroforese capilar. A- Padrão Tipo 1: três alelos que apresentam sinais de intensidades desiguais; B- Padrão Tipo 2: três alelos com a mesma intensidade de sinal. Retirado do site http://www.cstl.nist. gov/biotech/*STR*base/tri_tab.htm acessado em dezembro de 2013.

**PADRÃO TRI-ALÉLICO DO LÓCUS DA PEROXIDASE DA TIREÓIDE (LÓCUS *TPOX*)**

No interior do íntron 10 do gene que codifica para a peroxidase da tireóide humana (*TPOX*), se encontra um marcador do tipo STR (Anker *et al.*, 1992). Os marcadores do tipo STRs, também chamados de microssatélites, são sequências curtas de nucleotídeos, geralmente de 2 a 6 pares de base (pb) que aparecem repetidas até dezenas de vezes e que geram múltiplos alelos, os quais são identificados pelo número de repetições que apresentam (Butler, 2005).

O lócus *TPOX* é um exemplo de onde são observados casos de trialelia. O tipo mais frequente de padrão tri-alélico associado ao lócus *TPOX* é, aparentemente, do Tipo 2, no qual os três alelos apresentam a mesma intensidade de sinal (Crouse *et al.*, 1999). O Tipo 2, porém, foi descrito como sendo decorrente de rearranjos cromossômicos presentes em todas as células. No entanto, nos indivíduos onde se verifica a trialelia, nunca foi investigado se o terceiro alelo do lócus *TPOX* decorrente de rearranjos cromossômicos, de uma repetição (duplicação) de um gene *TPOX* completo que estaria extra no genoma, ou apenas trata-se de uma pequena sequência de DNA repetida que é positiva para o anelamento dos *primers TPOX* presente no cromossomo 2 ou dispersa em qualquer outro cromossomo. A última opção é a mais plausível, dado que uma alteração de dose do gene da *TPOX* (a duplicação do gene *TPOX* completo) seria acompanhada por características fenotípicas clínicas, próprias da ação da tireóide peroxidase (Endo *et al.*, 1995).

No lócus *TPOX* há diferentes números de repetições em tandem dos tetranucleotídeos $[AATG]_n$, repetidos de seis até 14 vezes (alelo 6 a alelo 14) (Anker *et al.*, 1992). Noventa por cento dos alelos encontrados no lócus *TPOX* têm 8, 9 ou 11 repetições da unidade de quatro nucleotídeos (alelo 8, alelo 9 e

alelo 11) (Nata *et al.*, 1999). Por estar incluído no CODIS (o já citado sistema de identificação molecular humana da base de dados dos Estados Unidos da América) o lócus *TPOX* é amplamente usado para testes moleculares de identificação pessoal e de paternidade (Lane, 2008).

Segundo a base de dados de STR do CODIS (*STR*base: http://www. cstl.nist.gov/biotech/*STR*base), uma frequência de genótipos tri-alélicos pode ser esperada quando se realizam estudos populacionais de larga escala. Isto é, dado que tem sido percebida a presença de trialelia em diferentes lócus, é natural que se identifique um terceiro e raro alelo para os marcadores STR, independente da população em estudo, quando se trabalha conjuntamente com milhões de análises. Embora os genótipos tri-alélicos sejam, de fato, raros, a base STRbase demonstra que estes são mais frequentes nos lócus *FGA, D18S51, VWA, D21S11* e *TPOX* quando comparados aos outros lócus do CODIS ( Tabela 1).

**Tabela 1:** Representação dos 13 lócus do CODIS e o número de combinações diferentes já observadas nas repetições de STRs de seus respectivos padrões tri-alélicos (modificado do site <http://www.cstl.nist.gov/div831/*STR*base/tri_tab.htm> acessado em 15 de janeiro de 2014).

| **Core STR Lócus** | **Padrões Tri-alélicos**[1] |
|---|---|
| *CSF1PO* | (8) |
| *FGA* | (35) |
| *TH01* | (4) |
| *TPOX* | (18) |
| *VWA* | (24) |
| *D3S1358* | (11) |
| *D5S818* | (8) |
| *D7S820* | (18) |
| *D8S1179* | (19) |
| *D13S317* | (15) |
| *D16S539* | (12) |
| *D18S51* | (35) |
| *D21S11* | (24) |

[1] Entre parênteses está o número de combinações diferentes das repetições de STRs dos padrões tri-alélicos encontradas para cada lócus STR.

Há casos onde a trialelia do lócus *TPOX* é observada devido à ocorrência de aneuploidias, como as do tipo trissomia parciais do cromossomo 2 (as trissomias totais do cromossomo 2 não são compatíveis com a vida). No entanto, como já foi mencionado, há a grande possibilidade de o terceiro alelo ser apenas uma sequência repetida de DNA inserida nas proximidades do lócus *TPOX* no cromossomo 2 ou, ainda, inserida em outro local do genoma, como no cromossomo X seguindo a sugestão de Lane (2008).

**TRISSOMIAS DO 2p**

A síndrome da trissomia parcial do 2p foi descrita pela primeira vez por Francke (1978). As características desta síndrome incluem retardo no crescimento pré e pós-natal, deficiência psicomotora, microcefalia, testa proeminente, hipertelorismo, estrabismo, miopia, micrognatia, doenças congênitas do coração e outros fenótipos mais raros (Aviram-Goldring *et al.*, 2000) (Tabela 2). Embora aproximadamente 60 casos da trissomia parcial do 2p tenham sido registrados até hoje (Siffroi *et al*., 1994; Lurie *et al.*, 1995; Chen *et al*., 1996; Mérgarbané *et al*., 1997; Winsor *et al*., 1997; Magge *et al*., 1998), a maioria deles foram produtos de translocações desbalanceadas (Aviram-Goldring *et al.*, 2000). Foi também registrado que a trissomia parcial do 2p pode ser compatível com um desenvolvimento biológico normal (Mérgarbané *et al*., 1997; Al-Saffar *et al.*, 2000; Bakker *et al*., 2001).

**Tabela 2:** Sumário dos casos de Trissomia parcial de 2p e seus fenótipos, segundo Mégarbané *et al.*, 1997; Aviram-Goldring *et al.*, 2000; Al-Saffar *et al.*, 2000.

| **Trissomia 2p** | **p13-pter** | **p21-p25** | **p21-p25** | **p25.1-p25.3** | **p14-p23** | **p21-p24.2** | **p12-p21** | **p13-p24** | **p13-p21** | **p13.1-p21** | **p21-p22** | **p21-p22** | **p21-p23** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Anomalias gatrointestinais | NR | NR | NR | NR | - | + | NR | + | - | + | - | - | + |
| Anomalias oculares | NR | NR | NR | NR | - | + | NR | + | + | NR | + | - | + |
| Anomalidades genitais | - | - | - | - | + | - | NR | NR | + | NR | NR | - | NR |
| Anomalidades urogenitais | NR | NR | NR | NR | + | - | NR | NR | + | + | NR | - | + |
| Clavículas curtas | NR | NR | NR | NR | NR | NR | + | NR | NR | NR | NR | NR | NR |
| Defeitos cardíacos | + | + | - | NR | + | - | + | + | - | + | - | - | + |
| Hérnia diafragmática | - | - | + | - | - | - | NR | NR | NR | NR | NR | NR | NR |
| Hipertelorismo | + | + | + | + | + | + | NR | + | + | NR | - | + | + |
| Hipoplasia maxilar | + | NR | - | - | - | + | NR | NR | NR | NR | NR | NR | NR |
| Hipoplasia pulmonar | + | + | + | - | - | - | NR | NR | - | + | - | - | NR |
| Hipotonia | - | - | - | - | - | + | NR | NR | - | NR | - | + | NR |
| Implantação baixa da orelha | NR | NR | NR | NR | + | + | + | - | NR | + | - | + | + |
| Longos dedos das mãos e dos pés | NR | + | + | + | - | - | NR | NR | - | NR | NR | + | + |
| Mandibulas pequenas | NR | - | NR | NR | + | + | + | NR | NR | NR | NR | NR | + |
| Microcefalia | NR | - | - | NR | + | + | NR | + | + | - | - | + | + |
| Moleira aberta | NR | NR | + | NR | + | + | NR | NR | - | NR | NR | NR | NR |
| Osso occipital achatado | - | - | NR | - | - | - | NR | NR | NR | NR | NR | + | NR |
| Pescoço curto | + | NR | - | - | - | + | NR | NR | + | NR | NR | NR | NR |
| Ponte nasal achatada | + | + | + | - | + | + | + | + | + | NR | - | + | + |
| Retardo psicomotor | NR | NR | NR | - | + | + | + | - | + | NR | + | + | + |

Legenda: p: braço curto do cromossomo 2; ter: região terminal do braço curto do cromossomo 2; +: presença; -: ausência; NR: não registrado.

**O GENE QUE CODIFICA PARA A PEROXIDASE DA TIREÓIDE**

A peroxidase da tireóide humana (*TPOX*) é uma proteína glicosilada de membrana plasmática, localizada na membrana apical das células foliculares da tireóide (Park & Chatterjee, 2005), a qual está envolvida na biossíntese do hormônio da tireóide. A *TPOX* humana tem 933 aminoácidos com um domínio transmembrana e cinco potenciais sítios de glicosilação (Libert *et al.*, 1987) e desempenha uma função importante na glândula da tireóide: ela catalisa duas reações da síntese de hormônio da tireóide, isto é, a iodinação dos resíduos de tirosina na tiroglobulina e a formação de éster fenoxi entre os pares iodinados para gerar os hormônios da tireóide Triiodotironina (T3) e (T4) Tiroxina (Endo *et al.*, 1995). O gene que codifica para a peroxidase da tireóide humana [lócus 2p25 (Endo *et al.*, 1995); GenBank NC_000002.10; OMIM 606765; GeneID 7173], possui uma extensão gênica de 150 kilobases (kb) que compreende uma região de 17 éxons e 16 íntrons (Kimura *et al.*, 1989) (Figura 3). O tamanho total do mRNA codificado a partir do *TPOX* é de 3kb (Kimura *et al.*, 1987; Seto *et al.*, 1987).

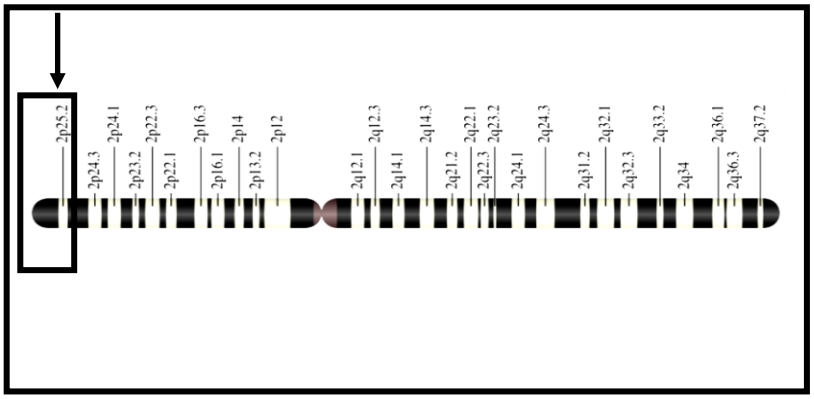



**Figura 3:** Ideograma do cromossomo 2 com o lócus gênico *TPOX*, na posição 2p25.2 indicado na seta (NCBI:Map Viewer, em http://www.ncbi.nlm.nih.gov, acessado em dezembro de 2013).

**MARCADORES COMERCIAIS PARA IDENTIFICAÇÃO MOLECULAR HUMANA**

Segundo Kimpton *et al*. (1993), marcadores de STR (*short tandem repeats*) foram descritos pela primeira vez como ferramentas efetivas para teste de identificação humana no início da década de 1990 (Edwards *et al.*, 1991; Edwards *et al.*, 1992). Logo, o Serviço de Ciência Forense dos Estados Unidos da América iniciou uma pesquisa por novos lócus e estudos de variações populacionais com um número de STRs candidatos (Kimpton *et al.*, 1993). Já se passou mais de uma década desde que os 13 marcadores genéticos de STRs, no ano 1997, foram então selecionados para formar o *Combined DNA Index System* (CODIS) do *Federal Bureau of Investigation* dos Estados Unidos da América (FBI-USA). Devido ao seu uso na Base de Dados Nacional de DNA dos Estados Unidos da América, bem como por outras bases de dados da justiça criminal em todo o mundo, estes lócus de STR são os mais usados para caracterizar indivíduos humanos (Gill, 2002; Butler, 2005).

Os 13 lócus do CODIS usados pela base de dados dos Estados Unidos da América são: *CSF1PO*, *FGA*, *TH01*, *TPOX* (ou *TPO*), *VWA*, *D3S1358*, *D5S818*, *D7S820*, *D8S1179*, *D13S317*, *D16S539*, *D18S51*, e *D21S11* (Budowle *et al.*, 1998) (Figura 4).

