

REDE PRÓ CENTRO-OESTE
UNIVERSIDADE FEDERAL DE GOIÁS
Programa de Pós-Graduação em Biotecnologia e Biodiversidade

DESENVOLVIMENTO DE UM SISTEMA MULTIPLEX DE
***LOCI STRs* AUTOSSÔMICOS POLIMÓRFICOS PARA A**
IDENTIFICAÇÃO HUMANA

Ricardo Goulart Rodovalho

Goiânia
Goiás – Brasil
2017

**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS
DE TESES E
DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG**

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: Dissertação Tese

2. Identificação da Tese ou Dissertação:

Nome completo do autor: Ricardo Goulart Rodovalho

Título do trabalho: DESENVOLVIMENTO DE UM SISTEMA MULTIPLEX DE *LOC* STRS AUTOSSÔMICOS POLIMÓRFICOS PARA A IDENTIFICAÇÃO HUMANA

3. Informações de acesso ao documento:

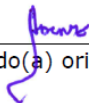
Concorda com a liberação total do documento SIM NÃO¹

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.



Assinatura do(a) autor(a)²

Ciente e de acordo:


Assinatura do(a) orientador(a)²

Data: 11 / 10 / 2017

¹ Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente
- Submissão de artigo em revista científica
- Publicação como capítulo de livro
- Publicação da dissertação/tese em livro

²A assinatura deve ser escaneada.

Ricardo Goulart Rodovalho

**DESENVOLVIMENTO DE UM SISTEMA MULTIPLEX DE
LOCI STRs AUTOSSÔMICOS POLIMÓRFICOS PARA A
IDENTIFICAÇÃO HUMANA**

Tese apresentada à Universidade Federal de Goiás como parte das exigências do Programa de Pós-Graduação em Biotecnologia e Biodiversidade, para obtenção do título de *Doctor Scientiae*.

Orientador: Prof. Dr. Aparecido Divino da Cruz
Co-Orientador: Prof. Dr. Rinaldo Wellerson Pereira

Goiânia
Goiás – Brasil
2017

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Goulart Rodovalho, Ricardo
DESENVOLVIMENTO DE UM SISTEMA MULTIPLEX DE LOCI STRs AUTOSSÔMICOS POLIMÓRFICOS PARA A IDENTIFICAÇÃO HUMANA [manuscrito] / Ricardo Goulart Rodovalho. - 2017.
xv, 101 f.: il.

Orientador: Prof. Dr. Aparecido Divino da Cruz; co-orientador Dr. Rinaldo Wellerson Pereira.

Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Patologia Tropical e Saúde Pública (IPTSP), Programa de Pós graduação em Biotecnologia e Biodiversidade, Goiânia, 2017.

Bibliografia. Anexos. Apêndice.

Inclui siglas, abreviaturas, tabelas, lista de figuras, lista de tabelas.

1. Identificação Humana. 2. Microsatélites. 3. Paternidade. 4. Multiplex. 5. Frequência Alélica. I. Divino da Cruz, Aparecido, orient. II. Título.

CDU 575



UNIVERSIDADE FEDERAL DE GOIÁS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E
BIODIVERSIDADE
Rua 235, S/N - Setor Universitário - Goiânia/GO CEP 74605-050
Fone (62) 3209.6362
email: pgbb.goias@gmail.com

ATA DA REUNIÃO DA BANCA EXAMINADORA DA DEFESA DE TESE DE RICARDO GOULART RODOVALHO - Aos cinco dias do mês de setembro do ano de 2017 (05/09/2017), às 14:00 horas, reuniram-se os componentes da Banca Examinadora: Profs. Drs. APARECIDO DIVINO DA CRUZ, THAÍS CIDÁLIA VIEIRA, FLÁVIA MELO RODRIGUES, KÁTIA KARINA VEROLLI DE OLIVEIRA MOURA e WALTER PINTO JUNIOR, para, sob a presidência da primeira, e em sessão pública realizada na PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS, procederem à avaliação da defesa de tese intitulada: **"DESENVOLVIMENTO DE UM SISTEMA MULTIPLEX DE LOCI STRS AUTOSSÔMICOS POLIMÓRFICOS PARA A IDENTIFICAÇÃO HUMANA."**, em nível de DOUTORADO, área de concentração em BIOTECNOLOGIA, de autoria de RICARDO GOULART RODOVALHO, discente do PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E BIODIVERSIDADE, da Universidade Federal de Goiás. A sessão foi aberta pelo orientador do discente, Prof. Dr. APARECIDO DIVINO DA CRUZ, que fez a apresentação formal dos membros da Banca e orientou o Candidato sobre como utilizar o tempo durante a apresentação de seu trabalho. A palavra a seguir, foi concedida ao autor da tese que, em 30 minutos procedeu à apresentação de seu trabalho. Terminada a apresentação, cada membro da Banca arguiu o Candidato, tendo-se adotado o sistema de diálogo sequencial. Terminada a fase de arguição, procedeu-se à avaliação da defesa. Tendo-se em vista o que consta na Resolução nº. 1181/2013 do Conselho de Ensino, Pesquisa, Extensão e Cultura (CEPEC), que regulamenta o Programa de Pós-Graduação em Biotecnologia e Biodiversidade a Banca, em sessão secreta, expressou seu julgamento, considerando o candidato **Aprovado** ou **Reprovado**:

Banca Examinadora

Prof. Dr. Aparecido Divino Da Cruz

Profa. Dra. Thaís Cidália Vieira

Profa. Dra. Flávia Melo Rodrigues

Profa. Dra. Kátia Karina Verolli de Oliveira Moura

Prof. Dr. Walter Pinto Junior

Aprovado / Reprovado

João Aprovado
Aparecido
Flávia
Kátia Aprovado
Walter Aprovado

Em face do resultado obtido, a Banca Examinadora considerou o candidato HABILITADO, (Habilitado ou não Habilitado), cumprindo todos os requisitos para fins de obtenção do título de DOUTOR EM BIOTECNOLOGIA E BIODIVERSIDADE, na área de concentração em BIOTECNOLOGIA, pela Universidade Federal de Goiás. Cumpridas as formalidades de pauta, às ___h ___min, a presidência da mesa encerrou esta sessão de defesa de tese e para constar eu, ZHARA HELOU RIBEIRO DE CASTILHO, secretária do Programa de Pós-Graduação em Biotecnologia e Biodiversidade lavrei a presente Ata que depois de lida e aprovada, será assinada pelos membros da Banca Examinadora e por mim em duas vias de igual teor. A Banca Examinadora aprovou a seguinte alteração no título da Tese:

Prof. Dr. Aparecido Divino da Cruz (PUC/GO) João
Profa. Dra. Thaís Cidália Vieira (UEG/GO) Thaís
Profa. Dra. Flávia Melo Rodrigues (UEG/GO) Flávia
Profa. Dra. Kátia Karina Verolli de Oliveira Moura (PUC/GO) Kátia
Prof. Dr. Walter Pinto Junior (Unicamp) Walter
Secretário da Pós-Graduação: Zhara Helou R. de Castilho

*Dedico esta tese à instituição que julgo ser
base da minha vida:
Minha Família!*

AGRADECIMENTOS

Primeiramente agradeço à Deus, por todas as experiências de crescimento e por me guiar para seguir adiante com meus objetivos. Agradeço à Ele por sempre manter minha família ao meu lado.

À minha esposa Daniella, pelo amor, carinho, e imenso apoio durante a realização deste projeto. Pela tolerância nos diversos momentos em que estive ausente e, sobretudo, por ser um exemplo de mãe e por se dedicar aos nossos filhos de maneira tão cuidadosa, carinhosa e amorosa.

Aos meus filhos João Pedro e Isabela, “anjos da guarda”, exemplos de amor incondicional e pureza! É por vocês!

Agradeço de coração aos meus pais, Auro e Luciana, pela confiança, amor, dedicação e por investirem e acreditarem tanto em mim. Agradeço também pela transmissão de todos os ensinamentos e valores que me acompanharão para o resto da vida.

Aos meus queridos irmãos Luciano, Aurélio e Murilo, pelo carinho especial, apoio constante, companheirismo e pelos momentos de descontração e felicidade que me fortaleceram para a conclusão desta grande etapa.

Agradeço de coração ao professor, orientador e, principalmente, amigo, Dr. Aparecido Divino da Cruz, pela enorme influência na minha formação, pelas importantes orientações, pelo cuidado paternal eterno, por sua disponibilidade e, acima de tudo, por sua irrestrita atenção nos inúmeros momentos de dificuldade. Agradeço muito por sempre ter acreditado em mim!

Agradeço especialmente ao meu grande amigo Eduardo Leal Rodrigues, pela parceria constante, pelos ensinamentos e conselhos e, principalmente, por sua inestimável disponibilidade em me orientar nos diversos momentos de dificuldade ao longo da execução do trabalho. Muito obrigado!

Agradeço ao Biocroma, em especial ao José Lobo e à Luzia, por possibilitarem que esse projeto fosse executado, além de compreender meus momentos de ausência do laboratório para dedicação aos estudos. Também estendo estes agradecimentos aos colegas de trabalho, pelo apoio e pela compreensão.

Aos membros da banca examinadora, que se sacrificaram para estar presentes e se dispuseram a avaliar e contribuir com importantes sugestões para o trabalho.

Aos meus colegas de curso e do Núcleo de Pesquisas Replicon da PUC Goiás, pelo apoio. Agradeço em especial ao meu amigo Marc Gigonzac, pela colaboração, companheirismo e pela amizade, e à Fernanda Godoy, pela inestimável ajuda em um momento de grande aflição e dificuldade.

À FAPEG, pelo apoio para a realização do projeto.

SUMÁRIO

Lista de Figuras	xi
Lista de Tabelas	xiii
Lista de Abreviaturas	xv
Resumo	xvii
Abstract	xviii
1. INTRODUÇÃO E JUSTIFICATIVA	1
2. REFERENCIAL TEÓRICO	3
2.1 Identificação humana: Um breve histórico	3
2.2 Organização do genoma humano	4
2.2.1 O uso de sequências de DNA repetitivo em identificação humana .	6
2.2.2 Princípios biológicos e matemáticos subjacentes às estimativas de	
vínculo genético	12
2.2.3 Sistemas comerciais de marcadores STRs autossômicos	
atualmente disponíveis para identificação humana	14
3. OBJETIVOS	21
4. ARTIGO CIENTÍFICO	22
5. MATERIAL E MÉTODOS	32
5.1 Teste dos primers individuais e composição do conjunto multiplex	32
5.1.1 Validação das sequências sintetizadas	32
5.1.2 Validação dos primers em reações multiplex	33
5.2 Determinação do custo de amplificação do multiplex desenvolvido	35
5.3 Comprovação da eficiência do sistema multiplex	35
5.3.1 Estudo de casos	35
5.3.2 Estudo de casos complexos que demandam ampliação do painel	
STR	36

6. RESULTADOS E DISCUSSÃO	37
6.1 <i>Loci</i> selecionados	37
6.2 Criação do layout espacial e espectral do painel multiplex	40
6.3 Determinação dos custos de Amplificação e comparação com kits comerciais	41
6.4 Verificação da eficácia do multiplex em situações reais de investigação de vínculo genético	42
6.4.1 Estudo de casos	42
6.4.2 Estudo de casos complexos que demandam ampliação do painel STR	46
6.4.2.a. Reconstrução genética	46
6.4.2.b. Detecção de mutação germinativa	48
6.4.2.c. Identificação de Alelo <i>Dropout</i>	49
7. CONCLUSÕES GERAIS	53
8. REFERÊNCIAS BIBLIOGRÁFICAS	55
9. APÊNDICES	71
Apêndice I: Tabela suplementar - frequência alélica obtida para os 21 <i>loci</i> autossômicos que compõem o sistema multiplex desenvolvido	72
Apêndice II: Termo de consentimento livre e esclarecido	74
Apêndice III: Tabelas alélicas – Estudos de casos	78
Apêndice IV: Produção Científica	92
Apêndice V: Justificativa pela ausência da comparação entre bancos de frequências alélicas	98
10. ANEXOS	100
Anexo I: Quadro com a descrição dos índices estatísticos normalmente aplicados em investigação de vínculo genético	101

LISTA DE FIGURAS

Figura 1. Fluxograma da proporcionalidade da organização do genoma humano, segundo a natureza funcional e a estrutura das sequências de nucleotídeos atualmente conhecidas	5
Figura 2. Ilustração esquemática do modelo clássico de formação do polimorfismo de STRs provocado por <i>slippage</i> da DNA-polimerase durante a replicação de regiões de DNA contendo cernes de repetição em tandem	10
Figura 3. Ilustração de fragmentos moleculares obtidos a partir da amplificação de loci STRs e submetidos à eletroforese convencional (coloração por prata) e eletroforese capilar (utilização de fluorescência)	15
Figura 4. Linha do tempo mostrando as principais metodologias utilizadas no processo de identificação humana pelo DNA	16
Figura 5. Separação dos primers em grupos a partir da marcação específica por fluoróforo	35
Figura 6. Distribuição dos loci STRs nos cromossomos autossômicos selecionados para compor um painel multiplex destinado a identificação humana no Brasil	39
Figura 7. Disposição dos loci na dimensão espacial e no eixo espectral do painel multiplex desenvolvido	41
Figura 8. Eletroferogramas apresentando os genótipos para o <i>locus</i> D8S1179, obtidos a partir da amplificação com o kit comercial PowerPlex® Fusion e com o sistema multiplex desenvolvido	50

Figura 9. Eletroferogramas apresentando os genótipos para o *locus* D5S818, obtidos a partir da amplificação com o kit comercial PowerPlex® Fusion e com o sistema multiplex desenvolvido 51

LISTA DE TABELAS

Tabela 1. Informações gerais referentes aos <i>loci</i> STR presentes inicialmente no CODIS/FBI, incluindo suas localizações cromossômicas e a variação de tamanhos esperada na população	17
Tabela 2. Principais kits comerciais disponíveis atualmente, com seus <i>loci</i> STRs, fluorescências adotadas e as respectivas localizações cromossômicas	19
Tabela 3. Condições de termociclagem utilizadas para a validação das sequências dos primers desenvolvidos para um sistema multiplex de marcadores STRs autossômicos para a identificação humana	33
Tabela 4. Proporções de PCR utilizadas para validação das sequências de primers sintetizados	33
Tabela 5. Condições de termociclagem aplicadas na PCR para amplificação multiplex de um sistema de marcadores STRs autossômicos desenvolvido para fins de identificação humana.	34
Tabela 6. Conjunto de <i>loci</i> STRs autossômicos pré-selecionados e Amelogenina com suas respectivas localizações cromossômicas para a composição de um sistema multiplex para identificação humana	37
Tabela 7. Informações dos microssatélites autossômicos selecionados para compor um painel de marcadores em multiplex e o marcador de determinação do sexo proposto para estudos de vínculo genético em populações humanas	39
Tabela 8. Custo por amplificação do multiplex desenvolvido e dos principais sistemas comerciais disponíveis no Brasil	41

Tabela 9. Estudo de casos e comparação dos resultados obtidos com o PowerPlex® Fusion e o multiplex desenvolvido	44
Tabela 10. Estudo de casos para avaliação da eficácia do painel multiplex de STRs autossômicos como ferramenta complementar para a identificação humana	47

LISTA DE SIGLAS E ABREVIATURAS

- APO** – Alelo paterno obrigatório
- CA** – Estado da Califórnia
- CODIS** – Sistema combinado de índices de DNA (do inglês, *combined DNA index system*)
- DNA** – Ácido desoxirribonucleico
- FBI** – Departamento Federal de Investigação dos Estados Unidos (do inglês, *federal Bureau of investigation*)
- FTA** – Cartão destinado ao armazenamento de sangue e células bucais (do inglês, *fast transient analysis*)
- HE** – Heterozigose Esperada
- HO** - Heterozigose Observada
- HLA** – Antígenos leucocitários de histocompatibilidade
- ICP** – Índice combinado de paternidade
- ICVG** – Índice combinado de vínculo genético
- IFSG** – Sociedade Internacional de Genética Forense (do inglês, *International Society for Forensic Genetics*)
- IP** – Índice de paternidade
- IVG** – Índice de vínculo genético
- LINE** – Elemento espaçador grande (do inglês, *long interspersed elements*)
- MLP** – sondas multi-locais (do inglês, *multi-locus probe*)
- NCBI** – Centro Nacional de Informação Biotecnológica (do inglês, *national center for biotechnology information*)
- NDIS** – Sistema nacional de índice de DNA dos Estados Unidos da América (do inglês, *national DNA index system*)
- NDNAD** – Banco nacional de dados de DNA do Reino Unido (do inglês, *national DNA database*)
- NRW** – Estado de Renânia do Norte-Vestfália
- Pb** – Pares de bases
- PCR** – Reação em cadeia da polimerase (do inglês, *polymerase chain reaction*)
- PD** – Poder de discriminação

PE – Poder de exclusão

PEC – Probabilidade combinada de exclusão

PIC – Conteúdo de informação polimórfica

pmol – Picomoles

PVG – Probabilidade de vínculo genético a posteriori

RFLP – Polimorfismo do comprimento de fragmento de restrição (do inglês, *restriction fragment length polymorphism*)

SINE – Elemento espaçador curto (do inglês, *short interspersed elements*)

SLP – sondas *locus* único (do inglês, *single locus probe*)

STR – Sequencia com repetições curtas em tandem (do inglês, *short tandem repeats*)

TCLE – Termo de consentimento livre e esclarecido

USA – Estados Unidos da América

VNTR – Sequencia com número variável de repetições em tandem (do inglês, *variable number of tandem repeats*)

WI – Estado de Wisconsin

1. INTRODUÇÃO E JUSTIFICATIVA

A análise do DNA com fins de identificação humana tornou-se uma ferramenta importante para revelar a culpabilidade de criminosos, evidenciar inocentes, identificar corpos e restos humanos em desastres em massa, determinar o vínculo parental, além de ser útil na detecção de erros técnicos causados por rotulação incorreta de material biológico, cometidos em laboratórios de patologia clínica (Sarmiento, 2006; Buckleton *et al.*, 2005; Ruitberg *et al.*, 2001).

O desenvolvimento de sistemas com múltiplos STRs (do inglês, *short tandem repeats*), que podem ser amplificados simultaneamente, possibilitou às instituições que realizam procedimentos de identificação humana, alcançar um elevado poder de discriminação entre os investigados, aliado ao baixo consumo de amostra biológica e insumos (Zhao *et al.*, 2015). Dessa forma, os sistemas multiplex permitem a obtenção de um perfil genético eficaz, passível de ser comparado e, portanto, discriminar as amostras biológicas analisadas (Zhang *et al.*, 2015; Butler, 2006; Wallin *et al.*, 2002). A análise simultânea de vários STRs, em paralelo com as tecnologias de detecção por fluorescência multicores, contribuiu para consolidar princípios técnicos importantes subjacentes à determinação da identidade genética (Wang, *et al.*, 2011; Hill *et al.*, 2008; Buckleton *et al.*, 2005).

Segundo Weigand *et al.* (2006), existem situações reais de alta complexidade analítica que exigem um conjunto maior de marcadores moleculares para garantir um poder de discriminação individual que aproxime a verdade biológica da verdade dos vínculos investigados. Nestas situações, o uso de marcadores que permitem aumentar o poder de discriminação dos perfis genéticos colabora para emissão de pareceres conclusivos (Tsai *et al.*, 2013). Por outro lado, a indisponibilidade destes recursos acarreta na inviabilização da investigação (Gill *et al.*, 2006; Gill *et al.*, 1991).

Para Poetsch *et al.* (2013), o número de regiões hipervariáveis disponível para um processo de investigação de vínculo genético está diretamente associado à probabilidade de certeza obtida durante a análise da investigação (Zhang *et al.*, 2015; Borsting, *et al.*, 2008). Sendo assim, é fundamental que os painéis de STRs tenham poder suficiente para estabelecer uma conclusão precisa e segura acerca dos vínculos investigados. Contudo, é evidente a existência de uma correlação muito próxima entre a quantidade e qualidade dos *loci* polimórficos presentes em um sistema STR e o grau de

precisão da metodologia. Assim, quanto maior o número de STRs analisados, maior será o poder de resolatividade do estudo, considerando que as regiões hipervariáveis escolhidas sejam polimórficas na população a que pertencem os investigados (Wiegand, 2006).

Em contraste, as empresas responsáveis pela comercialização dos sistemas moleculares de análise de STRs para identificação humana antagonizam este cenário, uma vez que normalmente a quantidade de *loci* STR presentes em um sistema de amplificação simultânea implica, diretamente, no preço a ser cobrado por este produto. Dessa forma, quanto maior o número de marcadores disponíveis nos kits comerciais de amplificação molecular, maior será seu preço de comercialização.

Todavia, considerando-se o contexto social do Brasil, percebe-se uma incompatibilidade de valores, provocada pela sobreposição dos interesses altamente capitalistas sobre a conduta adequada e segura em uma rotina técnica de identificação humana pelo DNA. Sendo assim, os efeitos desta contraposição estabelecida entre ciência e comércio, no âmbito atual da investigação de vínculo genético, dificulta a administração adequada dos recursos disponíveis no mercado e, conseqüentemente, a oferta do serviço com equidade social.

O desenvolvimento de um sistema multiplex próprio, capaz de discriminar indivíduos de uma população com eficiência e segurança, cria mecanismos que viabilizam uma maior amplitude de oferta, em função da discrepância de custos operacionais. Por consequência, além de proporcionar mais segurança para a análise genética, disponibiliza metodologias de identificação humana e investigação de vínculo biológico para determinados grupos sociais que atualmente não tem acesso aos serviços.

Adiante, estão descritos o desenvolvimento de um sistema multiplex de *loci* STRs, capaz de amplificar simultaneamente 21 *loci* polimórficos autossômicos e um *locus* para a determinação de sexo. Os resultados obtidos mostram que o painel proposto pode ser aplicado para resolução de casos simples e complexos de investigação de vínculo genético e confirmam sua eficiência em estratégias de identificação humana e investigação de vínculo genético pela análise do DNA. A presente tese está dividida em seções contendo um referencial teórico, objetivos, material e métodos, resultados e discussão, um artigo publicado com os achados mais relevantes e conclusão.

2. REFERENCIAL TEÓRICO

2.1 Identificação humana: Um breve histórico

O termo “identidade” refere-se ao conjunto de sinais, propriedades e características, próprias da individualidade de uma pessoa, tornando-a distinta das demais dentro de uma população. Por outro lado, “identificação” corresponde aos diversos processos e métodos responsáveis pelo reconhecimento de uma pessoa, legitimando-a para os devidos fins (Figini *et al.*, 2003).

Para Gastaldo (2012), em seu primórdio, o processo de investigação de vínculo genético se baseava na comparação de características fisionômicas e antropométricas entre os indivíduos em questão. Por ser um método bastante arcaico, apoiado em evidências substancialmente subjetivas, propiciava a perpetuação da dúvida ao longo de muitos anos.

Com a descrição dos tipos sanguíneos do sistema ABO, a partir do início do século XX, passou a ser reconhecida uma das primeiras características humanas herdáveis. Com isso, passou-se a empregar os tipos sanguíneos como marcadores biológicos para a determinação de vínculo por herança genética, sendo realizada com maior precisão e útil para investigações de parentesco, forenses e antropológicas em diferentes populações (Gastaldo, 2012). Porém, segundo Bugert *et al.* (2012), o uso do sistema ABO de forma independente não apresentava padrões informativos e seguros para a investigação de vínculo materno e paterno. Dessa forma, para a obtenção de resultados mais confiáveis, seria necessário adotá-la como alternativa complementar aos demais métodos associados à análise dos sistemas Rh, MNS, *Duffy*, *Kidd* e *Kell* (Kouniaki *et al.*, 2015).

A identificação de um sistema de histocompatibilidade, aplicada à análise de produtos gênicos do complexo HLA (do inglês, *histocompatibility leucocyte antigen*), na década de 50, incrementou a resolutividade das ferramentas aplicadas à Medicina Forense e marcou a segunda era na evolução da Imunogenética (Jobim *et al.*, 1996.) No entanto, a primeira estratégia capaz de excluir acima de 90% dos supostos pais testados surgiu somente em meados da década de 1970, por meio da associação dos haplótipos leucocitários humanos com os genótipos dos outros sistemas sanguíneos (Gastaldo,

2012). Porém, devido ao desequilíbrio de ligação e a predominância de certos alelos do HLA em alguns grupos étnicos, na maioria dos casos, os *loci* HLA não eram suficientemente informativos, exigindo-se assim a necessidade de exploração e proposição de métodos mais seguros e com elevado poder de exclusão para reduzir as chances de resultados falsos-positivos nas análises de vínculo genético entre as pessoas (Ikeda *et al.*, 2015).

Por volta dos anos 80, aproximadamente 100 polimorfismos de proteínas já haviam sido estudados e descritos. No entanto, apresentavam pouca relevância para a aplicação em identificação humana (NIJ, 2000).

A descoberta da estrutura da molécula de DNA, em 1953, por J. Watson e F. Crick, e sua implicação como componente responsável pelo patrimônio genético dos seres vivos, transformou a Biologia e promoveu mudanças importantes em diversas áreas do conhecimento científico. A partir de então, foi possível desenvolver técnicas capazes de caracterizar as particularidades no DNA de cada indivíduo (Bucleton *et al.*, 2005).

Em 1985, o geneticista inglês Dr. Alec Jeffreys iniciou o processo revolucionário para a ciência da identificação humana, ao descobrir algumas regiões do DNA que continham sequências nucleotídicas com repetições consecutivas e que o número destas repetições variava de indivíduo para indivíduo (Jeffreys *et al.*, 1985). Neste contexto, o avanço biotecnológico permitiu subsequentemente a proposição de estratégias de identificação do perfil genético de cada indivíduo, cujas ferramentas tornaram-se produtos bastante inovadoras e confiáveis na prática cotidiana da identificação individual a partir do genoma humano (Zhang *et al.*, 2015; Matanna *et al.*, 2012; Brevnov *et al.*, 2009; Budowle & Van Daal, 2009).

2.2 Organização do genoma humano

A descoberta da estrutura molecular do DNA, em 1953, foi responsável por um expressivo avanço tecnológico em pesquisas envolvendo essa molécula, possibilitando mais tarde, o advento do Projeto Genoma Humano, que proporcionou uma melhor compreensão da organização molecular do genoma humano (Primorac *et al.*, 2000).

O genoma humano nuclear diploide é constituído por 23 pares de cromossomos, cujas sequências correspondem à aproximadamente três bilhões de pares de bases (Levy

et al., 2007). Porém, apenas 2% de todo o genoma humano possuem sequências codificantes para transcrição de proteínas, cerca de 20.000 a 25.000 genes aproximadamente (Clamp *et al.*, 2007; Levy *et al.*, 2007). Para o Bennet (2000), pseudogenes, fragmentos gênicos, regiões intrônicas e promotoras correspondem à 27% do genoma, enquanto, os demais 70% restantes são formados por DNA extragênico. Destes 70%, cerca de 15%, representam uma classe de DNA repetitivo, conforme apresentado na figura 1.

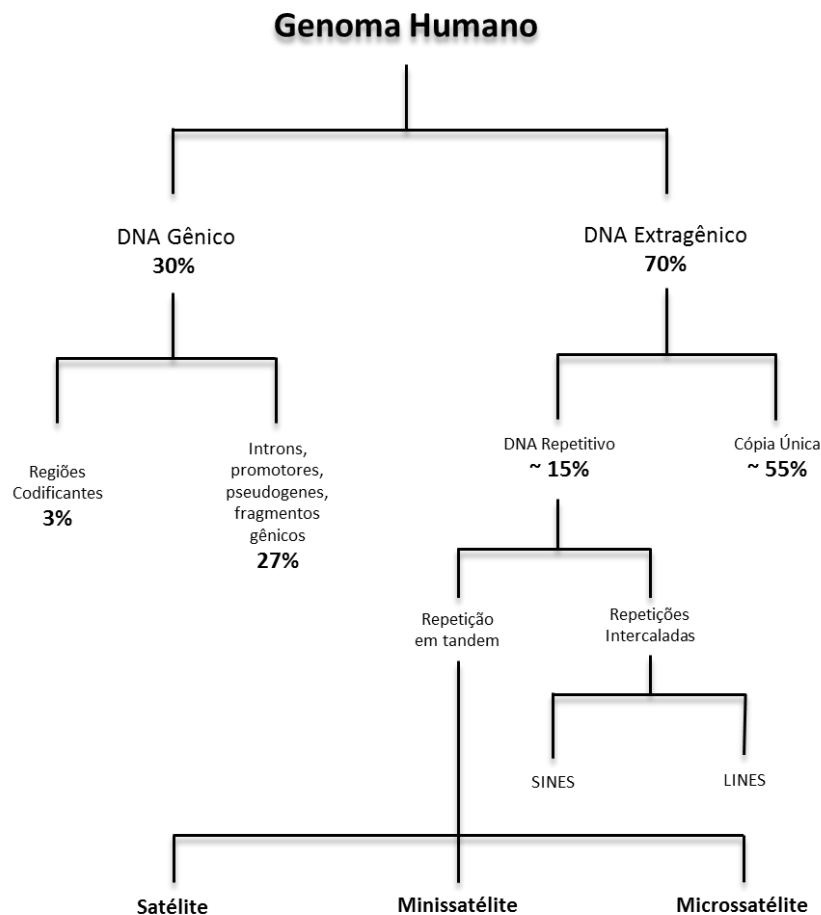


Figura 1. Fluxograma da proporcionalidade da organização do genoma humano, segundo a natureza funcional e a estrutura das sequências de nucleotídeos atualmente conhecidas.

A presença de DNA em células humanas, bem como a alta estabilidade química da molécula, mesmo após um longo período de tempo exposto às intempéries do ambiente, contribuem para que o DNA seja explorado para fins de determinação da identidade genética (Costa *et al.*, 2011; Butler, 2006; Pena, 2005). Para Pena (2005), as propriedades físico-químicas do DNA possibilitam a manipulação das moléculas,

permitindo que o DNA seja extraído, isolado, estocado e, quando necessário, reutilizado. Assim, potencializando o seu uso e oportunizando o desenvolvimento de testes precisos e exatos, que trabalham com o DNA como material biológico.

Ao longo do genoma humano existem segmentos de DNA responsáveis pela variabilidade intraespecífica, capaz de distinguir geneticamente indivíduos de uma população com um elevado grau de certeza (Butler, 2005; Jobling & Gill, 2004). Por estarem, em geral, localizados entre os genes, a variação dos segmentos de DNA não altera o equilíbrio gênico do indivíduo, não interferindo, portanto, na fisiologia e nem no ciclo vital das células (Butler, 2012; Butler, 2005; Ellegren, 2004).

Para Kashyap *et al.* (2004), uma identidade genética pode ser determinada mediante a análise de regiões polimórficas presentes ubiquamente no genoma de uma pessoa (Catasti *et al.*, 1999; Fowler *et al.*, 1988). As regiões polimórficas apresentam variações alélicas interindividuais, com frequência populacional igual ou superior a 1%. Neste sentido, para que um *locus* seja considerado polimórfico, é necessário que seus alelos raros estejam presentes numa frequência maior ou igual a 1% da população (Decanine, 2016).

Segundo Butler & Hill (2012), o polimorfismo genético pode ser classificado em dois grupos: polimorfismo de sequência e de comprimento. O primeiro tipo, polimorfismo de sequência, é caracterizado pela diferente composição nucleotídica em uma determinada localização genômica, sendo normalmente resultado de mutações pontuais ou de pequenas inserções ou de deleções na sequência de nucleotídeos do DNA (Weedn & Swarnen, 1998). Já os polimorfismos de comprimento são variações do número de repetições de uma sequência cerne. Ambos são bastante explorados nos processos de identificação individual (Butler & Hill, 2012; Ellegren, 2004; Nakamura *et al.*, 1987).

