

UNIVERSIDADE FEDERAL DO PARANÁ

JESSICA NATALIA LOYOLA



**VALIDAÇÃO INTERNA DO SISTEMA COMERCIAL PROMEGA POWER
PLEX Y23 DE AMPLIFICAÇÃO DE MARCADORES STRs DO
CROMOSSOMO Y**

CURITIBA, 2018

JESSICA NATALIA LOYOLA

**VALIDAÇÃO INTERNA DO SISTEMA COMERCIAL PROMEGA POWER
PLEX Y23 DE AMPLIFICAÇÃO DE MARCADORES STRs DO
CROMOSSOMO Y**

Monografia apresentada ao curso de Graduação em Ciências Biológicas, Setor de Ciências Biológicas, Universidade Federal do Paraná, como requisito parcial a obtenção do título de Bacharel em Biologia.

Orientadora: Prof. Dra. Danielle Malheiros

Co-orientador: Dr. Marcelo Malaghini

CURITIBA, 2018

AGRADECIMENTOS

À minha família que me prestou apoio durante a execução do trabalho e estiveram comigo quando precisei durante todo o curso de Biologia.

Aos meus amigos que estiveram presente em todos os momentos críticos e não me deixaram desistir de um sonho e me encorajaram a continuar em frente.

Ao Laboratório de Genética Molecular Forense, aos Peritos Criminais do DNA e Equipe Técnica, que me ensinaram tudo o que foi necessário para a execução do trabalho e para o crescimento profissional.

As minhas companheiras de bancada, estagiárias Aline Farias e Mary Guedes as quais me ensinaram bastante do que aprendi no Lab. de Genética Molecular Forense e me auxiliaram ainda além.

A minha orientadora, que me aceitou como aluna e me prestou todo o apoio necessário na escrita e apresentação, Danielle Malheiros.

Ao meu co - orientador e perito chefe do LGMF, Marcelo Malaghini, por toda a sua paciência e críticas necessárias durante o desenvolvimento do trabalho e acompanhamento nas etapas que foram fundamentais para o desenvolvimento do mesmo.

Ao Departamento de Genética, ao qual fui estagiária desde o início do curso e durante ao menos dois anos, recebi bolsa para aprender sobre o assunto ao qual me inspirou a ingressar no curso e sub área onde desenvolvi minha monografia.

Aos professores de toda a graduação, os quais me fizeram aprender e crescer a cada semestre e se hoje eu possuo outra maneira de ver o mundo, foi graças a contribuição de todos vocês.

A Pró Reitoria de Assuntos Estudantis (PRAE) a qual fui beneficiada desde o segundo ano de curso com o auxílio permanência e refeição, foi fundamental para a minha estadia na universidade.

A Universidade Federal do Paraná e todas as pessoas que fazem parte dessa incrível Universidade a qual eu tive o privilégio de fazer parte e aprender nessa jornada pelo curso de ciências biológicas, onde a cada aprendizado foi um degrau acima para a visão dos complexos processos que regem os seres vivos.

“Eu apelo para o futuro; eu apelo para uma época em que o ódio e a crueldade não mais controlarão os corações dos homens. É época onde poderemos aprender através da razão, do bom senso, do entendimento e da fé que cada vida vale a pena ser salva e que a compaixão é o maior atributo do homem.”

- Clarence Darrow (1857 – 1938)