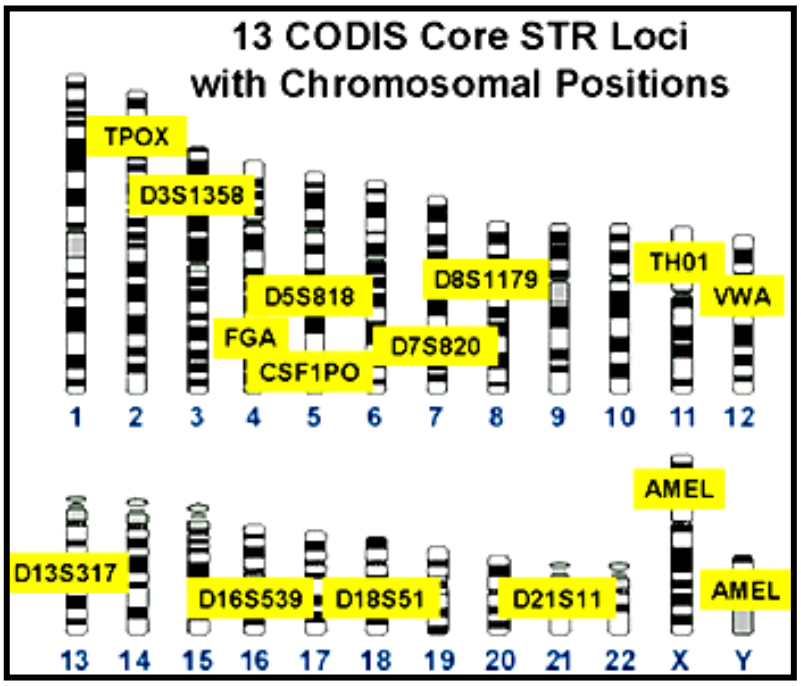



**Figura 4:** Marcadores de identificação do CODIS, e suas respectivas posições nos cromossomos. Retirado do site http://www.cstl.nist.gov/biotech/*STR*base/fbicore.htm, acessado em dezembro de 2013.

O Reino Unido e grande parte da Europa incluem marcadores adicionais como *D2S1338* e *D19S433* junto com oito lócus que sobrepõem os lócus *FGA*, *TH01*, *VWA*, *D3S1358*, *D8S1179*, *D16S539*, *D18S51* e *D21S11*. A análise desses lócus para os testes de identificação molecular humana e de paternidade tem se tornado comum em diversos países devido ao seu fácil uso na forma de *kits* comerciais. Isso permitiu, portanto, o intercâmbio de dados entre os países (Butler, 2006).

Em Maio de 2010, o FBI formou um grupo de trabalho "CODIS Core Loci" para revisar os 13 marcadores existentes do CODIS e determinar se a adição de novos marcadores de STR seria vantajoso. Tendo uma base de dados com quase dez milhões de perfis, a decisão foi manter os mesmos 13 lócus do CODIS, dividindo-os em grupo A e grupo B, e adicionando novos marcadores, aumentando assim o poder discriminatório dos conjuntos destes marcadores quando utilizados em um novo STR *kit* de tipagem de DNA.

A implementação e a validação deste novos kits estão em processo de desenvolvimento e o grupo de trabalho do "CODIS Core Loci"destaca que antes de finalizar esta expansão do CODIS, versões de protótipos dos novos painéis de STRs serão fabricadas e testadas em laboratórios, para uma posterior avaliação e deteminação do novo conjunto de marcadores de STR "CODIS Core Loci" a ser utilizado pela comunidade Forense (Hares, 2012).

## FREQUÊNCIA DO PADRÃO TRI-ALÉLICO DO GENE DA PEROXIDASE DA TIREÓIDE (*TPOX*)

A frequência do padrão tri-alélico no lócus *TPOX* varia muito entre as diferentes populações. Estudos populacionais baseados nos dados de rotina de testes de paternidade, realizados em 32.800 indivíduos da Bósnia, Kosovo e Sérvia, encontraram somente um caso de genótipo tri-alélico (8, 11, 12) no lócus *TPOX* (Huel et al., 2007). Em contraste, Crouse et al. (1999) registraram 18 casos de genótipos tri-alélicos no lócus *TPOX*, em 10.000 indivíduos escolhidos aleatoriamente no Estado do Alabama, USA, correspondendo a uma frequência de 0,18% da população a apresentar essa característica. Paralelamente, cerca de 2,4% dos nativos sul-africanos têm três ao invés de dois alelos no *TPOX* (Lane, 2008). Esse último estudo revelou que o alelo extra era, em geral, o alelo 10 e que ele segregava independentemente daqueles do lócus principal do *TPOX* (localizado no cromossomo 2). Foi ainda estimado que o terceiro alelo estaria duas vezes mais frequente em mulheres do que em homens, o que permitiu sugerir que o alelo extra estivesse no cromossomo X. Essa sugestão foi suportada por uma análise em um pequeno grupo de sujeitos, onde homens com três alelos *TPOX* transmitem constantemente dois alelos para suas filhas e somente um para seus filhos homens (Lane, 2008).

O padrão tri-alélico do lócus *TPOX* possui uma variabilidade de combinações de genótipos encontrados aleatoriamente durante as análises forenses, nos casos de identificação molecular de pessoas ou durante os testes moleculares de paternidade. A Tabela 3 relaciona os diferentes genótipos tri-alélicos do lócus *TPOX* já identificados, registrando o padrão tri-alélico e o número de vezes em que o genótipo foi detectado pela análise dos genótipos de STRs. Em destaque está o padrão mais frequente para o lócus *TPOX* (8, 10, 11), o qual já foi identificado num total de 30 casos independentes. Esses dados estão disponíveis na Base de Dados de STRs:<http://www.cstl.nist.gov/div831/STRbase/> e são atualizados frequentemente pelo Instituto Nacional da Justiça dos Estados Unidos da América. No Brasil, os casos relatados nesta base de dados de STR foram: 10, sendo seis no Estado de São Paulo e quatro no Estado do Rio de Janeiro.

**Tabela 3:** Representação dos genótipos tri-alélicos do lócus *TPOX* e suas respectivas frequências. Em destaque, está o genótipo mais frequente. (Modificado do site <http://www.cstl.nist.gov/ STRbase/var_*TPOX*.htm#Tri> acessado em 15 de janeiro de 2014).

| **Genótipos tri-alélicos *TPOX*** | **Número de genótipos[2]** |
|---|---|
| 6, 8, 10 | 4 |
| 6, 9, 10 | 7 |
| 6, 10, 11 | 5 |
| 6, 10,12 | 2 |
| 7, 8, 10 | 2 |
| 7, 9, 10 | 2 |
| 7, 10, 11 | 3 |
| 8, 9, 10 | 20 |
| 8, 9, 11 | 5 |
| 8, 10, 11 | 30 |
| 8, 10, 12 | 6 |
| 8, 11, 12 | 3 |
| 8, 11, 14.3 | 1 |
| 9, 10, 11 | 15 |
| 9, 10, 12 | 2 |
| 9, 11, 12 | 1 |
| 10, 10, 11 | 1 |
| 10, 11, 12 | 4 |

[2] Representa o número de vezes que o genótipo foi detectado pela análise dos genótipos de STR*s* em testes forenses ou até mesmo os desenvolvidos em laboratório (*In-house multiplex*).

## O ESTUDO DA FAMÍLIA LTMV E DE OUTRAS FAMILIAS COM CASOS DE TRIALELIA NO LÓCUS *TPOX*

Em 2004 foi realizada a confirmação de identidade de uma vítima de morte súbita, proveniente de Belém do Pará e encaminhada ao Estado do Rio Grande do Sul embalsamada e em caixão lacrado. No Rio Grande do Sul, Estado de origem da vítima, como rotina do Laboratório de DNA do Instituto-Geral de Perícias (IGP), foram coletadas amostras de sangue total dos possíveis sujeitos aparentados à vítima (esposa, filha e filho), com o objetivo de identificá-la e, finalmente, dar os encaminhamentos oficiais de óbito e sepultamento. A vítima foi confirmada sendo familiar dos sujeitos testados e, ao observar os perfis genéticos dos membros da família, foi verificado paralela e coincidentemente a presença de trialelia para o lócus *TPOX* nas amostras de DNA da esposa e da filha da vítima. Essas duas mulheres apresentavam desenvolvimento biológico compatível com a normalidade. Baseados nos dados preliminares da família LTMV, e nas publicações apresentadas no referencial teórico deste trabalho, sugeriu-se que o terceiro alelo do lócus *TPOX* não fosse um terceiro segmento completo de DNA que codificaria para a proteína *TPOX*, e sim apenas uma duplicação de um segmento do lócus *TPOX* localizado em qualquer ponto do genoma. Os integrantes da família LTMV demonstraram interesse em desvendar a natureza do terceiro alelo do lócus *TPOX* participando de uma investigação que foi proposta por nosso grupo. A primeira parte do estudo incluiu o levantamento de dados pessoais de toda a Família LTMV, para posteriores hipóteses das questões a serem avaliadas na pesquisa. A segunda parte da investigação abordou a coleta de material biológico para a realização das análises moleculares da família em questão usando marcadores moleculares. Foram coletados dados de saúde e moleculares de mais de 40 indivíduos pertencentes a família LTMV.

Com a finalidade de identificar o padrão de transmissão do terceiro alelo do lócus *TPOX*, nossa pesquisa buscou em centros de investigação de paternidade, dados sobre trialelia avaliando mais de 75 mil famílias atendidas nos Estados de São Paulo e Rio Grande do Sul.

## 2 OBJETIVOS

### Objetivo Geral

Estudar as trialelias no marcador forense *TPOX*, identificar qual é o terceiro alelo e indicar a sua possível localização.

### Objetivos Específicos

1. Analisar os membros da família LTMV em relação a trialelia do lócus *TPOX* e ao padrão de herança do terceiro alelo.
2. Analisar casos de trialelia do lócus *TPOX* e o padrão de herança do terceiro alelo em outras famílias brasileiras.
3. Comparar as frequências alélicas do lócus *TPOX* entre sujeitos bi-alélicos *versus* tri-alélicos a fim de sugerir qual seria o terceiro alelo (se é um único e quantas repetições STR tem).
4. Indicar e/ou sugerir qual é a possível localização do terceiro alelo do lócus *TPOX*.

# CAPITULO 2

Gene 535 (2014) 353–358

Contents lists available at ScienceDirect

Gene

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

# Tri-allelic pattern at the TPOX locus: A familial study

Juliane Bentes Picanço [a], Paulo Eduardo Raimann [a], Giorgio Adriano Paskulin [b], Luís Alvarez [c], António Amorim [c,d], Sidney Emanuel Batista dos Santos [e], Clarice Sampaio Alho [a,*]

[a] Faculdade de Biociências, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil
[b] Universidade Federal de Ciências da Saúde de Porto Alegre — UFCSPA e Complexo Hospitalar Santa Casa de Porto Alegre — CHSCPA, Porto Alegre, RS, Brazil
[c] IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal
[d] Department of Biology, Faculty of Sciences of the University of Porto, Porto, Portugal
[e] Universidade Federal do Pará, Centro de Ciências Biológicas, Laboratório de Genética Humana e Médica, Belém, Pará, Brazil

ARTICLE INFO

*Article history:*
Accepted 10 October 2013
Available online 18 October 2013

*Keywords:*
Family pedigree
STR
Tri-allelic pattern
TPOX

ABSTRACT

Alleles at the TPOX STR locus have 6–14 different numbers of a four-nucleotide (AATG) repeat motif arranged in tandem. Although tri-allelic genotypes are generally rare, the TPOX tri-allelic pattern has a higher frequency, varying widely among populations. Despite this, there are few accurate reports to disclose the nature of the TPOX third allele. In this work we present data obtained from 45 individuals belonging to the same pedigree, in which there are cases of tri-allelic TPOX genotypes. The subjects were apparently healthy with a normal biological development. We noticed six tri-allelic cases in this family, and all of them were women. Karyotype analysis showed no occurrence of partial 2p trisomy. All the tri-allelic cases had the genotype 8–10–11, probably due to three copies of the TPOX STR sequence in all cells (Type 2 tri-allelic pattern). Based on previous data we assumed the allele 10 as the TPOX third allele. The pedigree analyses show evidences that the TPOX extra-allele was the allele10, it is placed far from the main TPOX locus, and that there is a potential linkage of the TPOX extra-allele-10 with Xq. This was the first study that included a large pedigree analysis in order to understand the nature TPOX tri-allelic pattern.



## 1. Introduction

A forensic paternity test follows simple Mendelian inheritance where the child inherits one allele from the mother and another from the father at each locus. Rare events, such as point mutations in primer binding regions, slippage mutations, and other events such as gene conversion and copy number variations (CNVs) (Freeman et al., 2006) can occasionally cause an aberrant result, including an abnormal number of alleles, which seemingly break the rules of Mendelian inheritance (Lukka et al., 2006). Clayton et al. (2004) have distinguished two types of tri-allelic pattern. Type 1: when after PCR amplification two alleles have different intensity of the third allele and Type 2: when the three peaks have similar intensity. Type 1 is believed to be the result of a mutation in an early somatic cell while Type 2 is thought to represent a constitutional chromosomal rearrangement. Although tri-allelic genotypes are generally rare, data presented on the STRBase website (http://www.cstl.nist.gov/biotech/strbase) indicate that short tandem repeats (STR) tri-allelic genotypes can be unusually frequent in the TPOX human STR locus. Alleles at the TPOX locus have different numbers of a four-nucleotide repeat motif arranged in tandem (Anker et al., 1992). Therefore, TPOX polymorphism is widely used for paternity testing and personal identification, and is one of the FBI's CODIS STR loci (http://www.cstl.nist.gov/strbase/fbicore.htm). The TPOX tri-allelic profiles are due to some duplication of the STR segment and flanking regions recognized by the TPOX primers.