Várias técnicas moleculares têm sido utilizadas para a análise de polimorfismos do DNA. Porém, cada uma apresenta suas características próprias e estão vinculadas com a natureza das sequências de estudo (Pena, 2005; Kashyap, 2004).

2.2.1 O uso de sequências de DNA repetitivo em identificação humana

O DNA repetitivo é disperso pelo genoma e se caracteriza por envolver unidades de sequências repetitivas de tamanho variados (Weber, 1990), sendo os elementos de

repetição curtos e longos mais estudados, pertencentes às famílias SINEs (do inglês, *short interspersed elements*) e LINEs (do inglês, *long interspersed elements*), respectivamente (Bennett, 2000; Deininger & Batzer, 1999). SINEs, são elementos espaçadores curtos, que, assim como muitas regiões intercaladas, foram propagados por retrotransposons (Bennett, 2000). Essas sequências, representadas principalmente pelos elementos ALU, apresentam tamanho inferior à 400 pb (GU *et al.*, 2000; Pavlícek *et al.*, 2001). LINEs são sequências longas, variando de 6 a 8 kb, sendo representados principalmente pelos elementos L1, que corresponde à cerca de 15% do genoma humano (Smit, 1996).

O DNA repetitivo em tandem, com localização variada no genoma, apresenta sequências bastante variáveis, sugerindo uma enorme aplicação em análises de paternidade/maternidade biológicas (Butler & Hill, 2012; Lins *et al.*, 1996; Smouse *et al.*, 1986; Gill, *et al.*, 1985). Considerando a ocorrência de variação dos tamanhos médios das unidades de repetições nucleotídicas, denominadas “motivos” ou cernes, o DNA repetido em tandem pode ser dividido em três classes distintas: DNA satélite, minissatélite e microssatélite (Bennett, 2000).

A primeira sequência de DNA em tandem descoberta foi o DNA satélite (Britten & Kohne, 1968). Estas regiões são denominadas regiões satélites, devido a observação de uma ou mais bandas menores (bandas satélites) em experimentos que envolviam equilíbrio de densidade por centrifugação em gradiente (Butler, 2012).

O DNA satélite é encontrado nas regiões de heterocromatina, principalmente próximas ao centrômero. O DNA satélite não é transcrito e apresenta em sua composição extensos blocos de segmentos repetidos (Garrido and Rodrigues, 2015; Bennett, 2000). Seu cerne é formado por sequências simples ou moderadamente complexas, com variação de 5 a 250 pb e atinge um tamanho de expansão total superior à 100 kb (Strachan and Read, 2011; Bennett, 2000). Devido ao seu tamanho, localização restrita no genoma e reduzida variabilidade interindividual, essa classe de DNA repetitivo não é um recurso genético útil nos estudos de identificação humana e ligação genética (Bennett, 2000).

Os minissatélites são encontrados dispersos ao longo do genoma nuclear humano e são caracterizados por conter sequências de DNA hipervariáveis (altamente polimórficas), repetidas em tandem, de tamanho moderado, cujo cernes variam de 8 a 100 pares de bases (Jeffreys *et al.*, 1985a).

O primeiro método de detecção de polimorfismo de DNA, descrito por Wyman e White no início da década de 80 para diagnosticar doenças genéticas, sugeriu existir

regiões no genoma humano que continham muitas variantes (Wyman and White, 1980; Kan and Dozy, 1978). Em função destas regiões gerarem um grande número possível de perfis de DNA, foram consideradas essenciais para análise de identificação humana, sendo, por volta de 1984, utilizada por A. Jeffreys como primeiro teste aplicado para esta finalidade (Baird *et al.*, 1996). Inicialmente, a estratégia proposta por A. Jeffreys baseou-se no uso de sondas multi-locais, MLP (do inglês, *multi-locus probe*), obtidas pela técnica de RFLP (do inglês, *restriction fragment length polymorphism*), que permitiam a análise de muitas sequências de DNA de minissatélites que se repetem consecutivamente, conhecidas por VNTRs (do inglês, *variable number of tandem repeats*) (Garrido, 2009; Jeffreys, 2005; Pena, 2005; Baird *et al.*, 1996; Jeffreys *et al.*, 1985).

O polimorfismo detectado pela técnica de RFLP ocorre em decorrência da existência, em indivíduos geneticamente distintos, de diferentes sequências nucleotídicas ao longo da cadeia de DNA (Balazs *et al.*, 1989). A técnica é baseada na aplicação de enzimas de restrição capazes de clivar o DNA em sítios específicos ao redor das sequências VNTRs (Baird *et al.*, 1996). Em consequência, produz resultados popularmente conhecidos por “impressões digitais” de DNA (Butler, 2005 Nakamura *et al.*, 1987).

Em função da limitação quanto à dificuldade de criação de bancos de dados, de análise e interpretação estatística dos padrões moleculares gerados e de situações envolvendo análise de mistura de padrões genéticos, as sondas multi-locais foram substituídas por sondas de *locus* único, SLP (do inglês, *single locus probe*) (Sharma and Tale, 2016; Besselink, 2003). De acordo com Baird *et al.* (1996), os alelos correspondentes à estas regiões são resultantes do tamanho dos fragmentos, que afetam sua migração em um gel submetido a um campo elétrico constante e uniforme, que promove a separação das bandas em função do tamanho do segmento de DNA nela contido (Butler, 2012; Balazs *et al.*, 1989).

A técnica VNTR/RFLP estava diretamente associada à quantidade e integridade de DNA disponível na amostra biológica analisada. Por outro lado, esta estratégia era inviável para a automação dos procedimentos. Assim, foi rapidamente substituída pela técnica associada à manipulação de regiões microssatélites, por volta de 1993 (Decanine, 2016; Butler, 2005; Thomson *et al.*, 1999).

As regiões microssatélites são caracterizadas por apresentarem sequências repetitivas em tandem de tamanho curto, conhecidas por STRs (do inglês, *short tandem*

repeats) e representam a maioria das sequências hipervariáveis no genoma humano (Butler, 2005; Thomson *et al.*, 1999; Nakamura *et al.*, 1987). Para Butler (2005), os STRs apresentam número de repetições altamente variável entre os indivíduos, sendo sequências muito similares aos minissatélites, diferindo quanto ao tamanho e comprimento das unidades de repetição em tandem, já que estas se caracterizam por sequências menores (Zhou *et al.*, 2017; Butler, 2012; Thomson *et al.*, 1999).

Os marcadores STRs podem ser analisados a partir de uma reação de amplificação por PCR (do inglês, *polymerase chain reaction*), responsável por multiplicar *in vitro* os fragmentos moleculares previamente selecionados (Lane, 2013). Neste sentido, a reação em cadeia da polimerase revolucionou a biologia molecular, pois permitiu o uso de amostra biológica proveniente de qualquer tecido, mesmo estando com DNA parcialmente degradado (Amory *et al.*, 2012; Chemello, 2007).

A principal propriedade da PCR é sua capacidade em obter relativamente grandes quantidades de fragmentos moleculares a partir de pequenas quantidades e concentrações de DNA (Butler, 2012). O aumento da sensibilidade e especificidade, a diminuição do tempo de procedimento, bem como a facilidade de automação, são vantagens que essa técnica apresenta ao compará-la com as tecnologias associadas à utilização de RFLP (Thomson *et al.*, 1999).

Acredita-se que a variação polimórfica em sequências STRs são originadas em decorrência principalmente de falhas enzimáticas provocadas durante o processo de replicação do DNA, conhecidas como *DNA-Polymerase Slippage* (escorregão da polimerase) (Jin *et al.*, 2016; Zhao *et al.*, 2015; Boyer, 2008; Ellegren, 2004). O evento de *slippage* é provocado pelo desligamento da cadeia molde durante o processo de replicação, seguido pela renaturação da cadeia com formação de uma alça. Segundo Lovett *et al.* (1993), esse processo ocorre em função da alta instabilidade genômica dos microssatélites, o que provoca o deslizamento da cadeia de replicação ocasionado pelo aumento ou a diminuição do número de unidades de repetição, como mostrado na Figura 2. As alterações resultantes do processo de *slippage* polimerase podem ser transmitidas para os descendentes caso o erro de replicação ocorra na gametogênese e o sistema de reparo não corrija a falha de pareamento (Bennett, 2000).

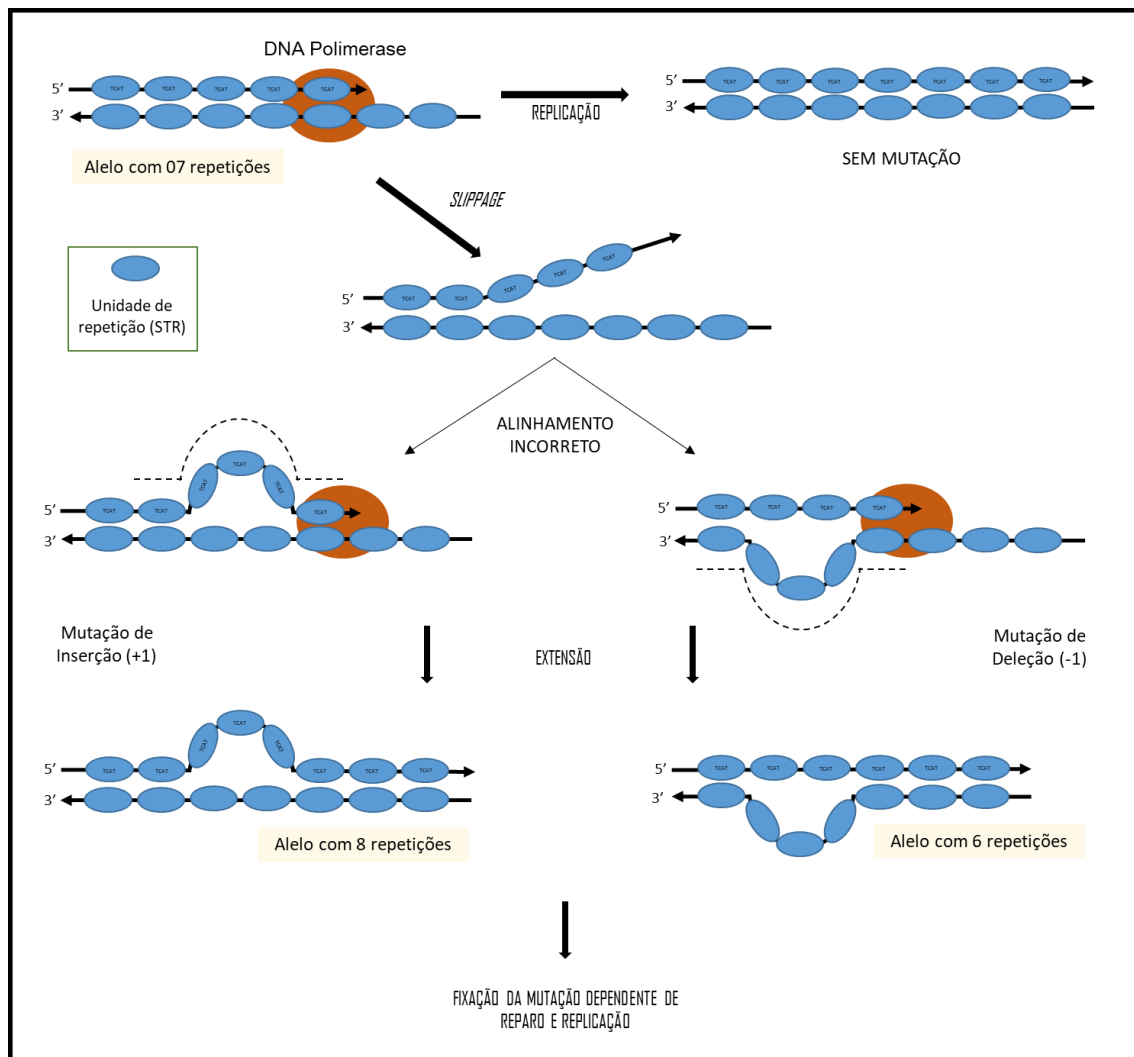


Figura 2. Ilustra o esquem tica do modelo cl ssico de forma o do polimorfismo de STRs provocado por *slippage* da DNA-polimerase durante a replica o de regi es de DNA contendo cernes de repeti o em tandem.

Para Fan and Chu (2007), a recombina o meiotica desigual   outro poss vel mecanismo associado   origem do polimorfismo dos microssat lites. Por m, segundo o mesmo autor, este evento parece n o ser o processo predominante na gera o de variabilidade de STRs. A alta taxa de muta o das sequ ncias microssat lites possibilita in meras aplica es envolvendo mapeamento gen tico, an lise populacional e filogen tica, bem como estudo de identifica o humana (Goldstein and Schlotterer 1999).

A varia o gen tica observada pela an lise de microssat lites entre os indiv duos de uma popula o est  associada ao n mero de vezes em que a unidade de repeti o aparece repetida (Chaudhary *et al.*, 2013). Portanto, a diversidade al lica resultante dessa varia o, a possibilidade de desenvolvimento de sistemas multiplex, a capacidade de

obtenção de fragmentos menores, a praticidade técnica de operacionalização e a robustez dos resultados são os motivos responsáveis pela intensa aplicação dos STRs como marcadores úteis para a identificação humana (Castañeda *et al.*, 2013; Silva *et al.*, 2010; Wilkening *et al.*, 2006; Butler, 2005; Lynch and Ritland, 1999).

As sequências STRs consistem em blocos de repetições que variam de 2 a 6 pares de bases e podem ser nomeadas considerando-se o comprimento da unidade de repetição (Willems *et al.*, 2014; Butler, 2012; Thomson *et al.*, 1999). Dessa forma, repetições dinucleotídicas apresentam dois nucleotídeos repetidos em blocos, enquanto trinucleotídicas apresentam três nucleotídeos (Butler, 2012). O mesmo se aplica para sequências que apresentam quatro, cinco e seis nucleotídeos, sendo nomeadas como tetra, penta e hexanucleotídicas, respectivamente. Teoricamente, existem 4, 16, 64, 256, 1024 e 4096 possíveis motivos de repetição para sequências mononucleotídicas, dinucleotídicas, trinucleotídicas, tetranucleotídicas, pentanucleotídicas e hexanucleotídica, respectivamente (Jin *et al.*, 1994).

Considerando a classificação pelo padrão de repetição, os STRs podem ser divididos em vários grupos, conforme a complexidade de suas sequências (Butler, 2012; Buckleton *et al.*, 2005). Microssatélites com tipo de repetição simples é formado por regiões que contêm unidades de tamanho e sequências idênticas. Já os STRs com repetições compostas são caracterizados por apresentarem duas ou mais repetições simples adjacentes; repetições complexas podem conter diversos blocos de repetição de unidades de tamanho variável entre as sequências; STRs com repetições hipervariáveis complexas apresentam numerosos alelos podendo diferir tanto em tamanho quanto em sequências (Urquhart *et al.*, 1994).

Segundo Butler (2005), alguns *loci* apresentam alelos formados por uma unidade de repetição “incompleta” localizados entre blocos de repetições completos. Estes alelos são chamados de microvariantes alélicas e surgem pela falta de um ou mais nucleotídeos para completar o bloco de repetição. Para Puers *et al.* (1993), a microvariante mais comum nas populações corresponde ao alelo 9.3 do *locus* STR tetranucleotídico TH01, resultante da combinação de nove blocos de repetição completos e um bloco com apenas três nucleotídeos, faltando um único nucleotídeo de adenina para completar a unidade de repetição AATG.

2.2.2 Princípios biológicos e matemáticos subjacentes às estimativas de vínculo genético

A investigação genética parental assumiu um lugar de destaque na ciência forense, tanto no âmbito cível, quanto na jurisdição criminal (Morling *et al.*, 2002). Diante disso, os princípios abordados, as metodologias adotadas, os procedimentos executados, bem como a dinâmica de interpretação e apresentação dos resultados, são fatores que devem estar em constante avaliação, uma vez que a tecnologia empregada pode se tornar obsoleta futuramente e, neste sentido, estabelecer novos paradigmas para esta ciência (Morling *et al.*, 2002).

Testes de paternidade/maternidade, atualmente realizados, consideram a exclusividade genética em determinadas regiões do DNA de cada indivíduo (exceto em gêmeos monozigóticos) e se baseiam nos princípios mendelianos da hereditariedade, que determina a dinâmica de transmissão genética entre os genitores e seus filhos (Zhou *et al.*, 2017; Weber-Lehmann *et al.*, 2014; Zhang, *et al.*, 2014). Neste contexto, a Sociedade Internacional de Genética Forense (ISFG), em 2002, estabeleceu os parâmetros mínimos necessários para a adequada investigação de vínculo biológico usando o DNA dos investigados (Gjertson *et al.*, 2007; Morling *et al.*, 2002).

As situações mais comuns envolvendo investigação de paternidade são realizadas a partir da análise mínima de 15 loci STRs distribuídos pelos cromossomos, o que seria suficiente para a obtenção de uma probabilidade de paternidade superior ou igual à 99,9% ou detecção de 4 ou mais exclusões alélicas (GEKO, 2013; Carboni *et al.*, 2011; Junge *et al.*, 2006). Porém, para Carboni *et al.* (2011), em algumas situações, a análise de apenas 15 marcadores moleculares pode não ser suficiente para a conclusão com segurança acerca do vínculo genético entre as partes investigadas. A detecção de poucas inconsistências observadas entre os indivíduos testados e a ausência da participação no exame da mãe biológica do suposto filho, são situações complexas que requerem a ampliação do número dos marcadores STRs para se atingir as probabilidades adequadas para os casos estudados (Akhteruzzaman *et al.*, 2012; Carboni *et al.*, 2011; Li *et al.*, 2011).

Seguindo-se a recomendação da ISFG, a evidência genética deve ser aferida com base nos princípios da razão de verossimilhança, que testa duas hipóteses mutuamente excludentes para obtenção do índice de paternidade (IP): a hipótese de que o homem

testado é o pai biológico (H1) “x” a hipótese de que o pai biológico é um homem aleatório na população (H2) (Gjertson *et al.*, 2007; Morling *et al.*, 2002). Neste sentido, a investigação de vínculo genético, *a priori*, é realizada com base na análise comparativa dos perfis genéticos dos indivíduos testados, a partir da identificação dos APOs, alelos paternos/maternos obrigatórios (Akhteruzzaman *et al.*, 2012). Portanto, em situações onde se identifica correspondência genética, para estimar o “peso” da evidência genética observada, os índices de vínculo genético são estimados considerando a configuração genotípica dos indivíduos e a frequência alélica populacional (Baird *et al.*, 1996). Os índices normalmente aplicados em investigação de vínculo genético pela análise do DNA estão descritos no Anexo I.

Por outro lado, a ausência de correspondência alélica entre os indivíduos testados, exceto pela ocorrência de mutação, caracteriza-se como exclusão de vínculo genético (Baird *et al.*, 1996). Para GEKO (2013), a exclusão do vínculo biológico direto pode ser aferida com base na detecção de no mínimo 4 *loci* inconsistentes entre os perfis genéticos analisados dos indivíduos investigados.

A identificação de mutações germinativas em marcadores STRs, provocadas pela perda ou ganho de uma unidade de repetição, é um fenômeno bastante comum em testes de identificação humana pela análise do DNA e tem sido amplamente reportado (Jin *et al.*, 2016; Mortera *et al.*, 2016; Forster *et al.*, 2015; Zhang *et al.*, 2014; Sun *et al.*, 2012; Li *et al.*, 2011; Jacewics *et al.*, 2004).

Alguns autores relataram a detecção de dois ou mais eventos de mutações germinativas em testes de paternidade. Em todas as situações descritas a paternidade biológica foi confirmada a partir de análises complementares (Lindner *et al.*, 2014; Sun *et al.*, 2012; Carboni *et al.*, 2011; Li *et al.*, 2011; Jacewics *et al.*, 2004). Li *et al.* (2011) observaram que 98,2% das mutações identificadas foram provocadas pela adição ou diminuição de apenas uma unidade de repetição, enquanto as mutações relacionadas com 2 ou 3 unidades correspondem a 1,6% e 0,2%, respectivamente. Neste sentido, a observação de eventos de mutação germinativa durante uma análise de vínculo genético deve ser sempre procedida pela ampliação dos *loci* utilizados para a interpretação adequada dos resultados (Sun *et al.*, 2015; Lindner *et al.*, 2014; Yuan *et al.*, 2012). Adicionalmente, uma análise criteriosa da natureza da incoerência alélica e o APO observado na criança deve ser realizada, uma vez que as mutações germinativas

envolvem, na maioria das vezes, a alteração de uma única unidade de repetição do STR (Sun *et al.*, 2012).

Segundo Borovko *et al.* (2015), a análise complementar com *loci* adicionais também pode ajudar a superar problemas relacionados a falhas de amplificação alélica. A eficiência da amplificação alélica pode ser influenciada em função de inúmeros fatores (Moretti *et al.*, 2016; Zhang *et al.*, 2016; Westen *et al.*, 2014). Mutações nas regiões flangeadoras do *locus*, presença de inibidores de PCR na reação, qualidade e quantidade de DNA disponível para a PCR, são eventos que podem implicar na inibição da amplificação do fragmento molecular e provocar o efeito de alelo *dropout* (Li *et al.*, 2017; Gjertson *et al.*, 2007; Grubwieser *et al.*, 2006). Portanto, a amplificação do mesmo *locus* por primers que se anelam em sequências diferentes pode contribuir para confirmação do efeito da mutação (Zhang *et al.*, 2016; BUDOWLE 2000).

2.2.3 Sistemas comerciais de marcadores STRs autossômicos atualmente disponíveis para identificação humana

Desde a década de 1990, já era notória a evolução dos métodos aplicados à análise de regiões STRs para identificação humana (Edwards *et al.*, 1991). A amplificação destes marcadores progrediu de reações com *loci* individuais (*singleplex*) para reações multiplex, com capacidade de amplificação de um conjunto superior a 20 *loci* simultaneamente dentro do mesmo tubo, como mostrado na Figura 3 (Butler, 2012). Para Butler (2010), as tecnologias empregadas na detecção dos *amplicons* produzidos passou de métodos manuais radioativos, a obtenção de “manchas” de prata, a detecção automatizada com o uso de fluoróforos. Avanços biotecnológicos e da bioengenharia promoveram progresso expressivo nas metodologias e possibilitou o desenvolvimento de uma ferramenta genética poderosa para os estudos de vínculo genético nas populações (Butler, 2010). O desenvolvimento de novas técnicas aplicadas à identificação humana pela análise do DNA permitiu, desde então, o arquivamento de informações genéticas de indivíduos e de seus familiares.

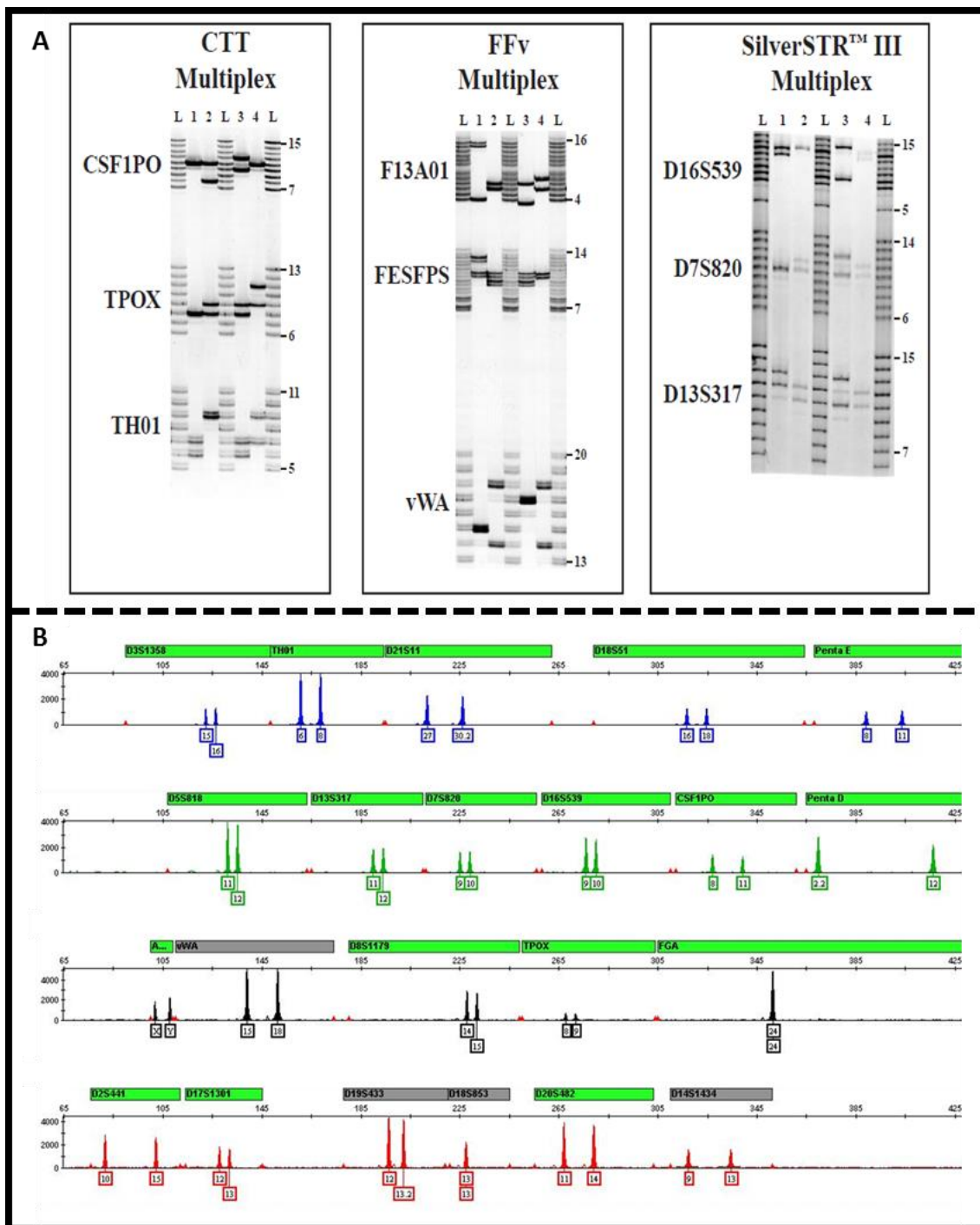


Figura 3. Ilustração de fragmentos moleculares obtidos a partir da amplificação de *loci* STRs e submetidos à eletroforese. Em **A**: Gel de poliacrilamida convencional, corado com nitrato de prata, mostrando a amplificação de 3 sistemas tríplex (3 *loci* STR amplificados simultaneamente). **B**: Sistema multiplex com amplificação simultânea de 21 *loci* STRs autossômicos com eletroforese capilar e análise a partir de software computacional. Fonte: Adaptado de LINS *et al.*, 1996.

A incontestável evolução das metodologias aplicadas à identificação pela análise do DNA, possibilitou que as informações genéticas, obtidas de *loci* presentes em todos os cromossomos autossômicos e sexuais, fossem aplicadas em pesquisas com diversos

propósitos, como por exemplo, construção e manutenção de bancos forenses, análise de ancestralidade e de diversidade genética populacional (Vieira *et al.*, 2014). A Figura 4 ilustra a linha do tempo com as principais metodologias utilizadas no processo de identificação humana envolvendo a análise do DNA.

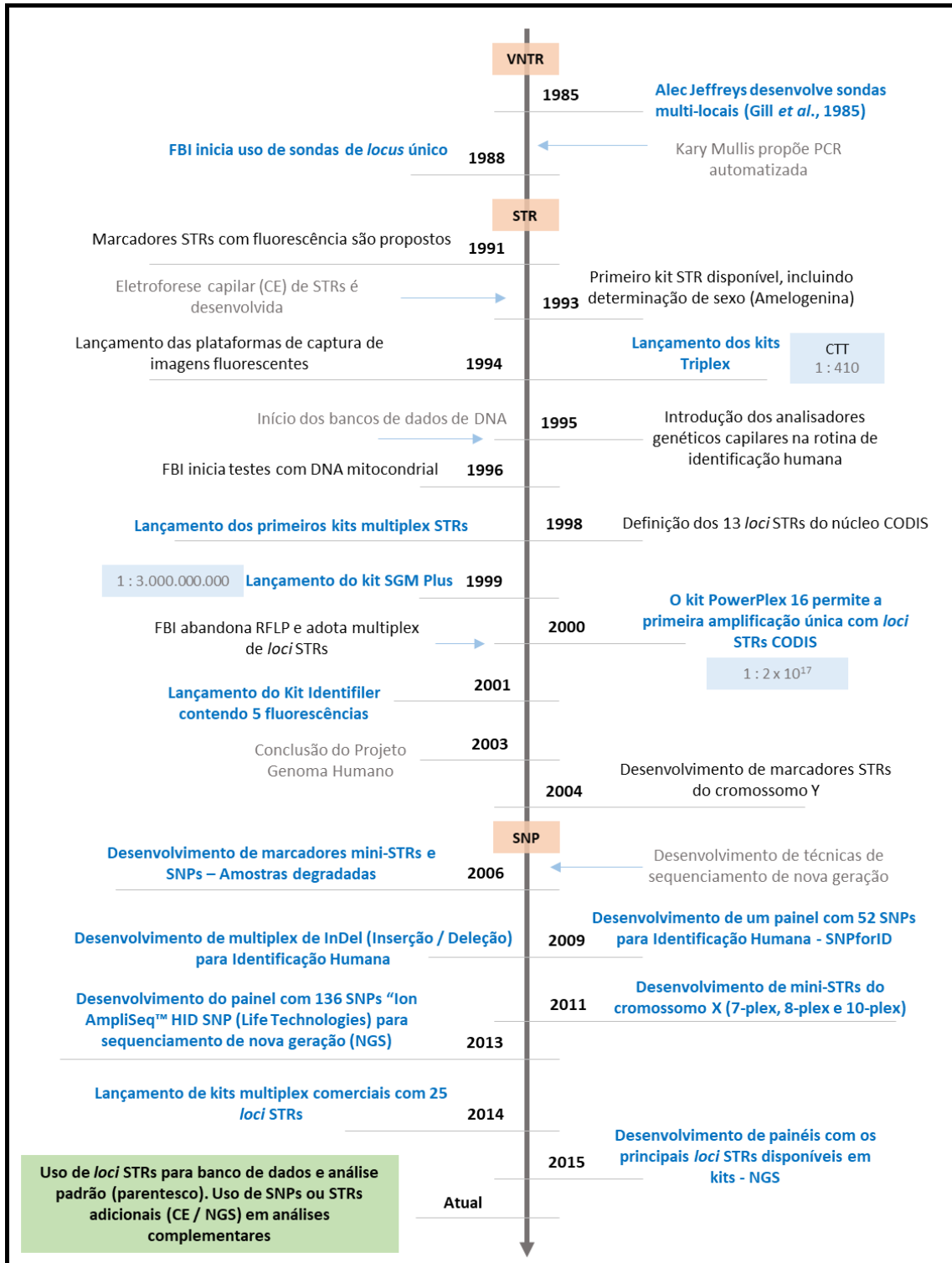


Figura 4. Linha do tempo mostrando as principais metodologias utilizadas no processo de identificação humana pelo DNA

Como pioneiros da constituição dos bancos de dados para auxílio em investigação criminal, o Reino Unido e os Estados Unidos da América, em 1994, criaram, respectivamente o NDNAD (do inglês, *National DNA Database*) e o NDIS (do inglês, *National DNA Index System*), sendo este último estabelecido pelo FBI (*Federal Bureau of Investigation*) (Garrido and Rodrigues, 2015; Butler, 2012). Os bancos de dados de perfis genéticos agregam enorme contribuição para a resolução de crimes, sobretudo homicídios e estupros, criando inovadoras formas de investigação.