RESUMO

Os laboratórios de genética forense das polícias científicas utilizam conjuntos comerciais para a geração de perfis genéticos, os quais são amplamente utilizados para identificação humana e aplicados na resolução de diferentes tipos de crimes. Para garantir a confiabilidade dos perfis genéticos, os laboratórios devem seguir padrões internacionais de qualidade, os quais preconizam a validação interna dos métodos e conjuntos comerciais utilizados. Sendo assim, o objetivo deste trabalho foi promover a validação interna de um método de amplificação do sistema comercial Promega Power Plex Y23 que recentemente começou a ser utilizado pelo Laboratório de Genética Molecular Forense (LGMF) de Curitiba. A análise de marcadores do cromossomo Y é complementar à amplificação de cromossomos autossômicos e auxilia na elucidação de crimes por meio da identificação do perfil masculino, bem como na determinação de sua linhagem parental. Os perfis são gerados pelo conjunto de genótipos de marcadores STR (*Short Tandem Repeats*) oriundos de análise de tamanhos de fragmentos em sequenciador de DNA. Em termos técnicos, a atribuição dos genótipos é feita a partir da interpretação dos picos de fluorescência gerados nos eletroferogramas. Os dados obtidos do sequenciador passaram por análise e foram excluídos todos os picos que ultrapassaram 300 RFUs (*Relative Fluorescence Unit*). Após esse tratamento, os dados filtrados foram submetidos a análises estatísticas onde por dois métodos os valores de AT (*Analytical Threshold*) foram determinados. No primeiro método padronizou-se um valor em comum para todos os marcadores e no segundo, um valor específico para cada marcador. Concluímos que para fins de confiabilidade do método na caracterização de um perfil, o segundo método demonstrou maior intervalo de confiabilidade.

ABSTRACT

The forensic genetics laboratories of the scientific police use commercial sets for the generation of genetic profiles, which are widely used for human identification and applied in solving different types of crimes. To ensure the reliability of genetic profiles, laboratories must follow international quality standards, which advocate the internal validation of the methods and commercial assemblies used. Therefore, the objective of this work was to promote the internal validation of an amplification method of the Promega Power Plex Y23 commercial system that has recently started to be used by the Laboratory of Forensic Molecular Genetics (LGMF) in Curitiba. The analysis of markers of the Y chromosome is complementary to the amplification of autosomal chromosomes and helps in the elucidation of crimes by means of the identification of the masculine profile, as well as in the determination of its parental lineage. The profiles are generated by the set of STR (Short Tandem Repeats) markers derived from fragment sizes analysis in DNA sequencer. In technical terms, the attribution of the genotypes is made from the interpretation of the fluorescence peaks generated in the electropherograms. Data from the sequencer were analyzed and all peaks exceeding 300 RFUs (Relative Fluorescence Unit) were excluded. After this treatment, the filtered data were submitted to statistical analysis where by two methods the AT (Analytical Threshold) values were determined. In the first method a common value was standardized for all markers and in the second, a specific value for each marker. We conclude that for the reliability of the method in the characterization of a profile, the second method demonstrated a greater reliability interval.

Sumário

1	INTRODUÇÃO	10
2	REVISÃO BIBLIOGRÁFICA	12
2.1	MARCADORES MOLECULARES NA IDENTIFICAÇÃO HUMANA	12
2.2	STRS DO CROMOSSOMO Y	14
2.3	SISTEMA COMERCIAL PROMEGA POWER PLEX Y23	17
2.4	VALIDAÇÃO INTERNA	19
2.5	THRESHOLDS	19
2.5.1	PICOS STUTTER	20
2.5.2	PERDA ALÉLICA	20
2.5.3	SINAL MUITO INTENSO (OFFSCALE)	21
3	OBJETIVOS	21
4	MATERIAIS E MÉTODOS	22
4.1	AMOSTRAS	22
a.	TESTE DE SENSIBILIDADE E ESTOCASTICIDADE	23
4.2.1	QUANTIFICAÇÃO	23
3.3	PCR E SEQUENCIAMENTO POR GENOTIPAGEM	24
3.3.	ANÁLISE ESTATÍSTICA	26
5	RESULTADOS E DISCUSSÃO	26
4.1	ANÁLISE DE PICOS NO PROGRAMA GeneMapperIDX 1.4	26
a.	MÉTODOS DA ANÁLISE ESTATÍSTICA	28
4.2.1	MÉTODO 1	28
4.2.2	MÉTODO 2	29
5	CONCLUSÕES	31
6	REFERÊNCIAS BIBLIOGRÁFICAS	31

1 INTRODUÇÃO

Diante da rotina dos laboratórios de genética forense, o processo de validação interna dos sistemas comerciais é importante para a consistência do método de análise de DNA (ácido desoxirribonucleico) na elucidação de perfis genéticos. Os dados gerados pelo laboratório necessitam de padronização e verificação do método de genotipagem, principalmente quando estes serão utilizados para alimentar a plataforma de bancos de dados tal como o CODIS (*Combined DNA Index System*) e o RIBPG (Rede Integrada de Bancos de Perfis Genéticos). Ainda, os laboratórios devem seguir a Direção de Controle da Garantia de Padrões do FBI (QAS), a qual preconiza validação interna dos métodos.