For more than a decade, TPOX tri-allelic genotypes have been reported with a widely varied frequency among human populations. Crouse et al., 1999 reported 18 tri-allelic genotypes in a sample of over 10,000 individuals drawn from Alabama, USA, which is equivalent to a frequency of 0.18% (18/10,000). Later, Huel et al. (2007) typed 32,800 individuals from Bosnia, Kosovo, and Serbia, but found only one subject with a TPOX tri-allelic genotype (1/32,800; 0.003%), and Fridman et al. (2008) analyzing 561 unrelated individuals (410 females and 151 males) from Brazil also observed one single occurrence of a tri-allelic pattern at the TPOX locus (1/561; 0.2%). In contrast, more than 2% of indigenous South Africans exhibit tri-allelic TPOX genotypes: 165 tri-allelic genotypes (116 females and 49 males) were found among a total of 6827 black South Africans (3399 females and 3428 males) (Lane, 2008). And in a larger study from Brazil, Poiares et al. (2010) typing 12,886 unrelated individuals were unable to find any TPOX tri-allelic genotype.

*Abbreviations:* CNVs, copy number variations; STR, short tandem repeat; TPOX, thyroid peroxidase; FBI, Federal Bureau Investigation; CODIS, Combined DNA Index System; DNA, deoxyribonucleic acid; GTG, G-banded chromosome; pb, base pair; PCR, polymerase chain reaction; X-STR, short tandem repeat of X-chromosome; IBGE, Instituto Brasileiro de Geografia e Estatística.

* Corresponding author at: Faculdade de Biociências, Pontifícia Universidade Católica do Rio Grande do Sul. Av. Ipiranga, 6681, 90619-900 Porto Alegre, RS, Brazil. Tel.: +55 51 33203545.
*E-mail address:* csalho@pucrs.br (C.S. Alho).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.10.019

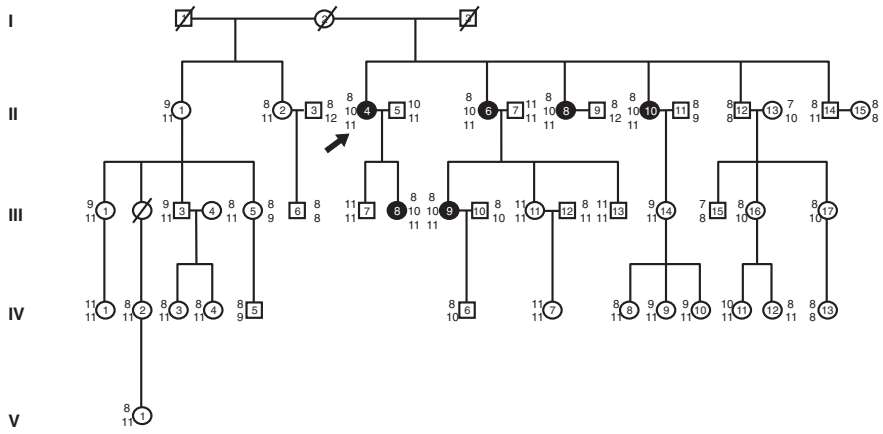



**Fig. 1.** Family pedigree with the DNA profiles of the TPOX locus. Roman numerals show the five generations of this family. Females are represented by circles. Males are represented by squares. Shaded circles indicate women with tri-allelic TPOX genotype. Numbers within the symbols (circles and squares) indicate each subject. Numbers outside the symbols indicate the individual's genotype for the TPOX STR locus.

In the African population (Lane, 2008), Dr. Lane was able to suggest that the TPOX extra allele were the allele 10. Indeed, according to STRBase website (with 102 TPOX tri-allelic reported subjects) the allele 10 is present in 90% of individuals reported with TPOX tri-allelic pattern, and all other alleles were present in lower percentage. In this database, comparing TPOX allele frequencies among bi and tri-allelic subjects, the frequency of the allele 10 in tri-allelic subjects is around five times higher than in bi-allelic subjects. Thus, it was possible to strengthen the hypothesis that the third-extra allele would be the allele 10. Only ten tri-allelic subjects from the STRBase without the allele 10 were reported. All of them had the allele 9 or 11 (genotypes were: 8–9–11; 8–11–12; 8–11–14.3; 9–11–12), which could denote slippage mutation from an ancestral third-extra allele 10.

As pointed out by Clayton et al. (2004), a Type 2 tri-allelic pattern may result from a chromosomal duplication and, in this case, it may be associated with severe clinical syndromes (Lukka et al., 2006). However, the partial maternal isodisomy for chromosome 2p (2pter–2p12) that also produces a tri-allelic pattern might be compatible with a minimal influence on normal development (Bakker et al., 2001). In this context, Muna Al-Saffar et al. (2000), have shown a 7-month-old well-developed girl with the karyotype 46, XX, der(13) t(2;13)(p23; p11.2).ish der(13)(wcp2+) de novo. A chromosome painting strategy confirmed that the additional segment of chromosome 2 was on 13p, resulting in trisomy of 2p23-2pter. Megarbane et al. (1997) also demonstrated that the partial trisomy of 2p was compatible with adulthood. Another possibility is that the third TPOX allele may not be linked to chromosome 2. Lane showed that two thirds of the TPOX tri-allelic adults were females, and TPOX tri-allelic fathers only transmitted the TPOX tri-allelic genotype to their daughters (Lane, 2008). With this evidence, he suggested that the inserted allele was on an X chromosome. This report was enhanced by Díaz et al. (2009).

Despite the substantial frequency of the TPOX tri-allelic pattern, the nature of the third allele is still poorly understood. In this work we present data obtained from 45 individuals of the same family where there were cases of TPOX tri-allelic pattern, and employ these data to investigate the underlying inheritance of the third allele detected at this locus.

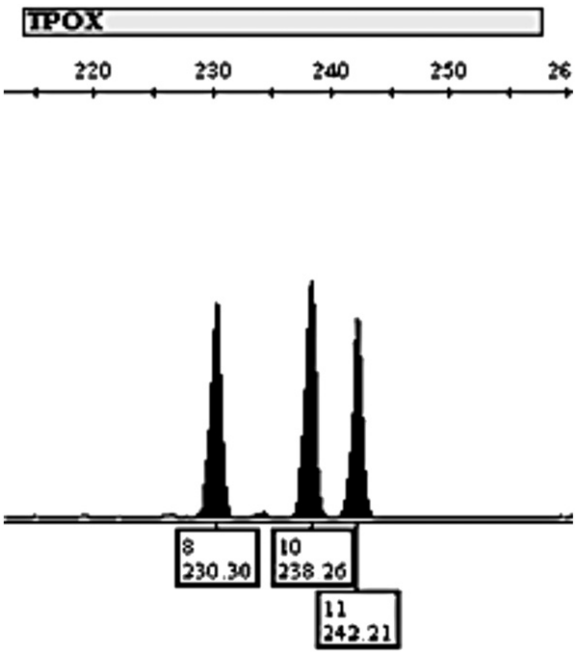



**Fig. 2.** Tri-allelic TPOX genotype as depicted by the ABI GeneMapper3.2 software. The three TPOX STR peaks represent alleles 8, 10, and 11.

## 2. Material and methods

The 45 studied subjects, belonging to the same family pedigree, were born and are living in four different cities in the Rio Grande do Sul State, southern Brazil. The proband (or propositus) was a female called LTMV, she was the first tri-allelic family member who received attention for her genetic characteristic. The DNA profiles of her and her daughter were analyzed in a kinship lab routine. Then, a large number of her relatives were invited to participate in this study. This project was approved by the Research Ethics Committee of Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) (Protocol #09/04688-OfCEP943/29); and the informed written consent and assent to participate were obtained from all subjects or their surrogates.

All subjects participated in an interview on their own residence when they response a questioner about health and development conditions. Data about important medical disorders, syndromic conditions, typical diseases, pathological sickness, abnormal features, atypical appearances, distinctive symptoms, intellectual capacities, and other characteristics were enquired. No medical exams (biochemical or images) were evaluated.

Blood samples were collected on FTA cards and DNA was extracted from blood spots using the manufacturer's protocols. In all 45 subjects, we confirmed kinship between them, by using autosomal STRs (AmpF/STR® Identifiler™ PCR Amplification Kit; Applied Biosystems; Life Technologies, USA), with paternity index values above 10,000. A total of 0.5–1.0 ng of DNA was used to amplify 15 autosomal STR loci (D8S1179,

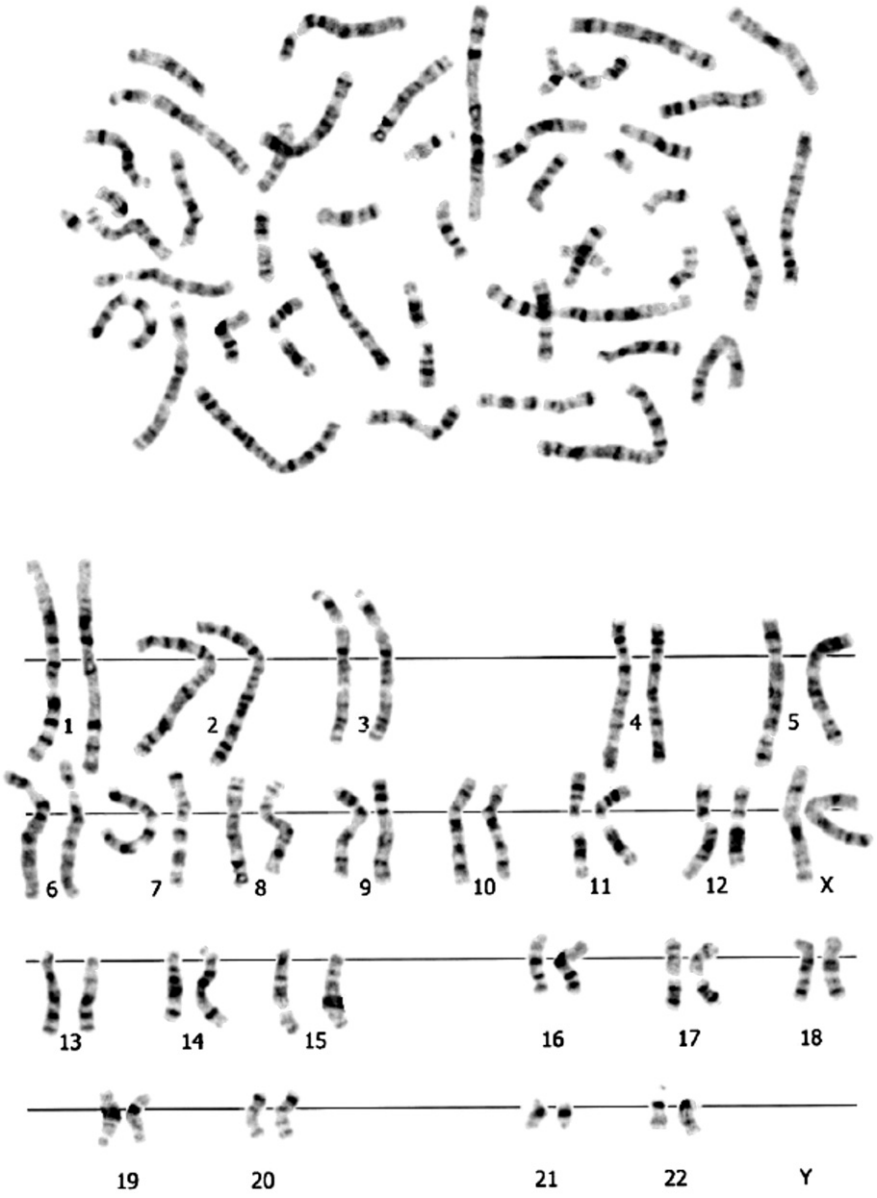



**Fig. 3.** High-resolution karyotype analysis (>550 bands) of lymphocytes from female II-4 stained with Trypsin and Giemsa.

D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, and FGA) with the amplification kit, following the manufacturer's instructions, followed by analysis on a capillary ABI 3130*xl* Genetic Analyzer. All tri-allelic subjects were also tested by PowerPlex®21 System kits, to confirm the tri-allelic pattern in TPOX locus. To avoid the possibility of a false tri-allelic genotype caused by micro-variants (shorter alleles) of the adjacent bins (e.g.: D18S51 — AmpFlSTR® Identifiler® or the FGA- PowerPlex®21), all TPOX tri-allelic subjects had their DNA amplified by the singleplex PCR-based analysis of the TPOX-locus; using primer forward 5′-GCACAGAAC AGGCACTTAGG-3′ and primer reverse 5′-CGCTCAAACGTGAGGTTG-3′ they had the TPOX tri-allelic pattern confirmed.