Por volta de 1998, foi implantado pelo FBI, um *software* para comparar os perfis genéticos de amostras questionadas com outros perfis já pré-existentes no NDIS (Butler, 2005). O *software*, batizado por CODIS (do inglês, *combined DNA index system*), usou inicialmente um conjunto padrão de treze marcadores STRs autossômicos tetranucleotídicos altamente polimórficos (Garrido and Rodrigues, 2014; Katsanis and Wagner, 2013; Butler and Hill, 2012). Este conjunto de marcadores do CODIS/FBI, cujos *loci* estão descritos na Tabela 1, foi universalmente adotado como o mínimo de marcadores para estudos de identificação humana, o que favoreceu a padronização dos procedimentos e o compartilhamento de dados em âmbito internacional (Garrido and Rodrigues, 2014; Katsanis and Wagner, 2013).

Tabela 1. Informações gerais referentes aos *loci* STR presentes inicialmente no CODIS/FBI, incluindo suas localizações cromossômicas e a variação de tamanhos esperada na população.

<i>Locus</i>	Localização Cromossômica	Tamanho (Mb)	Diversidade Alélica
D3S1358	3p21.31	45,543	6 – 20
FGA	4q28	155,866	14 – 46,2
D5S818	5q23.2	123,139	5 – 19
CSFIPO	5q33.1	149,436	5 – 19
D7S820	7q21.11	83,433	6 – 17
D8S1179	8q24.13	125,976	3 – 18
TH01	11p15.5	2,149	4 – 17
vWA	12p13.31	5,963	8 – 25
D13S317	13q31.1	81,620	4 – 16
D16S539	16q24.1	84,944	3 – 17
D18S51	18q21.33	59,100	5 – 27
D21S11	21q21.1	19,476	24 – 38
TPOX	2p25.3	1,472	5 – 17

O *locus* Amelogenina não está incluído. A diversidade alélica foi baseada na análise de recentes estudos de frequência alélica para composição de bancos de dados populacionais. Fonte: Modificado de Ge *et al.*, 2012.

Segundo Garrido and Rodrigues (2014), o padrão CODIS para obtenção de perfil genético, recomendado pelo FBI, é suficiente para a individualização de uma pessoa e para a investigação de vínculo parental. Esta estimativa é baseada na probabilidade de ocorrência de um perfil genético, construído com os genótipos mais frequentes de cada um dos 13 *loci*, correspondente a 1 em 160 bilhões (Garrido and Rodrigues, 2014). Entretanto, conforme Wilkening *et al.* (2006), a obtenção da estimativa de parentesco com precisão suficiente está intrinsicamente relacionada com as características dos marcadores moleculares analisados. Neste sentido, o número de STRs polimórficos, bem como o número de alelos por marcador e a frequência alélica deste marcador na população são fatores críticos que devem ser considerados durante a realização de uma investigação genética.

Um estudo, por exemplo, envolvendo uma análise comparativa de perfil genético, utilizando o sistema comercial *AmpFISTR® Identifiler® PCR Amplification Kit* (Thermo Fisher Scientific, Foster City, CA, United States) com 10.000 tios e sobrinhos, pares de “tio-sobrinhos”, mostrou que 3,3% destes pares apresentam genótipo completamente compatível entre si (Thomson *et al.*, 2001).

Zhang *et al.* (2014), analisando transmissão genética em exames de investigação de vínculo genético realizados com o sistema comercial *AmpFISTR® Identifiler® PCR Amplification Kit* observou que, ao retirar o genótipo da mãe em um caso de exclusão de paternidade de configuração TRIO (mãe, suposto filho e suposto pai), o suposto pai apresentou correspondência genética com o suposto filho em todos os *loci* analisados. Ao descrever este evento o autor evidencia dois grandes elementos indispensáveis para realização de investigação de vínculo genético: a implicação da limitação quanto ao número de marcadores disponíveis, e sobretudo, a importância da participação da mãe para composição do quadro alélico e discriminação do perfil genético de origem paterna.

Neste sentido, em 2010, o núcleo de desenvolvimento do CODIS, visando minimizar a probabilidade de *match* ao acaso, resolveu aumentar o poder discriminatório para situações de identificação humana e de vínculo genético, além de intensificar a compatibilidade para correspondência internacional de dados, concluiu que a ampliação dos painéis STRs seria de extrema necessidade (Li *et al.*, 2017; Hares, 2015; Katsanis and Wagner, 2013; Ge *et al.*, 2012; Hares, 2012a).

Concomitante a isso, vários critérios foram adotados para a seleção dos novos STRs, dentre os quais se destacam: a não associação das características de interesse

médico; alto poder de discriminação; baixa taxa de mutação; ausência de *linkage*, utilização pelas comunidades internacionais e cumprimento dos padrões de qualidade adotados pelo FBI (Garrido and Rodrigues, 2014; Ge *et al.*, 2012).

Tabela 2. Principais kits comerciais disponíveis atualmente, com seus *loci* STRs, fluorescências adotadas e as respectivas localizações cromossômicas.

PROMEGA®		LIFE TECHNOLOGIES®		QIAGEN®		
PowerPlex Fusion 6C		GlobalFiler		Investigator 24plex QS Kit		
<i>locus</i>	Rótulo Fluorescência	<i>locus</i>	Rótulo Fluorescência	<i>locus</i>	Rótulo Fluorescência	localização Cromossômica
D3S1358	FL-6C	D3S1358	6-FAM	D3S1358	6-FAM	3p21.31
D1S1656	FL-6C	D1S1656	SID	D1S1656	BTG	1q42
D2S441	FL-6C	D2S441	NED	D2S441	BTR2	2p14
D10S1248	FL-6C	D10S1248	SID	D10S1248	BTY	10q26.3
D13S317	FL-6C	D13S317	TAZ	D13S317	BTP	13q31.1
Penta E	FL-6C	-	-	-	-	15q26.2
D16S539	JOE-6C	D16S539	6-FAM	D16S539	BTP	16q24.1
D18S51	JOE-6C	D18S51	VIC	D18S51	BTR2	18q21.33
D2S1338	JOE-6C	D2S1338	SID	D2S1338	BTY	2q35
CSF1PO	JOE-6C	CSF1PO	6-FAM	CSF1PO	BTP	5q33.1
Penta D	JOE-6C	-	-	-	-	21q22.3
TH01	TMR-6C	TH01	NED	TH01 (TC11)	6-FAM	11p15.5
vWA	TMR-6C	vWA	6-FAM	vWA	6-FAM	12p13.31
D21S11	TMR-6C	D21S11	VIC	D21S11	6-FAM	21q21.1
D7S820	TMR-6C	D7S820	TAZ	D7S820	BTP	7q21.11
D5S818	TMR-6C	D5S818	TAZ	D5S818	BTP	5q23.2
TPOX	TMR-6C	TPOX	6-FAM	TPOX	BTG	2p25.3
D8S1179	CXR-6C	D8S1179	VIC	D8S1179	BTY	8q24.13
D12S391	CXR-6C	D12S391	SID	D12S391	BTG	12p12
D19S433	CXR-6C	D19S433	NED	D19S433	BTY	19q12
SE33	CXR-6C	SE33	TAZ	SE33 (ACTBP2)	BTG	6q14
D22S1045	CXR-6C	D22S1045	TAZ	D22S1045	BTY	22q12.3
DYS391	TOM-6C	DYS391	VIC	DYS391	BTG	Y
FGA	TOM-6C	FGA	NED	FGA (FIBRA)	BTR2	4q28
DYS576	TOM-6C	-	-	-	-	Y
DYS570	TOM-6C	-	-	-	-	Y
-	-	Y indel	VIC	-	-	Y
-	-	-	-	QS1	BTP	
-	-	-	-	QS2	BTP	
Amelogenina	FL-6C	Amelogenina	VIC	Amelogenina	6-FAM	Xp22.1 - 22.3 e Y

Fonte: Adaptado de Oostdik *et al.*, 2014; Hennessy *et al.*, 2014; Kraemer *et al.*, 2017.

Para atendimento ao padrão de ampliação proposto pelo sistema CODIS, as indústrias de biotecnologia estenderam o conjunto de *loci* disponíveis nos seus kits comerciais, agregando assim enorme poder de discriminação e robustez aos sistemas utilizados para estratégias de identificação humana (Kraemer *et al.*, 2017; Hares, 2015;

Wang *et al.*, 2015; Oostik *et al.*, 2014; Hares, 2012b). Neste contexto, visando uma unificação prevista pelo CODIS para compartilhamento de dados internacionais, os *loci* componentes dos painéis atualmente utilizados apresentam grande sobreposição entre os kits comercialmente disponíveis (Tabela 2), incluindo aqueles comercializados no território nacional (Hares, 2015). Dessa forma, um sistema *in house* desenvolvido para identificação humana que contemple outros *loci* polimórficos não presentes nestes conjuntos comerciais, poderia acrescentar recursos consideráveis, tendo notável valor para situações envolvendo investigações complexas que necessitem explorar informações genéticas adicionais (Zhu *et al.*, 2014).

Muitos cientistas forenses e especialistas em identificação humana pelo DNA raramente utilizam sistemas desenvolvidos *in house*, optando sempre por adquirir kits comerciais. Para Butler (2012), esta medida se justifica pelo fato destes produtos apresentarem qualidade controlada, sendo certificados e validados internacionalmente. Porém, alguns estudos já demonstraram que, a partir da aplicação de testes de sensibilidade e estudos de concordância em sistemas com microssatélites desenvolvidos *in house*, os resultados obtidos foram satisfatórios, confirmando assim sua robustez e eficácia para estratégias de identificação humana e investigação de vínculo genético (Li *et al.*, 2017; Zhang *et al.*, 2015; Zhu *et al.*, 2015; Pacheco, 2010).

3. OBJETIVOS

3.1 Objetivo Geral

Desenvolver um sistema STR multiplex, que inclua 21 *loci* polimórficos autossômicos altamente informativos e um marcador para determinação sexual, que seja eficiente para estudos de identificação humana.

3.2 Objetivos específicos

- Selecionar os *loci* a serem incluídos no sistema;
- Propor oligonucleotídeos iniciadores passíveis de serem combinados em uma PCR multiplex;
- Padronizar as condições de amplificação e genotipagem para o referido conjunto multiplex;
- Avaliar a sensibilidade e concordância da amplificação de DNA com a PCR multiplex;
- Implantar o sistema multiplex em rotina laboratorial de identificação humana e investigação de vínculo genético e apresentar resultados que demonstre a eficiência do sistema multiplex em situações reais de investigação de vínculo genético;
- Construir um banco de frequências alélicas correspondente aos *loci* do sistema multiplex, a fim de se obter padrões estatísticos que se aproximem da realidade populacional;
- Determinar o custo de amplificação do multiplex desenvolvido e comparar com sistemas comerciais atualmente disponíveis no Brasil.

4. ARTIGO CIENTÍFICO



Development of a polymorphic short tandem repeat locus multiplex system for efficient human identification

R.G. Rodovalho^{1,2}, E.L. Rodrigues⁷, G.S. Santos², L.M. Cavalcanti², P.R. Lima², A.G. Rodovalho³, R.G. Vital⁴, M.A.D. Gigonzac^{1,5,6} and A.D. da Cruz^{1,5,6}

Genet. Mol. Res. 16 (2): gmr16029557

Received November 29, 2016

Accepted March 8, 2017

Published April 5, 2017

DOI <http://dx.doi.org/10.4238/gmr16029557>

Development of a polymorphic short tandem repeat locus multiplex system for efficient human identification

R.G. Rodovalho^{1,2}, E.L. Rodrigues³, G.S. Santos², L.M. Cavalcanti², P.R. Lima², A.G. Rodovalho³, R.G. Vital⁴, M.A.D. Gigonzac^{1,5,6} and A.D. da Cruz^{1,5,6}

¹Programa de Pós-Graduação em Biotecnologia e Biodiversidade, Rede Pró Centro-Oeste, Goiânia, GO, Brasil

²Laboratório Biocroma - Clínica de Exames de DNA, Goiânia, GO, Brasil

³Programa de Pós-Graduação em Gestão Organizacional - Mestrado Profissional, Universidade Federal de Goiás, Catalão, GO, Brasil

⁴Instituto Federal de Educação, Ciência e Tecnologia Goiano, Campus Rio Verde, Rio Verde, GO, Brasil

⁵Departamento de Biologia, Núcleo de Pesquisas Replicon, Pontifícia Universidade Católica de Goiás, Goiânia, GO, Brasil

⁶Laboratório de Citogenética Humana e Biologia Molecular, Secretaria Estadual de Saúde, Goiânia, GO, Brasil

⁷Promega Corporation, Departamento de Identificação Humana, São Paulo, SP, Brasil

Corresponding author: R.G. Rodovalho
E-mail: ricardogrodovalho@hotmail.com

Genet. Mol. Res. 16 (2): gmr16029557

Received November 29, 2016

Accepted March 8, 2017

Published April 5, 2017

DOI <http://dx.doi.org/10.4238/gmr16029557>

Copyright © 2017 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

ABSTRACT. This study aimed to develop a short tandem repeat (STR) multiplex system, made up of 22 highly informative STR loci, for application in forensic genetics. The system comprised 21 polymorphic autosomal loci (D3S1358, TH01, D21S11, D18S51, Penta E, D5S818,

2015). Meanwhile, several other highly polymorphic STR loci, neither a subset of CODIS nor commercially available, were capable of considerably raising the power of discrimination in analysis (Opel et al., 2007; Hill et al., 2009).

According to Wiegand et al. (2006), several real situations of high analytical complexity require a molecular arsenal focused on the power of individual discrimination. In these situations, the application of methods that increase the power of discrimination corroborates the issuance of conclusive opinions. On the other hand, the unavailability of these resources prompts the unfeasibility of analysis (Gill et al., 1991, 2006).

In this sense, Poetsch et al. (2013) concluded that the number of hypervariable loci available for a genetic linkage analysis is directly associated with the percentage of certainty established during the assessment of the analysis (Børsting et al., 2008; Rodvalho et al., 2015; Zhang et al., 2015). Therefore, it is critical that STR panels provide sufficient security, permitting an accurate and unfailing conclusion regarding the genetic linkage analysis. However, both a very close correlation between the quantity and quality of the polymorphic loci present in an STR system, and the degree of accuracy of the methodology, are evident. Thus, the greater the number of STRs analyzed, the greater would be the resolving power of the study (Wiegand et al., 2006).

This study aimed to develop a new STR multiplex system comprising CODIS and non-CODIS highly polymorphic loci for simultaneous analysis, and to assess its efficiency for forensic application. The loci introduced into the system are distinguished from commercial systems, to allow the introduction of new markers in complex genetic analyses that require complementation. Thus, this panel comprised 18 loci commonly present in commercial systems, allowing for the complementary analysis required in situations that include null alleles.

MATERIAL AND METHODS

Selection of the STR loci

This study selected 22 STR loci coupled with the layout of the desired multiplex project, based on literature reviews from the National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov>) and STRBase (National Institute of Standards and Technology, Chemical Science & Technology Laboratory; <http://www.cstl.nist.gov/strbase/>), with the following criteria: location on different chromosomes (physical distance of at least 10 Mb from another locus if on the same chromosome), power of discrimination, allelic diversity, amplicons size, and presence in compatible databases.

Obtaining sequencing primers

The sequencing primers for the selected loci were obtained using the online Primer3 software (<http://bioinfo.ut.ee/primer3-0.4.0/>) (Rozen and Skaletsky, 2000), so that the size of the amplified products would be within the range established for the entire planned system. Then, the AutoDimer software (<https://www-s.nist.gov/dnaAnalysis/>) (Vallone and Butler, 2004) was used to assess sequences that could potentially self-anneal to form primer dimers, an event known as hairpin formation. Genome alignment analysis was performed using Primer-BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>), an NCBI tool used to identify locations in the genome where spurious annealing may occur beyond the target DNA.

Obtaining primers and developing the multiplex assay

The allelic variations for each STR locus were analyzed on the basis of literature reviews related to the dimensions that each molecular marker may take because of insertions of rare alleles, which are at the ends of the STR locus. These precautions were taken to avoid possible overlaps between neighboring loci and components of the same fluorescence channels.

The locus organization in the panel was designed by considering the schematic multiplex arrangement with the candidate positions for each locus. The architecture of the system layout allowed for estimation of the spacing between the loci and calculation of the size required for each PCR product according to its reference sequence.

All primers were synthesized by Applied Biosystems (Foster City, CA, USA), and the forward sequence of each STR was labeled with one of the following dye sets: 6-FAM™, VIC®, NED™, and PET® (Applied Biosystems), as shown in Table 1.

Amplification using PCR and detection of genetic profiles

The PCRs for all the developed STRs were performed in singleplex assays to assess the behavior of the primer and ensure that the amplicon corresponded to the expected size. Upon validation of all the individual STRs, amplification in a single-multiplex PCR was carried out. For this, a primer mix was prepared according to the individual concentrations of each primer, established on the basis of the joint amplification signal strength. For the multiplex PCR, a Qiagen® Multiplex PCR Kit (Qiagen, Hilden, Germany) was used together with the primer mix and 5.0 ng genomic DNA, in a final reaction volume of 25 µL. The thermocycling conditions were as follows: an initial incubation at 95°C for 5 min; 30 cycles of 94°C for 30 s, 60°C for 90 s, and 72°C for 60 s; and a final extension at 60°C for 30 min.

Capillary electrophoresis of the PCR products preceded the preparation step with 1 µL of amplified product, 8.5 µL Hi-Di™ formamide (Applied Biosystems), and 0.5 µL GeneScan™ - 600 LIZ® Size Standard v2.0 (Applied Biosystems). The separation and detection of fragments were performed with the ABI 3500 Genetic Analyzer (Applied Biosystems) using POP4 (Applied Biosystems). The samples were injected electrokinetically at 1.2 kV for 15 s. The electrophoresis run time for allele separation was 1210 s at 15 kV and 60°C. The allelic fragments obtained were analyzed with GeneMapper®ID-X version 1.2 (Applied Biosystems). The overall performance of the multiplex system was assessed using the 2800M-positive control DNA (Promega Corporation, Madison, WI, USA).

Sensitivity and concordance test

The sensitivity of the multiplex system was assessed by decreasing the amount (10, 5, 3, 1, and 0.5 ng) of the 2800M-positive control DNA (Promega Corporation) in the PCR. The sensitivity limit was based on the amplification having at least 80 relative fluorescence units (RFU). The DNA from blood samples, buccal swabs, and hair was extracted using the QIAamp® DNA Investigator Kit (Qiagen), and amplified to assess the sensitivity and efficiency of the multiplex for different samples. The direct amplification from blood samples was also assessed for the developed multiplex using the Whatman® FTA® card.

The PowerPlex® Fusion System (Promega Corporation) was used to perform concordance tests between the common loci in the two systems (D3S1358, TH01, D21S11,

D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, amelogenin, vWA, D8S1179, TPOX, FGA, D2S441, and D19S433). The other non-coincident loci (D17S1301, D18S853, D20S482, and D14S1434) were not assessed, because they were not present in any kit currently commercially available in Brazil.

Population analysis of STR polymorphism

Peripheral blood was collected by digital puncture from 450 unrelated individuals living in the State of Goiás, Brazil. The blood was impregnated on Whatman® FTA® cards for further amplification with the newly developed multiplex set and the PowerPlex® Fusion System.

The collection of biological material began after the Research Ethics Committee of the Pontifical Catholic University of Goiás had approved the research (CAAE: 51483415.9.0000.0037 and opinion number 1.721.890).

Statistical analyses

The statistical parameters for the allele frequencies were obtained using Genetix version 4.05 and PowerStats softwares (Tereba, 1999; Belkhir et al., 2004; da Costa Francez et al., 2011). The alleles were tested for Hardy-Weinberg equilibrium using Genepop version 4.1.4 and Arlequin version 3.1 softwares (Excoffier et al., 2007; da Costa Francez et al., 2011). Other statistical parameters associated with the study population, such as the probability of identity, power of discrimination, polymorphism information content, power of exclusion, and typical paternity index, were calculated using PowerStats (Tereba, 1999).

RESULTS AND DISCUSSION

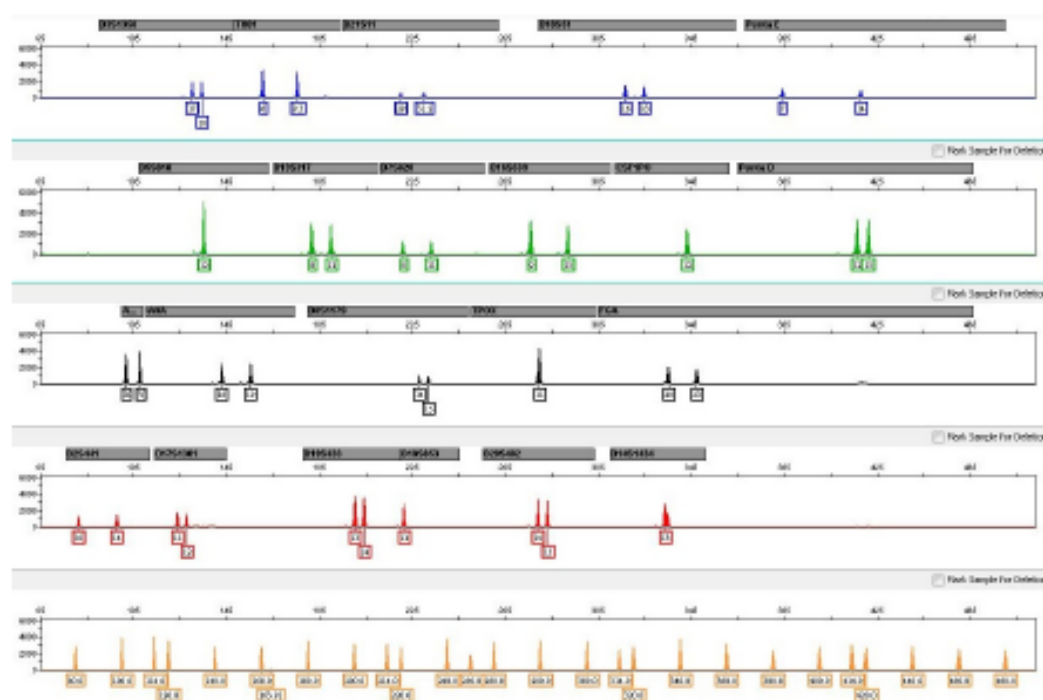
Following the criteria for selecting STRs, 22 polymorphic loci were selected to make up the multiplex panel (Table 1). Of these 22 loci, only TPOX and D2S441 (chromosome 2), D5S818 and CSF1PO (chromosome 5), D18S51 and D18S853 (chromosome 18), and D21S11 and Penta D (chromosome 21) were located on the same chromosome, at a minimum distance of 10 Mb. The developed multiplex set comprised 15 CODIS loci and 6 non-CODIS loci (Penta E, Penta D, D17S1301, D18S853, D20S482, and D14S1434).

To assess the performance of the synthesized primers and identify the annealing temperature common to all of them, the STRs were first assessed individually. Since no locus presented amplification failure and/or artifact excess that could compromise the yield expected for the system, the primers were subsequently combined in a multiplex reaction. As expected, the preliminary tests of the multiplex combination showed a large imbalance between the involved primers, probably due to some factors that interfered with the reaction, such as different annealing temperatures. To balance the combined primers, the amplification intensity was assessed individually to adjust their specific concentrations, assigning higher concentrations to the primers that obtained lower yield and, conversely, lower concentrations for the primers that showed allelic peaks with high RFU. The fluorescence labels, specific concentrations of each primer, and other information pertinent to the loci included in the multiplex are described in Table 1.

After optimization of the PCR in the multiplex system, the set of selected primers was successfully amplified in a single reaction, obtaining a harmonic and noise-free final panel. Figure 1 shows the allelic profile obtained for the 2800M-positive control DNA.

Table 1. Data for the short tandem repeat loci present in the developed multiplex system.

Dye label	Locus	Chromosomal location	Repeat unit	Primer concentration (μM)	Fragment size range (bp)
M TM	D3S1358	3p21.31	TCTG/TCTA	0.10	91-143
6-FAM TM	TH01	11p15.5	AATG	0.20	156-205
6-FAM TM	D21S11	21q21.1	TCTA/TCTG	0.30	215-271
6-FAM TM	D18S51	18q21.33	AGAA	0.20	282-362
6-FAM TM	Penta E	15q26.2	AAAGA	0.20	369-494
VIC [®]	D5S818	5q23.2	AGAT	0.20	111-155
VIC [®]	D13S317	13q31.1	TATC	0.10	164-204
VIC [®]	D7S820	7q21.11	GATA	0.20	215-259
VIC [®]	D16S539	16q24.1	GATA	0.25	268-312
VIC [®]	CSF1PO	5q33.1	AGAT	0.15	321-369
VIC [®]	Penta D	21q22.3	AAAGA	0.30	381-489
NED TM	Amelogenin	Xp22.1-22.3 and Y	-	0.10	106-112
NED TM	vWA	12p13.31	TCTG/TCTA	0.20	115-183
NED TM	D8S1179	8q24.13	TCTA/TCTG	0.50	195-251
NED TM	TPOX	2p25.3	AATG	0.08	258-302
NED TM	FGA	4q31.3	TTTC	0.15	314-444
PET [®]	D2S441	2p14	TCTA	0.15	78-106
PET [®]	D17S1301	17q25.1	AGAT	0.50	110-150
PET [®]	D19S433	19q12	AAGG	0.25	184-212
PET [®]	D18S853	18p11.31	ATA	0.40	215-242
PET [®]	D20S482	20p13	AGAT	0.15	259-311
PET [®]	D14S1434	14q32.13	CTGCTAT	0.20	321-369

**Figure 1.** Genetic profile of the 2800M-positive control DNA obtained with the newly developed multiplex system.

The multiplex amplification assessment of the different concentrations of the 2800M-positive control DNA showed 100% allelic identity in the reactions with DNA ranging from 1 to 10 ng. However, DNA at the concentration of 0.5 ng did not present complete amplification for some loci in the system, revealing an incomplete genotypic profile with approximately 90% allelic identity. Direct amplification of blood samples impregnated on FTA® cards and of DNA extracted from peripheral blood, buccal swab, and hair samples showed 100% allelic identity for all samples tested.

A comparison of the allelic profiles obtained by direct amplification of the 450 blood samples impregnated on FTA® cards, using both the newly developed multiplex system and the PowerPlex® Fusion System, showed genotype matching for all loci analyzed.

The allelic frequencies and statistical parameters related to the genetic and forensic linkage analysis, obtained for the 21 autosomal multiplex loci for the population of the State of Goiás, Brazil, are described in Table S1 and Table 2, respectively.

Table 2. Statistical parameters of autosomal multiplex loci for the population of the State of Goiás, Brazil.

Locus	%Hets	MP	PD	PIC	PE	TPI	N
D3S1358	80.9	0.104	0.896	0.73	0.616	2.62	450
TH01	80.2	0.077	0.923	0.76	0.603	2.53	450
D21S11	87.1	0.047	0.953	0.82	0.737	3.88	450
D18S51	87.8	0.026	0.974	0.87	0.750	4.09	450
Penta E	90.7	0.019	0.981	0.90	0.809	5.36	450
D5S818	74.0	0.115	0.885	0.69	0.493	1.92	450
D13S317	76.0	0.075	0.925	0.76	0.527	2.08	450
D7S820	77.8	0.075	0.925	0.76	0.558	2.25	450
D16S539	79.8	0.077	0.923	0.76	0.595	2.47	450
CSF1PO	70.4	0.111	0.889	0.69	0.435	1.69	450
Penta D	85.8	0.035	0.965	0.85	0.710	3.52	450
vWA	77.8	0.072	0.928	0.77	0.558	2.25	450
D8S1179	78.4	0.063	0.937	0.78	0.570	2.32	450
TPOX	67.3	0.143	0.857	0.65	0.388	1.53	450
FGA	88.2	0.028	0.972	0.87	0.759	4.25	450
D2S441	76.4	0.085	0.915	0.74	0.535	2.12	450
D17S1301	66.2	0.166	0.834	0.62	0.372	1.48	450
D19S433	79.3	0.050	0.950	0.80	0.587	2.42	450
D18S853	74.4	0.105	0.895	0.70	0.500	1.96	450
D20S482	69.1	0.146	0.854	0.64	0.415	1.62	450
D14S1434	69.3	0.127	0.873	0.67	0.418	1.63	450

%Hets: percentage heterozygosity; MP: match probability; PD: power of discrimination; PIC: polymorphism information content; PE: power of exclusion; TPI: typical paternity index; N: study population.