As técnicas utilizadas na obtenção de perfis genéticos vêm adquirindo cada vez mais relevância e espaço nas investigações criminais que auxiliam diversas áreas das ciências forenses como a biologia, medicina e a antropologia a obter respostas com bons índices de consistência em seus resultados. A análise de DNA é aplicada a casos de confronto genético para crimes ou identificações onde não é possível a elucidação por outros métodos.

O DNA humano está presente em células nucleadas onde estão presentes os cromossomos autossômicos e sexuais. Ao que diz respeito aos cromossomos, existem regiões variantes chamadas STRs (do inglês, *Short Tandem Repeats*) ou repetições curtas em tandem. Tais variações, em sua maioria, são polialélicas e polimórficas, ou seja, apresentam mais de dois alelos, nenhum dos quais com frequência superior a 99% na população em estudo. Assim, essas variações, em conjunto podem ser diferenciar indivíduos e se tornam marcadores moleculares ideais para técnicas de genotipagem de DNA para fins de identificação humana (MALAGHINI et al., 2006).

Os cromossomos sexuais também possuem variantes STRs, porém o cromossomo Y tem menor variabilidade, já apresenta baixa taxa de mutação. A consequência da baixa taxa mutacional é que na maioria das vezes o processo de discriminação de indivíduos aparentados não é possível. Ou seja, onde as variações pouco mudam, todos os homens geneticamente relacionados em uma família acabam compartilhando o mesmo haplótipo do cromossomo Y.

Porém, a detecção dos marcadores do Y auxilia a identificar células de origem masculina coletadas em uma cena de crime ou em casos de crimes sexuais onde é essencial determinar o perfil do criminoso em um perfil de mistura com o da vítima. Também é muito importante na determinação do grau de parentesco de uma linhagem paterna em indivíduos masculinos para casos de identificação de cadáveres em acidentes onde o reconhecimento visual da vítima não mais é possível. Assim o emprego do cromossomo Y em análises forenses é um auxílio ao que se obtém com cromossomos autossômicos na elucidação de um perfil e é essencial em crimes sexuais (PRINZ, M. et al. 1997, SWGDAM 2014).

Nos laboratórios de genética forense, é imprescindível que os sistemas comerciais disponíveis para genotipagem passem por uma validação interna. Esta consiste em acumular dados obtidos em testes de laboratório a fim de demonstrar a eficiência e funcionamento dos métodos e procedimentos estabelecidos. Também auxilia na definição de uma margem das condições para se obter resultados confiáveis em um laboratório e é recomendado pela empresa responsável pelo sistema comercial.

O método de análise empregado pelos sistemas comerciais é a análise de tamanho de fragmentos em sequenciador de DNA. Sendo assim, os resultados das genotipagens são visualizados através de picos de fluorescência em eletroferogramas. Os principais problemas associados a análise de fragmentos por eletroforese capilar e que precisam ser identificados e corrigidos são os picos resultantes de artefatos (*stutter*), perda alélica (*allele drop-out*) e sinal muito intenso (*offscale*). Se estes não forem devidamente corrigidos pode-se originar um perfil errado. Os chamados *thresholds* (limiares de detecção) permitem avaliar a qualidade dos dados. Por exemplo, as alturas dos picos nos eletroferogramas se correlacionam com a intensidade do sinal ou unidades fluorescentes relativas (RFUs). Existem dois tipos básicos de *thresholds* usados na análise de DNA forense - analítica e estocástica. Um *threshold* analítico é o menor valor de RFU no qual o DNA pode ser distinguido do ruído. O sinal acima pode ser um pico de DNA ou um artefato; sinal abaixo é geralmente desconsiderado. Para estabelecer limiares analíticos, se mede o ruído de fundo da química e do instrumento. O objetivo deste trabalho é avaliar

a sensibilidade do sistema comercial Y23 da Promega quanto a sua efetividade em amplificar amostras com baixas concentrações de DNA, visto que em amostras forenses a quantidade é sempre limitada. A quantidade mínima de material amplificado necessita ser padronizada em virtude de resultados positivos e que estejam de acordo com os outros métodos de etapas prévias do laboratório de Genética Molecular Forense (LGMF). Este será testado em comparação ao sistema já validado Yfiler da Life Technologies a fim de se avaliar os resultados de picos verdadeiros de STR na elucidação de um perfil obtido e diferenciar estes picos alélicos de possíveis ruídos e artefatos de PCR (do inglês, *Polymerase Chain Reaction*) captados pelo sequenciador na genotipagem e DNA através da determinação do *threshold* analítico.