A high-resolution karyotype (>550 bands) was obtained from lymphocytes through the adapted from Yunis' technique (Yunis, 1981). Briefly, this includes a cell culture for 72 h of lymphocytes stimulated with phytohemagglutinin, synchronization with the use of methotrexate/thymidine and stained with Trypsin and Giemsa, producing G-banded chromosomes (GTG). Twenty five metaphase plates per case were analyzed with the aid of a Zeiss Axioskop microscope.

Based on Lane's (2008) suggestion, we considered the TPOX extra-allele as allele10 and cogitated that it was in X chromosome. To verify if there was linkage disequilibrium between the TPOX extra-allele (allele10) and the X-STR markers, we used a multiplex system that allows the simultaneous analysis of 12 STRs associated to the X chromosome; a multiplex PCR (Loci DXS9895, DXS7132, DXS6800, DXS9898, DXS6789, DXS7133, GATA172D5, DXS7130, HPRTB, GATA 31E08, DXS7423, and DXS10011) was performed based on Ribeiro-Rodrigues et al. (2008, 2009).

## 3. Results and discussion

The 45 studied subjects were apparently healthy and normally developed. Fig. 1 shows the family pedigree with the DNA profiles of the TPOX STR locus. The proband on this pedigree was the female LTMV (II-4) and it is noted with an arrow. In order to show the Clayton category of the tri-allelic pattern, we presented in Fig. 2 a TPOX STR profile of this II-4 woman (depicted by the GeneMapper® ID version 3.2 software of the ABI 3130*xl* Genetic Analyzer). The three peaks

**Table 1**
Frequency of each allele in bi- and tri-allelic subjects in our pedigree and in Brazil, and tri-allelic/bi-allelic frequencies Ratio.

| TPOX allele | Allelic frequency in tri-allelic subjects in the Pedigree | Allelic frequency in bi-allelic subjects in the Pedigree | Relative Ratio tri-allelic/ bi-allelic in the Pedigree | Allelic frequency in bi-allelic subjects in Brazil | Relative Ratio tri-allelic subjects in the Pedigree/bi-allelic subjects in Brazil |
|---|---|---|---|---|---|
| | Subjects N = 6<br>TPOX alleles N = 18 | Subjects N = 39<br>TPOX alleles N = 78 | | Subjects N = 123,102[a]<br>TPOX alleles N = 246,204 | |
| 6 | 0 | 0 | 0 | 0.018 | 0 |
| 7 | 0 | 0.026 | 0 | 0.008 | 0 |
| 8 | 0.333 | 0.359 | 0.928 (0.333/0.359) | 0.456 | 0.730 (0.333/0.456) |
| 9 | 0 | 0.115 | 0 | 0.124 | 0 |
| 10 | 0.333 | 0.089 | 3.741 (0.333/0.089) | 0.067 | 4.970 (0.333/0.067) |
| 11 | 0.333 | 0.385 | 0.865 (0.333/0.385) | 0.275 | 1.211 (0.333/0.275) |
| 12 | 0 | 0.026 | 0 | 0.048 | 0 |
| 13 | 0 | 0 | 0 | 0.002 | 0 |

[a] Aguiar et al. (2012): 123,102 individuals; N = 246,204 alleles.

shown here and in the other five tri-allelic subjects have similar areas which are compatible with Type 2 as defined by Clayton et al. (2004), which reflects the presence of three copies of the STR sequence in all cells, and not to the Type 1 pattern, that could result from a mutation in sporadic somatic cells. The age of all pedigree subjects ranged from 6 to 73 years. We noticed a total of six tri-allelic subjects in this family, and all of them were women. Analyzing the pedigree, it was observed that only the descendants of the generation I, who are located on the right side of the genealogy, had inherited of TPOX third allele.

Lukka et al. (2006) show evidence that the tri-allelic pattern should be due to the duplication in chromosome 2 containing the TPOX STR Locus. Thus, since partial isodisomy for chromosome 2p might be compatible with minimal influence on normal development and/or compatible with adulthood (Al-Saffar et al., 2000; Bakker et al., 2001; Megarbane et al., 1997), we carried out a karyotype analysis in the cells of female II-4 (proband) to test a probable partial trisomy of chromosome 2. A partial 2p duplication, located on the same chromosome 2 or elsewhere in the genome, should be clearly viewable with a karyotype assessment. However, no partial trisomy of the 2p segment was apparent in the GTG-Banded karyotype of this female (Fig. 3).

Forty five family members are no large enough to make some conclusions with statistical significance, nevertheless, when we compared the allele frequencies among bi and tri-allelic subjects, we observed that the frequencies of the alleles 8 and 11 were similar, but the allele 10 frequency in tri-allelic subjects was four (3.741) and five (4.970) times higher than in bi-allelic subjects in pedigree members' and in Brazilian global population, respectively (Table 1). With these results, it was possible to strengthen the hypothesis that the third-extra allele would be the allele 10.

Our six tri-allelic females had alleles 8–10–11. Due to the presence of those alleles we focused our investigation on the analysis of the transmission and the segregation of them. Based on previous suggestion (Lane, 2008 and STRBase data), we presumed the third TPOX extra-allele as the allele 10. Analyzing the descendants III-7 (11–11; son of II-4 and II-5), III-13 and III-11 (11–11; descendants of II-6 and II-7), and III-14 (9–11; daughter II-10 and II-11) we observed that the allele 11 was transmitted as a regular allele; i.e. as an ordinary TPOX allele on chromosome 2. These results also indicated that this allele 11 segregated independently from the supposed extra-allele-10. About the allele 8, observing the man II-12 (8–8; brother of II-4, II-6, II-8, and II-10), we detected that allele 8 was also transmitted as a regular allele, as well as it had an independent segregation from the extra-allele-10. We believe that our data were able to maintain the supposition previously assumed for allele 10 as the TPOX extra-allele. We examined the DNA profile of subject IV-6 since this boy could have three TPOX alleles with a possible concomitant homozygosis of alleles 8 or 10. Fig. 4 shows that his DNA yielded two even TPOX peaks which indicate that he has equal doses of alleles 8 and 10 and not a double dose of either. It is possible that his mother III-9 transmitted allele 8 and the father transmitted his normal allele 10. All these data may indicate that the TPOX extra-allele segment is not necessarily closely linked to a chromosome 2 regular TPOX locus. This data corroborate with Lane (2008) and Díaz et al. (2009).

Lane's study suggested that the extra-allele-10 was inserted on an X chromosome (Lane, 2008). However the results in this pedigree neither support nor refute this hypothesis, since there is the possibility of segregation of the extra-allele-10 either as X-linked or as an autosomal chromosome. If the extra-allele-10 is located on an X chromosome then Fig. 1 suggests that the father I-3 may have been the transmitter parent because all his daughters have the tri-allelic genotype and neither of his two sons has. To support or reject the hypothesis that the third extra-allele is X-linked, we conducted a study with specific X-linked markers. In an attempt to identify linkage disequilibrium between the extra-allele-10 and X-chromosome markers, we evaluated 12 X-STRs in all subjects of the II-4 female family. Based on X-STRs present in individuals of generation II, it was possible to reconstruct the X-haplotype of the father I-3 (Fig. 5). Considering the occurrence of events of crossing-over between X chromosomes in females from generations III and IV, we evaluated if the transmission of the extra-allele-10 would be linked to some of the 12 X-STR alleles. We performed an association analysis with each pedigree member presented on Fig. 5 considering the presence/absence of each allele per loci and presence/absence of extra-allele 10; the analysis showed that only the allele 36 of the DXS10011 locus was statically significant ($p = 0.0022$; Fisher exact test, two-tailed) and it would be linked to extra-allele-10. In this analysis, none of the other alleles were associated with extra-allele-10.

However, it was not possible to reach a conclusion about linkage disequilibrium in loci DXS7423, DXS9898 and HPRTB loci because there was homozygosis. According to the result we can infer that the extra-allele-10 could be located near the telomeric region of this Xq.

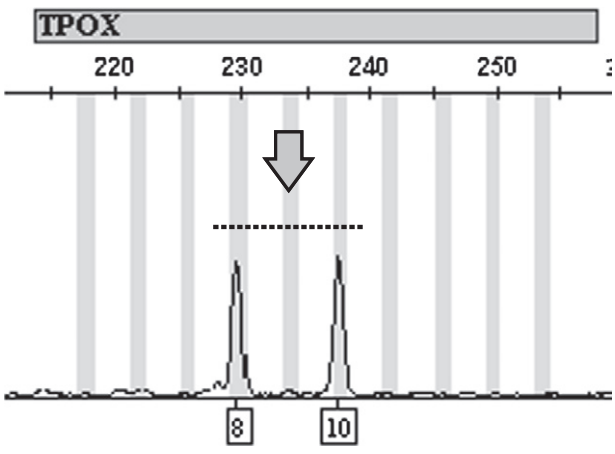



**Fig. 4.** Normal bi-allelic TPOX genotype of boy IV-6 depicted by ABI GeneMapper3.2 software. No homozygosis in alleles 8 or 10 was observed.

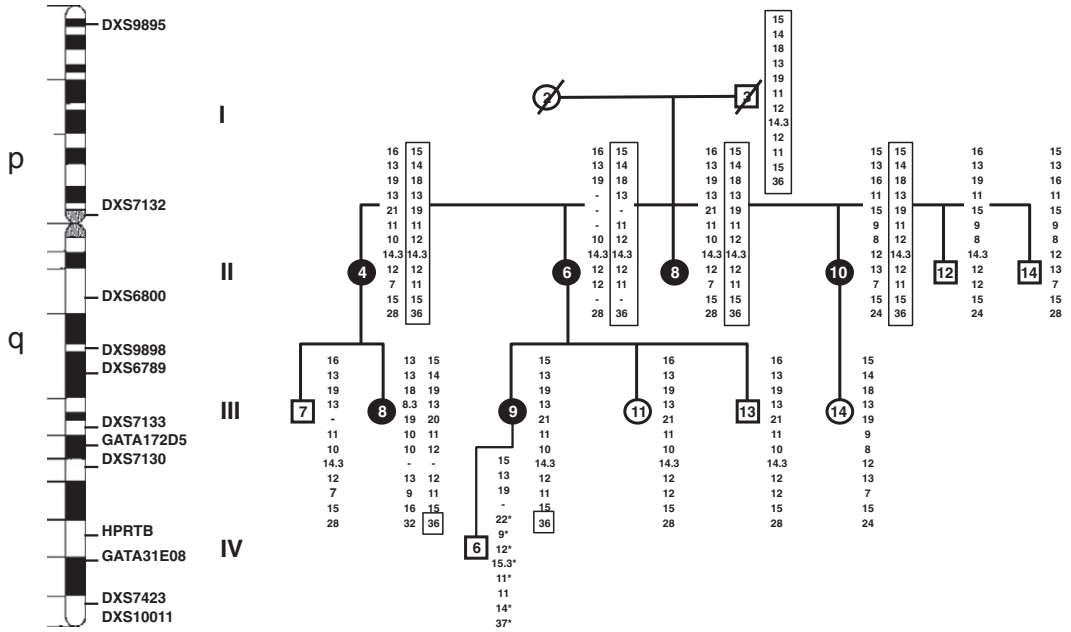



**Fig. 5.** Family pedigree with the DNA profiles of the X-STR loci. Roman numerals show the five generations of this family. Females are represented by circles. Males are represented by squares. Shaded circles indicate women with tri-allelic TPOX genotype. Numbers within the symbols (circles and squares) indicate each subject. Numbers outside the symbols (left) indicate the X-STR haplotypes. The symbol "-" indicates locus where the allele designation was not possible. Based on the X-STR haplotype in subjects of generation II, it was possible to reconstruct the haplotype of the X chromosome of the father I-3, which is shown within the box. In women III-9, III-11, and III-14 parental haplotypes were removed to clarify the transmission of maternal X-chromosome. The alleles marked with asterisk (boy IV-6) are confirmed in the X-maternal of his grandfather. The allele 36 (locus DXS10011; inside the box in generation III), has been shown in linkage disequilibrium with allele-extra-10 present in tri-allelic females (black circles).

With our study on this pedigree, we cannot speculate about the origin of the TPOX extra-allele, but the possible mechanism for the origin of the third allele was discussed by Lane (2008) suggesting that a duplication/insertion could take place before the start of the Bantu expansion; it might be possible that slaves, who were drawn from both Western and Eastern Bantu groups, led the African tri-allelic genotype with them to the New World. So, according to Lane, the relatively common X-linked extra-allele-10 found in South African populations must have arisen before the start of the Bantu expansion. The same case was verified in other more recent population originated from African slaves (Díaz et al., 2009). Although the Southern Brazilian population has some African genetic component, the woman I-2 seems to have a full European origin. Her ethnic characteristic was typically Caucasian (e.g., thin nose, and white skin, eyes, and hair). The subjects of generation II also reported having Italian Matrilineal Ancestry, but there is no information about the father I-3. Brazil was discovered and colonized by the Portuguese at the beginning of the 16th century. In the same century, the slavery of African (mainly Bantu and Yoruba) individuals was introduced. After the eighteenth century, other people have migrated to Brazil, mainly from Portugal, Italy, and Germany. Nowadays, the general Brazilian genetic structure is considered to be quite complex and different population groups can be classified according to their ethnicity in European-derived, African-derived, Brazilian Mulattos and Asian-derived (IBGE, 2000). The admixture process happened diversely in different geographic regions, with a most pronounced Native American contribution in the North, a high African contribution in the Northeast and a relatively low Native American and African contributions in the South.