In general, the multiplex system showed large allelic variations, ranging from 7 alleles (TH01, D13S317, and D17S1301) to 24 (FGA), and a total of 239 alleles for the whole set. Deviation from Hardy-Weinberg equilibrium was observed for TPOX ($P=0.01016$), D17S1301 ($P=0.00319$), and D14S1434 ($P=0.03695$). According to Hardy (1908), this suggests that the population is under the influence of some factors, such as mutation rate, migration, selective pressure of the environment, or natural selection. By comparing the observed and expected heterozygosities, it was possible to find more homozygotes than expected. As such, amplification failure from a mutational event in the 3' region of the hybridization site of the primers had resulted in null alleles (Gomes et al., 1999; Dakin and Avise, 2004). Leibelt et al. (2003) described a mutation at the hybridization site in the D8S1179 locus in the population of Guam Island. Their study population demonstrated an excess of homozygotes at this locus when genotyped using AmpFISTR Profiler Plus™. Similar results were found by Budowle

- Gill P, Evett IW, Woodroffe S, Lygo JE, et al. (1991). Databases, quality control and interpretation of DNA profiling in the Home office Forensic Science Service. *Electrophoresis* 12: 204-209. <http://dx.doi.org/10.1002/elps.1150120217>
- Gill P, Fereday L, Morling N and Schneider PM (2006). The evolution of DNA databases--recommendations for new European STR loci. *Forensic Sci. Int.* 156: 242-244. <http://dx.doi.org/10.1016/j.forsciint.2005.05.036>
- Gomes I, Collins A, Lonjou C, Thomas NS, et al. (1999). Hardy-Weinberg quality control. *Ann. Hum. Genet.* 63: 535-538. <http://dx.doi.org/10.1046/j.1469-1809.1999.6360535.x>
- Hardy GH (1908). Mendelian proportions in a mixed population. *Science* 28: 49-50. <http://dx.doi.org/10.1126/science.28.706.49>
- Hares DR (2012). Expanding the CODIS core loci in the United States. *Forensic Sci. Int. Genet.* 6: e52-e54. Erratum in *Forensic Sci. Int. Genet.* 6: e135.
- Hares DR (2015). Selection and implementation of expanded CODIS core loci in the United States. *Forensic Sci. Int. Genet.* 17: 33-34. <http://dx.doi.org/10.1016/j.fsigen.2015.03.006>
- Hill CR, Kline MC, Coble MD and Butler JM (2008). Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. *J. Forensic Sci.* 53: 73-80. <http://dx.doi.org/10.1111/j.1556-4029.2008.00595.x>
- Hill CR, Butler JM and Vallone PM (2009). A 26plex autosomal STR assay to aid human identity testing. *J. Forensic Sci.* 54: 1008-1015. <http://dx.doi.org/10.1111/j.1556-4029.2009.01110.x>
- Leibelt C, Budowle B, Collins P, Daoudi Y, et al. (2003). Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. *Forensic Sci. Int.* 133: 220-227. [http://dx.doi.org/10.1016/S0379-0738\(03\)00035-5](http://dx.doi.org/10.1016/S0379-0738(03)00035-5)
- Mattana C, Fernandes MM, Tinoco R, Oliveira RN, et al. (2012). Forensic importance of DNA and the dentist participation. *Braz. J. Forensic Sci. Med. Law Bioethics* 2: 65-82. <http://dx.doi.org/10.17063/bjfs2-1-y201265>
- Opel KL, Chung DT, Drabek J, Butler JM, et al. (2007). Developmental validation of reduced-size STR Miniplex primer sets. *J. Forensic Sci.* 52: 1263-1271. <http://dx.doi.org/10.1111/j.1556-4029.2007.00584.x>
- Poetsch M, Preusse-Prange A, Schwark T and von Wurmb-Schwark N (2013). The new guidelines for paternity analysis in Germany-how many STR loci are necessary when investigating duo cases? *Int. J. Legal Med.* 127: 731-734. <http://dx.doi.org/10.1007/s00414-013-0867-y>
- Rodvalho RG, Santos GS, Cavalcanti LM, Moura BF, et al. (2015). Allele frequencies of combined DNA index system (CODIS) and non-CODIS short tandem repeat loci in Goiás, Central Brazil. *Genet. Mol. Res.* 14: 7310-7314. <http://dx.doi.org/10.4238/2015.July.3.6>
- Rozen S and Skaletsky H (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132: 365-386.
- Ruitberg CM, Reeder DJ and Butler JM (2001). STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res.* 29: 320-322. <http://dx.doi.org/10.1093/nar/29.1.320>
- Sarmento FJQ (2006). Modelagem de um ambiente para análise de DNA em genética forense. Dissertação de Mestrado, Universidade Federal de Alagoas, Campus Maceió, Maceió.
- Tereba A (1999). Tools for analysis of population statistics. *Profiles DNA* 2: 14-16.
- Vallone PM and Butler JM (2004). AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques* 37: 226-231.
- Vanderheyden N, Mai A, Gilissen A, Cassiman JJ, et al. (2007). Identification and sequence analysis of discordant phenotypes between AmpFISTR SGM Plus and PowerPlex 16. *Int. J. Legal Med.* 121: 297-301. <http://dx.doi.org/10.1007/s00414-007-0167-5>
- Wang DY, Chang CW, Lagacé RE, Oldroyd NJ, et al. (2011). Development and validation of the AmpFISTR® Identifier® Direct PCR Amplification Kit: a multiplex assay for the direct amplification of single-source samples. *J. Forensic Sci.* 56: 835-845. <http://dx.doi.org/10.1111/j.1556-4029.2011.01757.x>
- Wiegand P, Klein R, Braunschweiger G, Hohoff C, et al. (2006). Short amplicon STR multiplex for stain typing. *Int. J. Legal Med.* 120: 160-164. <http://dx.doi.org/10.1007/s00414-005-0048-8>
- Zhang S, Bian Y, Tian H, Wang Z, et al. (2015). Development and validation of a new STR 25-plex typing system. *Forensic Sci. Int. Genet.* 17: 61-69. <http://dx.doi.org/10.1016/j.fsigen.2015.03.008>

Supplementary material

Table S1. Allelic frequency obtained for the 21 autosomal loci comprising the multiplex system.

5. MATERIAL E MÉTODOS

Essa seção descreve parte da metodologia adotada para execução do projeto. As demais informações pertinentes a este capítulo encontram-se no artigo científico “*Development of a polymorphic short tandem repeat locus multiplex system for efficient human identification*”, que compõe esta tese.

5.1 Teste dos primers individuais e composição do conjunto multiplex

O processo de validação visou determinar as condições e limitações estabelecidas pelas metodologias propostas, sendo caracterizado pela avaliação da eficácia e confiabilidade dos métodos aplicados e resultados obtidos (DNA Advisory Board, 2000). Para Myers (2012) o processo de validação interna é fundamental para determinar as condições operacionais do sistema e demonstrar suas limitações.

5.1.1 Validação das sequências sintetizadas

Para a execução dos testes propostos, as reações de amplificação foram realizadas utilizando-se o controle positivo de DNA 2800M (Promega Corporation, Madison, WI, United States) em caráter *single-plex* (apenas um par de iniciadores por reação), visando certificar que o amplicon obtido corresponderia ao tamanho esperado.

A PCR foi realizada em um termociclador *Veriti[®] Thermal Cycler* (Thermo Fisher Scientific, Foster City, CA, United States), sendo aplicado gradiente para a temperatura de anelamento, com variação térmica de 54°C a 64°C (Tabela 3). Esta variação permitiu determinar uma temperatura de anelamento ideal comum a todos os conjuntos de primers utilizados, contribuindo para estabelecer a temperatura adequada para reação multiplex.

Para a PCR, utilizou-se o *TopTaq[™] Master Mix Kit* (Qiagen, Hilden, NRW, Germany) e a proporção de seus reagentes foi ajustada seguindo-se o protocolo proposto pelo fabricante (Tabela 4).

Tabela 3. Condições de termociclagem utilizadas para a validação das sequências dos primers desenvolvidos para um sistema multiplex de marcadores STRs autossômicos para a identificação humana.

Temperatura (°C)	Tempo	Nº Ciclos	Etapa
94	3 minutos	-	Desnaturação Inicial
94	30 segundos	30	Desnaturação
54	30 segundos		Anelamento (Gradiente)
56			
58			
60			
62			
64			
72	60 segundos		Extensão
72	10 minutos	-	Extensão Final
4	∞	-	-

Tabela 4. Proporções de PCR utilizadas para validação das sequências de primers sintetizados.

Componentes	Volume (µL)	Concentração Final
10x TopTaq PCR Buffer	5,00	1 x
dNTP mix	1,00	200 µM cada DNTP
Primer Foward	1,00	0,2 µM
Primer Reverse	1,00	0,2 µM
TopTaq DNA Polymerase	0,25	1,25 unidades
Água	14,75	-
DNA <i>Template</i>	2,00	1 µg
Volume Final	25,00	-

Os produtos amplificados foram submetidos à eletroforese convencional em gel de agarose a 4% em TBE 1X, a um campo elétrico uniforme com voltagem constante de 10V/cm por 2 horas. Em seguida, os fragmentos separados foram corados em solução de brometo de etídeo (5mg/mL) e visualizados no sistema de vídeo-documentação *Gel Doc™ XR+System* (Bio-Rad, Hercules, CA, United States), sendo subsequentemente realizado o registro fotográfico das “bandas” obtidas. Os amplicons obtidos foram pareados com as referências alélicas para cada *locus* amplificado a fim de se confirmar que os produtos de PCR apresentam os tamanhos esperados.

5.1.2 Validação dos primers em reações multiplex

Após a confirmação da amplificação *single-plex* de acordo com os parâmetros esperados, foi analisado o comportamento dos primers em reações de amplificação

multiplex, dando início ao processo de “multiplexação” dos marcadores genéticos. Para estabelecer o conjunto multiplex, os conjuntos de primers foram selecionados considerando-se o intervalo de tamanho esperado após a amplificação de cada *locus*. O conjunto de 21 marcadores foi, assim, distribuído em 4 grupos de PCR multiplex. Para a PCR utilizou-se o Qiagen® Multiplex PCR Kit (Qiagen, Hilden, NRW, Germany) e a proporção de seus reagentes foi ajustada seguindo-se o protocolo proposto pelo fabricante, para um volume final de 25µL.

A partir dos testes realizados em PCR gradiente, foi possível definir uma única temperatura de anelamento comum para todos os primers, referente ao conjunto de STRs para a composição do sistema multiplex proposto (Tabela 5).

Tabela 5. Condições de termociclagem aplicadas na PCR para amplificação multiplex de um sistema de marcadores STRs autossômicos desenvolvido para fins de identificação humana.

Temperatura (°C)	Tempo	Nº Ciclos	Etapa
94	5 minutos	-	Desnaturação Inicial
94	30 segundos	30	Desnaturação
60	90 segundos		Anelamento
72	60 segundos		Extensão
68	60 minutos	-	Extensão Final
4	∞	-	-

O destaque em negrito representa a temperatura de anelamento comum a todos os primers

Os produtos amplificados foram submetidos à eletroforese em gel de agarose a 2% em TBE 1X, em um campo elétrico uniforme com voltagem constante de 10V/cm por 2 horas. Em seguida, o gel foi corado em solução de brometo de etídeo (5mg/mL) e visualizados no sistema de vídeo-documentação *Gel Doc™ XR+System*. As imagens foram capturadas para fins de registro.

Após o estabelecimento dos grupos multiplex, prosseguiu-se para a próxima etapa, com o objetivo de combinar todos os primers em um único “*Mix Primer*”. Para a separação dos amplicons, adotou-se a eletroforese capilar para aumentar o poder de resolução dos alelos amplificados.

A metodologia de eletroforese capilar exige que os primers estejam “marcados” com uma fluorescência específica de acordo com o grupo ao qual este iniciador pertence. Assim, foi necessário escolher as possíveis fluorescências a serem utilizadas na marcação dos primers. A escolha dos fluorocromos é dependente da disposição dos filtros presentes

no analisador genético ABI 3500[®] (Thermo Fisher Scientific, Foster City, CA, United States). A marcação com fluoróforo foi realizada na sequência *forward* dos pares de primers, seguindo os rótulos propostos pela combinação dos grupos específicos. Utilizou-se então as seguintes fluorescências: 6-FAM[™], VIC[®], NED[™], PET[®] e para a marcação do *Size Standard* foi utilizado o LIZ[®] (Figura 4).

Grupo Primers	Fluoróforo
1	6-FAM [™]
2	VIC [®]
3	NED [™]
4	PET [®]
<i>Size Standard</i>	LIZ [®]

Figura 5. Separação dos primers em grupos a partir da marcação específica por fluoróforo.

5.2 Determinação do custo de amplificação do multiplex desenvolvido

O custo da reação de amplificação do painel multiplex foi obtido considerando-se as despesas referentes a aquisição do Qiagen[®] Multiplex PCR Kit (Qiagen, Hilden, NRW, Germany) e da síntese e marcação dos primers pela empresa Applied Biosystems (Thermo Fisher Scientific, Foster City, CA, United States). A síntese dos primers foi realizada considerando a escala de 80.000 pmol.

Para comparação dos custos de amplificação entre o multiplex e os kits comerciais disponíveis atualmente no Brasil, foi feita uma pesquisa envolvendo os sistemas: *PowerPlex[®] Fusion* e *PowerPlex[®] 16 HS* (Promega Corporation, Madison, WI, United States); *AmpFLSTR[™] Identifier[™] PCR Amplification* e *GlobalFiler[™] PCR Amplification* (Thermo Fisher Scientific, Foster City, CA, United States); *Investigator 24plex* e *Investigator IDplex Plus* (Qiagen, Hilden, NRW, Germany).

5.3 Comprovação da eficiência do sistema multiplex

5.3.1 Estudo de casos

Amostras biológicas de sangue e swab bucal, provenientes de exames de rotina do Laboratório Biocroma (Goiânia – GO), relacionados à três situações envolvendo

investigação de vínculo genético, foram submetidas à análise pelo ensaio multiplex STR, a fim de comprovar a eficiência do sistema desenvolvido.

Os resultados obtidos, a partir da amplificação direta em cartão FTA[®] (Whatman, GE Healthcare Life Sciences, Maidstone, Kent, United Kingdom), foram comparados aos resultados gerados a partir da análise destes mesmos exames utilizando o sistema comercial *PowerPlex[®] Fusion*.

As situações de investigação de paternidade selecionadas foram classificadas como:

- TRIO: Investigação de paternidade envolvendo um suposto pai, um suposto filho e a mãe biológica do suposto filho;
- DUO: Investigação de paternidade envolvendo apenas um suposto pai e um suposto filho;
- RECONSTRUÇÃO GENÉTICA: Investigação de paternidade envolvendo a mãe e o pai biológicos do suposto pai (ausente), um Suposto filho e a mãe biológica do suposto filho.

5.3.2 Estudo de casos complexos que demandam ampliação do painel STR

A fim de avaliar a performance do multiplex desenvolvido como ferramenta de ampliação de informação genética para a investigação, foram selecionados 6 casos reais do Laboratório Biocroma (Goiânia – GO) que, após avaliação prévia com utilização do sistema comercial *PowerPlex[®] Fusion*, necessitou de análises complementares com outros sistemas. As amostras de sangue dos exames selecionados foram submetidas à análise pelo ensaio multiplex STR para confirmar os genótipos referentes aos *loci* sobrepostos com o kit *PowerPlex[®] Fusion*, além de ampliar a análise genética a partir da dos *loci* STRs não sobrepostos com os sistemas comerciais.

6. RESULTADOS E DISCUSSÃO

Essa seção descreve parte dos resultados obtidos durante a execução do projeto. Os demais resultados encontram-se no artigo científico “*Development of a polymorphic short tandem repeat locus multiplex system for efficient human identification*”.

6.1 Loci selecionados

Para a composição do painel multiplex de identificação humana, um total de 43 *loci* STRs foi pré-selecionado, conforme apresentado na Tabela 6.

Tabela 6. Conjunto de *loci* STRs autossômicos pré-selecionados e Amelogenina com suas respectivas localizações cromossômicas para a composição de um sistema multiplex para identificação humana.

<i>locus</i>	Localização Cromossômica	<i>locus</i>	Localização Cromossômica
D1GATA113	1p36.23	D10S1248	10q26.3
D1S1656	1q42	D10S2325	10p12
F13B	1q31-32.1	TH01	11p15.5
D2S441	2p14	VWA	12p13.31
D2S1338	2q35	D12S391	12p13.2
D2S1360	2p24-p22	D13S317	13q31.1
TPOX	2p25.3	D14S1434	14q32.13
D3S1358	3p21.31	FESFPS	15q25-qter
D3S1744	3p24	Penta E	15q26.2
FGA	4q31.3	D16S539	16q24.1
D4S2366	4p16-15.2	D17S1301	17q25.1
D5S818	5q23.2	D18S51	18q21.33
CSF1PO	5q33.1	D18S853	18p11.31
D5S2500	5q11.2	D19S433	19q12
F13A01	6p24-p25	D20S482	20p13
D6S474	6q21-22	D20S1082	20q13.2
SE33	6q14.2	D21S11	21q21.1
D7S820	7q21.11	D21S2055	21q22
D7S1517	7q31.33	Penta D	21q22.3
D8S1179	8q24.13	D22S1045	22q12.3
LPL	8p22	Amelogenina	Xp22.1-22.3 e Y
D8S1132	8q23.1		

Fonte: PowerPlex® CS7 System; Investigator® Human Identification PCR Kits – Qiagen; NCBI; STRBase.

A construção de um sistema multiplex para a identificação humana que seja informativo e tenha um elevado poder discriminatório dos indivíduos de uma população, requer marcadores STRs polimórficos, cujos alelos possuam intervalos de tamanho que permitem a combinação entre si, sem prejudicar ou confundir interpretações de resultados (Butler, 2005). Considerando estes critérios e os requisitos descritos anteriormente adotados para obtenção dos marcadores genéticos, foram selecionados 21 *loci* polimórficos autossômicos e 1 *locus* de determinação sexual, sendo eles: D2S441, TPOX, D3S1358, FGA, D5S818, CSF1PO, D7S820, D8S1179, TH01, VWA, D13S317, D14S1434, Penta E, D16S539, D17S1301, D18S51, D18S853, D19S433, D20S482, D21S11, Penta D e Amelogenina. Os demais *loci* pré-selecionados (D1GATA113, D1S1656, F13B, D2S1338, D2S1360, D3S1744, D4S2366, D5S2500, F13A01, D6S474, SE33, D7S1517, LPL, D8S1132, D10S1248, D10S2325, D12S391, FESFPS, D20S1082, D21S2055 e D22S1045) não incluídos no sistema por não cumprirem os requisitos estabelecidos para a seleção de *loci*, ou por estarem próximos à microssatélites de interesse, ou por fazerem parte de sistemas de ampla comercialização no cenário nacional.

A inclusão de *loci* adicionais no conjunto dos 22 já selecionados foi inviabilizada em decorrência da impossibilidade de acomodá-los na composição do painel, em função do intervalo de tamanho esperado dos amplicons. Qualquer inclusão de outro marcador interferia nos intervalos de segurança entre os marcadores STR do painel. Conseqüentemente, poderia provocar sobreposição dos alelos localizados nas extremidades dos intervalos.

A distância entre os *loci* selecionados que se encontram no mesmo cromossomo foi uma condição fundamental para a composição do sistema multiplex. O cumprimento deste requisito é de extrema importância para garantir a segregação independente de todos os alelos presentes nos *loci* estudados. Os marcadores TPOX e D2S441 (cromossomo 2), D5S818 e CSF1PO (cromossomo 5), D18S51 e D18S853 (cromossomo 18) e D21S11 e Penta D (cromossomo 21) compreendem os grupos de *loci* localizados nos mesmos cromossomos. Porém, a distância de segurança entre eles foi devidamente respeitada (Bright *et al.*, 2014). No gráfico 01 é possível observar a distribuição dos *loci* selecionados pelos cromossomos.



Figura 6. Distribuição dos *loci* STRs nos cromossomos autossômicos selecionados para compor um painel multiplex destinado a identificação humana no Brasil.

As informações dos microssatélites selecionados no multiplex, referentes ao cromossomo que pertence, à classificação pelo tamanho e padrão da unidade de repetição, quanto à presença no sistema CODIS e os fluoróforos usados na composição do painel, estão descritas na Tabela 7.

Tabela 7. Informações dos microssatélites autossômicos selecionados para compor um painel de marcadores em multiplex e o marcador de determinação do sexo proposto para estudos de vínculo genético em populações humanas.

<i>Locus</i>	Cromossomo	Unidade de Repetição	Padrão de Repetição	Sistema	Fluoróforo
D2S441	2	Tetranucleotídica	Simple	CODIS	PET®
TPOX	2	Tetranucleotídica	Simple	CODIS	NED™
D3S1358	3	Tetranucleotídica	Complexo	CODIS	6-FAM™
FGA	4	Tetranucleotídica	Simple	CODIS	NED™
D5S818	5	Tetranucleotídica	Simple	CODIS	VIC®
CSFIPO	5	Tetranucleotídica	Simple	CODIS	VIC®
D7S820	7	Tetranucleotídica	Simple	CODIS	VIC®
D8S1179	8	Tetranucleotídica	Complexo	CODIS	NED™
TH01	11	Tetranucleotídica	Simple	CODIS	6-FAM™
VWA	12	Tetranucleotídica	Complexo	CODIS	NED™
D13S317	13	Tetranucleotídica	Simple	CODIS	VIC®
D14S1434	14	Tetranucleotídica	Complexo	não-CODIS	PET®
Penta E	15	Pentanucleotídica	Simple	não-CODIS	6-FAM™
D16S539	16	Tetranucleotídica	Simple	CODIS	VIC®
D17S1301	17	Tetranucleotídica	Simple	não-CODIS	PET®
D18S51	18	Tetranucleotídica	Simple	CODIS	6-FAM™
D18S853	18	Trinucleotídica	Simple	não-CODIS	PET®
D19S433	19	Tetranucleotídica	Complexo	CODIS	PET®
D20S482	20	Tetranucleotídica	Simple	não-CODIS	PET®
D21S11	21	Tetranucleotídica	Simple	CODIS	6-FAM™
Penta D	21	Pentanucleotídica	Simple	não-CODIS	VIC®
Amelogenina	X / Y	-	-	CODIS	NED™

Considerando a classificação dos microssatélites quanto ao tamanho da unidade de repetição (UR), o *locus* D18S853 é classificado como STR trinucleotídico, por apresentar unidade repetitiva com três nucleotídeos. Neste mesmo contexto, os *loci*

D2S441, TPOX, D3S1358, FGA, D5S818, CSF1PO, D7S820, D8S1179, TH01, vWA, D13S317, D14S1434, D16S539, D17S1301, D18S51, D19S433, D20S482 e D21S11 são compostos por UR tetranucleotídica. Já os marcadores Penta E e Penta D, como o próprio nome sugere, são formados unidade de repetição pentanucleotídica. O *locus* Amelogenina não apresenta unidade de repetição, pois o que diferencia o alelo “X” do alelo “Y” é uma deleção para o cromossomo X do segmento AAAGTG.

Conforme a classificação pelo padrão de repetição, pôde-se estabelecer dois grupos entre os STR selecionados. Destacam-se como STRs de repetições simples os *loci*: D2S441, TPOX, FGA, D5S818, CSF1PO, D7S820, TH01, D13S317, Penta E, D16S539, D17S1301, D18S51, D18S853, D20S482, D21S11 e Penta D. Os *loci* D3S1358, D8S1179, VWA, D14S1434 e D19S433 podem ser classificados como STRs de repetições complexas. Para o *locus* Amelogenina aplica-se a mesma condição estabelecida conforme a classificação STR pelo tamanho da UR.

O conjunto STR criado apresenta sobreposição em 15 *loci* com o sistema CODIS expandido (Tabela 7). Esta homologia é cientificamente importante, uma vez que um perfil genético obtido a partir de 15 *loci* STRs comuns aos demais sistemas frequentemente utilizados, garante uma compatibilidade de dados genéticos gerados universalmente, tanto para aplicação em futuras análises envolvendo identificação humana, quanto para fins de pesquisa associadas às construções de bancos de dados populacionais. Os demais *loci* componentes deste sistema, por não apresentarem homologia com o sistema CODIS, foram classificados como “não-CODIS”. A presença destes marcadores no conjunto selecionado também assume fundamental importância para o processo de identificação humana, pois seu uso não é explorado no atual contexto nacional. Para Tsai *et al.*, (2013), em inúmeras situações reais de investigação de vínculo genético envolvendo alta complexidade, a aplicação dos marcadores não-CODIS agrega informação genética extra ao processo discriminatório, ausente nos sistemas comerciais disponíveis atualmente.

6.2 Criação do layout espacial e espectral do painel multiplex

É relevante ressaltar a importância da proposição de eletroforese convencional dos amplicons obtidos a partir da utilização de primers inicialmente não marcados por fluorescência. Esta estratégia se justifica pelo alto custo financeiro vinculado à marcação

dos primers frente a aquisição destes não marcados. Assim, a alternativa de validar as sequências sintetizadas por metodologias convencionais, atendeu às expectativas impostas para a realização desta importante etapa do projeto.

No intuito de evitar sobreposição entre *loci* vizinhos no eixo espectral do painel, o agrupamento dos primers foi estabelecido considerando as características específicas referentes à cada *locus* selecionado, como: tamanho dos amplicons produzidos e intervalo de variação alélica (Figura 5).

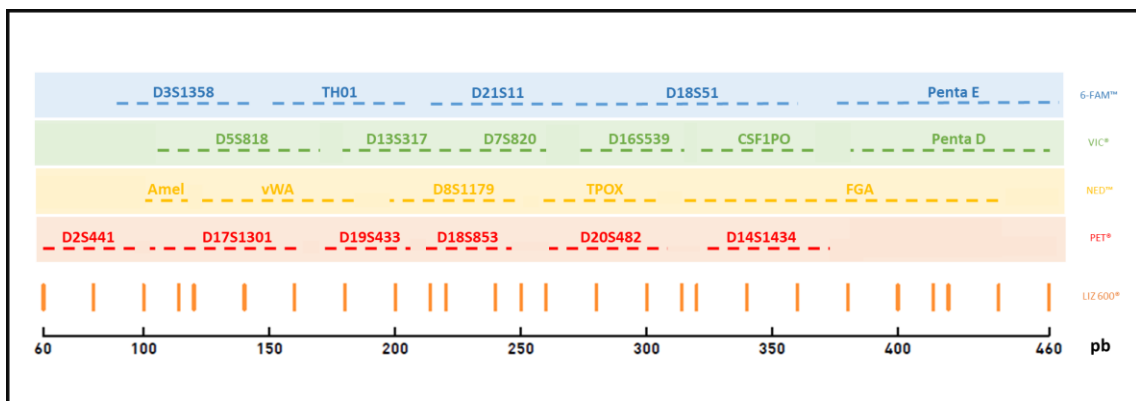


Figura 7. Disposição dos *loci* na dimensão espacial e no eixo espectral do painel multiplex desenvolvido.

6.3 Determinação dos custos de Amplificação e comparação com kits comerciais

Na tabela 8 está descrito o custo de amplificação estabelecido para o multiplex proposto, bem como para os demais kits comerciais compatíveis, atualmente disponíveis no Brasil.

Tabela 8. Custo por amplificação do multiplex desenvolvido e dos principais sistemas comerciais disponíveis no Brasil.

Empresa	Sistema	Nº <i>loci</i> Autossômicos	Custo por Amplificação (R\$)
Promega Corporation	PowerPlex® Fusion	22	79,50
	PowerPlex® 16 HS	15	43,48
Thermo Fisher Scientific	AmpFLSTR™ Identifiler™ PCR Amplification	15	88,67
	GlobalFiler™ PCR Amplification	21	86,10
Qiagen	Investigator 24plex	21	94,38
	Investigator IDplex Plus	15	67,42
-	Sistema Multiplex Desenvolvido	21	5,30

Os sistemas comerciais *PowerPlex® 16 HS*, *AmpFLSTR™ Identifiler™ PCR Amplification* e *Investigator IDplex Plus* possibilitam a amplificação de apenas 15 *loci*

STR autossômicos, enquanto que, com os kits *GlobalFiler™ PCR Amplification* e *Investigator 24plex* é possível obter amplificação de 21 *loci* STR autossômicos. Já o *PowerPlex® Fusion* possui 22 STRs autossômicos disponíveis em seu painel multiplex. Todos os kits comerciais aqui reportados apresentam atualmente ampla comercialização em todo território nacional.

O kit comercial com maior preço de amplificação por amostra é o *Investigator 24plex* (R\$94,38), correspondendo à aproximadamente 1.800% do custo de amplificação do multiplex desenvolvido (R\$5,30). No mesmo sentido, o *PowerPlex® 16 HS* apresenta o menor custo de amplificação por amostra entre os kits comerciais pesquisados (R\$43,48), correspondendo à aproximadamente 820% do custo estimado de amplificação do multiplex. Neste contexto, é possível constatar que o custo de amplificação do multiplex é substancialmente menor dos que os demais kits comercializados atualmente no Brasil, o que viabiliza sua implantação e justifica seu desenvolvimento.

Assim sendo, o desenvolvimento de um sistema eficiente, confiável e prático viabiliza estratégias comerciais que permitem aumentar a amplitude da oferta. Por consequência, neste cenário, garante-se inclusão e a acessibilidade para um estrato social que atualmente não tem acesso a esta metodologia de investigação, além de proporcionar mais segurança para a análise. Assim, a implementação do multiplex proposto colaborará para que a metodologia seja oferecida com equidade a todos que dela necessitam.

6.4 Verificação da eficácia do multiplex em situações reais de investigação de vínculo genético

6.4.1 Estudo de casos

Treze estudos de vínculo genético provenientes da rotina do Laboratório Biocroma foram amplificados usando-se o sistema multiplex desenvolvido. Destes casos, um exame corresponde à reconstrução genética, nove estão relacionados a análises de parentesco com configuração de TRIO e três foram realizados apenas entre o suposto filho e o suposto pai, sendo classificados como DUO. O tipo de amostra biológica analisada, o resultado final do processo de investigação, os painéis STRs utilizados, os índices estatísticos obtidos, bem como o número de incoerências alélicas observadas estão descritos na Tabela 9. As composições alélicas e os índices de paternidade obtidos para

todos os *loci* analisados dos dois sistemas utilizados referentes aos casos selecionados, encontram-se no Apêndice III.

Conforme demonstrado na Tabela 10, os casos 2, 6, 7, 10 e 13 foram realizados a partir da amplificação direta de swab bucal, enquanto dos demais casos, 1, 3, 4, 5, 8, 9, 11 e 12 foram coletadas amostras de sangue em cartão FTA[®] para realização da investigação. A amplificação direta simplifica o processamento das amostras e, por consequência, oferece grande praticidade à rotina de obtenção de perfil genético, uma vez que elimina procedimentos de extração convencional e quantificação do DNA (Flores *et al.*, 2014; Guo *et al.*, 2014; Myers *et al.*, 2012). Por apresentar robustez, sensibilidade e baixo custo, a amplificação direta a partir de amostras de sangue e swab bucal impregnados em cartão FTA[®], é recomendada para estudos envolvendo identificação humana e investigação de paternidade (Hennessy *et al.*, 2014; Laurin *et al.*, 2012).

Para todas as amostras analisadas houve amplificação satisfatória de todos os *loci* componentes dos painéis STRs adotados, corroborando com os resultados dos testes de sensibilidade do multiplex, descritos neste trabalho.

Os casos 1, 5, 6, 7, 8, 9, 10, 11 e 13 correspondem à exames cujo resultado final permitiu a conclusão pela inclusão do vínculo genético investigado. As análises estatísticas adotadas para aferir vínculo parental revelaram que as evidências genéticas encontradas reforçam a hipótese de que existe uma maior probabilidade de o indivíduo investigado ser o pai biológico do suposto filho do que outro indivíduo aleatório da população (“H1”) (Gjertson *et al.*, 2007; Junge *et al.*, 2006; Morling *et al.*, 2002).

O caso 5 apresentou os maiores índices estatísticos obtidos para os dois sistemas testados. O Índice de Paternidade (IP) obtido com o sistema comercial *PowerPlex[®] Fusion* corresponde à aproximadamente $1,36E+17$, cuja Probabilidade de Paternidade (PP) aferida foi de aproximadamente 99,99999999999999%. Já com o multiplex desenvolvido, foi obtido um IP de aproximadamente $1,08E+15$, com PP correspondente a aproximadamente 99,99999999999999%.

A situação de inclusão com menores índices estatísticos observados corresponde ao caso 1, conforme esperado por se tratar de uma reconstrução genética. Os IPs obtidos correspondem à $2,19E+07$ e $9,95E+06$, para os painéis *PowerPlex[®] Fusion* e multiplex próprio, respectivamente. Estes índices revelam probabilidades de paternidade de aproximadamente 99,99999% e 99,99999%, respectivamente.

Tabela 9. Estudo de casos e comparação dos resultados obtidos com o *PowerPlex® Fusion* e o multiplex desenvolvido para estudo de vínculo genético.

Código caso	Investigação vínculo genético	Indivíduos envolvidos	Amostra biológica	<i>Powerplex® Fusion</i>		Resultado	Multiplex desenvolvido		Resultado
				I.C.P	P.P		I.C.P	P.P	
1	Reconstrução Genética	MSP + PSP + SF + MSF	Sangue em FTA	2,19E+07	99,999999%	INCLUSÃO	9,95E+06	99,999999%	INCLUSÃO
2	TRIO	MSF + SF + SP	Swab bucal	14 incoerências alélicas		EXCLUSÃO	14 incoerências alélicas		EXCLUSÃO
3	TRIO	MSF + SF + SP	Sangue em FTA	10 incoerências alélicas		EXCLUSÃO	08 incoerências alélicas		EXCLUSÃO
4	TRIO	MSF + SF + SP	Sangue em FTA	04 incoerências alélicas		EXCLUSÃO	05 incoerências alélicas		EXCLUSÃO
5	TRIO	MSF + SF + SP	Sangue em FTA	1,36E+17	99,999999999999%	INCLUSÃO	1,08E+15	99,999999999999%	INCLUSÃO
6	TRIO	MSF + SF + SP	Swab bucal	1,05E+11	99,999999999%	INCLUSÃO	7,86E+08	99,9999999%	INCLUSÃO
7	TRIO	MSF + SF + SP	Swab bucal	2,79E+11	99,999999999%	INCLUSÃO	3,69E+10	99,9999999%	INCLUSÃO
8	TRIO	MSF + SF + SP	Sangue em FTA	6,15E+12	99,999999999%	INCLUSÃO	2,94E+10	99,9999999%	INCLUSÃO
9	TRIO	MSF + SF + SP	Sangue em FTA	1,61E+10	99,999999999%	INCLUSÃO	2,87E+09	99,9999999%	INCLUSÃO
10	TRIO	MSF + SF + SP	Swab bucal	1,09E+13	99,99999999999%	INCLUSÃO	7,97E+11	99,9999999999%	INCLUSÃO
11	DUO	SF + SP	Sangue em FTA	09 incoerências alélicas		EXCLUSÃO	08 incoerências alélicas		EXCLUSÃO
12	DUO	SF + SP	Sangue em FTA	6,34E+08	99,999999%	INCLUSÃO	1,12E+10	99,9999999%	INCLUSÃO
13	DUO	SF + SP	Swab bucal	1,79E+09	99,9999999%	INCLUSÃO	1,33E+09	99,9999999%	INCLUSÃO

MSP = Mãe biológica do Suposto Pai Ausente; PSP = Pai biológico do Suposto Pai Ausente; SF = Suposto Filho; MSF = Mãe biológica do Suposto Filho; SP = Suposto Pai; ICP = Índice Acumulado de Paternidade; PP = Probabilidade de Paternidade.