## 4. Conclusion

We presented a genealogy with TPOX tri-allelic subjects where the tri-allelic pattern is due to three copies of STR sequence in all cells (Type 2) and is not the result of overt visible partial 2p trisomy. Six tri-allelic subjects were women and all had genotype 8–10–11. We presented evidences to assume the TPOX third extra-allele as the allele 10. The alleles 8 and 11 appeared segregating independently from this extra-allele. Thus, our results allowed to support that insertion of extra-allele-10 is placed far from the main TPOX locus. We show evidences of a potential linkage of the TPOX extra-allele-10 with Xq.

## Conflict of interest

The authors hereby declare that there are no conflicts of interest which may have affected the results and discussion provided herein.

## Acknowledgments

We thank AB Lane for the suggestions, and for Tricia K. Albuquerque on behalf of the "Instituto Geral de Perícias do Rio Grande do Sul (IGP)" and the Applied Biosystems Company for supporting the analysis.

## References

Aguiar, V.R., et al., 2012. Updated Brazilian STR allele frequency data using over 100,000 individuals: an analysis of CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, Penta D, Penta E, TH01, TPOX and vWA loci. Forensic Sci. Int. Genet. 504–509.

Al-Saffar, M., Lemyre, E., Koenekoop, R., Duncan, A.M.V., Der Kaloustian, V.M., 2000. Phenotype of a patient with pure partial trisomy 2p(p23-pter). Am. J. Med. Genet. 94, 428–432.

Anker, R., Steinbrueck, T., Donis-Keller, H., 1992. Tetranucleotide repeat polymorphism at the human thyroid peroxidase (hTPO) locus. Hum. Mol. Genet. 1, 137.

Bakker, B., Bikker, H., Hennekam, R.C.M., Lommen, E.J.P., Schipper, M.G.J., Vulsma, T., 2001. Maternal isodisomy for chromosome 2p causing severe congenital hypothyroidism. J. Clin. Endocrinol. Metab. 86, 1164–1168.

Clayton, T.M., Guest, J.L., Urquhart, A.J., Gill, P.D., 2004. A genetic basis for anomalous band patterns encountered during DNA STR profiling. J. Forensic Sci. 49, 1207–1214.

Crouse, C.A., Rogers, S., Amiott, E., Gibson, S., Masibay, A., 1999. Analysis and interpretation of short tandem repeat microvariants and three banded allele patterns using multiple allele detection systems. J. Forensic Sci. 44, 87–94.

Díaz, V., Rivas, P., Carracedo, A., 2009. The presence of tri-allelic TPOX genotypes in Dominican population. Forensic Sci. Int. Genet. (Suppl. 2), 371–372.

Freeman, J.L., Perry, G.H., Feuk, L., Redon, R., McCarroll, S.A., Altshuler, D.M., 2006. Copy number variation: new insights in genome diversity. Genome Res. 16, 949–961.

Fridman, C., Santos, P.C.C., Kohler, P., Garcia, C.F., Lopez, L.F., Massad, E., 2008. Brazilian population profile of 15 STR markers. Forensic Sci. Int. Genet. 2, e1–e4.

Huel, R.M., Basic, L., Madacki-Todorovic, K., Smajlovic, L., Eminovic, I., Berbic, I., 2007. Variant alleles, tri-allelic patterns, and point mutations observed in nuclear short tandem repeat typing of populations in Bosnia and Serbia. Croat. Med. J. 48, 494–502.

Instituto Brasileiro de Geografia Estatística (IBGE), 2000. BRASIL 500 Anos de Povoamento. IBGE, Rio de Janeiro.

Lane, A.B., 2008. The nature of tri-allelic TPOX genotypes in African populations. Forensic Sci. Int. Genet. 2, 134–137.

Lukka, M., Tasa, G., Ellonen, P., Moilan, K., Vassiljev, V., Ulmanen, I., 2006. Triallelic patterns in STR loci used for paternity analysis: evidence for a duplication in chromosome 2 containing the TPOX STR locus. Forensic Sci. Int. 164, 3–9.

Megarbane, A., Souraty, N., Prieur, M., Theophile, D., Chedid, P., Auge, J., 1997. Interstitial duplication of the short arm of chromosome 2: report of a new case and review. J. Med. Genet. 34, 783–786.

Poiares, L.A., Osorio, P.S., Spanhol, F.A., Coltre, S.C., Rodenbusch, R., Gusmão, L., 2010. Allele frequencies of 15 STRs in a representative sample of the Brazilian population. Forensic Sci. Int. Genet. 4, e61–e63.

Ribeiro Rodrigues, E.M., Leite, F.P., Hutz, M.H., Palha, T.J., Ribeiro dos Santos, A.K., dos Santos, S.E., 2008. A multiplex PCR for 11 X chromosome STR markers and population data from a Brazilian Amazon Region. Forensic Sci. Int. Genet. 2, 154–158.

Ribeiro-Rodrigues, E.M., et al., 2009. An INDEL polymorphism at the X-STR GATA172D05 flanking region. Int. J. Legal Med. 123, 89–94.

Yunis, J.J., 1981. New chromosome techniques in the study of human neoplasia. Hum. Pathol. 12, 540–549.

## CAPITULO 3

# IDENTIFICATION OF THE THIRD-EXTRA ALLELE FOR FORENSIC APPLICATION IN CASES WITH *TPOX* TRI-ALLELIC PATTERN.

### INTRODUCTION

Genotyping of polymorphic short tandem repeats (STRs) loci is widely used in forensic DNA analysis, in familial relationship testing, and for genetic identification of individuals in paternity disputes. Tri-allelic patterns are categorized as a genotyping irregularity that can be found in STR profiling [1, 2]. Clayton and co-workers [3] have distinguished two types of tri-allelic patterns. Type 1 has two alleles with a different intensity (after PCR amplification) of a third allele and Type 2 can be manifested as three peaks with the same height. Type 1 is believed to be the result of a mutation in an early somatic cell, while Type 2 is thought to represent a constitutional chromosomal rearrangement. Although tri-allelic genotypes are generally rare, data presented on the STRBase website (http://www.cstl.nist.gov/biotech/ STRbase) indicate that tri-allelic STR genotypes can be unusually found.

Alleles of the *TPOX* STR locus have different numbers of a four nucleotide repeat motif arranged in tandem [4] and, for more than a decade, *TPOX* tri-allelic genotypes have been reported with a widely varied frequency among human populations. The frequencies to the tri-allelic pattern varies; 0.18% in 10,000 subjects for Alabama, USA [5], 0.003% in 32,800 individuals from Bosnia, Kosovo and Serbia [6], 2% in 6,827 South Africans [7], 0.2% in 561 subjects from Brazil [8]. On the other hand, Poiares et al. [9] typed 12,886 unrelated subjects from Brazil and were unable to find any *TPOX* tri-allelic genotype. In the African population [7], Dr. Lane was able to demonstrate that two thirds of the *TPOX* tri-allelic adults were females, and *TPOX* tri-allelic fathers only had transmitted the *TPOX* tri-allelic genotype to their daughters. With this evidence, he suggested that the extra allele (usually an allele 10) would be inserted in the X chromosome. This report was enhanced by Díaz et al. [10].

Despite the substantial frequency of the *TPOX* tri-allelic pattern, the nature of the third-extra allele is still poorly understood. In this work, we present data obtained from 66 families with cases of *TPOX* tri-allelic pattern (105 tri-allelic subjects). We used these data to investigate which would be the third-extra allele and in which chromosome it would be.

## MATERIAL AND METHODS

Data were obtained from routine paternity tests performed in the states of Rio Grande do Sul, Brazil (~4,000 cases/year from Fundação Estadual de Produção e Pesquisa em Saúde, FEPPS) and São Paulo (~12,000 cases/year from Instituto de Medicina Social e de Criminologia de São Paulo, IMESC), from 2008 to 2012. In this routine, the genomic DNA was purified from dried blood samples preserved in FTA cards following the manufacturer's protocols. A total of 0.5–1.0ng of DNA was used to amplify the STR loci (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, *TPOX*, D18S51, D5S818, and FGA) with the AmpFℓSTR® Identifiler™ Amplification kit (Applied Biosystems, Foster City, USA), using the manufacturer's instructions, followed by analysis on a capillary ABI 3130*xl* Genetic Analyzer. Familias Software v.1.81 [11] was used to evaluate the statistical parameters and to calculate the combined paternity/maternity index in all cases considering the Combined Paternity Index (IPC) of at least 10,000. All families that presented cases of *TPOX* tri-allelic pattern were also typed using a different system for human identification (PowerPlex®21 System, Promega, Madison, WI) to confirm the tri-allelic pattern.

In order to prevent the possibility of a false tri-allelic genotype caused by micro-variants (shorter alleles) of the adjacent bins (eg: D18S51- AmpFlSTR® Identifiler® or FGA- PowerPlex®21) occupying the bins of the *TPOX* locus, 10% of our tri-allelic subjects had their *DNA* amplified by the singleplex PCR-based analysis of the *TPOX* locus. In all cases, the presence of three alleles was confirmed based on fragment analysis by capillary electrophoresis.

To avoid that the true tri-allelic pattern was confused with a tri-allelic pattern-like (caused by mixtures of DNA samples) we included in this study subjects whose three alleles were observed only in *TPOX*-locus, i.e., all other loci (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, D18S51, D5S818, and FGA) presented a normal bialellic pattern. All Peak Height Ratio (PHR) from the other loci were confirmed to be bi-allelic pattern.

This project was approved by the Research Ethics Committee of Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) (Protocol #10/05317).

## RESULTS AND DISCUSSION

We looked for tri-allelic subjects in 75,113 families (54,863 from IMESC, and 20,250 from FEPPS). Considering only the parental generation (mother+father) we had 150,226 unrelated subjects evaluated. From this total, we found 88 unrelated subjects with tri-allelic pattern in *TPOX* locus,

which is equivalent to the frequency of 0.06% (88/150,226). Analyzing the electropherograms of *TPOX* locus in subjects with tri-allelic pattern we checked each peak and its PHR. We found 73 (83%; 73/88) subjects with the classic Type 2 tri-allelic pattern, i.e. with three peaks with the same height. In 15 (17%; 15/88) subjects we observed that there was heterozygosis peak imbalance, showing that one peak was the highest (double dose homozygote) and the second one had the regular sized (Fig. 1). This showed a tri-allelic profile hidden by the heterozygote peak imbalance. With this observation, we used Clayton's Type 2 category (Fig. 2-a) to create three devived categories from Clayton's Type 2 tri-allelic pattern (Fig. 2-b). This result indicates that around 17% of *TPOX* tri-allelic genotypes would remain unknown if the heterozygosis peak imbalance were disregarded.

Sixty-six trios (75%; 66/88) were families with true biological relationships between mother and father with the child. These 66 families, where the mother or the father had tri-allelic pattern in *TPOX* locus, are presented in Figure 3. In 39 families (59%; 39/66) the third *TPOX* allele was transmitted from the mother or from the father to the child. Thus, a total of 105 (66+39) tri-allelic subjects were studied here. The *TPOX* alleles present in all 105 tri-allelic subjects were: 6, 7, 8, 9, 10, 11, 12, and 13. To identify which one could be the third-extra allele, we performed two analyses. In the first analysis, we evaluated the possible transmission of the third-extra allele. Evaluating the 39 families where the third-extra *TPOX* allele was transmitted, we observed that the alleles 6, 8, 9, 11, 12, and 13 could be excluded as being the third allele, because the parents had the allele, but it was absent on the child. Thus, only the alleles 7 and 10 could not be excluded as being the extra allele (Table 1). Considering the alleles 7 and 10 as possible third allele, we analyzed the 66 families and observed that the allele 7 appeared as possible third allele in only one family (FAM002), but an allele 10 was present as a possible third allele in 100% of the families (66/66). Besides, in 51.5% (34/66) of families, an allele 10 was mandatorily transmitted by the tri-allelic parent.

Based on the premise described by Lane [7], that the third-extra allele would be an allele 10, our result led us to accept that the third-extra allele should be an allele 10. This hypothesis was confirmed by the second analysis made by the investigation of the absolute and relative frequencies of each *TPOX* allele in the tri-allelic subjects. We analyzed our 66 families plus the tri-allelic individuals already reported in STRBase. Table 2 shows how many times each *TPOX* allele appears in each tri-allelic subject. According to Table 2, an allele 10 was present in 100% of our tri-allelic subjects and in 90% of individuals reported on the tri-allelic STRBase. All other alleles were present in lower percentage. The 10 (9.8%; 10/102) tri-allelic subjects from the STRBase without an allele 10 had always the alleles 9 or 11 (genotypes were: 8-9-11; 8-11-12; 8-11-14.3; 9-11-12), what could denote slippage mutation from an ancestral third-extra allele 10. The comparison between each allele frequency in bi- and tri-allelic subjects and the ratio of these frequencies is showed in Table 3.

Comparing the allele frequencies among bi- and tri-allelic subjects, we observed that the frequency of the allele 10 in tri-allelic subjects was around five times higher than in bi-allelic subjects in both Brazilian and Global populations; the binomial test showed a significant highest frequency of the allele 10 in tri-allelic subjects than a theoretically expected distribution of observations in bi-allelic subjects in both Brazilian and Global populations ($P < 0.0001$; Binomial Test). With these results, it was possible to strengthen the hypothesis that the third-extra allele would be an allele 10.

Additionally, we noted that in 100% of cases from 27 families where there was no transmission of the third-extra allele from the parent to the child, an allele 10 was absent in the bi-allelic offspring.

Finally, we tested where the third-extra allele 10 would be. We noticed that in 100% of cases the tri-allelic father transmitted the third-extra allele 10 to his daughter, but never to his son (Fig. 3: E-F). Moreover, the tri-allelic mothers transmitted the third-extra allele 10 in around 50% of cases, both to her daughter and to her son (Fig.3: A-D). This last result allowed us to believe that the third-extra allele 10 would be on the X chromosome as previously suggested by Dr. Lane [7].

In our study with 75,113 Brazilian families any tri-allelic case was detected only in the offspring (child), i.e. here we did not detect any "de novo" *TPOX* tri-allelic mutation.

When STR loci present tri-allelic pattern it causes doubt on the tri-allele locus calculation, possibly reducing the total profile analysis strength. Knowing that the majority of the third-extra *TPOX* allele is probably an allele 10 and that it would be on X chromosome, it may allow excluding only this third-extra allele from genetic profile, thus preserving the standard *TPOX* alleles in forensic analyses. Additionally, the third-extra allele transmission from the father to the daughter could improve paternity analysis; it may be important in some special cases. The tri-allelic pattern study has important consequences for forensic applications, including the criteria for locus interpretation in cases of admixtures.

## CONCLUSIONS

In our study with Brazilian families the frequency of tri-allelic subjects in *TPOX* locus was 0.06% (88/150,226), and 17% (15/88) of *TPOX* tri-allelic genotypes presented heterozygosis peak imbalance. Our data support the previous hypothesis that the third-extra *TPOX* allele was an allele 10, and that it is probably localized on X chromosome. These data improve the knowledge about tri-allelic pattern of *TPOX* CODIS' locus allowing the use of *TPOX* profile in forensic analysis even when with tri-allelic pattern.

**REFERENCES**


[1] J.M. Butler, Short tandem repeat typing technologies used in human identity testing, Biotechniques, 43 (2007) 2–5.

[2] J.M. Butler, Genetics and genomics of core *STR* loci used in human identity testing, J. Forensic Sci., 51 (2006) 253–265.

[3] T.M. Clayton, J.L. Guest, A.J. Urquhart, P.D. Gill, A genetic basis for anomalous band patterns encountered during *DNA STR* profiling, *J. Forensic Sci.* 49 (2004) 1207–1214.

[4] R. Anker, T. Steinbrueck, H. Donis-Keller, Tetranucleotide repeat polymorphism at the human thyroid peroxidase (hTPO) locus, Hum. Mol. Genet. 1 (1992) 137.

[5] C.A. Crouse, S. Rogers, E. Amiott, S. Gibson, A. Masibay, Analysis and interpretation of short tandem repeat microvariants and three banded allele patterns using multiple allele detection systems, J. Forensic Sci. 44 (1999) 87–94.

[6] R.M. Huel, L. Basic, K. Madacki-Todorovic, L. Smajlovic, I. Eminovic, I. Berbic, A. Miloš, T. J. Parsons, Variant alleles, tri-allelic patterns, and point mutations observed in nuclear short tandem repeat typing of populations in Bosnia and Serbia, Croat. Med. J. 48 (2007) 494–502.

[7] A.B. Lane, The nature of tri-allelic *TPOX* genotypes in African populations, Forensic Sci. Int: Genet. 2 (2008) 134-137.

[8] C. Fridman, P. C. C. Santos, P. Kohler, C. F. Garcia, L. F. Lopez, E. Massad, G.J. Gattás, Brazilian population profile of 15 *STR* markers, Forensic Sci. Int: Genet. 2 (2008) e1–e4.

[9] L. A. Poiares, P. S. Osorio, F. A. Spanhol, S. C. Coltre, R. Rodenbusch, L. Gusmão, A. Largura, F. Sandrini, C. M. da Silva, Allele frequencies of 15 *STR*s in a representative sample of the Brazilian population. Forensic Sci. Int: Genet 4 (2010) e61–e63.

[10] V. Días, P. Rivas, A. Carracedo, The presence of tri-allelic *TPOX* genotypes in Dominican Population. Forensic Sci. Int. Gene. Suppl. 2 (2009) 371-372.

[11] T. Egeland, P. Mostad, B. Mevåg, M. Stenersen, Beyond traditional paternity and identification cases. Selecting the most probable pedigree, Forensic Sci. Inter. 110 (2000) 47-59.

[12] V.R. Aguiar, E.V. Wolfgramm, F.S. Malta, A.G. Bosque, A.C. Mafia, V.C. Almeida, F. A. Caxito, V.C. Pardini, A.C. Ferreira, I.D. Louro, Updated Brazilian *STR* allele frequency data using over 100,000 individuals: an analysis of CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, Penta D, Penta E, TH01, *TPOX* and vWA loci, Forensic Sci. Int. Genet. 6 (2012) 504-9.

[13] J.M Butler, R. Schoske, P.M. Vallone, J.W. Redman, M.C. Kline, Allele frequencies for 15 autosomal *STR* loci on U.S. Caucasian, African American, and Hispanic populations, J. Forensic Sci. 48 (2003) 908-11.

**Figure 1.** Heterozygote peak imbalance in *TPOX* locus. Fifteen subjects are identified by the family numbers and their electropherograms to *TPOX* locus are showed. Circle: female; Square: male; M: mother; C: child. The numbers below the subjects' symbols represent the alleles of *TPOX* locus. The 15 subjects presented a higher peak (double dose homozygous) and a regular sized peak. The Peak Height Ratio (PHR) is shown: PHR Mean is equivalent to PHR average of all regular heterozygous locus in the subject; PHR *TPOX* is the PHR value to *TPOX* locus.

**Figure 2. (a)** Clayton et al. (2004) tri-allelic pattern categories (Type 1 and Type 2); **(b)** Our suggestion for three categories derived from Clayton's Type 2 tri-allelic pattern. In these new derived Type 2 categories, the tri-allelic pattern may looks as three peaks with similar area (Type 2-A), as two peaks with different high (Type 2-B; one regular and one doble), or as one single peak with triplicate size (Type 2-C). The last category (Type 2-C) has not been observed, but it is a possibillity.

**Figure 3.** Sixty six families with subjects with *TPOX* tri-allelic pattern. Circle: female; Square: male; M: mother; F: father; C: child. The numbers below the subjects' symbols represent the alleles of *TPOX* locus. The families are organized by tri-allelic parent (grey symbol); in 49 families the mother was tri-allelic (A, B, C, D), and in 17 families the father was tri-allelic (E, F). The mother transmitted the third-extra allele to her child (son or daughter equally) in 61% (30/49) of families. The father transmitted the third-extra allele only when the child was a daughter.

**Table 1.** Transmission analysis of the third allele in 39 families to identify which allele was not transmitted from tri-allelic parent to tri-allelic child.

| Alelle | Families that exclude the transmission of this allele as third/extra allele* |
|---|---|
| 6 | FAM002, FAM024, FAM050 |
| 7 | - |
| 8 | FAM026, FAM003, FAM027, FAM034, FAM053 |
| 9 | FAM004, FAM030, FAM031 |
| 10 | - |
| 11 | FAM008, FAM052 |
| 12 | FAM051 |

***** The number indicates the family that is shown in Figure 3.

**Table 2.** Absolute frequency of each allele in tri-allelic subjects in our 66 families (with 105 tri-allelic subjects) and in STRBase (with 102 tri-allelic reported subjects).

| Alelle | Our 66 families (N=105) | *STR*Base (N=102)* |
|---|---|---|
| 6 | 0.086 [09/105] | 0.156 [16/102] |
| 7 | 0.038 [04/105] | 0.058 [06/102] |
| 8 | 0.667 [70/105] | 0.637 [65/102] |
| 9 | 0.248 [26/105] | 0.450 [46/102] |
| 10 | 1.00 [105/105] | 0.901 [92/102] |
| 11 | 0.638 [67/105] | 0.598 [61/102] |
| 12 | 0.171 [18/105] | 0.176 [18/102] |
| 13 | 0.009 [01/105] | ---------- |
| 14.3 | ---------- | 0.009 [01/102] |

*From: http://www.cstl.nist.gov/*STR*base/ accessed on December 2012.

**Table 3.** Frequency of each allele in bi- and tri-allelic subjects in our 66 families and in STRBase, and tri-allelic/bi-allelic frequencies Ratio in Brazil and Global populations.

| Population | *TPOX* Alelle | Allelic frequency in tri-allelic subjects | Allelic frequency in bi-allelic subjects | Relative tri-allelic/ bi-allelic Ratio | P value* (*p* > *k*) |
|---|---|---|---|---|---|
| **Our cases** | | **Subjects N=105** | **Subjects N=123,102[a]** | | |
| | | ***TPOX* alleles N=315** | ***TPOX* alleles N=246,202** | | |
| | 6 | 0.028 [09/315] | 0.018 | 1.555 | 0.0613 |
| | 7 | 0.012 [04/315] | 0.008 | 1.500 | 0.1107 |
| | 8 | 0.248 [79/315] | 0.456 | 0.544 | 1.0000 |
| | 9 | 0.088 [28/315] | 0.124 | 0.710 | 0.9687 |
| | 10 | 0.339 [108/315] | 0.067 | 5.060 | <0.0001 |
| | 11 | 0.216 [69/315] | 0.275 | 0.785 | 0.9861 |
| | 12 | 0.059 [19/315] | 0.048 | 1.230 | 0.1261 |
| | 13 | 0.003 [01/315] | 0.002 | 1.500 | 0.1318 |
| **Global** | | **Subjects N=102** | **Subjects N=700[b]** | | |
| | | ***TPOX* alleles N=306** | ***TPOX* alleles N=1400** | | |
| | 6 | 0.052 [16/306] | 0.039 | 1.333 | 0.0935 |
| | 7 | 0.019 [06/306] | 0.008 | 2.375 | 0.0124 |
| | 8 | 0.212 [65/306] | 0.463 | 0.458 | 1.0000 |
| | 9 | 0.150 [46/306] | 0.138 | 1.087 | 0.2364 |
| | 10 | 0.303 [93/306] | 0.064 | 4.734 | <0.0001 |
| | 11 | 0.199 [61/306] | 0.241 | 0.826 | 1.0000 |
| | 12 | 0.058 [18/306] | 0.046 | 1.261 | 0.1164 |
| | 14.3 | 0.003 [01/306] | - | - | - |

[a] Aguiar et al, 2012: 123,102 individuals; N=246,202 alleles [12]; [b] Butler et al, 2003: 700 individuals; N=1,400 alleles (U.S. Caucasian, African American, and Hispanic populations) [13]. * Binomial Test to *p* > *k*.