Foi possível constatar que, quanto à obtenção dos índices de vínculo genético, ambos os sistemas STRs utilizados apresentaram resultados satisfatórios em todas as situações envolvendo inclusão de paternidade. Assim, considerando os critérios recomendados internacionalmente para análise de vínculo genético (Junge *et al.*, 2006; Morling *et al.*, 2003), foi possível concluir com segurança acerca do vínculo biológico pesquisado.

Os casos 2, 3, 4 e 11 apresentaram resultados que reforçam a hipótese de que é mais provável que o pai biológico não é o homem testado, e sim, um homem aleatório na população e não relacionado ao indivíduo analisado. Neste sentido, esses exames, pela falta de compatibilidade genética entre o suposto filho e o suposto pai, foram caracterizados como exclusão do vínculo genético investigado.

Com o sistema *PowerPlex® Fusion*, nos casos 2, 3 e 11 foram encontradas, respectivamente, 14, 10 e 9 incoerências alélicas durante as análises, correspondendo a erros mendelianos na transmissão genética. Os resultados do sistema multiplex próprio para os mesmos casos, revelaram 14, 8 e 11 erros mendelianos, respectivamente. Com poder resolutivo semelhante ao *PowerPlex® Fusion*, o sistema multiplex de STRs autossômicos desenvolvido atendeu satisfatoriamente às recomendações propostas pela IFSG para conclusão pela exclusão da paternidade biológica investigada (Jha *et al.*, 2013; Gjertson *et al.*, 2007; Morling *et al.*, 2003; Baird *et al.*, 1996).

No caso 4 foram observadas 4 incoerências alélicas após a análise com o *PowerPlex® Fusion*. Tais resultados dificultaram a conclusão do exame, frente à possível ambiguidade de resultados, uma vez que já existe registro de 3 erros mendelianos em paternidade biológica comprovada (Sun *et al.*, 2012; Jacewics *et al.*, 2004). Quando os dados do caso 4 foram produzidos com o multiplex de STRs autossômicos desenvolvido, foram identificadas 5 incoerências alélicas, que reforçou a conclusão pela exclusão do vínculo genético investigado entre o suposto pai e a criança investigada.

Em todas as situações analisadas, envolvendo reconstrução genética, trios e duos, tanto para evidências de inclusão, quanto para os casos de exclusão de vínculo genético, o sistema multiplex desenvolvido apresentou grande coerência com o sistema comercial *PowerPlex® Fusion* e se mostrou eficiente para resolver os casos da rotina do Laboratório Biocroma de Goiânia (Goiás).

6.4.2 Estudo de casos complexos que demandam ampliação do painel STR

Deteção de mutação germinativa, observação de evento de *dropout*, reconstruções genéticas, bem como TRIOS ou DUOS com índices estatísticos obtidos abaixo dos padrões recomendados ou com poucas incoerências alélicas identificadas, são situações complexas que demandam ampliação do painel STR utilizado para atingir resultados mais confiáveis (Sun *et al.*, 2015; Lindner *et al.*, 2014; Tsai *et al.*, 2013; Akhteruzzaman *et al.*, 2012; Li *et al.*, 2012; Carboni *et al.*, 2011; Li *et al.*, 2011). No entanto, os kits comerciais normalmente apresentam grande sobreposição dos *loci* STRs, não podendo, em algumas situações, aumentar a eficiência quando utilizados em caráter de complemento (Zhu *et al.*, 2014; Yuan *et al.*, 2012).

Assim, para avaliação da eficácia do painel multiplex desenvolvido como ferramenta complementar de identificação humana, foram selecionados 5 casos complexos de investigação de paternidade da rotina do Laboratório Biocroma, conforme descrito na Tabela 10 (Goiânia-GO).

Todos os exames foram previamente analisados com emprego de pelo menos um sistema comercial autossômico, não sendo suficiente para se obter uma conclusão segura sobre o vínculo genético entre os investigados. Portanto, as amostras biológicas foram submetidas à amplificação pelo sistema multiplex de STRs autossômicos desenvolvido no intuito de complementar as análises, fornecendo 4 *loci* STRs adicionais (D17S1301, D18S853, D20S482 e D14S1434) e 17 *loci* sobrepostos com o sistema comercial *PowerPlex® Fusion*.

6.4.2.a. Reconstrução genética

O caso 14 consiste na reconstrução do perfil genético do Suposto Pai Falecido por meio de seus genitores. O Índice Acumulado de Paternidade (ICP), obtido a partir da análise dos sistemas comerciais *PowerPlex® Fusion* e *PowerPlex® Y23* (Promega Corporation, Madison, WI, United States), não obedeceu aos padrões mínimos para sua conclusão, sendo necessário ampliar os *loci* de estudo para outro painel multiplex. O ICP obtido com emprego dos sistemas comerciais adotados foi de 3,40E+03. Esse valor aumentou consideravelmente com a ampliação de mais 4 *loci* STRs complementares presentes no multiplex, alcançando um ICP correspondente à 1,49E+04 e uma Probabilidade de Paternidade (PP) de aproximadamente 99,99%, o que foi fundamental para viabilizar a conclusão do parentesco investigado.

Tabela 10. Estudo de casos para avaliação da eficácia do painel multiplex de STRs autossômicos como ferramenta complementar para a identificação humana.

Código caso	Investigação vínculo genético	Indivíduos envolvidos	Amostra biológica	Motivo Ampliação	Powerplex® Fusion		Resultado	Multiplex desenvolvido		Resultado
					I.C.P	P.P		I.C.P	P.P	
14	Reconstrução Genética	MSF - SF - PSP	Sangue FTA	ICP relativamente baixo	3,40E+03	99,9%	Inclusão	1,49E+04	99,99%	Inclusão
15	TRIO	MSF - SF - SP	Swab Bucal	Deteção mutação germinativa	6,82E+07	99,999999%	Inclusão	3,73E+08	99,999999%	Inclusão
16	DUO	SF - SP	Sangue FTA	Deteção mutação germinativa	1,66E+06	99,9999%	Inclusão	4,94E+07	99,999999999999%	Inclusão
17	DUO	SF - SP	Sangue FTA	Suspeita alelo dropout	-	-	-	-	-	-
18	TRIO	MSF - SF - SP	Sangue FTA	Suspeita alelo dropout	-	-	-	-	-	-

MSP = Mãe biológica do Suposto Pai Ausente; PSP = Pai biológico do Suposto Pai Ausente; SF = Suposto Filho; MSF = Mãe biológica do Suposto Filho; SP = Suposto Pai; ICP = Índice Acumulado de Paternidade; PP = Probabilidade de Paternidade.

6.4.2.b. Detecção de mutação germinativa

O tipo de mutação mais comum observado em *loci* microssatélites é provocado pela variação da quantidade das unidades de repetição (Brinkmann *et al.*, 1998) e sua ocorrência em análise de parentesco tem sido amplamente reportada (Sun *et al.*, 2015; Zhang *et al.*, 2014; Sun *et al.*, 2012; Carboni *et al.*, 2011; Li *et al.*, 2011; Jacewics *et al.*, 2004). Para Lindner *et al.* (2014), em situações com constatação de incoerência alélica envolvendo um ou dois *loci* STRs, recomenda-se a análise suplementar com marcadores STRs adicionais para confirmação do vínculo pesquisado. Sendo assim, dois casos com registros de detecção de apenas um erro mendeliano foram selecionados para ampliação dos *loci* com o sistema multiplex desenvolvido. Os exames 15 e 16 correspondem, respectivamente, a um trio e a um duo, sendo submetidos a análise genética preliminar com o kit *PowerPlex® Fusion*.

No caso 15 foi detectada uma inconsistência alélica no *locus* Penta E, cuja taxa de mutação foi incorporada ao cálculo do índice de paternidade e obteve-se um ICP de 6,82E+07. O emprego do multiplex para ampliação da análise genética contribuiu para confirmação da incoerência no *locus* Penta E. Os dados gerados com os marcadores adicionais não-CODIS corroboraram a hipótese de que o erro mendeliano foi provocado por uma mutação germinativa na linhagem paterna. O ICP obtido após análise do multiplex correspondeu a 3,73E+08 e a Probabilidade de Paternidade foi de aproximadamente 99,999999%. Seguindo a recomendação proposta em Lindner *et al.* (2014), a comparação dos *loci* STRs presentes no cromossomo Y do suposto filho e do suposto pai foi realizada para confirmar o vínculo genético pesquisado. Os haplótipos foram obtidos a partir da amplificação de 23 *loci* STRs componentes do kit comercial *PowerPlex® Y23*.

Para o caso 16, uma incoerência alélica foi verificada no *locus* vWA entre os perfis genéticos analisados para o alelo obrigatório. Por se tratar de um duo, é fundamental que os painéis STRs analisados ofereçam segurança suficiente para uma conclusão inequívoca, uma vez que, teoricamente, o risco de inclusão ao acaso de um indivíduo aleatório da população, é maior que para situações envolvendo análise de trio. Nesse sentido, a necessidade de ampliação do conjunto de STRs fez-se necessária pelos mesmos motivos apresentados para o caso 15. Os ICPs obtidos antes e após a amplificação com o multiplex desenvolvido foi de 1,66E+06 e 4,94E+07, respectivamente, o que resultou em uma PP de aproximadamente 99,99999999999999%.

Os haplótipos do cromossomo X, analisados a partir da amplificação de 12 *loci* STRs presentes no kit comercial Argus X QS (Qiagen, Hilden, NRW, Germany), confirmaram o vínculo genético pesquisado.

6.4.2.c. Identificação de Alelo *Dropout*

A eficiência da amplificação alélica pode ser influenciada em função de inúmeros fatores (Moretti *et al.*, 2016; Zhang *et al.*, 2016; Westen *et al.*, 2014). Mutações nas regiões flanqueadoras do *locus*, presença de inibidores de PCR na reação, qualidade e quantidade de DNA disponível para a PCR, são eventos que podem implicar na inibição da amplificação do fragmento molecular e provocar o efeito de alelo *dropout* (Li *et al.*, 2017; Gjertson *et al.*, 2007; Grubwieser *et al.*, 2006). Portanto, a amplificação do mesmo *locus* por primers que se anelam em sequências diferentes pode contribuir para confirmação do efeito da mutação (Zhang *et al.*, 2016; BUDOWLE 2000).

Os casos 17, 18 e 19, apresentaram incoerências alélicas em *loci* individuais resultando em genótipos homocigotos. Assim, estes casos foram selecionados para verificação de suposto evento envolvendo alelo *dropout*. O uso do multiplex desenvolvido como conjunto alternativo de marcadores STRs, com *loci* sobrepostos com o sistema *PowerPlex® Fusion*, detectou os fragmentos não amplificados inicialmente e confirmou o evento de *dropout* alélico envolvendo três *loci* STRs do kit *PowerPlex® Fusion* (Borovko *et al.*, 2015; Gjertson *et al.*, 2007).

O caso 17 trata-se de uma investigação de paternidade classificada como um trio. A paternidade biológica foi confirmada pela constatação de correspondência genética em todos os *loci* analisados do kit comercial *PowerPlex® Fusion*. Porém, no *locus* D8S1179 foi observada uma incoerência genética envolvendo o suposto filho e sua mãe biológica.

Como pode ser visto na Figura 6 (A), o perfil alélico da mãe para o *locus* D8S1179, amplificado com o sistema *PowerPlex® Fusion*, corresponde ao genótipo homocigoto para o alelo 10, enquanto seu filho biológico possui apenas o alelo 16, revelando possíveis perfis falso homocigotos. No entanto, com a amplificação pelo painel multiplex desenvolvido, nota-se que o perfil alélico heterocigoto para as duas amostras, tendo a mãe os alelos 10 / 13 e seu filho, os alelos 13 / 16 (Figura 6 B).

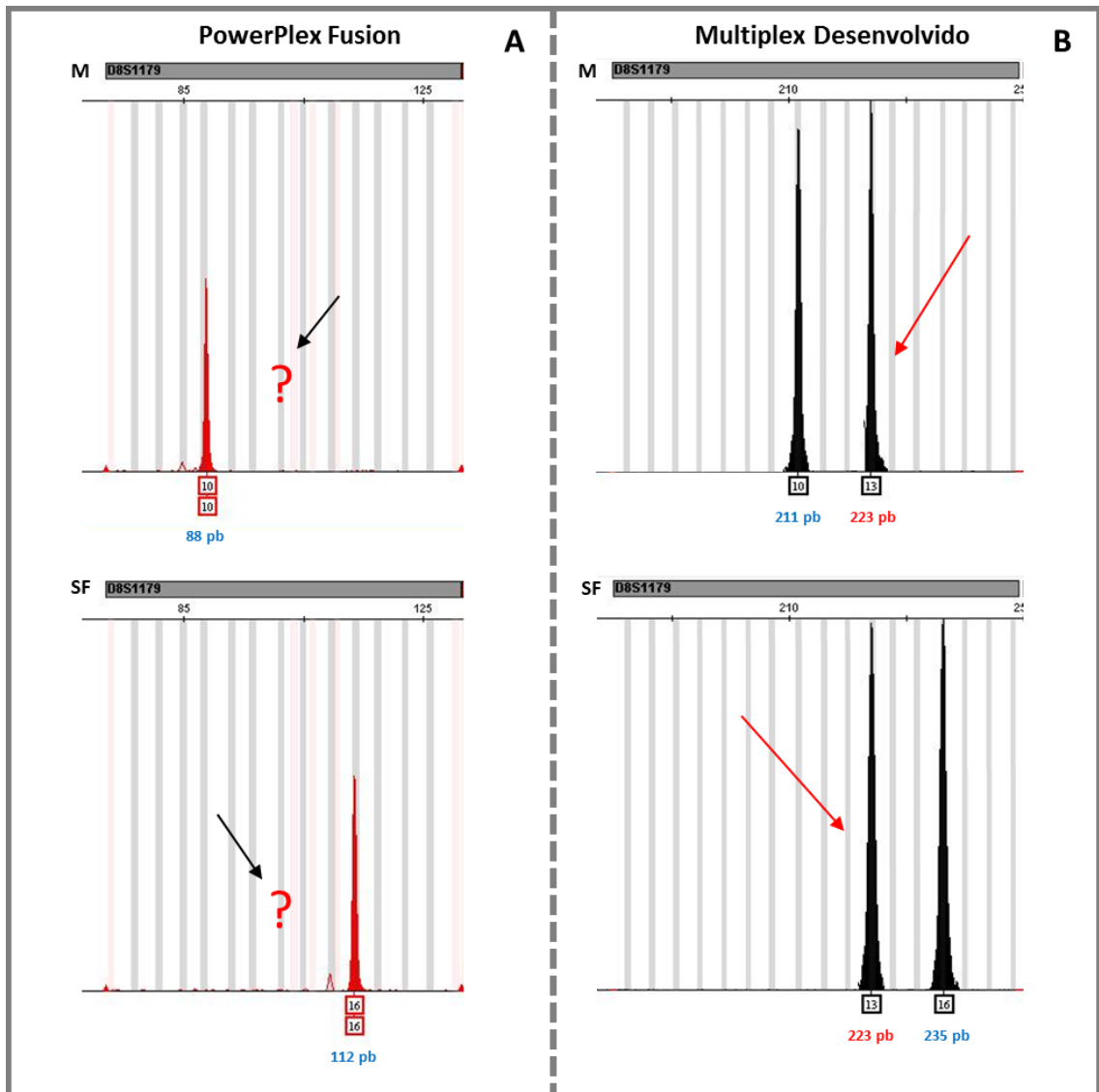


Figura 8. Eletroferogramas apresentando os genótipos para o *locus* D8S1179. (A) Perfil alélico obtido com a utilização do kit comercial *PowerPlex® Fusion*, mostrando a ausência da amplificação para o alelo 13, possivelmente provocada em decorrência de *dropout* alélico. (B) Perfil alélico heterozigoto obtido por amplificação com o multiplex desenvolvido, mostrando que houve amplificação normal para o alelo 13.

O alelo 13 do *locus* D8S1179, ausente nas amostras do caso 17 quando utilizado o kit comercial *PowerPlex® Fusion*, compreende um fragmento de aproximadamente 100 pb. Porém, quando a mesma amostra foi submetida à amplificação pelos iniciadores referentes ao mesmo *locus*, presentes no sistema multiplex desenvolvido, o alelo com 13 núcleos de repetição STR, foi amplificado, gerando um fragmento de aproximadamente 223 pb.

O evento de alelo *dropout* foi confirmado para o caso 18. Porém, nesse exame o efeito de *dropout* alélico envolveu o *locus* D5S818. A Figura 7 mostra os

eletroferogramas obtidos com a amplificação dos dois painéis, *PowerPlex® Fusion* e multiplex desenvolvido. O alelo 10 do *locus* D5S818 amplificado com o sistema comercial *PowerPlex® Fusion* corresponde a um fragmento de aproximadamente 338 pb, enquanto que a amplificação para o mesmo *locus* de um alelo com 10 núcleos de repetição, amplificado a partir dos *primers* presentes no multiplex desenvolvido, gerou um fragmento de aproximadamente 127 pb. A amplificação do alelo 10 do *locus* D5S818 foi observada a partir da utilização do sistema multiplex desenvolvido, enquanto que, quando a mesma amostra foi submetida à amplificação por *PowerPlex® Fusion*, não houve amplificação para o alelo em questão, revelando um *dropout* alélico.

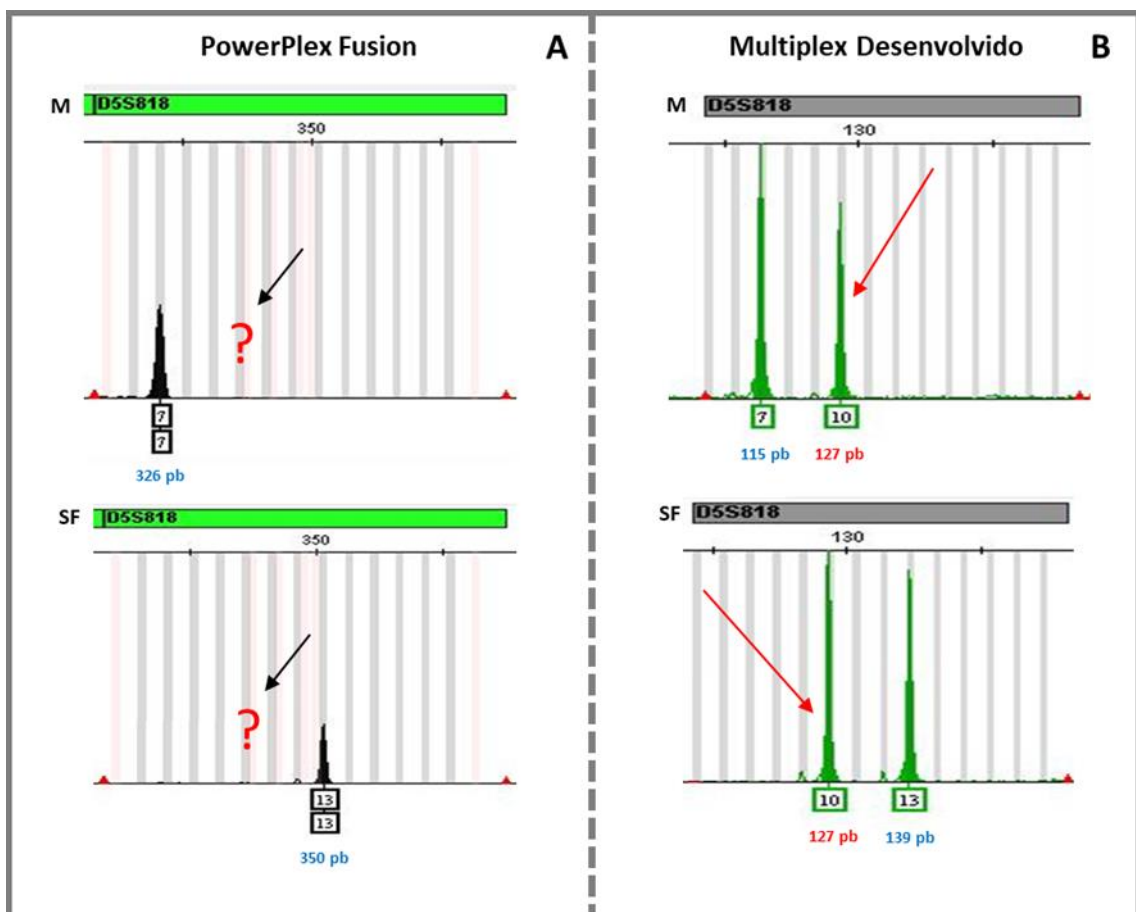


Figura 9. Eletroferogramas apresentando os genótipos para o *locus* D5S818. (A) Perfil alélico obtido com a utilização do kit comercial *PowerPlex® Fusion*, mostrando a ausência da amplificação para o alelo 10, provocada em decorrência de *dropout* alélico. (B) Perfil alélico heterozigoto obtido por amplificação com o multiplex desenvolvido, mostrando que houve amplificação normal para o alelo 10.

É importante ressaltar que, conforme demonstrado com exemplos reais, a possibilidade de indivíduos com perfis heterozigotos apresentarem fenótipos alélicos homozigotos, cria um cenário de potencial risco para determinação de falsa exclusão de

vínculo genético em análise de parentesco (Dakin and Avise, 2004). Neste sentido, para reduzir a possibilidade de interpretação equivocada, a existência de *dropout* alélico e outras variações na estrutura genômica, deve ser sempre considerada durante a análise genética por microssatélites (Zhou *et al.*, 2017; Zhang *et al.*, 2016; Lane, 2013).

A utilização do multiplex em situações especiais mostra que os *loci* STRs presentes no painel desenvolvido contribuíram para a conclusão de investigação de paternidade a partir de reconstrução genética complexa e, também, para a redução da possibilidade de equívoco provocado em função da detecção de mutação germinativa. Por último, os iniciadores presentes no multiplex desenvolvido foram capazes de revelar alelos *dropout* para os *loci* D8S1179 e D5S818, amplificados inicialmente com o sistema comercial *PowerPlex® Fusion*. Portanto, os resultados obtidos nessa seção evidenciam a eficácia do multiplex como ferramenta complementar para aplicação em situações complexas de identificação humana e investigação de vínculo genético.

7. CONCLUSÕES GERAIS

Em conformidade com o objetivo principal do estudo, foi desenvolvido um sistema multiplex composto por 21 *loci* STRs autossômicos e 1 *locus* para determinação sexual. Neste sentido:

- Em função da limitação quanto à quantidade de *loci* microssatélites presentes em um sistema multiplex, a partir de 43 *loci* pré-selecionados, 21 *loci* STRs autossômicos e 1 *locus* para determinação de sexo foram incluídos para composição do conjunto desenvolvido. Os *loci* STRs foram selecionados considerando seu polimorfismo e a variação de tamanho dos fragmentos esperados. O conjunto selecionado foi constituído por 15 *loci* STR sobrepostos ao sistema CODIS expandido e 6 *loci* não-CODIS.
- Foram determinados os oligonucleotídeos iniciadores passíveis de combinação em PCR para composição do multiplex. Os primers foram obtidos considerando a organização dos *loci* STR na dimensão espacial em função da separação dos fragmentos mediante eletroforese capilar, evitando sobreposição de tamanhos entre os marcadores
- As condições de amplificação e genotipagem dos microssatélites foram padronizadas para serem realizadas em uma reação multiplex e separação de fragmentos mediante eletroforese capilar.
- A avaliação da sensibilidade da amplificação de DNA em reação de PCR multiplex foi testada em partir de diferentes tipos de amostras biológicas. As amplificações foram eficientes a partir do DNA extraído de amostras de sangue, swab bucal e cabelo, que são as amostras mais comumente usadas para os estudos de vínculo genético com fins de identificação humana.
- Os testes de concordância revelaram que os genótipos obtidos por amplificação com o sistema multiplex e com o kit comercial *PowerPlex® Fusion*, apresentaram o mesmo perfil alélico para os *loci* sobrepostos dos dois conjuntos analisados.
- As frequências alélicas e os parâmetros forenses foram obtidos a partir da análise genotípica dos *loci* presentes no multiplex desenvolvido, para um grupo de 450 indivíduos não relacionados geneticamente e residentes no Estado de Goiás.

- O custo de amplificação por amostra foi determinado para o sistema multiplex e comparado com os principais kits comerciais atualmente disponíveis no Brasil. O custo de amplificação com o multiplex desenvolvido foi de R\$5,30 por amostra e corresponde à aproximadamente 6% do custo médio para os kits comerciais *PowerPlex® Fusion*, *GlobalFiler™ PCR Amplification* e *Investigator 24plex*.
- O sistema multiplex foi implantado em rotina laboratorial de identificação humana e investigação de vínculo genético, sendo sua eficiência avaliada em situações reais de investigação de vínculo biológico. Em todos os casos analisados, envolvendo baixa e alta complexidade, o sistema multiplex apresentou grande coerência com o kit comercial *PowerPlex® Fusion*. Os resultados obtidos mostraram a eficácia do multiplex para como ferramenta complementar para aplicação em situações complexas de investigação de vínculo genético, além de se mostrar eficiente para resolver os casos da rotina do Laboratório Biocroma de Goiânia (Goiás).

O sistema multiplex de STRs autossômicos desenvolvido apresentou-se sensível para a identificação humana a partir de diferentes tipos de amostras biológicas, incluindo amplificação direta em FTA, proporcionando agilidade para a obtenção de perfil genético na rotina laboratorial. Os valores obtidos para os parâmetros estatísticos sugerem que o sistema multiplex desenvolvido é polimórfico e apresenta um elevado poder de discriminação, o que revela sua eficiência e justifica a aplicação em estratégias envolvendo identificação humana e investigação de vínculo genético.

A relevância desse estudo, voltado para o desenvolvimento de uma ferramenta biotecnológica acessível, fica evidente ao constatar a performance satisfatória no auxílio à estudos envolvendo diversidade genética por microssatélites, além de, sobretudo, fornecer um mecanismo robusto e confiável capaz de discriminar com eficácia indivíduos entre si.

8. REFERÊNCIAS BIBLIOGRÁFICAS

Akhteruzzaman S, Majumder AK, Ferdous A, Ali ME (2012) False paternity with one or two mismatches using commercial STR kits. *Australian Journal of Forensic Sciences* 44(3), 253-259.

Amory S, Huel R, Bilić A, Loreille O, Parsons TJ (2012) Automatable full demineralization DNA extraction procedure from degraded skeletal remains. *Forensic Science International: Genetics* 6(3), 398-406.

Belkhir K, Borsa P, Chiki L, Raufaste N, Bonhomme FG (2004) Genetix 4.05, logiciel sous windows TM pour la genetique des populations. Laboratoire Genome, Populations, Interactions, CNRS UMR 5000. Universite de Montpellier II. Available at [<http://kimura.univ-montp2.fr/genetix/>].

Bennett P (2000) Demystified... microsatellites. *Journal of Clinical Pathology* 53(4), 177.

Besselink SBWVE (2003) Current and Future Developments in Forensic DNA Typing. Unisersiteit van Amsterdam. Justitie Nederlands Forensisch Instituut.

Borovko S, Shyla A, Korban V, Borovko A (2015) Amelogenin test abnormalities revealed in Belarusian population during forensic DNA analysis. *Forensic Science International: Genetics* 15, 98-104.

Børsting C, Sanchez JJ, Hansen HE, Hansen AJ, Bruun HQ, Morling N (2008) Performance of the *SNPforID* 52 SNP-plex assay in paternity testing. *Forensic Science International: Genetics* 2(4), 292-300.

Boyer JC, Hawk JD, Stefanovic L, Farber RA (2008) Sequence-dependent effect of interruptions on microsatellite mutation rate in mismatch repair-deficient human cells. *Mutation Research* 640(1), 89-96.

Brevnov MG, Pawar HS, Mundt J, Calandro LM, Furtado MR, Shewale JG (2009) Developmental validation of the PrepFiler™ forensic DNA extraction kit for extraction of genomic DNA from biological samples. *Journal of forensic Sciences* 54(3), 599-607.

Bright JA, Hopwood A, Curran JM, Buckleton JS (2014) A guide to forensic DNA interpretation and linkage. Promega Corporation, Updated.

Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *The American Journal of Human Genetics* 62(6), 1408-1415.

Britten RJ, Kohne DE (1968) Repeated sequences in DNA. *Science* 161(3841), 529-540.

Buckleton JS, Triggs CM, Walsh SJ (2004) Forensic DNA evidence interpretation. CRC Press.

Budowle B (2000) STR allele concordance between different primer sets—a brief summary. *Profiles DNA* 3(10).

Budowle B, Van Daal A (2009) Extracting evidence from forensic DNA analyses: Future molecular biology directions. *Biotechniques* 46(5), 339.

Bugert P, Rink G, Kemp K, Klüter H (2012) Blood group ABO genotyping in paternity testing. *Transfusion Medicine and Hemotherapy* 39(3), 182-186.

Butler JM (2012) Short tandem repeat (STR) loci and kits. *Advanced topics in forensic DNA typing: Methodology*. Elsevier Academic Press, San Diego, CA.

Butler JM (2007) Short tandem repeat typing technologies used in human identity testing. *Biotechniques* 43(4), 2-5.

Butler JM (2006) Genetics and genomics of core short tandem repeat loci used in human identity testing. *Journal of forensic Sciences* 51(2), 253-265.

Butler JM (2005) a. Constructing STR multiplex assays. *Methods in Molecular Biology*, 297, 53–66.

Butler JM (2005) b. *Forensic DNA typing: biology, technology, and genetics of STR markers*. 2^a ed. Elsevier Academic Press, San Diego, CA.

Butler JM, Hill CR (2012) Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic Science Review*. 24(1):15-26.

Carboni I, Iozzi S, Nutini AL, Macrì PG, Torricelli F, Ricci U (2011) 87 DNA markers for a paternity testing: Are they sufficient?. *Forensic Science International: Genetics Supplement Series*, 3(1), 552-553.

Castañeda M, Odriozola A, Gómez J, Zarrabeitia MT (2013) Development and validation of a multiplex reaction analyzing eight miniSTRs of the X chromosome for identity and kinship testing with degraded DNA. *International Journal of Legal Medicine*, 127(4), 735-739.

Chaudhary R, Dogra D, Sairkar M, Pathak N, Batav N (2013) The use STR loci in human identification. *Science Secure Journal of Biotechnology*, V2. (4), 92-98.

Chemello E (2007) *Ciência Forense: Exame de ADN*. *Revista Química Virtual*, 1-10.