**Figure 1**

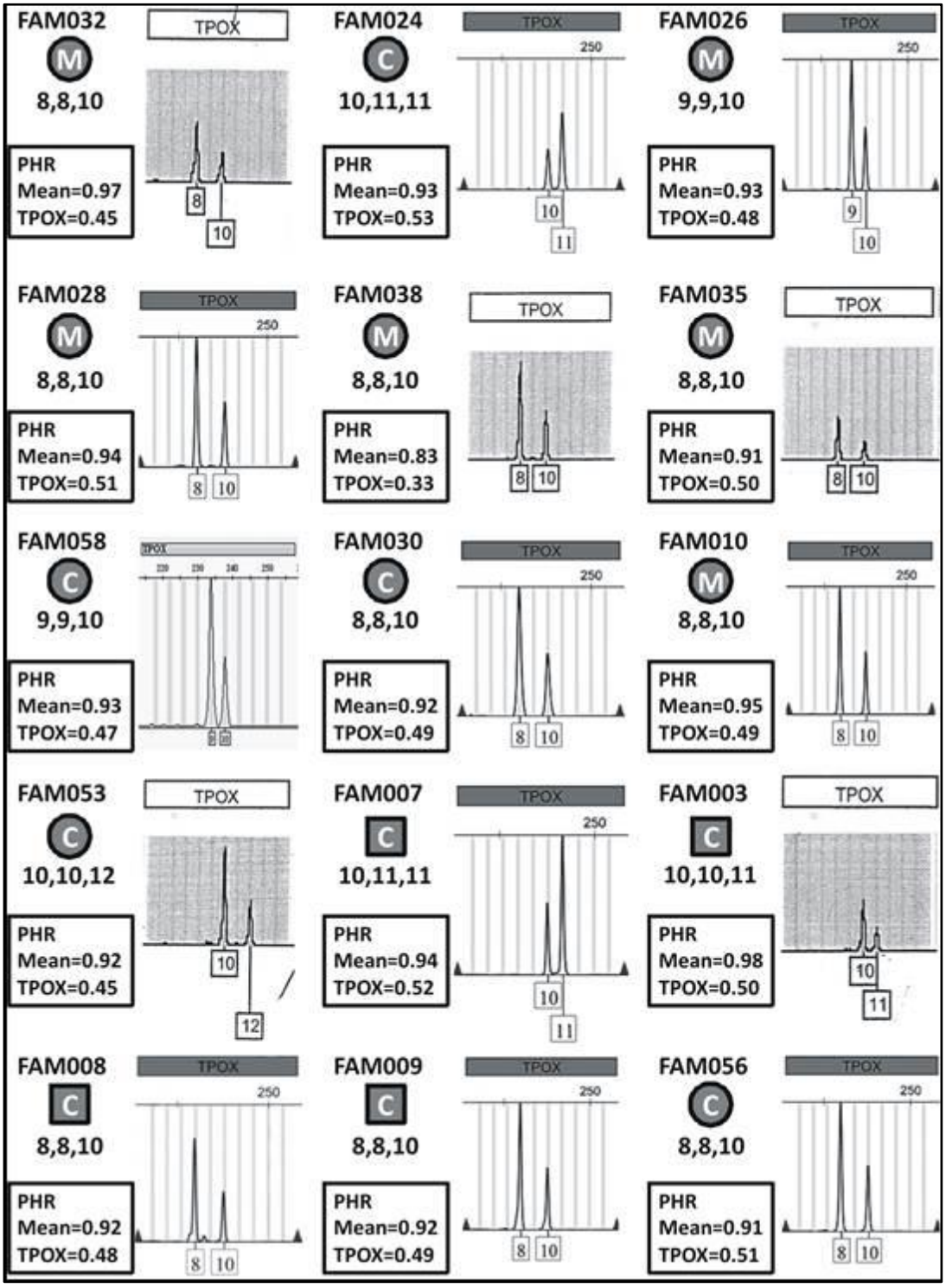

FAM032
M
8,8,10
PHR
Mean=0.97
TPOX=0.45
TPOX
8
10
FAM024
C
10,11,11
PHR
Mean=0.93
TPOX=0.53
TPOX
250
10
11
FAM026
M
9,9,10
PHR
Mean=0.93
TPOX=0.48
TPOX
250
9
10
FAM028
M
8,8,10
PHR
Mean=0.94
TPOX=0.51
TPOX
250
8
10
FAM038
M
8,8,10
PHR
Mean=0.83
TPOX=0.33
TPOX
8
10
FAM035
M
8,8,10
PHR
Mean=0.91
TPOX=0.50
TPOX
8
10
FAM058
C
9,9,10
PHR
Mean=0.93
TPOX=0.47
TPOX
FAM030
C
8,8,10
PHR
Mean=0.92
TPOX=0.49
TPOX
250
8
10
FAM010
M
8,8,10
PHR
Mean=0.95
TPOX=0.49
TPOX
250
8
10
FAM053
C
10,10,12
PHR
Mean=0.92
TPOX=0.45
TPOX
10
12
FAM007
C
10,11,11
PHR
Mean=0.94
TPOX=0.52
TPOX
250
10
11
FAM003
C
10,10,11
PHR
Mean=0.98
TPOX=0.50
TPOX
10
11
FAM008
C
8,8,10
PHR
Mean=0.92
TPOX=0.48
TPOX
250
8
10
FAM009
C
8,8,10
PHR
Mean=0.92
TPOX=0.49
TPOX
250
8
10
FAM056
C
8,8,10
PHR
Mean=0.91
TPOX=0.51
TPOX
250
8
10

**Figure 2**

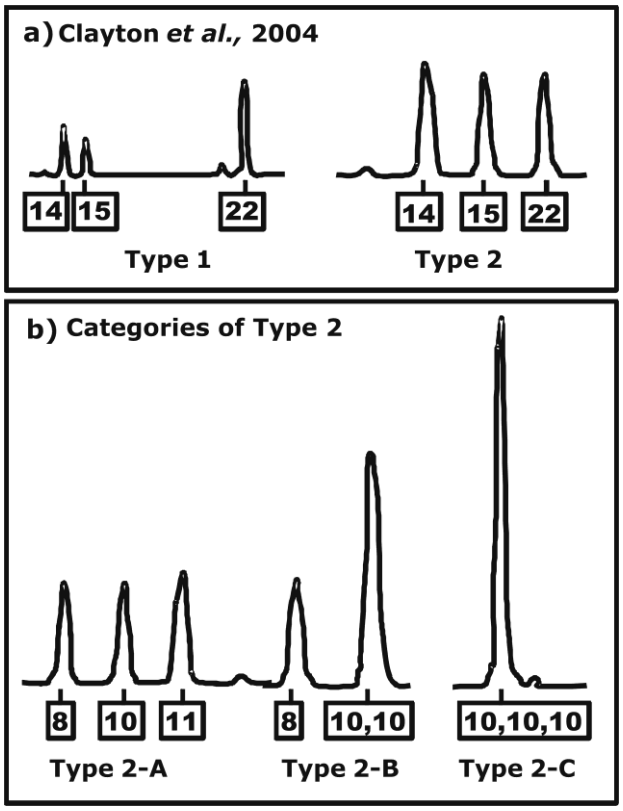
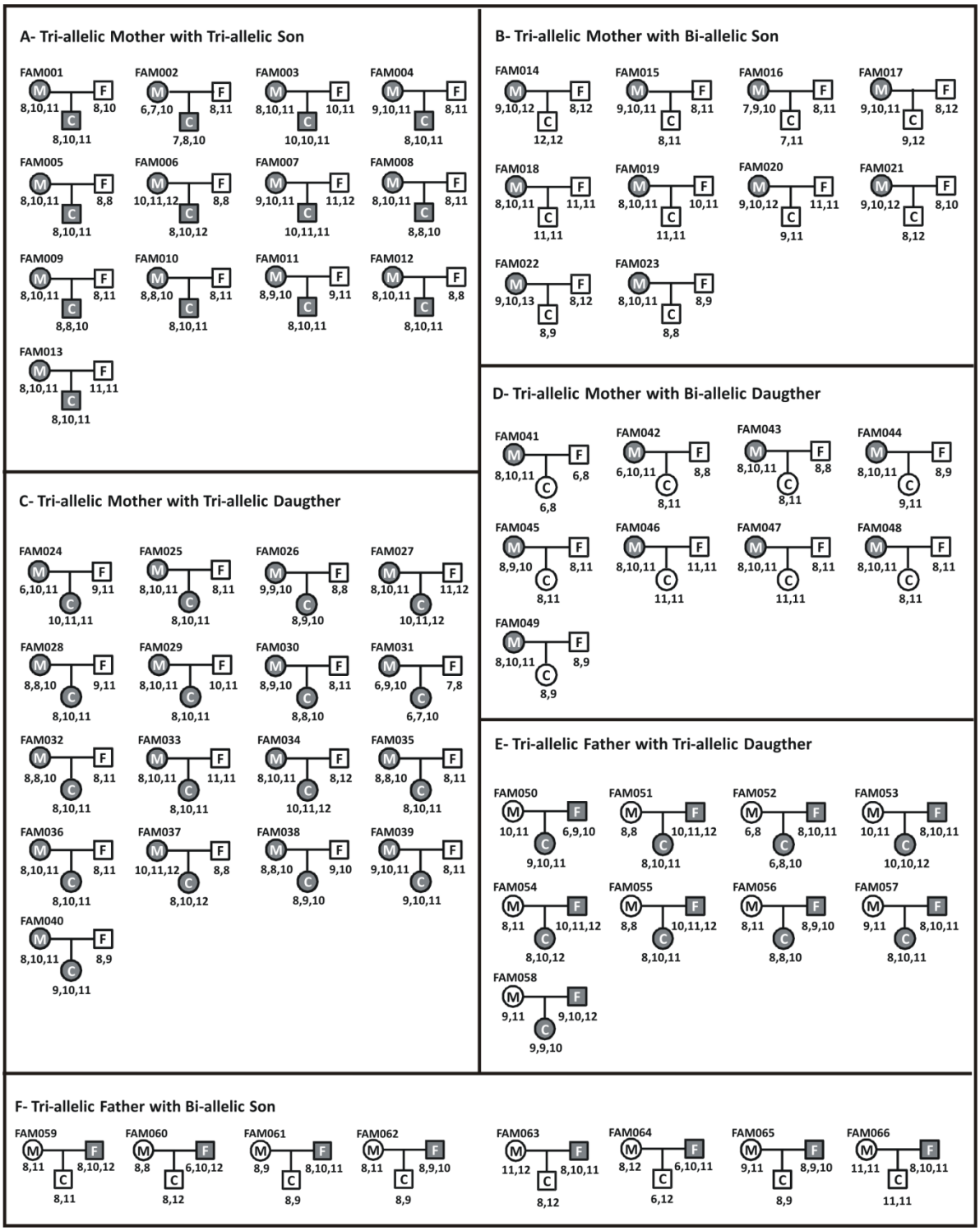

a) Clayton *et al.*, 2004
14
15
22
Type 1
14
15
22
Type 2
b) Categories of Type 2
8
10
11
Type 2-A
8
10,10
Type 2-B
10,10,10
Type 2-C

**Figure 3**


A- Tri-allelic Mother with Tri-allelic Son
FAM001 M 8,10,11 F 8,10 C 8,10,11
FAM002 M 6,7,10 F 8,11 C 7,8,10
FAM003 M 8,10,11 F 10,11 C 10,10,11
FAM004 M 9,10,11 F 8,11 C 8,10,11
FAM005 M 8,10,11 F 8,8 C 8,10,11
FAM006 M 10,11,12 F 8,8 C 8,10,12
FAM007 M 9,10,11 F 11,12 C 10,11,11
FAM008 M 8,10,11 F 8,11 C 8,8,10
FAM009 M 8,10,11 F 8,11 C 8,8,10
FAM010 M 8,8,10 F 8,11 C 8,10,11
FAM011 M 8,9,10 F 9,11 C 8,10,11
FAM012 M 8,10,11 F 8,8 C 8,10,11
FAM013 M 8,10,11 F 11,11 C 8,10,11
B- Tri-allelic Mother with Bi-allelic Son
FAM014 M 9,10,12 F 8,12 C 12,12
FAM015 M 9,10,11 F 8,11 C 8,11
FAM016 M 7,9,10 F 8,11 C 7,11
FAM017 M 9,10,11 F 8,12 C 9,12
FAM018 M 8,10,11 F 11,11 C 11,11
FAM019 M 8,10,11 F 10,11 C 11,11
FAM020 M 9,10,12 F 11,11 C 9,11
FAM021 M 9,10,12 F 8,10 C 8,12
FAM022 M 9,10,13 F 8,12 C 8,9
FAM023 M 8,10,11 F 8,9 C 8,8
C- Tri-allelic Mother with Tri-allelic Daugther
FAM024 M 6,10,11 F 9,11 C 10,11,11
FAM025 M 8,10,11 F 8,11 C 8,10,11
FAM026 M 9,9,10 F 8,8 C 8,9,10
FAM027 M 8,10,11 F 11,12 C 10,11,12
FAM028 M 8,8,10 F 9,11 C 8,10,11
FAM029 M 8,10,11 F 10,11 C 8,10,11
FAM030 M 8,9,10 F 8,11 C 8,8,10
FAM031 M 6,9,10 F 7,8 C 6,7,10
FAM032 M 8,8,10 F 8,11 C 8,10,11
FAM033 M 8,10,11 F 11,11 C 8,10,11
FAM034 M 8,10,11 F 8,12 C 10,11,12
FAM035 M 8,8,10 F 8,11 C 8,10,11
FAM036 M 8,10,11 F 8,11 C 8,10,11
FAM037 M 10,11,12 F 8,8 C 8,10,12
FAM038 M 8,8,10 F 9,10 C 8,9,10
FAM039 M 9,10,11 F 8,11 C 9,10,11
FAM040 M 8,10,11 F 8,9 C 9,10,11
D- Tri-allelic Mother with Bi-allelic Daugther
FAM041 M 8,10,11 F 6,8 C 6,8
FAM042 M 6,10,11 F 8,8 C 8,11
FAM043 M 8,10,11 F 8,8 C 8,11
FAM044 M 8,10,11 F 8,9 C 9,11
FAM045 M 8,9,10 F 8,11 C 8,11
FAM046 M 8,10,11 F 11,11 C 11,11
FAM047 M 8,10,11 F 8,11 C 11,11
FAM048 M 8,10,11 F 8,11 C 8,11
FAM049 M 8,10,11 F 8,9 C 8,9
E- Tri-allelic Father with Tri-allelic Daugther
FAM050 M 10,11 F 6,9,10 C 9,10,11
FAM051 M 8,8 F 10,11,12 C 8,10,11
FAM052 M 6,8 F 8,10,11 C 6,8,10
FAM053 M 10,11 F 8,10,11 C 10,10,12
FAM054 M 8,11 F 10,11,12 C 8,10,12
FAM055 M 8,8 F 10,11,12 C 8,10,11
FAM056 M 8,11 F 8,9,10 C 8,8,10
FAM057 M 9,11 F 8,10,11 C 8,10,11
FAM058 M 9,11 F 9,10,12 C 9,9,10
F- Tri-allelic Father with Bi-allelic Son
FAM059 M 8,11 F 8,10,12 C 8,11
FAM060 M 8,8 F 6,10,12 C 8,12
FAM061 M 8,9 F 8,10,11 C 8,9
FAM062 M 8,11 F 8,9,10 C 8,9
FAM063 M 11,12 F 8,10,11 C 8,12
FAM064 M 8,12 F 6,10,11 C 6,12
FAM065 M 9,11 F 8,9,10 C 8,9
FAM066 M 11,11 F 8,10,11 C 11,11

## CAPITULO 4

### CONSIDERAÇÕES FINAIS

A genotipagem de *short tanden repeats* ( STRs ) é amplamente utilizada na análise de DNA forense, contudo eventuais ocorrências de trialelias podem reduzir o poder da análise. Apesar de o padrão tri-alélico no lócus *TPOX* ter uma frequência importante, há poucos estudos para entender a natureza do terceiro alelo. Com este estudo nós indicamos um alelo 10 do lócus *TPOX* como sendo o alelo extra, e que o mesmo está provavelmente ligado ao cromossomo X. Acreditamos que nossos resultados poderão auxiliar a contornar o fator complicador gerado pela trialelia do *TPOX* dado que, isolando-se o terceiro alelo, pode-se identificar mais facilmente a presença e a transmissão dos alelos *TPOX* convencionais. Também chamamos a atenção para uma categoria derivada do padrão Tipo 2 clássico para trialelias, no qual não há os três picos com área similar, mas sim apenas dois picos, sendo um deles duplicado (em homozigose). Este achado nos leva a sugerir que alguns casos de trialelia podem estar ocultos nos levantamentos de ocorrência deste evento.

Como perspectivas para ampliar esse estudo, planeja-se realizar análises em amostras com trialelia usando a técnica de FISH (*fluorescence* in situ *hybridization*), a qual é um técnica citogenética usada para detectar e localizar a presença ou a ausência de determinadas sequências de DNA em cromossomos. Este mapeamento físico utiliza cromossomos metafásicos, sondas fluorescentes que se ligam somente às partes do cromossomo a que elas apresentam um elevado grau de complementaridade de sequência, e usa da microscopia de fluorescência para a sua visualização. Algumas barreiras podem limitar este estudo, contudo, dado que duplicações de curta extensão podem não ter sinal visivelmente detectável considerando a sensibilidade do experimento, e duplicações em tandem podem não ser suficientemente diferentes do sinal original não duplicado (Crotwell e Hoyme 2012). Embora estes limites existam, nosso plano atual inclui realizar o estudo do mapeamento físico com o uso inicial de uma sonda para o lócus *TPOX* juntamente com marcadores fluorescentes para os cromossomos X e 2. Construimos a sonda para o lócus *TPOX* e estamos atualmente realizando tais experimentos.

## REFERÊNCIAS BIBLIOGRÁFICAS


Anker R., Steinbrueck T., Donis-Keller H., **Tetranucleotide repeat polymorphism at the human thyroid peroxidase (h*TPOX*) locus.** Hum Mol Gen v.1, 2 (1992) 137.

Al-Saffar, M., Lemyre, E., Koenekoop, R., Duncan, A. M. V. & Der Kaloustian, V. M. **Phenotype of a Patient With Pure Partial Trisomy 2p(p23-pter)**. Am. J. Med. Genet. 94, 428–432 (2000).

Aviram-Goldring A., Fritz B., Bartsch C., Steuber E., Daniely M., Lev D., Chaki R., Barkai G., Frydman M., Rehder H., **Molecular Cytogenetic Studies in Three Patients With Partial Trisomy 2p, Including CGH From Paraffin-Embedded Tissue.** Am J Med Genet 91 (2000) 74–82.

Bakker, B., Bikker, H., Hennekam, R. C. M., Lommen, E. J. P., Schipper, M. G. J., Vulsma, T. et al. **Maternal Isodisomy for Chromosome 2p Causing Severe Congenital Hypothyroidism**. J. Clin. Endocrinol. Metab. 86, 1164–1168 (2001).

Brinkmann B., Klintschar M., Neuhuber F., Huhne J., Rolf B., **Mutation rate in human microsatellite: influence of the structure and length of the tandem repeat.** Am J Hum Genet 62 (1998) 1408–1415.

Butler J. M., **Forensic *DNA* typing: biology and technology behind *STR* markers.** London: Academic Press, 2001.

Butler J. M., **Forensic *DNA* typing: biology, technology, and genetics of *STR* markers**. 2 ed., New York: Elsevier, 2005. 660p.

Butler J. M., **Genetics and Genomics of Core Short Tandem Repeat Loci Used in Human Identity Testing.** J Forensic Sci 51 (2006) 253-265.

Budowle B., Moretti T. R., Niezgoda S. J., Brown B. L., ***CODIS* and *PCR*-based short tandem repeat loci**: **law enforcement tools.** Proceedings of the Second European Symposium on Human Identification, Innsbruck, Austria, June 1998. Madison, WI: PromegaCorporation.1998;73-88;http://www.promega.com/geneticidproc/eusymp2pro/17.pdf

Chen C. P., Liu F. F., Jan S. W., Lin S. P., Lan C. C., **Prenatal diagnosis of partial monosomy 3p and partial trisomy 2p in a fetus associated with shortening of the long bones and a single umbilical artery.** Prenat Diagn 16 (1996) 270–275.

Clayton T. M., Guest J. L., Urquhart A. J., Gill P. D., **A genetic basis for anomalous band patterns encountered during *DNA STR* profiling.** J Forensic Sci 49 (2004) 1207–1214.

Crotwell, P. L., Hoyme, H. E., **Advances in Whole-Genome Genetic Testing: From Chromosomes to Microarrays.** Curr Probl Pediatr Adolesc Health Care 42 (2012) 47–72.

Crouse C. A., Rogers S., Amiott E., Gibson S., Masibay A., **Analysis and interpretation of short tandem repeat microvariants and three-banded allele patterns using multiple allele detection systems.** J Forensic Sci 44 (1999) 87–94.

Edwards A., Civitello A., Hammond H. A., Caskey C. T., ***DNA* typing and genetic mapping with trimeric and tetrameric tandem repeats.** Am J Hum Genet 49 (1991) 746–56.

Edwards A., Hammond H. A., Jin L., Caskey C. T., Chakraborty R., **Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups.** Genomics 12 (1992) 241–53.

Endo Y., Onogi S., Umeki K., Yamamoto I., Kotani T., Ohtaki S., Fujita T., **Regional localization of the gene for thyroid peroxidase to human chromosome 2p25 and mouse chromosome 12C.** Genomics 25 (1995) 760–761.

Freitas N.S.C., Resque R.L., Ribeiro-Rodrigues E.M., Guerreiro J.F., Santos N.P.C., Ribeiro-dos-Santos A., Santos S., **X-linked insertion/deletion polymorphisms: forensic applications of a 33-markers panel.** Int J Legal Med (2010).

Gill P., **Role of short tandem repeat *DNA* in forensic casework in the UK— past, present, and future perspectives.** BioTechniques 32 (2002) 366–72.

Hares D.R., **Expanding the *CODIS* core loci in the United States**. Forensic
Sci Int Genet: Gen 6 (2012) e52–e54.

Huel R. M., Basic L., Madacki-Todorovic K., Smajlovic L., Eminovic I., Berbic I., Milos A., Parsons T. J., **Variant alleles, tri-allelic patterns, and point mutations observed in nuclear short tandem repeat typing of populations in Bosnia and Serbia.** Croat Med J 48 (2007) 494–502.

Kimpton C. P., Gill P., Walton A., Urquhart A., Millican E. S., Adams M., **Automated *DNA* profiling employing multiplex amplification of short tandem repeat loci.** *PCR* Meth Appl 3 (1993)13–22.

Kimura S., Kotani T., McBride O. W., Umeki K., Hirai K., Nakayama T., Ohtaki S., **Human thyroid peroxidase: complete c*DNA* and protein sequence, chromosome mapping, and identification of two alternately spliced mRNAs.** Proc Natl Acad Sci USA 84 (1987) 5555–5559.

Kimura S., Hong Y. S., Kotani T., Ohtaki S., Kikkawa F., **Structure of the human thyroid peroxidase gene: comparison and relationship to the human myeloperoxidase gene.** Biochemi*STR*y 28 (1989) 4481–4489.

Lane A. B., **The nature of tri-allelic *TPOX* genotypes in African populations.** Forensic Sci Int: Genetics 2 (2008) 134–137.

Lewis R., **Genética Humana: conceitos e aplicações.** 5 ed.,Rio de Janeiro: Guanabara Koogan, 2004. 453p.

Libert F., Ruel J., Ludgate M., Swillens S., Alexander N., Vassart G., Dinsart C., **Thyroperoxidase, an autoantigen with a mosaic structure made of nuclear and mitochondrial gene modules.** EMBO J 6 (1987) 4193–4196.

Lukka M., Tasa G., Ellonen P., Moilanen K., Vassiljev V., Ulmanen I., **Triallelic patterns in *STR* loci used for paternity analysis: Evidence for a duplication in chromosome 2 containing the *TPOX STR* locus.** Forensic Sci Int 164 (2006) 3–9.

Lurie I. W., Ilyina H. G., Gurevich D. B., Rumyantseva N. V., Naumchik I. V., Castellan C., Hoeller A., Schinzel A., **Trisomy 2p: analysis of unusual phenotypic findings.** Am J Med Genet 55 (1995) 229–236.

Mégarbané, A., Souraty, N., Prieur, M., Theophile, D., Chedid, P., Auge, J. et al. **Interstitial duplication of the short arm of chromosome 2: report of a new case and review**. J. Med. Genet. 34, 783-786 (1997).

Magee A. C., Humphreys M. W., McKee S., Stewart M., Nevin N. C., **De novo direct duplication 2 (p12→p21) with paternally inherited pericentric inversion 2p11.2 2q12.2.** Clin Genet 54 (1998) 65–69.

Nata M., Kimura T., Hashiyada M., He P., Yan W., Li X., Funayama M., Sagisaka K., **Allele frequencies of eight *STR*s in Japanese and Chinese.** Int J Legal Med 112 (1999) 396-399.

Park S. M., Chatterjee V. K. K., **Genetics of congenital hypothyroidism.** J Med Genet 42 (2005) 379–389.

Rodrigues E.M.R., Leite F.P.N, Hutz M.H., Palha T.J.B.F., Santos A.K.C.R., Santos S.E.B., **A multiplex *PCR* for 11 X chromosome *STR* markers and population data from a Brazilian Amazon Region.** Forensic Sci Int: Gen 2 (2008) 154–158.

Ribeiro E.M.R.,Santos N.P.C., Ribeiro-dos-Santos A.K.C., Pereira R., Amorim A., Gusmão L., Zago M.A., Santos S.E.B., **Assessing interethnic admixture using an X-linked insertion-deletion multiplex.** Am J Hum Biol 1 (2009):1–3.

Schaffner S.F., **The X chromosome in population genetics.** Nat Rev Genet 5 (2004): 43–51.

*STR*achan T., Read A. P., **Genética Molecular Humana.** 2 ed., Porto Alegre: Artemed, 2002. 543p.

Seto P., Hirayu H., Magnusson R. P., Gestantas J., Portmann L., DeGroot L. J., Rapoport B., **Isolation of a complementary *DNA* clone for thyroid microsomal antigen. Homology with the gene for thyroid peroxidase.** J Clin Invest 80 (1987) 1205–1208.

Siffroi J. P., Molina-Gomez D., Viguie F., Nessmann C., Dadoune J. P., **Prenatal diagnosis of partial 2p trisomy by de novo duplication 2p (13.1→21).** Confirmation by FISH [letter]. Prenat Diagn 14 (1994) 1097–1099.

Winsor S. H., McGrath M. J., Khalifa M., Duncan A. M., **A report of recurrent anencephaly with trisomy 2p23-2pter: additional evidence for the involvement of 2p24 in neural tube development and evaluation of the role for cytogenetic analysis.** Prenat Diagn 17 (1997) 665–669.

Zamir A., Shipitzen M., Oz C., Motro U., Meiner V., Gafny R., **Presentation of a three-banded pattern—analysis and interpretation.** J Forensic Sci 47 (2002) 824–826.