Clamp M, Fry B, Kamal M, Xie X, Cuff J, Lin MF, Kellis M, Lindblad-Toh K, Lander ES (2007) Distinguishing protein-coding and noncoding genes in the human genome. *Proceedings of the National Academy of Sciences*, 104(49), 19428-19433.

Costa EOA, Silva DDM, de Melo AV, Godoy FR, Nunes HF, Pedrosa ER, Flores BC, Rodovalho RG, da Silva CC, da Cruz AD (2011) The effect of low-dose exposure on germline microsatellite mutation rates in humans accidentally exposed to caesium-137 in Goiânia. *Mutagenesis*, 26(5), 651-655.

Crouse CA, Rogers S, Amriott E, et al. Analysis and interpretation of short tandem repeat microvariants and three-banded allele patterns using multiple allele detection systems. *J Forensic Sci.* 1999;44(1):87-94.

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity*, 93(5), 504.

Decanine D (2016) O papel de marcadores moleculares na genética forense. *Revista Brasileira de Criminalística*, 5(2), 18-27.

DNA Advisory Board (2000) Quality Assurance Standards for Forensic DNA Testing Laboratories and Quality Assurance Standards for Convicted Offender DNA Databasing Laboratories, *Forensic Science Communications*, 2(3).

Deininger PL, Batzer MA (1999) Alu repeats and human disease. *Molecular genetics and metabolism*, 67(3), 183-193.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1: 47-50.

Edwards AL, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics*, 12(2), 241-253.

Edwards AL, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics*, 49(4), 746.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, 5(6), 435-445.

Fan H, Chu JY (2007) A brief review of short tandem repeat mutation. *Genomics, Proteomics & Bioinformatics*, 5(1), 7-14.

Figini AR, Silva JRL, Jobim LF, Silva M (2003) Identificação Humana. 2ª Ed. Millennium, Campinas, SP.

Forster P, Hohoff C, Dunkelmann B, Schürenkamp M, Pfeiffer H, Neuhuber F, Brinkmann B (2015) Elevated germline mutation rate in teenage fathers. In Proceedings of the Royal Society B, 282(1803), 20142898.

Francez PADC, Rodrigues EMR, Frazão GF, Borges NDDR, Santos SEBD (2011) Allelic frequencies and statistical data obtained from 12 codis STR loci in an admixed population of the Brazilian Amazon. Genetics and Molecular Biology, 34(1), 35-39.

Garrido RG (2009) Evolução dos Processos de Identificação Humana: das características antropométricas ao DNA. Genética na Escola, 5, 38-40.

Garrido RG, Rodrigues EL (2015) O Banco de Perfis Genéticos Brasileiro Três Anos após a Lei nº 12.654. Revista de Bioética y Derecho, (35), 94-107.

Garrido RG, Rodrigues EL (2014) Ciência Forense: Da cena do crime ao laboratório de DNA. Ed. Projeto Cultural/FAPERJ, Rio de Janeiro, RJ.

Gastaldo AZ (2012) Frequência alélica de sete marcadores *short tandem repeat* em indivíduos do Estado do Rio Grande do Sul, Brasil. Rio Grande do Sul, Brasil, (M.Sc. Dissertation. Pontifícia Universidade Católica do Rio Grande do Sul).

Ge J, Eisenberg A, Budowle B (2012) Developing criteria and data to determine best options for expanding the core CODIS loci. Investigative Genetics, 3(1), 1.

Gill P, Evett IW, Woodroffe S, Lygo JE, Millican E, Webster M (1991) Databases, quality control and interpretation of DNA profiling in the Home Office Forensic Science Service. Electrophoresis, 12(2-3), 204-209.

Gill P, Fereday L, Morling N, Schneider PM (2006) The evolution of DNA databases—recommendations for new European STR loci. *Forensic Science International*, 156(2), 242-244.

Gill P, Jeffreys AJ, Werrett DJ (1985) Forensic application of DNA 'fingerprints'. *Nature*, 318(6046), 577-579.

Gjertson DW, Brenner CH, Baur MP, Carracedo A, Guidet F, Luque JA, Lessig R, Mayr WR, Pascali VL, Prinz M, Schneider PM, Morling N (2007) ISFG: recommendations on biostatistics in paternity testing. *Forensic Science International: Genetics*, 1(3), 223-231.

Goldstein DB, Schlotterer C (1999) *Microsatellites: Evolution and applications*. Oxford University Press, New York, 343.

Grubwieser P, Mühlmann R, Berger B, Niederstätter H, Pavlic M, Parson W (2006) A new "miniSTR-multiplex" displaying reduced amplicon lengths for the analysis of degraded DNA. *International Journal of Legal Medicine*, 120(2), 115-120.

Gu Z, Wang H, Nekrutenko A, Li WH (2000) Densities, length proportions, and other distributional features of repetitive sequences in the human genome estimated from 430 megabases of genomic sequence. *Gene*, 259(1), 81-88.

Guo F, Shen H, Tian H, Jin P, Jiang X (2014) Development of a 24-locus multiplex system to incorporate the core loci in the Combined DNA Index System (CODIS) and the European Standard Set (ESS). *Forensic Science International: Genetics*, 8(1), 44-54.

Hares DR (2012) a. Addendum to expanding the CODIS core loci in the United States. *Forensic Science International: Genetics*, 6(5), 135.

Hares DR (2012) b. Expanding the CODIS core loci in the United States. *Forensic Science International: Genetics*, 6(1), 52-54.

Hares DR (2015) Selection and implementation of expanded CODIS core loci in the United States. *Forensic Science International: Genetics*, 17, 33-34.

Hennessy LK, Mehendale N, Chear K, Jovanovich S, Williams S, Park C, Gangano S (2014) Developmental validation of the GlobalFiler® express kit, a 24-marker STR assay, on the RapidHIT® System. *Forensic Science International: Genetics*, 13, 247-258.

Hill CR, Kline MC, Coble MD, Butler JM (2008) Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. *Journal of Forensic Sciences*, 53(1), 73-80.

Holt CL, Wallin JM, Lazaruk KD, Walsh PS, Nguyen TH (2002) Constructing universal multiplex PCR systems for comparative genotyping. *Journal of Forensic Science*, 47(1), 52-65.

Ikedo N, Kojima H, Nishikawa M, Hayashi K, Futagami T, Tsujino T, Kusunoki Y, Fujii N, Suegami S, Miyazaki Y, Middleton D, Tanaka H, Saji H (2015) Determination of HLA-A,-C,-B,-DRB1 allele and haplotype frequency in Japanese population based on family study. *Tissue Antigens*, 85(4), 252-259.

Jacewicz R, Berent J, Prośniak A, Dobosz T, Kowalczyk E, Szram S (2004) Non-exclusion paternity case with a triple genetic incompatibility. *International Congress Series*, 1261, 511-513.

Jeffreys A (2005) Genetic fingerprinting. *Nature Medicine*, 11(10), 14-18.

Jeffreys AJ, Brookfield JF, Semeonoff R (1985) Positive identification of an immigration test-case using human DNA fingerprints. *Nature*, 317(6040), 818-819.

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. *Nature*, 314(6006), 67-73.

Jha DK, Rijal JP, Tuladhar BS, Pokharel BR (2013) Mutations or Exclusion: An Uncommon Parentage Assessment Case. *Scientific World*, 11(11), 74-76.

Jin B, Su Q, Luo H, Li Y, Wu J, Yan J, Hou Y, Liang W, Zhang L (2016) Mutational analysis of 33 autosomal short tandem repeat (STR) loci in southwest Chinese Han population based on trio parentage testing. *Forensic Science International: Genetics*, 23, 86-90.

Jobim LF, Horta MF, Mur MC, Jobim MR (1996) Perícias médicas em investigações de paternidade pelos principais sistemas genéticos. *Revista HCPA*, 16, 13-20.

Jobling MA, Gill P (2004) Encoded evidence: DNA in forensic analysis. *Nature Reviews Genetics*, 5(10), 739-751.

Junge A, Brinkmann B, Fimmers R, Madea B (2006) Mutations or exclusion: an unusual case in paternity testing. *International Journal of Legal Medicine*, 120(6), 360-363.

Kashyap VK, Sitalaximi T, Chattopadhyay P, Trivedi R (2004) DNA profiling technologies in forensic analysis. *International Journal of Human Genetics*, 4(1), 11.

Katsanis SH, Wagner JK (2013) Characterization of the standard and recommended CODIS markers. *Journal of Forensic Sciences*, 1(58), 169-172.

Kouniaki DI, Papasteriades C, Tsirogianni A (2015) Short tandem repeats loci in parentage testing. *Hospital Chronicles*, 10(2), 83.

Kraemer M, Prochnow A, Bussmann M, Scherer M, Peist R, Steffen C (2017) Developmental validation of QIAGEN Investigator[®] 24plex QS Kit and Investigator[®] 24plex GO! Kit: Two 6-dye multiplex assays for the extended CODIS core loci. *Forensic Science International: Genetics*, 29, 9-20.

Lane AB (2013) STR null alleles complicate parentage testing in South Africa. *SAMJ: South African Medical Journal*, 103(12), 1004-1008.

Laurin N, DeMoors A, Frégeau C (2012) Performance of Identifiler Direct and PowerPlex 16 HS on the Applied Biosystems 3730 DNA Analyzer for processing biological samples archived on FTA cards. *Forensic Science International: Genetics*, 6(5), 621-629.

Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Haung J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AWC, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers Y, Frazier ME, Scherer EW, Strausberg RL, Venter JC (2007). The diploid genome sequence of an individual human. *PLoS biology*, 5(10), 254.

Li HX, Tong DY, Lu HL, Ou XL, Chen WJ, Zhang YM, Liu SJ, Chen Y, Sun HY (2011) Mutation analysis of 24 autosomal STR loci using in paternity testing. *Forensic Science International: Genetics Supplement Series*, 3(1), 159-160.

Li L, Ge J, Zhang S, Guo J, Zhao S, Li C, Tang H, Davis C, Budowle B, Hou Y, Liu Y (2012) Maternity exclusion with a very high autosomal STRs kinship index. *International journal of legal medicine*, 126(4), 645-648.

Li S, Chen L, Wang Y, Xu Q, Liu H, Li Y, Liu C (2017) Developmental validation of a 6-dye STR kit with 27 loci. *International Journal of Legal Medicine*, 1-8.

Lindner I, von Wurmb-Schwark N, Meier P, Fimmers R, Büttner A (2014) Usefulness of SNPs as supplementary markers in a paternity case with 3 genetic incompatibilities at autosomal and Y chromosomal loci. *Transfusion Medicine and Hemotherapy*, 41(2), 117-121.

Lins AM, Sprecher CJ, Puers C, Schumm JW (1996) Multiplex sets for the amplification of polymorphic short tandem repeat loci-silver stain and fluorescence detection. *BioTechniques*, 20(5), 882-889.

Lovett ST, Drapkin PT, Sutera VA, Gluckman-Peskind TJ (1993). A Sister-Strand Exchange Mechanism For recA-Independent Deletion of Repeated DNA Sequences in *Escherichia coli*. *Genetics*. 135 (3): 631–642.

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. *Genetics*, 152(4), 1753-1766.

Martínez G, Borosky A, Corach D, Llull C, Locarno L, Lojo M., Marino M, Miozzo MC, Modesti N, Pacharoni C, Pilili JP, Ramella MI, SalaA, Schaller C, Vullo C, Toscanini U (2017) Investigator[®] HDplex (Qiagen) reference population database for forensic use in Argentina. *Forensic Science International: Genetics*, 26, 91-95.

Mattana C, Fernandes MM, Tinoco R, Oliveira RN, Mattana R, Rodrigues B (2012) Forensic importance of DNA and the dentist participation. *Brazilian Journal of Forensic Sciences, Medical Law and Bioethics* 2(1) 5:65-82.

Moretti TR, Moreno LI, Smerick JB, Pignone ML, Hizon R, Buckleton JS, Bright JA, Onorato AJ (2016) Population data on the expanded CODIS core STR loci for eleven populations of significance for forensic DNA analyses in the United States. *Forensic Science International: Genetics*, 25, 175-181.

Morling N, Allen RW, Carracedo A, Geada H, Guidet F, Hallenberg C, Martin W, Mayr WR, Olaisen B, Pascali VL, Schneider PM (2002) Paternity Testing Commission of the International Society of Forensic Genetics: recommendations on genetic investigations in paternity cases. *Forensic Science International*, 129(3), 148-157.

Mortera J, Vecchiotti C, Zoppis S, Merigioli S (2016) Paternity testing that involves a DNA mixture. *Forensic Science International: Genetics*, 23, 50-54.

Myers BA, King JL, Budowle B (2012) Evaluation and comparative analysis of direct amplification of STRs using PowerPlex[®] 18D and Identifiler[®] Direct systems. *Forensic Science International: Genetics*, 6(5), 640-645.

Nacional Institute of Justice (2000) The future of Forensic DNA testing. U.S Department of Justice.

Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kulmin E, White R. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science*, 235, 1616-1623.

Oostdik K, Lenz K, Nye J, Schelling K, Yet D, Bruski S, Strong J, Buchanan C, Sutton J, Linner J, Frazier N, Young H, Matthies L, Sage A, Hahn J, Wells R, Williams N, Price M, Koehler J, Staples M, Swango KL, Hill C, Oyerly K, Duke W, Katzilierakis L, Ensenberger MG, Bourdeau JM, Sprecher CJ, Krenke B, Frazier N (2014). Developmental validation of the PowerPlex® Fusion System for analysis of casework and reference samples: a 24-locus multiplex for new database standards. *Forensic Science International: Genetics*, 12, 69-76.

Pacheco AC (2010) Emprego de miniSTRs “non-CODIS” em amostras biológicas de DNA forense. São Paulo, Brasil, (M.Sc. Dissertation. Instituto Butantan. IPT).

Pavlíček A, Jabbari K, Pačes J, Pačes V, Hejnar J, Bernardi G (2001) Similar integration but different stability of Alus and LINEs in the human genome. *Gene*, 276(1), 39-45.

Pena SD (2009) O DNA como (única) testemunha em determinação de paternidade. *Revista Bioética*, 5(2).

Pena SD (2005) Segurança pública: determinação de identidade genética pelo DNA. Seminários Temáticos para a 3ª Conferência Nacional de C, T & I. *Parcerias Estratégicas*, (20), 447-460.

Poetsch M, Preusse-Prange A, Schwark T, von Wurmb-Schwark N (2013) The new guidelines for paternity analysis in Germany—how many STR loci are necessary when investigating duo cases? *International Journal of Legal Medicine*, 127(4), 731-734.

Primorac D, Schanfield MS, Primorac D (2000) Application of forensic DNA testing in the legal system. *Croatian Medical Journal*, 41(1), 32-46.

Puers C, Hammond HA, Jin L, Caskey CT, Schumm JW (1993) 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*, 53(4), 953.

Rozen S, Skaletsky H (1999) Primer3 on the www for general users and for biologist programmers. *Bioinformatics Methods and Protocols*, 365-386.

Ruitberg CM, Reeder DJ, Butler JM (2001) STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Research*, 29(1), 320-322.

Sarmiento FJQ (2006) Modelagem de um ambiente para análise de DNA em genética forense. Alagoas, Brasil (M.Sc. Dissertation. Universidade Federal de Alagoas).

Scott AC, Harding HW, Fowler JCS, Burgoyne LA (1988) Repetitive deoxyribonucleic acid (DNA) and human genome variation--a concise review relevant to forensic biology. *Journal of Forensic Science*, 33(5), 1111-1126.

Sharma S, Tale V (2016) Overview of various DNA profiling techniques used in forensics. *International Journal of Pharma and Bio Sciences*, 7(3), 1186-1192.

Silva MB, e Silva DDM, Rodovalho RG, Rios PA, Silva WM, da Cruz AD (2010) Allele frequencies of fifteen STR loci in a population from Central Brazil. *Forensic Science International: Genetics*, 4(5), 151-152.

Smouse PE, Chakraborty R (1986) The use of restriction fragment length polymorphisms in paternity analysis. *American journal of human genetics*, 38(6), 918.

Strachan T, Read A (2011) Human molecular genetics. 4^a ed. Nova Iorque: Garland Science.

Sun HY, Li HX, Zeng XP, Ren Z, Chen WJ (2012) A paternity case with mutations at three CODIS core STR loci. *Forensic Science International: Genetics*, 6(1), 61-62.

Sun M, Zhang X, Wu D, Shen Q, Wu Y, Fu S (2016) Mutations of short tandem repeat loci in cases of paternity testing in Chinese. *International journal of legal medicine*, 130(5), 1203-1204.

Tsai C W, Yang CH, Chou SL, Cheng SG, Pai CY (2013) Non-CODIS DNA markers could be more effective than CODIS-based STRs in problematic biological relationship cases. *Romanian Journal of Legal Medicine*, 21, 245-248.

Tereba A (1999) Tools for Analysis of Population Statistics. *Profiles in DNA* 2:14-16.

Thomson JA, Ayres KL, Pilotti V, Barrett MN, Walker JIH, Debenham PG (2001) Analysis of disputed single-parent/child and sibling relationships using 16 STR loci. *International Journal of Legal Medicine*, 115(3), 128-134.

Thomson JA, Pilotti V, Stevens P, Ayres KL, Debenham PG (1999) Validation of short tandem repeat analysis for the investigation of cases of disputed paternity. *Forensic Science International*, 100(1), 1-16.

Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques*, 37(2), 226-231.

Vieira TC, Gigonzac MAD, Silva DM, Rodovalho RG, Santos GS, da Cruz AD (2014) Y-STR haplotype diversity and population data for Central Brazil: implications for environmental forensics and paternity testing. *Genetics and Molecular Research* 13 (2): 3404-3410.

Vieira TC, Silva DM, Gigonzac MAD, Ferreira VL, Gonçalves MW, da Cruz AD (2013) Allelic frequencies and statistical data obtained from 15 STR loci in a population of the Goiás State. *Genetics and Molecular Research* 12 (1): 23-27.

Vorwort I (2013) Richtlinie der Gendiagnostik-Kommission (GEKO) für die Anforderungen an die Qualitätssicherung genetischer Analysen zu medizinischen Zwecken gemäß § 23 Abs. 2 Nr. 4 GenDG.

Wang DY, Chang CW, Lagacé RE, Oldroyd NJ, Hennessy LK (2011) Development and Validation of the AmpF ℓ STR[®] Identifiler[®] Direct PCR Amplification Kit: A Multiplex Assay for the Direct Amplification of Single-Source Samples. *Journal of Forensic Sciences*, 56(4), 835-845.

Wang DY, Gopinath S, Lagacé RE, Norona W, Hennessy LK, Short ML, Mulero JJ (2015) Developmental validation of the GlobalFiler[®] Express PCR Amplification Kit: a 6-dye multiplex assay for the direct amplification of reference samples. *Forensic Science International: Genetics*, 19, 148-155.

Weber JL (1990) Informativeness of human (dC-dA) n-(dG-dT) n polymorphisms. *Genomics*, 7(4), 524-530.

Weber-Lehmann J, Schilling E, Gradl G, Richter DC, Wiehler J, Rolf B (2014) Finding the needle in the haystack: differentiating “identical” twins in paternity testing and forensics by ultra-deep next generation sequencing. *Forensic Science International: Genetics*, 9, 42-46

Weedn VW, Swarnen SL (1998) Exames forenses de identificação por análises do DNA. *Henry, JB Diagnósticos clínicos e tratamento por métodos laboratoriais*, 19, 1427-1438.

Westen AA, Kraaijenbrink T, de Medina EAR, Harteveld J, Willemse P, Zuniga SB, Van Der Gaag KJ, Weiler NEC, Warnaar J, Kayser M, Sijen T, Knijff P (2014) Comparing

six commercial autosomal STR kits in a large Dutch population sample. *Forensic Science International: Genetics*, 10, 55-63.

Wiegand P, Klein R, Braunschweiger G, Hohoff C, Brinkmann B (2006) Short amplicon STR multiplex for stain typing. *International Journal of Legal Medicine*, 120(3), 160-164.

Willems T, Gymrek M, Highnam G, Mittelman D, Erlich Y, 1000 Genomes Project Consortium (2014) The landscape of human STR variation. *Genome Research*, 24(11), 1894-1904.

Wilkening S, Chen B, Hemminki K, Försti A (2006) STR markers for kinship analysis. *Human Biology*, 78(1), 1-8.

Yuan L, Ge J, Lu D, Yang X (2012) Population data of 21 non-CODIS STR loci in Han population of northern China. *International journal of legal medicine*, 126(4), 659-664.

Zhao Z, Zhang J, Wang H, Liu ZP, Liu M, Zhang Y, Sun L, Zhang H (2015) Mutation rate estimation for 15 autosomal STR loci in a large population from Mainland China. *Meta gene*, 5, 150-156.

Zhang, S., Bian, Y., Tian, H., Wang, Z., Hu, Z., & Li, C. (2015). Development and validation of a new STR 25-plex typing system. *Forensic Science International: Genetics*, 17, 61-69.

Zhang MX, Gao HM, Han SY, Liu Y, Tian YL, Sun SH, Xiao DJ, Li CT, Wang YS (2014) Risk analysis of duo parentage testing with limited STR loci. *Genetics and Molecular Research*. 13 (1): 1179-1186.

Zhang L, Wu H, Wang K, Fan A, Niu J, Zhu Z (2016) Identification of a discordant genotype at the D2S1338 locus between Identifiler Plus[®] 16 kit and Power Plex[®] 21 system kit. *Journal of forensic and legal medicine*, 44, 174-177.

Zhou S, Wang H, Wang QK, Wang P, Wang F, Xu C (2017) Loss of heterozygosity

detected at three short tandem repeat locus commonly used for human DNA identification in a case of paternity testing. *Legal Medicine*, 24, 7-11.

Zhu BF, Zhang YD, Shen CM, Du WA, Liu WJ, Meng HT, Wang HD, Yang G, Jin R, Yang CH, Yan JW, Bie XH (2015) Developmental validation of the AGCU 21+ 1 STR kit: a novel multiplex assay for forensic application. *Electrophoresis*, 36(2), 271-276.

9. APÊNDICES

APÊNDICE I

Tabela suplementar - frequência alélica obtida para os 21 loci autossômicos que compõem o sistema multiplex desenvolvido

(continua)

Allele	D3S1358	TH01	D21S11	D18S51	Penta E	D5S818	D13S317	D7S820	D16S539	CSF1PO	Penta D	vWA	D8S1179	TPOX	FGA	D2S441	D17S1301	D19S433	D18S853	D20S482	D14S1434
2.2	-	-	-	-	-	-	-	-	-	-	0,0433	-	-	-	-	-	-	-	-	-	-
3.2	-	-	-	-	-	-	-	-	-	-	0,0044	-	-	-	-	-	-	-	-	-	-
5	-	0,0033	-	-	0,0589	-	-	-	0,0011	-	0,0200	-	-	0,0011	-	-	-	-	-	-	-
6	-	0,2100	-	-	-	0,0011	-	-	0,0011	0,0011	-	-	-	0,0189	-	-	-	-	-	-	-
7	-	0,2433	-	-	0,1267	0,0278	-	0,0111	-	0,0089	0,0144	-	-	0,0033	-	-	-	-	-	-	-
8	-	0,1133	-	-	0,0589	0,0167	0,1011	0,1544	0,0244	0,0189	0,0400	-	0,0067	0,4444	-	0,0022	-	-	-	-	0,0022
9	-	0,1744	-	-	0,0222	0,0333	0,0978	0,1133	0,2111	0,0300	0,1889	-	0,0133	0,1256	-	0,0022	0,0067	-	0,0056	0,0044	0,2078
9.2	-	-	-	-	-	-	-	-	-	-	0,0067	-	-	-	-	-	-	-	-	-	-
9.3	-	0,2433	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10	-	0,0122	-	0,0111	0,0533	0,0700	0,0567	0,2911	0,0756	0,2633	0,1311	-	0,0667	0,0833	-	0,2022	0,0222	0,0022	0,0189	0,0011	0,0367
10.2	-	-	-	0,0011	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0011	-	-	-
10.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0011	-	-	-	-	-
11	0,0022	-	-	0,0111	0,1100	0,3089	0,2822	0,2467	0,3022	0,2822	0,1533	0,0044	0,0844	0,2833	-	0,3189	0,2922	0,0278	0,4111	0,0233	0,0467
11.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0533	-	-	-	-	-
12	0,0022	-	-	0,1011	0,1789	0,3656	0,3056	0,1633	0,2167	0,3278	0,1567	-	0,1289	0,0400	-	0,0656	0,4500	0,0944	0,1133	0,0289	0,3111
12.2	-	-	-	0,0011	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0133	-	-	-
12.3	-	-	-	-	-	-	-	-	-	-	-	-	0,0011	-	-	0,0033	-	-	-	-	-
13	0,0044	-	-	0,0889	0,0922	0,1656	0,1056	0,0178	0,1456	0,0567	0,1667	0,0044	0,2978	-	-	0,0378	0,1900	0,2444	0,1422	0,1956	0,3678
13.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0522	-	-	-
14	0,1011	-	-	0,1522	0,0656	0,0100	0,0511	0,0011	0,0222	0,0100	0,0489	0,0733	0,2444	-	-	0,2744	0,0378	0,2811	0,2344	0,4700	0,0211
14.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0422	-	-	-
15	0,3067	-	-	0,1656	0,0778	0,0011	-	0,0011	-	0,0011	0,0233	0,1256	0,1256	-	-	0,0378	0,0011	0,1311	0,0711	0,2178	0,0067
15.2	0,0011	-	-	0,0022	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0556	-	-	-
16	0,2833	-	-	0,1256	0,0422	-	-	-	-	-	0,0022	0,2844	0,0244	-	0,0022	0,0011	-	0,0356	0,0033	0,0578	-
16.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0189	-	-	-
17	0,1944	-	-	0,1267	0,0344	-	-	-	-	-	-	0,2700	0,0067	-	0,0022	-	-	-	-	0,0011	-
18	0,1000	-	-	0,1022	0,0278	-	-	-	-	-	-	0,1600	-	-	0,0100	-	-	-	-	-	-
18.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0067	-	-	-	-	-	-
19	0,0044	-	-	0,0544	0,0178	-	-	-	-	-	-	0,0522	-	-	0,0878	-	-	-	-	-	-

(continuação)

Allele	D3S1358	TH01	D21S11	D18S51	Penta E	D5S818	D13S317	D7S820	D16S539	CSF1PO	Penta D	vWA	D8S1179	TPOX	FGA	D2S441	D17S1301	D19S433	D18S853	D20S482	D14S1434
19.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0011	-	-	-	-	-	-
20	-	-	-	0,0311	0,0122	-	-	-	-	-	-	0,0244	-	-	0,0900	-	-	-	-	-	-
20.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0011	-	-	-	-	-	-
21	-	-	-	0,0133	0,0078	-	-	-	-	-	-	0,0011	-	-	0,1422	-	-	-	-	-	-
21.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0033	-	-	-	-	-	-
22	-	-	-	0,0089	0,0111	-	-	-	-	-	-	-	-	-	0,1589	-	-	-	-	-	-
22.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0078	-	-	-	-	-	-
23	-	-	-	0,0022	0,0022	-	-	-	-	-	-	-	-	-	0,1522	-	-	-	-	-	-
23.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0044	-	-	-	-	-	-
24	-	-	-	0,0011	-	-	-	-	-	-	-	-	-	-	0,1422	-	-	-	-	-	-
24.2	-	-	0,0022	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0967	-	-	-	-	-	-
26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0533	-	-	-	-	-	-
27	-	-	0,0278	-	-	-	-	-	-	-	-	-	-	-	0,0222	-	-	-	-	-	-
28	-	-	0,1789	-	-	-	-	-	-	-	-	-	-	-	0,0078	-	-	-	-	-	-
29	-	-	0,1956	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30	-	-	0,2533	-	-	-	-	-	-	-	-	-	-	-	0,0011	-	-	-	-	-	-
30.2	-	-	0,0244	-	-	-	-	-	-	-	-	-	-	-	0,0011	-	-	-	-	-	-
31	-	-	0,0789	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31.2	-	-	0,0922	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
32	-	-	0,0067	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
32.2	-	-	0,0789	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
33.2	-	-	0,0367	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
34	-	-	0,0011	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
34.2	-	-	0,0078	-	-	-	-	-	-	-	-	-	-	-	0,0011	-	-	-	-	-	-
35	-	-	0,0111	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
36	-	-	0,0044	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0011	-	-	-	-	-	-
43.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,0033	-	-	-	-	-	-
N	450	450	450	450	450	450	450	450	450	450	450	450	450	450	450	450	450	450	450	450	450

Fonte: Rodvalho RG, Rodrigues EL, Santos GS, Cavalcanti LM, Lima PR, Rodvalho AG, Vital RG, Gigonzac MAD, da Cruz, AD (2017). Development of a polymorphic short tandem repeat locus multiplex system for efficient human identification. Genetics and molecular research: GMR, 16(2). (Capítulo 4 da tese)

APÊNDICE II

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado (a) para participar, como voluntário (a), do Projeto de Pesquisa sob o título **Desenvolvimento de um sistema multiplex de loci STRs polimórficos eficiente para a identificação humana**. Meu nome é Ricardo Goulart Rodovalho, sou o pesquisador (a) responsável, doutorando em Biotecnologia e Biodiversidade. Após receber os esclarecimentos e as informações a seguir, no caso de aceitar fazer parte do estudo, este documento deverá ser assinado em duas vias e em todas as páginas, sendo a primeira via de guarda e confidencialidade do Pesquisador (a) responsável e a segunda via ficará sob sua responsabilidade para quaisquer fins. Em caso de recusa, você não será penalizado (a) de forma alguma. Em caso de dúvida sobre a pesquisa, você poderá entrar em contato com o (a) pesquisador (a) responsável Ricardo Goulart Rodovalho ou com o (a) orientador (a) da pesquisa Professor (a) Aparecido Divino da Cruz, nos telefones: (62) 8105-1155/ (62) 3945-8162, ou através do e-mail ricardogrovalho@gmail.com.

Em caso de dúvida sobre a ética aplicada a pesquisa, você poderá entrar em contato com o Comitê de Ética em Pesquisa da Pontifícia Universidade Católica de Goiás, telefone: (62) 3946-1512, localizado na Avenida Universitária, Nº 1069, Setor Universitário, Goiânia – Goiás.

Para garantir o princípio da autonomia ao participante, atendendo o disposto na Resolução 466/12, o esclarecimento referente à pesquisa será realizado antes da coleta de material biológico, nas dependências do Laboratório Biocroma Clínica de Exames de DNA. Sua participação na pesquisa se justifica porque o levantamento de perfil genético individual deve ser realizado ao acaso na população, tendo esta pesquisa como objetivo desenvolver uma ferramenta altamente eficaz para o processo de identificação humana pelo DNA.

Com a utilização de uma lanceta, será realizada uma coleta de sangue a partir de uma pequena punção (“furo”) no dedo anelar ou médio direito. Os riscos decorrentes da sua participação na pesquisa são mínimos, próprios de qualquer coleta de sangue, que são, dor no local e possível aparecimento de manchas roxas (hematomas). Caso ocorra qualquer intercorrência devido à coleta de sangue (crise nervosa, dificuldade respiratória, aumento da pressão arterial, sudorese intensa), o (a) Sr. (a) será encaminhado ao serviço de assistência à saúde gratuita e integral para dano direto ou indireto, imediato ou tardio

em decorrência da sua participação na pesquisa. Para minimizar os riscos durante a coleta de sangue, o coletador será um profissional qualificado e com experiência para esta coleta e ainda o participante do estudo será orientado a pressionar a região coletada por um minuto para evitar a formação de hematomas.

Em seguida, todos os procedimentos de manipulação serão aplicados ao DNA extraído desta amostra obtida. O DNA extraído então será submetido às metodologias próprias (PCR e eletroforese) para obtenção do perfil genético a partir da análise das regiões STRs hipervariáveis. Sendo assim, sua participação nesta pesquisa, no entanto, é de extrema importância, pois seus dados genéticos contribuirão para a validação do sistema a ser desenvolvido e para a composição do banco de frequência alélica (perfis de DNA que permitem estimar a diversidade genética de uma população) para sua região geográfica específica. O (a) Sr. (a) tem a opção de tomar conhecimento ou não dos resultados genéticos obtidos.

A sua participação nesta pesquisa é voluntária e a recusa em participar não acarretará qualquer penalidade ou perda de benefícios. O pesquisador irá tratar a sua identidade com padrões profissionais de sigilo e as informações colhidas serão utilizadas somente para fins científicos. Os resultados da pesquisa serão analisados e publicados, mas a sua identidade não será divulgada, sendo guardada em sigilo. Todos os dados que permitam a identificação pessoal serão mantidos em sigilo profissional e científico, sendo-lhe garantido que todos os resultados obtidos serão utilizados somente para estudos científicos e não irão prejudicar qualquer tratamento que o participante esteja sendo submetido (a).

É importante ressaltar que o(a) Sr.(a) não terá ônus financeiro pela sua participação nesta pesquisa e, por ser voluntária, também não haverá nenhuma remuneração financeira. Os gastos necessários para sua participação na pesquisa serão integralmente assumidos pelo pesquisador, sendo ressarcidas todas as despesas geradas pelo Sr. (a) e seus acompanhantes para esta participação. Fica também garantida indenização em casos de danos comprovadamente decorrentes da participação na pesquisa, conforme decisão judicial (justiça comum).

O (a) Sr. (a) poderá a qualquer momento, antes ou durante sua participação, sem penalidades ou prejuízo, desistir de participar desta pesquisa.

Eu _____, RG _____, abaixo assinado, discuti com o Sr. (Ricardo Goulart Rodovalho) sobre a minha decisão em participar nesse estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é isenta de despesas e que tenho garantia de assistência integral e gratuita por danos diretos e indiretos, imediatos ou tardios quando necessário. Concordo voluntariamente em participar deste estudo e poderei retirar o meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades ou prejuízo ou perda de qualquer benefício que eu possa ter adquirido, ou no meu atendimento neste serviço.

Goiânia, _____, de _____, de 201____.

Assinatura do participante

____/____/_____
Data

Assinatura do responsável pelo estudo

____/____/_____
Data

APÊNDICE III

TABELAS ALÉLICAS – ESTUDO DE CASOS

Para adequada interpretação das tabelas alélicas, a legenda abaixo deverá ser consultada.

CÓDIGO	DESCRIÇÃO
MSP	Mãe biológica do Suposto Pai Ausente
PSP	Pai biológico do Suposto Pai Ausente
SF	Suposto Filho
MSF	Mãe biológica do Suposto Filho
SP	Suposto Pai
IP	Índice de Paternidade

Caso 1

PowerPlex® Fusion									
Locus STR	MSP	PSP	SF	MSF	IP				
D3S1358	14	16	15	17	14	17	17	17	2,6837
D1S1656	15	16	15	16,3	11	15	11	13	3,1506
D2S441	11	14	12	12	10	14	10	10	0,9597
D10S1248	13	14	13	14	11	14	11	14	1,5934
D13S317	12	13	11	12	11	11	10	11	0,8469
Penta E	11	12	7	12	8	11	8	8	2,2664
D16S539	9	11	8	13	9	11	9	11	1,1136
D18S51	17	17	18	18	18	22	11	22	6,4466
D2S1338	16	17	17	21	17	20	17	20	1,5142
CSF1PO	12	12	12	12	12	12	10	12	3,3063
Penta D	9	12	11	11	9	11	8	11	1,3577
TH01	6	9,3	7	7	9,3	9,3	6	9,3	1,1054
vWA	16	20	16	17	17	20	16	17	17,6816
D21S11	29	30	28	32,2	31	32,2	31	31	3,3277
D7S820	8	10	10	11	10	11	8	11	1,818
D5S818	11	13	10	11	11	12	11	12	0,7403
TPOX	8	11	8	8	8	10	9	10	1,6452
D8S1179	14	14	11	14	11	14	14	15	3,2423
D12S391	17	24	18	19	17	21	19	21	2,6042
D19S433	14	16,2	13	15,2	15	15,2	14	15	4,8924
FGA	20	25	21	23	20	22	21	22	2,1948
D22S1045	15	16	15	16	14	15	14	18	1,387
Amelogenina	X	X	X	Y	X	Y	X	X	1,0000

Índice Acumulado de Paternidade: **2,19E+07**
 Probabilidade de Paternidade: **99,99999%**

Multiplex desenvolvido									
Locus STR	MSP	PSP	SF	MSF	IP				
D3S1358	14	16	15	17	14	17	17	17	2,6837
TH01	6	9,3	7	7	9,3	9,3	6	9,3	1,1054
D21S11	29	30	28	32,2	31	32,2	31	31	3,3277
D18S51	17	17	18	18	18	22	11	22	6,4466
Penta E	11	12	7	12	8	11	8	8	2,2664
D5S818	11	13	10	11	11	12	11	12	0,7403
D13S317	12	13	11	12	11	11	10	11	0,8469
D7S820	8	10	10	11	10	11	8	11	1,818
D16S539	9	11	8	13	9	11	9	11	1,1136
CSF1PO	12	12	12	12	12	12	10	12	3,3063
Penta D	9	12	11	11	9	11	8	11	1,3577
vWA	16	20	16	17	17	20	16	17	17,6816
D8S1179	14	14	11	14	11	14	14	15	3,2423
TPOX	8	11	8	8	8	10	9	10	1,6452
FGA	20	25	21	23	20	22	21	22	2,1948
D2S441	11	14	12	12	10	14	10	10	0,9597
D17S1301	11	12	13	14	11	12	11	12	0,6844
D19S433	14	16,2	13	15,2	15	15,2	14	15	4,8924
D18S853	11	13	11	14	11	14	11	13	1,0309
D20S482	15	15	15	15	13	15	13	14	3,3127
D14S1434	11	13	9	13	11	12	12	12	5,3305
Amelogenina	X	X	X	Y	X	Y	X	X	1,0000

Índice Acumulado de Paternidade: **9,95E+06**
 Probabilidade de Paternidade: **99,99999%**

Caso 2

PowerPlex® Fusion							
Locus STR	MSF		SF		SP		IP Exclusões
D3S1358	15	16	15	15	14	16	Exclusão
D1S1656	15	17,3	15,3	17,3	12	17,3	Exclusão
D2S441	12	15	11	15	11	14	1,7325
D10S1248	13	14	14	14	11	15	Exclusão
D13S317	11	12	11	12	11	11	1,6992
Penta E	13	16	9	13	8	12	Exclusão
D16S539	11	12	11	12	9	12	0,9179
D18S51	14	16	14	18	12	17	Exclusão
D2S1338	17	18	17	17	17	24	2,4355
CSF1PO	12	13	7	12	8	11	Exclusão
Penta D	10	13	12	13	10	13	Exclusão
TH01	6	9	9	9	9	9,3	3,0757
vWA	15	16	15	17	16	18	Exclusão
D21S11	30	31,2	30	30	28	31,2	Exclusão
D7S820	11	12	11	11	10	11	2,1154
D5S818	11	11	11	14	9	11	Exclusão
TPOX	8	10	10	11	8	9	Exclusão
D8S1179	12	15	12	14	15	15	Exclusão
D12S391	15	23	19	23	17	18	Exclusão
D19S433	14	14	13,2	14	13,2	15	9,3458
FGA	22	27	24	27	22	23	Exclusão
D22S1045	15	15	15	16	11	16	1,5211
Amelogenina	X	X	X	X	X	Y	1,0000
Quantidade incoerências alélicas					14		

Multiplex desenvolvido							
Locus STR	MSF		SF		SP		IP Exclusões
D3S1358	15	16	15	15	14	16	Exclusão
TH01	6	9	9	9	9	9,3	3,0757
D21S11	30	31,2	30	30	28	31,2	Exclusão
D18S51	14	16	14	18	12	17	Exclusão
Penta E	13	16	9	13	8	12	Exclusão
D5S818	11	11	11	14	9	11	Exclusão
D13S317	11	12	11	12	11	11	1,6992
D7S820	11	12	11	11	10	11	2,1154
D16S539	11	12	11	12	9	12	0,9179
CSF1PO	12	13	7	12	8	11	Exclusão
Penta D	10	13	12	13	10	13	Exclusão
vWA	15	16	15	17	16	18	Exclusão
D8S1179	12	15	12	14	15	15	Exclusão
TPOX	8	10	10	11	8	9	Exclusão
FGA	22	27	24	27	22	23	Exclusão
D2S441	12	15	11	15	11	14	1,7325
D17S1301	11	13	11	12	12	12	2,2222
D19S433	14	14	13,2	14	13,2	15	9,3458
D18S853	14	14	14	14	12	13	Exclusão
D20S482	15	15	13	15	11	14	Exclusão
D14S1434	9	13	12	13	13	13	Exclusão
Amelogenina	X	X	X	X	X	Y	1,0000
Quantidade incoerências alélicas					14		

Caso 3

PowerPlex® Fusion							
-------------------	--	--	--	--	--	--	--

<i>Locus STR</i>	MSF		SF		SP		IP Exclusões
D3S1358	15	16	15	16	15	17	0,8908
D1S1656	13	15	13	15	12	16,3	Exclusão
D2S441	10	10	10	11	11	14	1,7325
D10S1248	14	15	14	14	13	14	1,6404
D13S317	11	12	12	12	11	14	Exclusão
Penta E	5	21	8	21	13	15	Exclusão
D16S539	11	12	11	11	11	11	3,4640
D18S51	17	18	15	18	16	18	Exclusão
D2S1338	21	22	22	24	17	25	Exclusão
CSF1PO	7	13	12	13	10	11	Exclusão
Penta D	9	12	9	12	9	13	1,4303
TH01	6	7	7	7	7	9	2,0794
vWA	15	18	18	18	18	19	2,9385
D21S11	28	31,2	28	30	28	29	Exclusão
D7S820	12	12	11	12	11	12	2,1154
D5S818	11	12	11	12	12	12	1,4805
TPOX	8	9	8	11	9	12	Exclusão
D8S1179	13	14	14	16	10	13	Exclusão
D12S391	18	19	18	19	18	19,3	1,3974
D19S433	13	13,2	13	15,2	12	15,2	9,7847
FGA	21	24	21	24	23	31,2	Exclusão
D22S1045	11	16	11	15	11	15	1,3870
Amelogenina	X	X	X	Y	X	Y	1,0000

Quantidade incoerências alélicas

10

Multiplex desenvolvido							
<i>Locus STR</i>	MSF		SF		SP		IP Exclusões
D3S1358	15	16	15	16	15	17	0,8908
TH01	6	7	7	7	7	9	2,0794
D21S11	28	31,2	28	30	28	29	Exclusão
D18S51	17	18	15	18	16	18	Exclusão
Penta E	5	21	8	21	13	15	Exclusão
D5S818	11	12	11	12	12	12	1,4805
D13S317	11	12	12	12	11	14	Exclusão
D7S820	12	12	11	12	11	12	2,1154
D16S539	11	12	11	11	11	11	3,4640
CSF1PO	7	13	12	13	10	11	Exclusão
Penta D	9	12	9	12	9	13	1,4303
vWA	15	18	18	18	18	19	2,9385
D8S1179	13	14	14	16	10	13	Exclusão
TPOX	8	9	8	11	9	12	Exclusão
FGA	21	24	21	24	23	31,2	Exclusão
D2S441	10	10	10	11	11	14	1,7325
D17S1301	12	14	12	14	10	12	1,0251
D19S433	13	13,2	13	15,2	12	15,2	9,7847
D18S853	8	11	11	11	11	11	2,4324
D20S482	14	15	15	16	11	16	8,6538
D14S1434	12	13	12	13	12	13	1,4730
Amelogenina	X	X	X	Y	X	Y	1,0000

Quantidade incoerências alélicas

8

Caso 4

PowerPlex® Fusion

<i>Locus STR</i>	MSF		SF		SP		IP Exclusões
D3S1358	14	15	14	15	15	18	1,3016
D1S1656	14	17,3	14	15	15	17,3	3,1506
D2S441	10	11	11	14	10	14	1,9194
D10S1248	14	16	14	16	14	14	2,3787
D13S317	8	13	8	11	8	12	Exclusão
Penta E	10	13	10	16	15	17	Exclusão
D16S539	11	12	9	12	9	11	3,1192
D18S51	12	18	12	15	15	19	3,2849
D2S1338	17	23	19	23	18	25	Exclusão
CSF1PO	12	12	10	12	10	11	1,9057
Penta D	11	11	11	13	12	13	3,2395
TH01	6	8	6	6	6	9,3	2,367
vWA	16	17	16	19	14	19	7,6073
D21S11	28	29	28	30	27	30	2,1582
D7S820	10	12	10	11	10	11	2,1154
D5S818	12	12	11	12	11	12	1,5195
TPOX	8	11	8	8	8	11	1,0968
D8S1179	9	13	9	14	14	14	3,9861
D12S391	19	19	19	19	19	19	5,8858
D19S433	13	15	13	13	13	14	2,0048
FGA	21	23	23	24	19	21	Exclusão
D22S1045	11	11	11	17	11	17	4,1528
Amelogenina	X	X	X	Y	X	Y	1,0000

Quantidade incoerências alélicas

4

Multiplex desenvolvido							
<i>Locus STR</i>	MSF		SF		SP		IP Exclusões
D3S1358	14	15	14	15	15	18	1,3016
TH01	6	8	6	6	6	9,3	2,367
D21S11	28	29	28	30	27	30	2,1582
D18S51	12	18	12	15	15	19	3,2849
Penta E	10	13	10	16	15	17	Exclusão
D5S818	12	12	11	12	11	12	1,5195
D13S317	8	13	8	11	8	12	Exclusão
D7S820	10	12	10	11	10	11	2,1154
D16S539	11	12	9	12	9	11	3,1192
CSF1PO	12	12	10	12	10	11	1,9057
Penta D	11	11	11	13	12	13	3,2395
vWA	16	17	16	19	14	19	7,6073
D8S1179	9	13	9	14	14	14	3,9861
TPOX	8	11	8	8	8	11	1,0968
FGA	21	23	23	24	19	21	Exclusão
D2S441	10	11	11	14	10	14	1,9194
D17S1301	11	12	12	12	12	12	2,2936
D19S433	13	15	13	13	13	14	2,0048
D18S853	13	14	11	13	14	15	Exclusão
D20S482	14	15	12	14	14	16	Exclusão
D14S1434	9	10	9	13	13	13	2,7152
Amelogenina	X	X	X	Y	X	Y	1,0000

Quantidade incoerências alélicas

5

Caso 5

PowerPlex® Fusion

<i>Locus STR</i>	MSF		SF		SP		IP
D3S1358	16	18	14	16	14	18	5,3673
D1S1656	13	18,3	12	13	11	12	5,0100
D2S441	10	14	14	16	10	16	166,6667
D10S1248	13	15	14	15	12	14	1,6404
D13S317	9	11	11	12	8	12	1,7045
Penta E	8	10	8	14	14	14	14,3978
D16S539	11	11	11	11	11	12	1,7320
D18S51	15	18	15	18	14	15	2,1761
D2S1338	24	24	21	24	21	23	10,1626
CSF1PO	10	12	10	12	8	10	0,8852
Penta D	9	13	13	14	9	14	8,3433
TH01	7	9,3	7	8	8	8	6,6955
vWA	14	19	14	20	18	20	35,3632
D21S11	30	30	28	30	28	30	3,1374
D7S820	11	12	11	12	11	12	2,5142
D5S818	11	13	12	13	12	12	2,8870
TPOX	8	11	8	9	9	10	4,0285
D8S1179	10	13	13	16	16	16	29,7354
D12S391	18	20	20	22	20	22	7,9365
D19S433	11	15	14,2	15	13	14,2	10,9409
FGA	25	25	25	27	22	27	31,3421
D22S1045	16	16	16	16	11	16	1,5211
Amelogenina	X	X	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:							1,36E+17
Probabilidade de Paternidade:							99,9999999999999%

Multiplex desenvolvido							
<i>Locus STR</i>	MSF		SF		SP		IP
D3S1358	16	18	14	16	14	18	5,3673
TH01	7	9,3	7	8	8	8	6,6955
D21S11	30	30	28	30	28	30	3,1374
D18S51	15	18	15	18	14	15	2,1761
Penta E	8	10	8	14	14	14	14,3978
D5S818	11	13	12	13	12	12	2,8870
D13S317	9	11	11	12	8	12	1,7045
D7S820	11	12	11	12	11	12	2,5142
D16S539	11	11	11	11	11	12	1,7320
CSF1PO	10	12	10	12	8	10	0,8852
Penta D	9	13	13	14	9	14	8,3433
vWA	14	19	14	20	18	20	35,3632
D8S1179	10	13	13	16	16	16	29,7354
TPOX	8	11	8	9	9	10	4,0285
FGA	25	25	25	27	22	27	31,3421
D2S441	10	14	14	16	10	16	166,6667
D17S1301	11	12	12	12	12	13	1,1111
D19S433	11	15	14,2	15	13	14,2	10,9409
D18S853	11	14	14	15	12	15	7,0313
D20S482	13	14	13	14	14	15	0,7513
D14S1434	9	12	9	13	9	13	1,3595
Amelogenina	X	X	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:							1,08E+15
Probabilidade de Paternidade:							99,9999999999999%

Caso 6

PowerPlex® Fusion

<i>Locus STR</i>	MSF		SF		SP		IP
D3S1358	15	17	15	17	15	15	2,0232
D1S1656	13	14	11	14	11	15	11,8765
D2S441	14	15	12	15	11	12	6,8399
D10S1248	14	16	16	16	11	16	4,3253
D13S317	9	10	9	13	12	13	3,9550
Penta E	11	11	8	11	8	11	8,6226
D16S539	12	13	9	12	9	10	3,1192
D18S51	16	18	17	18	17	18	4,1012
D2S1338	20	21	19	20	19	23	3,9841
CSF1PO	10	11	10	12	12	13	1,6532
Penta D	5	9	9	9	5	9	2,7154
TH01	7	7	7	8	7	8	3,3478
vWA	15	19	15	16	16	17	1,9229
D21S11	28	35	28	30	28	30	2,1582
D7S820	10	11	10	11	11	13	0,9777
D5S818	12	12	12	12	12	12	2,8870
TPOX	8	8	8	9	8	9	4,0285
D8S1179	11	13	11	15	14	15	3,6145
D12S391	18	19	15	19	15	20	10,9649
D19S433	14	15	14	15	13,2	15	1,2475
FGA	22	24	22	24	22	26	1,5995
D22S1045	15	15	15	16	15	16	1,5211
Amelogenina	X	X	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:							1,05E+11
Probabilidade de Paternidade:							99,99999999%

Multiplex desenvolvido							
<i>Locus STR</i>	MSF		SF		SP		IP
D3S1358	15	17	15	17	15	15	2,0232
TH01	7	7	7	8	7	8	3,3478
D21S11	28	35	28	30	28	30	2,1582
D18S51	16	18	17	18	17	18	4,1012
Penta E	11	11	8	11	8	11	8,6226
D5S818	12	12	12	12	12	12	2,8870
D13S317	9	10	9	13	12	13	3,9550
D7S820	10	11	10	11	11	13	0,9777
D16S539	12	13	9	12	9	10	3,1192
CSF1PO	10	11	10	12	12	13	1,6532
Penta D	5	9	9	9	5	9	2,7154
vWA	15	19	15	16	16	17	1,9229
D8S1179	11	13	11	15	14	15	3,6145
TPOX	8	8	8	9	8	9	4,0285
FGA	22	24	22	24	22	26	1,5995
D2S441	14	15	12	15	11	12	6,8399
D17S1301	12	13	12	12	12	13	1,1111
D19S433	14	16	14	15	13,2	15	3,6023
D18S853	11	14	11	14	11	14	1,5491
D20S482	14	15	14	14	14	14	2,1277
D14S1434	9	13	9	9	9	12	2,4064
Amelogenina	X	X	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:							7,86E+08
Probabilidade de Paternidade:							99,9999999%

Caso 7

PowerPlex® Fusion

<i>Locus STR</i>	MSF	SF	SP	IP			
D3S1358	15	17	17	15	17	2,4597	
D1S1656	13	14	14	14	14	3,6819	
D2S441	12	15	11	15	11	3,4650	
D10S1248	14	15	14	15	13	0,9709	
D13S317	11	12	11	11	11	3,3878	
Penta E	7	10	7	8	8	8,6226	
D16S539	9	10	9	12	12	1,9528	
D18S51	14	15	15	16	16	3,4605	
D2S1338	17	24	17	20	20	4,0032	
CSF1PO	10	10	10	10	10	3,8113	
Penta D	11	11	11	11	9	3,1888	
TH01	6	7	7	7	7	2,0794	
vWA	15	18	16	18	16	1,9229	
D21S11	29	30	29	30	29	30	2,2740
D7S820	9	10	10	11	11	13	2,1154
D5S818	12	14	11	14	11	11	3,0389
TPOX	8	8	8	8	8	8	2,1936
D8S1179	13	14	13	16	15	16	14,8677
D12S391	15	18	16	18	16	21	14,1243
D19S433	13	13	13	14	14	14	3,8168
FGA	20	20	20	23	22	23	3,3493
D22S1045	15	16	15	15	15	15	2,7739
Amelogenina	X	X	X	X	X	Y	1,0000
Índice Acumulado de Paternidade:							2,79E+11
Probabilidade de Paternidade:							99,99999999%

Multiplex desenvolvido							
<i>Locus STR</i>	MSF	SF	SP	IP			
D3S1358	15	17	17	17	15	17	2,4597
TH01	6	7	7	7	7	8	2,0794
D21S11	29	30	29	30	29	30	2,2740
D18S51	14	15	15	16	16	18	3,4605
Penta E	7	10	7	8	8	16	8,6226
D5S818	12	14	11	14	11	11	3,0389
D13S317	11	12	11	11	11	11	3,3878
D7S820	9	10	10	11	11	13	2,1154
D16S539	9	10	9	12	12	13	1,9528
CSF1PO	10	10	10	10	10	10	3,8113
Penta D	11	11	11	11	9	11	3,1888
vWA	15	18	16	18	16	17	1,9229
D8S1179	13	14	13	16	15	16	14,8677
TPOX	8	8	8	8	8	8	2,1936
FGA	20	20	20	23	22	23	3,3493
D2S441	12	15	11	15	11	11	3,4650
D17S1301	12	13	11	13	11	12	1,7110
D19S433	13	13	13	14	14	14	3,8168
D18S853	12	12	12	13	13	13	7,0313
D20S482	15	15	13	15	13	14	2,5568
D14S1434	9	9	9	9	9	12	2,4064
Amelogenina	X	X	X	X	X	Y	1,0000
Índice Acumulado de Paternidade:							3,69E+10
Probabilidade de Paternidade:							99,99999999%

Caso 8

PowerPlex® Fusion

<i>Locus STR</i>	MSF	SF	SP	IP			
D3S1358	14	17	14	15	15	16	1,7183
D1S1656	14	18,3	18,3	19,3	18,3	19,3	138,8889
D2S441	10	14	12	14	11	12	6,8399
D10S1248	13	15	13	13	13	14	1,8512
D13S317	8	8	8	11	8	11	1,6939
Penta E	12	13	12	13	7	13	1,7792
D16S539	11	11	11	12	10	12	1,9528
D18S51	18	18	12	18	12	20	4,1282
D2S1338	17	21	17	21	17	17	3,9293
CSF1PO	10	11	11	13	10	13	7,8510
Penta D	10	13	10	14	14	14	16,6867
TH01	9,3	9,3	9,3	9,3	9,3	9,3	4,4217
vWA	16	16	15	16	15	17	3,6233
D21S11	30	30	30	30	29	30	2,1582
D7S820	9	12	9	10	10	10	3,6359
D5S818	11	12	11	11	11	11	3,0389
TPOX	9	11	8	11	8	8	2,1936
D8S1179	15	15	11	15	11	14	6,4846
D12S391	15	18	18	21	20	21	5,9952
D19S433	13	14	13	14	12	13	0,9777
FGA	19	22	19	20	20	26	4,3895
D22S1045	11	16	11	16	15	16	1,1631
Amelogenina	X	X	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:							6,15E+12
Probabilidade de Paternidade:							99,999999999%

Multiplex desenvolvido							
<i>Locus STR</i>	MSF	SF	SP	IP			
D3S1358	14	17	14	15	15	16	1,7183
TH01	9,3	9,3	9,3	9,3	9,3	9,3	4,4217
D21S11	30	30	30	30	29	30	2,1582
D18S51	18	18	12	18	12	20	4,1282
Penta E	12	13	12	13	7	13	1,7792
D5S818	11	12	11	11	11	11	3,0389
D13S317	8	8	8	11	8	11	1,6939
D7S820	9	12	9	10	10	10	3,6359
D16S539	11	11	11	12	10	12	1,9528
CSF1PO	10	11	11	13	10	13	7,8510
Penta D	10	13	10	14	14	14	16,6867
vWA	16	16	15	16	15	17	3,6233
D8S1179	15	15	11	15	11	14	6,4846
TPOX	9	11	8	11	8	8	2,1936
FGA	19	22	19	20	20	26	4,3895
D2S441	10	14	12	14	11	12	6,8399
D17S1301	12	13	12	12	12	12	2,2222
D19S433	13	14	13	14	12	13	0,9777
D18S853	11	14	11	11	11	12	1,2162
D20S482	13	14	14	15	15	15	4,5918
D14S1434	9	13	13	13	13	13	2,7190
Amelogenina	X	X	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:							2,94E+10
Probabilidade de Paternidade:							99,99999999%

Caso 9

PowerPlex® Fusion

<i>Locus STR</i>	MSF		SF		SP		IP
D3S1358	16	18	15	16	15	16	1,7183
D1S1656	12	12	12	18,3	17,3	18,3	11,0865
D2S441	10	11	10	11	11	11	1,8192
D10S1248	13	14	13	16	13	16	4,3253
D13S317	12	14	12	14	12	13	1,4638
Penta E	10	14	10	15	5	15	8,0700
D16S539	10	11	11	11	11	11	3,4640
D18S51	12	17	12	16	16	19	3,4605
D2S1338	20	21	17	21	16	17	2,4355
CSF1PO	10	11	10	12	7	12	1,6532
Penta D	11	12	12	13	5	13	3,2395
TH01	7	9,3	7	9	7	9	3,0757
vWA	16	18	16	16	15	16	1,9229
D21S11	30	32	31,2	32	30	31,2	5,5626
D7S820	11	12	12	12	11	12	3,0981
D5S818	9	12	12	12	12	12	2,8870
TPOX	9	12	11	12	8	11	1,8147
D8S1179	13	13	13	13	13	13	3,5886
D12S391	18	22	18	22	19	22	1,9928
D19S433	14	15	14	15	14	14	2,4950
FGA	23	26	20	23	20	25	4,3895
D22S1045	15	15	15	15	11	15	1,3870
Amelogenina	X	X	X	X	X	Y	1,0000
Índice Acumulado de Paternidade:	1,61E+10						
Probabilidade de Paternidade:	99,9999999%						

Multiplex desenvolvido							
<i>Locus STR</i>	MSF		SF		SP		IP
D3S1358	16	18	15	16	15	16	1,7183
TH01	7	9,3	7	9	7	9	3,0757
D21S11	30	32	31,2	32	30	31,2	5,5626
D18S51	12	17	12	16	16	19	3,4605
Penta E	10	14	10	15	5	15	8,0700
D5S818	9	12	12	12	12	12	2,8870
D13S317	12	14	12	14	12	13	1,4638
D7S820	11	12	12	12	11	12	3,0981
D16S539	10	11	11	11	11	11	3,4640
CSF1PO	10	11	10	12	7	12	1,6532
Penta D	11	12	12	13	5	13	3,2395
vWA	16	18	16	16	15	16	1,9229
D8S1179	13	13	13	13	13	13	3,5886
TPOX	9	12	11	12	8	11	1,8147
FGA	23	26	20	23	20	25	4,3895
D2S441	10	11	10	11	11	11	1,8192
D17S1301	11	13	11	12	11	12	1,1111
D19S433	14	15	14	15	14	14	2,4950
D18S853	13	14	12	13	12	14	4,4118
D20S482	13	14	14	16	12	16	8,6538
D14S1434	9	12	9	13	12	13	1,3595
Amelogenina	X	X	X	X	X	Y	1,0000
Índice Acumulado de Paternidade:	2,87E+09						
Probabilidade de Paternidade:	99,9999999%						

Caso 10

PowerPlex® Fusion

<i>Locus STR</i>	MSF	SF	SP	IP			
D3S1358	15	17	17	16	17	2,4597	
D1S1656	12	13	12	16,1	16,1	16,3	3,7651
D2S441	12	14	13	14	11,3	13	14,6628
D10S1248	13	13	13	13	13	14	1,8512
D13S317	9	11	9	9	9	12	6,1173
Penta E	12	12	5	12	5	15	8,2478
D16S539	9	11	11	12	11	12	1,9528
D18S51	13	18	13	18	12	18	2,6868
D2S1338	16	22	16	21	17	21	10,1626
CSF1PO	11	11	10	11	10	10	3,8113
Penta D	13	13	9	13	9	9	5,4309
TH01	6	8	6	7	7	9	2,0794
vWA	11	17	16	17	16	17	1,9229
D21S11	28	30	28	28	28	30	3,1374
D7S820	10	11	9	10	9	12	4,1483
D5S818	12	12	9	12	9	11	13,1154
TPOX	9	12	8	9	8	8	2,1936
D8S1179	10	14	14	15	15	15	7,2290
D12S391	19,1	21	19	21	19	19	5,8858
D19S433	12	13	12	14	11	14	1,9084
FGA	20	25	20	25	21	25	2,3009
D22S1045	11	17	15	17	15	15	2,7739
Amelogenina	X	X	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:							1,09E+13
Probabilidade de Paternidade:							99,9999999999%

Multiplex desenvolvido							
<i>Locus STR</i>	MSF	SF	SP	IP			
D3S1358	15	17	17	17	16	17	2,4597
TH01	6	8	6	7	7	9	2,0794
D21S11	28	30	28	28	28	30	3,1374
D18S51	13	18	13	18	12	18	2,6868
Penta E	12	12	5	12	5	15	8,2478
D5S818	12	12	9	12	9	11	13,1154
D13S317	9	11	9	9	9	12	6,1173
D7S820	10	11	9	10	9	12	4,1483
D16S539	9	11	11	12	11	12	1,9528
CSF1PO	11	11	10	11	10	10	3,8113
Penta D	13	13	9	13	9	9	5,4309
vWA	11	17	16	17	16	17	1,9229
D8S1179	10	14	14	15	15	15	7,2290
TPOX	9	12	8	9	8	8	2,1936
FGA	20	25	20	25	21	25	2,3009
D2S441	12	14	13	14	11,3	13	14,6628
D17S1301	12	12	12	13	10	13	2,6316
D19S433	12	13	12	14	11	14	1,9084
D18S853	11	14	11	14	14	14	1,5491
D20S482	11	13	11	16	15	16	8,6538
D14S1434	13	13	9	13	9	13	2,4064
Amelogenina	X	X	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:							7,97E+11
Probabilidade de Paternidade:							99,9999999999%

Caso 11

PowerPlex® Fusion

<i>Locus STR</i>	SF		SP		IP Exclusões
D3S1358	15	16	17	17	Exclusão
D1S1656	11	16	16	16,3	1,8825
D2S441	10	11	12	14	Exclusão
D10S1248	13	16	12	16	2,1626
D13S317	11	12	11	12	1,6992
Penta E	8	16	8	11	4,3113
D16S539	10	12	10	13	2,8279
D18S51	15	21	15	17	1,6424
D2S1338	17	17	21	22	Exclusão
CSF1PO	9	11	12	14	Exclusão
Penta D	10	12	5	12	1,5110
TH01	8	9	8	8	3,3478
vWA	16	17	15	19	Exclusão
D21S11	28	29	30	31,2	Exclusão
D7S820	8	12	10	11	Exclusão
D5S818	11	12	11	12	1,4815
TPOX	8	11	9	11	0,9073
D8S1179	14	15	11	14	0,9965
D12S391	15	20	18	23	Exclusão
D19S433	12	15	12,2	14	Exclusão
FGA	24	25	20	25	2,4177
D22S1045	15	17	15	16	0,6935
Amelogenina	X	X	X	Y	1,0000
Quantidade incoerências alélicas				9	

Multiplex desenvolvido					
<i>Locus STR</i>	SF		SP		IP Exclusões
D3S1358	15	16	17	17	Exclusão
TH01	8	9	8	8	3,3478
D21S11	28	29	30	31	Exclusão
D18S51	15	21	15	17	1,6424
Penta E	8	16	8	11	4,3113
D5S818	11	12	11	12	1,4815
D13S317	11	12	11	12	1,6992
D7S820	8	12	10	11	Exclusão
D16S539	10	12	10	13	2,8279
CSF1PO	9	11	12	14	Exclusão
Penta D	10	12	5	12	1,5110
vWA	16	17	15	19	Exclusão
D8S1179	14	15	11	14	0,9965
TPOX	8	11	9	11	0,9073
FGA	24	25	20	25	2,4177
D2S441	10	11	12	14	Exclusão
D17S1301	11	12	11	13	0,8555
D19S433	12	15	12,2	14	Exclusão
D18S853	11	11	14	15	Exclusão
D20S482	13	14	14	14	1,0638
D14S1434	12	13	9	13	0,6798
Amelogenina	X	X	X	Y	1,0000
Quantidade incoerências alélicas				8	

Caso 12

PowerPlex® Fusion					
<i>Locus STR</i>	SF		SP		IP
D3S1358	15	16	15	16	1,7840
D1S1656	15	16,3	14	16,3	4,9505
D2S441	10	15	10	11	0,9575
D10S1248	12	13	13	14	0,9256
D13S317	11	12	11	12	1,6992
Penta E	8	11	11	13	2,2664
D16S539	9	10	9	10	4,3875
D18S51	16	18	16	18	4,9536
D2S1338	16	19	19	20	1,9920
CSF1PO	8	11	7	8	12,4782
Penta D	2,2	12	2,2	11	7,4629
TH01	6	8	6	9,3	1,1835
vWA	15	18	15	16	1,8117
D21S11	27	30,2	30,2	30,2	17,5840
D7S820	9	10	10	11	0,9090
D5S818	11	12	12	13	0,7218
TPOX	8	9	8	8	1,0968
D8S1179	14	15	14	15	2,8038
D12S391	18	21	21	23	2,9976
D19S433	12	13,2	12	13	2,4900
FGA	24	24	23	24	3,3128
D22S1045	15	15	15	15	2,7739
Amelogenina	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:	6,34E+08				
Probabilidade de Paternidade:	99,999999%				

Multiplex desenvolvido					
<i>Locus STR</i>	SF		SP		IP
D3S1358	15	16	15	16	1,7840
TH01	6	8	6	9,3	1,1835
D21S11	27	30,2	30,2	30,2	17,5840
D18S51	16	18	16	18	4,9536
Penta E	8	11	11	13	2,2664
D5S818	11	12	12	13	0,7218
D13S317	11	12	11	12	1,6992
D7S820	9	10	10	11	0,9090
D16S539	9	10	9	10	4,3875
CSF1PO	8	11	7	8	12,4782
Penta D	2,2	12	2,2	11	7,4629
vWA	15	18	15	16	1,8117
D8S1179	14	15	14	15	2,8038
TPOX	8	9	8	8	1,0968
FGA	24	24	23	24	3,3128
D2S441	10	15	10	11	0,9575
D17S1301	12	13	12	13	1,8713
D19S433	12	13,2	12	13	2,4900
D18S853	10	13	10	10	26,4706
D20S482	11	14	11	14	11,2462
D14S1434	9	13	9	9	2,4064
Amelogenina	X	Y	X	Y	1,0000
Índice Acumulado de Paternidade:	1,12E+10				
Probabilidade de Paternidade:	99,99999999%				

Caso 13

PowerPlex® Fusion					
--------------------------	--	--	--	--	--

<i>Locus STR</i>	SF		SP		IP
D3S1358	15	16	14	15	0,8591
D1S1656	15	17,3	15	17,3	3,3922
D2S441	11	11,3	11,3	14	5,4945
D10S1248	14	15	14	14	1,6404
D13S317	9	11	9	9	6,1173
Penta E	7	12	5	7	2,2855
D16S539	12	13	9	12	0,9764
D18S51	14	14	12	14	3,3962
D2S1338	19	24	18	24	3,0414
CSF1PO	10	12	10	10	1,9057
Penta D	9	13	12	13	1,6197
TH01	6	9,3	6	9,3	2,2889
vWA	15	19	18	19	3,8037
D21S11	28	29	28	30	1,5687
D7S820	8	12	8	12	3,1406
D5S818	12	12	11	12	1,4435
TPOX	8	8	8	8	2,1936
D8S1179	10	13	12	13	0,8971
D12S391	17	19	17	22	2,6042
D19S433	14	15,2	14	14	1,9084
FGA	21	21,2	21,2	22	186,0119
D22S1045	15	15	15	16	1,3870
Amelogenina	X	Y	X	Y	1,0000

Índice Acumulado de Paternidade: 6,34E+08
Probabilidade de Paternidade: 99,999999%

Multiplex desenvolvido

<i>Locus STR</i>	SF		SP		IP Exclusões
D3S1358	15	16	14	15	0,8591
TH01	6	9,3	6	9,3	2,2889
D21S11	28	29	28	30	1,5687
D18S51	14	14	12	14	3,3962

Penta E	7	12	5	7	2,2855
D5S818	12	12	11	12	1,4435
D13S317	9	11	9	9	6,1173
D7S820	8	12	8	12	3,1406
D16S539	12	13	9	12	0,9764
CSF1PO	10	12	10	10	1,9057
Penta D	9	13	12	13	1,6197
vWA	15	19	18	19	3,8037
D8S1179	10	13	12	13	0,8971
TPOX	8	8	8	8	2,1936
FGA	21	21,2	21,2	22	186,0119
D2S441	11	11,3	11,3	14	5,4945
D17S1301	13	14	12	14	6,6176
D19S433	14	15,2	14	14	1,9084
D18S853	12	13	12	13	3,9637
D20S482	13	14	13	15	1,2784
D14S1434	13	13	9	13	1,3595
Amelogenina	X	Y	X	Y	1,0000

Índice Acumulado de Paternidade: 1,33E+09
Probabilidade de Paternidade: 99,999999%

APÊNDICE IV

PRODUÇÃO CIENTÍFICA

Artigo Publicado 1 (Capítulo 4):

Development of a polymorphic short tandem repeat locus multiplex system for efficient human identification. Genetics and molecular research: GMR, 16(2).

Rodvalho RG, Rodrigues EL, Santos GS, Cavalcanti LM, Lima PR, Rodvalho AG, Vital RG, Gigonzac MAD, da Cruz, AD.

Artigo Publicado 2:

Mutation rates in 21 autosomal short tandem repeat loci in a population from Goiás, Brazil: ELECTROPHORESIS. doi: 10.1002 / elps.201700192.

Vieira TC, Gigonzac MAD, Rodvalho RG, Cavalcanti LM, Minasi LB, Rodrigues FM, da Cruz AD.

Electrophoresis (ISSN 0173-0835)

ELECTROPHORESIS

© WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Editor-in-Chief: Ziad El Rassi

Impact Factor: 2.482

ISI Journal Citation Reports © Ranking: 2015: 26/75

(Chemistry Analytical); 38/77

(BIOCHEMICAL RESEARCH METHODS)

Online ISSN: 1522-2683

Associated Title(s): [Journal of Separation Science](#)



Thaís Cidália Vieira^{1,2,3}
 Marc Alexandre Duarte
 Gígonzac^{1,2,3}
 Ricardo Goulart Rodovalho⁴
 Luana Morais Cavalcanti⁴
 Lysa Bernardes Minasi³
 Flávia Melo Rodrigues^{2,3}
 Aparecido Divino da Cruz^{1,3}

¹LaGene-Laboratory of Human Cytogenetics and Molecular Genetics, Secretary of State for Health of Goiás (LACEN/SESGO), Goiânia, GO, Brazil

²State University of Goiás (UEG), Goiânia, GO, Brazil

³Postgraduate Program in Genetics (MGene)/Replicon Research Center, Pontifical Catholic University of Goiás (PUC-GO), Goiânia, GO, Brazil

⁴Biocroma Laboratory, DNA Analysis Laboratory, Goiânia, GO, Brazil

Received May 7, 2017

Revised June 20, 2017

Accepted July 29, 2017

Short Communication

Mutation rates in 21 autosomal short tandem repeat loci in a population from Goiás, Brazil

The appearance of new mutations in polymorphic markers plays a central role in a range of genetic applications, including dating phylogenetic events, informing disease studies, and evaluating forensic evidence. The present study estimated the mutation rates of 21 autosomal STR loci in a population from Central Brazil. We studied 15 046 paternity cases from Goiás, Brazil from August 2012 to February 2015. We identified 262 mutations in the 21 loci. The loci that presented more mutations were FGA and D18S51, with a total of 46 and 28 mutations, respectively. The results showed mutational rates ranging from 1.7×10^{-5} to 7.6×10^{-4} mutations per site/region and the overall mutational rate was 2.1×10^{-4} ; these values were within the expected values for the STR markers. The most common type of mutation was one-step mutation, which totaled 96.2%. We found a higher rate of mutations of paternal origin (67.6%) than of mutations of maternal origin. The occurrence of mutations in STRs has important consequences for human identification, including the definition of criteria for exclusion in paternity testing and interpretation of genetic profiles in criminal cases.

Keywords:

Germline mutation / Paternity testing / Short tandem repeats

DOI 10.1002/elps.201700192

Microsatellites or STRs are types of genetic markers that are widely used in the study of human populations for a variety of different purposes, including relationship tests for routine parentage testing and forensic applications. They are highly polymorphic markers and can be analyzed accurately, even with low or degraded DNA [1–3].

Knowledge about mutation rates and the mutational process of STR or microsatellite loci used in paternity testing and forensic analysis is crucial for the correct interpretation of resulting genetic profiles [4].

Previous studies have estimated the rates and patterns of de novo STR mutations by using CE genotyping of specialized sets of markers, combined DNA index system markers, specific Y-chromosome STRs as well as STRs of the X chromosome [4, 5, 7].

Although several researchers have estimated the rates of mutations in STRs in various populations, in Central Brazil, these types of studies are important despite peculiar genetic characteristics in the population of this region [8, 9]. The central region of Brazil has a mixed population owing to the historical process of colonization that contributed to the

existence of high genetic diversity due to the contribution of different ethnic groups, such as Europeans, Africans, and Amerindians [8, 9, 11].

Microsatellite mutation rates can vary between different populations, as well as being influenced by external factors such as ionizing radiation. In the study by da Cruz et al. [12], with 12 STR markers in families from Goiânia, Central Brazil, in which at least one of the parents was accidentally exposed to Cesium-137, a mutational rate of 1.1×10^{-2} was verified. With determination of the basal mutation rate for this population, it is now possible to calculate the actual increase in mutation rate determined by the effect of radiation in that population.

In this study, germline mutations in 21 STR loci were studied (FGA-Fibrinogen Alpha Chain, D18S51, VwA, D5S818, D13S317, D21S11, D8S1179, CSF1PO, Penta D, D3S1358, Penta E, D16S539, D7S820, TH01, TPOX, D12S391, D14S1434, D17S1301, D18S853, D1GATA113, and D20S1082). We analyzed 631 932 allelic transfers involving both biological parents and child, derived from 15 046 cases of paternity testing, with a combined paternity index $\geq 99.9999\%$, between August 2012 and February 2015 in Goiás, Central Brazil.

Genomic DNA was purified from blood samples preserved on FTA cards (Whatman Bioscience, Cambridge, UK), following the manufacturer's instructions. STRs were amplified by PCR using PowerPlex® (Promega Corporation®, USA), AmpFISTR® Identifier™ and NGM® (Applied Biosystems, USA) according to the manufacturers' recommendations to confirm the existence of the mutation.

Correspondence: Dr. Thaís Cidália Vieira, LaGene-Laboratory of Human Cytogenetics and Molecular Genetics, Secretary of State for Health of Goiás (LACEN/SESGO), Parque Santa Cruz, Goiânia, GO, Brazil

Fax: +55 62 34961443

E-mail: thaiscidalia@gmail.com

Table 1. Paternal, maternal, and total number of mutations, mutation rate, and confidence interval for the 21 STR Loci

Loci	Chromosome location	Paternal				Maternal				Paternal Total or maternal			
		No. of Meiosis mutations	Mutation rate	95% Confidence intervals	No. of Meiosis mutations	Mutation rate	95% Confidence intervals	No. of Meiosis mutations	Mutation rate	95% Confidence intervals	No. of Meiosis mutations	Mutation rate	95% Confidence intervals
FGA	4q28	30 092	0.0011963	0.0008–0.0017	30 092	8	0.0002658	0.0001–0.0005	2	60 184	46	0.000764323	0.0006–0.0010
D18S51	18q21.3	30 092	0.0005982	0.0004–0.0009	30 092	8	0.0002658	0.0001–0.0005	2	60 184	28	0.000465240	0.0003–0.0007
VWA	12p12-pter	30 092	0.0004652	0.0003–0.0008	30 092	2	0.0000664	0.0000–0.0002	7	60 184	23	0.000382161	0.0002–0.0006
D8S818	5q21-31	30 092	0.0003323	0.0002–0.0006	30 092	1	0.0000332	0.0000–0.0002	7	60 184	18	0.000299083	0.0002–0.0005
D13S17	13q22-31	30 092	0.0004320	0.0002–0.0007	30 092	1	0.0000332	0.0000–0.0002	1	60 184	15	0.000249236	0.0001–0.0004
D21S11	21q11.2-q21	30 092	0.0003655	0.0002–0.0007	30 092	8	0.0002658	0.0001–0.0005	1	60 184	20	0.000332314	0.0002–0.0005
D8S1179	8q24.1-24.2	30 092	0.0004652	0.0003–0.0008	30 092	0	0.0000000	0.0000–0.0000	0	60 184	14	0.000232620	0.0001–0.0004
CSF1PO	5q33.3-34	30 092	0.0003655	0.0002–0.0008	30 092	2	0.0000664	0.0000–0.0002	3	60 184	16	0.000265851	0.0002–0.0004
Penta D	21q22.3	30 092	0.0001994	0.0001–0.0004	30 092	1	0.0000332	0.0000–0.0002	2	60 184	9	0.000149541	0.0001–0.0003
D6S1358	3p21.31	30 092	0.0002991	0.0001–0.0006	30 092	1	0.0000332	0.0000–0.0002	1	60 184	11	0.000182773	0.0001–0.0003
Penta E	15q26.2	30 092	0.0002991	0.0001–0.0006	30 092	3	0.0000996	0.0000–0.0003	0	60 184	12	0.000199389	0.0001–0.0003
D16S539	16q24-pter	30 092	0.0001994	0.0001–0.0004	30 092	6	0.0001993	0.0001–0.0004	0	60 184	12	0.000199389	0.0001–0.0003
D7S820	7q11.21-22	30 092	0.0001662	0.0001–0.0004	30 092	2	0.0000664	0.0000–0.0002	2	60 184	9	0.000149541	0.0001–0.0003
TH01	11p15.5	30 092	0.0000000	0.0000–0.0001	30 092	2	0.0000664	0.0000–0.0002	0	60 184	2	0.000033231	0.0000–0.0001
TPOX	2p23-pter	30 092	0.0000332	0.0000–0.0001	30 092	0	0.0000000	0.0000–0.0000	0	60 184	1	0.000016616	0.0000–0.0001
D12S381	17q25.1	30 092	0.0000332	0.0000–0.0001	30 092	0	0.0000000	0.0000–0.0000	0	60 184	1	0.000016616	0.0000–0.0001
D14S1434	20p13	30 092	0.0000665	0.0000–0.0002	30 092	5	0.0001661	0.0001–0.0004	2	60 184	9	0.000149541	0.0001–0.0003
D17S1301	14q32.13	30 092	0.0001994	0.0001–0.0004	30 092	0	0.0000000	0.0000–0.0000	2	60 184	8	0.000132926	0.0001–0.0003
D18S853	1p36.23	30 092	0.0000667	0.0000–0.0003	30 092	0	0.0000000	0.0000–0.0000	0	60 184	3	0.000046847	0.0000–0.0001
D16A113	2q35	30 092	0.0000000	0.0000–0.0001	30 092	0	0.0000000	0.0000–0.0000	1	60 184	1	0.000016616	0.0000–0.0001
D20S1082	18p11.31	30 092	0.0000665	0.0000–0.0002	30 092	1	0.0000332	0.0000–0.0002	1	60 184	4	0.000066462	0.0000–0.0002

Overall mutation rate: 0.000207300

However, the same kit was used within families in determining mutation rates. The PCR products were separated by CE using the ABI 3500[®] USA Genetic Analyzer (Life Technology, Applied Biosystems, USA) to obtain the allelic profiles that were then analyzed using the GeneMapper[®] version 2.1 (Applied Biosystems, USA). All procedures were carried out according to the laboratory internal control standards and international consensus guidelines and recommendations.

Mutations were considered and scored following the stepwise mutation model, assuming a one-step mutation from parent to child as the most probable mutational event. All mutations were confirmed by reagentyping of both parents and offspring to reduce the impact of genotyping errors. All confidence intervals for the mutation rates were in accordance to Levin [13].

We identified 262 de novo mutations in the 21 STR loci that were analyzed from 631 932 allelic transfers. With respect to the parental origin of mutations, paternal mutations (67.6%) were 3.5-fold more frequent than maternal mutations (19.5%). On the other hand, 12.9% of the mutations could not be assigned the parental origin. Generally, mutation rates are about five times higher in paternal lineage cells due to the speed of cell divisions that occur in the precursor cells of male gametes during spermatogenesis [1]. Our observation of higher paternal contribution to the burden of mutations in the offspring was also supported by Dupuy et al. [14], Gusmão et al. [15], Andrade et al. [16], Chula et al. [17], Mardini et al. [6], Weng et al. [18], and Bo et al. [19]. In the study by Brinkmann et al. [20] with autosomal STRs, an association between father's age and mutation rate was observed. However studies of Y-chromosomal STR loci, [14, 21] found no correlation between mutation rate and paternal age. Because of anonymity, the influence of paternal age could not be evaluated in our study.

The one-step mutation model was more frequently observed, corresponding to 96.2% in the Brazilian group which is consistent with the general notion that the majority of mutations comprise single step repeat gain or loss due to strand slippage during replication [22]. Moreover, we found four two-step mutations in our group. Furthermore, mutations with gains in the STR core were more frequent than losses as was also observed by Gusmão et al. [15]. The results presented here are similar to those reported for the Chinese [19], Brazilian [6], American from Texas [22], and Polish populations [23]. FGA was the most frequently mutated locus, with 46 mutations (17.6%) and a mutation rate of 7.6×10^{-4} , followed by D18S51, with 28 mutations and a mutation rate of 4.7×10^{-4} . On the other hand, the STR markers TPOX, D12S39, and D1GATA113 showed the lowest mutation rates, each with a frequency of 1.7×10^{-5} . The overall germline mutation rate for the Goiás population was 2.1×10^{-4} . All mutation rates were normally distributed, calculated, and tested at a confidence interval of 95% using JavaStat (statpages.org/confint). Germline mutation rates ranged from 1.7×10^{-5} to 7.6×10^{-4} . Our data were in accordance with Butler et al. [24], Lu et al. [25], and

Mardini et al. [6] for the combined DNA index system locus, which was within the expected values for the STR markers (Table 1).

The rates at which new mutations arise play a central role in a range of genetic applications, including evolutionary and population events, disease studies, and evaluating forensic evidence. Thus, reliable estimates of mutation rates may be used in the quantitative assessment of inclusion or exclusion of an alleged parent during parentage testing and of convict felons in forensic cases. Additionally, germline mutation rates are useful for studying genetic and ancestry dynamics. Therefore, the use of advanced tools and consistent understanding of the STR mutation pathways will ensure reliable results when STR markers are used for human identification both in civil and forensic scenarios [21, 26, 27].

This work was supported by LaGene-Secretary of Goiás State for Public Health and Replicon Research Center, Pontifical Catholic University of Goiás (PUC-GO).

The authors have declared no conflict of interest.

References

- [1] Weber, J. L., Wong, C., *Hum. Mol. Genet.* 1993, 1, 1123–1128.
- [2] Willems, T., Gymrek, M., Poznik, G. D., Tyler-Smith, C., Erlich, Y., *Am. J. Hum. Genet.* 2016, 98, 919–933.
- [3] Qian, X. Q., Yin, C. Y., Ji, Q., Li, K., Fan, H. T., Yu, Y. F., Bu, F. L., Hu, L. L., Wang, J. W., Mu, H. F., Haigh, S., Chen, F., *Electrophoresis* 2015, 36, 1633–1639.
- [4] Decker, A. E., Kline, M. C., Redman, J. W., Reid, T. M., Butler, J. M., *Forensic Sci. Int.* 2008, 2, e31–e35.
- [5] Goedbloed, M., Vermeulen, M., Fang, R. N., Lembring, M., Wollstein, A., Ballantyne, K., Lao, O., Brauer, S., Krüger, C., Roewer, L., Lessig, R., *Int. J. Legal Med.* 2009, 123, 471–482.
- [6] Mardini, A. C., Rodenbusch, R., Schumacher, S., Chula, F. G. L., Michelon, C. T., Gastalgo, A. Z., Maciel, L. P., Almeida, S. E. M., Silva, C. M. D., *Int. J. Legal Med.* 2013, 127, 45–47.
- [7] Diegoli, T. M., Linacre, A., Schanfield, M. S., Coble, M. D., *Int. J. Legal Med.* 2014, 128, 579–587.
- [8] Vieira, T. C., Silva, D. M., Gigonzac, M. A., Ferreira, V. L., Gonçalves, M. W., da Cruz, A. D., *Genet. Mol. Res.* 2013, 12, 23–27.
- [9] Resque, R., Gusmão, L., Geppert, M., Roewer, L., Palha, T., Alvarez, L., Ribeiro dos Santos, A., Santos, S., *PLoS One* 2016, 11, e0152573.
- [10] Oliveira, S. F., Godinho, N. M., Amorim, C. E., Gontijo, C. C., Falcão-Alencar, G., Barcelos, R. S., Luizon, M. R., Klautau-Guimarães, M. N., Simões, A. L., Mendes-Junior, C. T., *Forensic Sci. Int.* 2009, 2, 475–476.
- [11] Vieira, T. C., Gigonzac, M. A., Silva, D. M., Rodovalho, R. G., Santos, G. S., Da Cruz, A. D., *Genet. Mol. Res.* 2014, 13, 3404–3410.
- [12] da Cruz, A. D., da Silva, D. D., da Silva, C. C., Nelson, R. J., Ribeiro, L. M., Pedrosa, E. R., Jayme, J. C., Curado, M.

- P., *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* 2008, 652, 175–179.
- [13] Levin, J., *Elementary Statistics in Social Research*, Harper & Row, New York, 1978.
- [14] Dupuy, B. M., Stenersen, M., Egeland, T., Olaisen, B. *Hum. Mutat.* 2004, 23, 117–124.
- [15] Gusmão, L., Sánchez-Diz, P., Calafell, F., Martín, P., Alonso, C. A., Álvarez-Fernández, F., Alves, C., Borjas-Fajardo, L., Bozzo, W. R., Bravo, M. L., Builes, J. J., *Hum. Mutat.* 2005, 26, 520–528.
- [16] De Andrade, E. S., Gomes, A. V., Raposo, G., Mauricio-da-Silva, L., dos Santos, S. R., *Forensic Sci. Int. Genet.* 2009, 3, 141–143.
- [17] Chula, F.G., Rodenbusch, R., Schumacher, S., Grandi, T., Michelin, C. T., Gastalgo A. Z., Costi, C., Carvalho, B., Silva, C. M. D., *Forensic Sci. Genet.* 2009, 3, e35–e38.
- [18] Weng, W., Liu, H., Li, S., Ge, J., Wang, H., Liu, C., *Int. J. Legal Med.* 2013, 127, 369–372.
- [19] Bo, J., Qin, S., Haibo, L., Yingbi, L., Jin, W., Jing, Y., Yiping, H., Weibo, L., Lin, Z., *Forensic Sci. Int. Genet.* 2016, 23, 86–90.
- [20] Brinkmann, B., Klitschar, M., Neuhuber, F., Hühne, J., Rolf, B., *Am. J. Hum. Genet.* 1998, 30, 1408–1415.
- [21] Ge, J., Budowle, B., Aranda, X. G., Planz, J. V., Eisenberg, A. J., Chakraborty, R., *Forensic Sci. Int.* 2009, 3, 179–184.
- [22] Wojtas, M., Piniewska, D., Polan'ska, N., Stawowiak, A., Sanak, M., *Forensic Sci. Int. Genet.* 2013, 7, 389–391.
- [23] Butler, J.M., *J. Forensic Sci.* 2006, 51, 253–265.
- [24] Lu, D., Liu, Q., Wu, W., Zhao, H., *Int. J. Legal Med.* 2012, 126, 331–335.
- [25] Kayser, M., Roewer, L., Hedman, M., Henke, L., Henke, J., Brauer, S., Krüger, C., Krawczak, M., Nagy, M., Dobosz, T., Szibor, R., *Am. J. Hum. Genet.* 2000, 66, 1580–1588.
- [26] Liu, Q. L., Chen, Y. F., Huang, X. L., Liu, K. Y., Zhao, H., *Int. J. Legal Med.* 2016, 1441, 1–4.
- [27] Qin, S., Bo, J., Haibo, L., Yingbi, L., Jin, W., Jing, Y., Yiping, H., Weibo, L., Lin, Z., *Forensic Sci. Int. Genet.* 2016, 44, 10–13.

APÊNDICE V

JUSTIFICATIVA PELA AUSÊNCIA DE COMPARAÇÃO ENTRE BANCOS DE FREQUÊNCIAS ALÉLICAS

Os alelos de um *locus* microssatélite geralmente são nomeados de acordo com a estrutura do seu núcleo de repetição e com a quantidade de repetições em série dentro da sequência. No entanto, como a detecção destes alelos é executada considerando o tamanho do fragmento amplificado e a migração no processo de eletroforese, sua nomenclatura ocorre através de uma relação com seu peso molecular e não com a composição da sua sequência nucleotídica. Sendo assim, variações nucleotídicas adjacentes ao núcleo de repetição podem influenciar a determinação de um alelo, prejudicando a nomeação de acordo com o número de repetições.

As sequências de iniciadores presentes em um conjunto multiplex podem influenciar a amplificação de um alelo em um determinado *locus* e, assim, provocar discordância entre resultados obtidos a partir de painéis multiplex compostos por iniciadores diferentes para o mesmo *locus*.

Uma vez que as regiões flanqueadoras do STR podem esconder polimorfismos presentes em uma população, recomenda-se, para estudo comparativo de frequências alélicas entre diferentes populações, o profundo conhecimento acerca de possíveis variações na estrutura genômica dos *loci* a serem confrontados. Neste sentido, como os iniciadores referentes aos *loci* de um sistema multiplex geralmente não são idênticos, a comparação de frequência alélica entre populações pode não oferecer resultados seguros, desde que os fragmentos a serem amplificados sejam previamente sequenciados em esfera populacional.

Como a estratégia de sequenciamento não foi adotada neste projeto e não existem bancos consistentes de frequências alélicas referentes aos *loci* adicionais do multiplex

desenvolvido para outras populações brasileiras, a análise comparativa entre bancos populacionais não foi realizada.

Moretti, T. R., Moreno, L. I., Smerick, J. B., Pignone, M. L., Hizon, R., Buckleton, J. S., Bright, J. A., & Onorato, A. J. (2016). Population data on the expanded CODIS core STR loci for eleven populations of significance for forensic DNA analyses in the United States. *Forensic Science International: Genetics*, 25, 175-181.

Zhang, L., Wu, H., Wang, K., Fan, A., Niu, J., & Zhu, Z. (2016). Identification of a discordant genotype at the D2S1338 locus between Identifiler Plus® 16 kit and Power Plex® 21 system kit. *Journal of forensic and legal medicine*, 44, 174-177.

Crouse CA, Rogers S, Amiott E, et al. Analysis and interpretation of short tandem repeat microvariants and three-banded allele patterns using multiple allele detection systems. *J Forensic Sci.* 1999;44(1):87-94.

Butler JM. Genetics and genomics of core short tandem repeat loci used in human identity testing. *J Forensic Sci.* 2006; 51(2):253-65.

10. ANEXO

ANEXO I

Quadro com a descrição dos índices estatísticos normalmente aplicados em investigação de vínculo genético.

Índice	Descrição
Índice de Paternidade	Corresponde à razão de verossimilhança entre duas hipóteses mutuamente excludentes, sendo calculado para cada <i>locus</i> individualmente. A hipótese primária é baseada na presunção de que o indivíduo testado é o (a) pai/mãe biológico do suposto filho, enquanto a hipótese alternativa se baseia na vaticinação de que o indivíduo testado não é o (a) pai/mãe biológico do suposto filho, sendo esse, um indivíduo com genótipo desconhecido, não relacionado com o (a) pai/mãe testado e selecionado randomicamente na população.
Índice Acumulado de Paternidade	É o produto dos IPs parciais. Este índice mede a força da evidência genética global, indicando se é mais aceita a hipótese de que os indivíduos testados apresentam vínculo genético ou a hipótese vinculada à suposição de que alguém randomicamente escolhido na população apresenta o vínculo genético investigado.
Probabilidade de Vínculo Genético <i>a priori</i>	Pode caracterizada por uma probabilidade marginal, sobre uma variável, na ausência de evidência genética, sendo baseada em toda e qualquer informação que não advenha da análise genética. Como normalmente o laboratório não dispõe do dado não genético, é usual aplicar um valor nulo, ou parcialmente nulo à ambas as hipóteses testadas. Neste sentido, convencionalmente, no intuito de manter o princípio da imparcialidade, a maioria dos laboratórios utilizam o valor de 0,5 para a probabilidade <i>a priori</i> , o que corresponde à uma razão de verossimilhança igual a 1.
Probabilidade de Vínculo Genético <i>a posteriori</i>	É assentada pela informação genética analisada e baseada no Teorema de Bayes. Para cálculo desta probabilidade é considerado o valor obtido na estimativa do IPC.
Pode de Exclusão	Esta estimativa corresponde ao poder com o qual um indivíduo selecionado aleatoriamente na população e não testado, seja excluído do vínculo genético investigado. É calculado desconsiderando a frequência alélica do indivíduo investigado, restringindo-se à combinação genotípica do genitor “conhecido” e do suposto filho.
Probabilidade de Exclusão	É baseada no Teorema de Bayes e calculada considerando os resultados obtidos a partir da combinação dos PEs individuais.

Fonte: Garrido and Rodrigues, 2014; Primorac *et al.*, 2000.