2 REVISÃO BIBLIOGRÁFICA

2.1 MARCADORES MOLECULARES NA IDENTIFICAÇÃO HUMANA

Os marcadores moleculares são amplamente empregados no ramo da genética e para a genética forense esses marcadores possuem aplicabilidade na determinação do perfil genético do indivíduo, os quais são baseados nos marcadores denominados STRs. Os STRs são sequências repetitivas não conservadas de tamanho curto no DNA, de 2 a 7 nucleotídeos e se repetem ao longo do genoma. São variáveis em uma população e entre indivíduos de uma mesma população com elevado grau de polimorfismo e polialelismo. Estes são amplificados pela técnica de PCR, o que torna essas variações efetivas para identificações humanas (BUTLER, 2012)

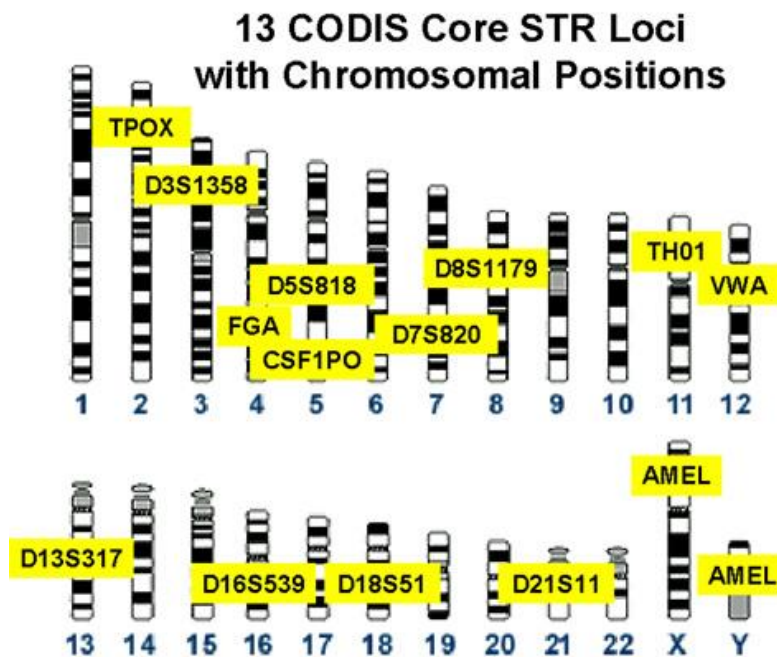
A escolha de um marcador está relacionada com a frequência daquela variante na população e a sua relevância que é medida pela taxa de mutação, ou seja, o quanto aquela variação ainda pode variar em um determinado espaço de tempo, se há elementos evolucionistas como a seleção natural ou a deriva genética atuando sobre aquele gene, delimitando um período de tempo

em que aquele loco pode ser utilizado. Há estudos de escalas temporais de marcadores que estimam o tempo em anos como eventos recentes que variam de 50.000 até 200.00 mil anos (RIBEIRO, 2009).

Os perfis genéticos amplamente utilizados na genética forense são compostos pelos genótipos de no mínimo 13 marcadores, estabelecidos inicialmente pelo Departamento Federal de Investigação dos estados Unidos (FBI, do inglês, *Federal Bureau of Investigation*). São estes o TPOX, D3S1358, D5S818, FGA, CSF1PO, D7S820, D8S1179, TH01, vWA, D13S317, D16S539, D18S51 e D21S1 e a amelogenina distribuídos pelo genoma (Figura 1). Com exceção da amelogenina, todos os demais STRs estão em cromossomos autossômicos. Os perfis são obtidos a partir da amplificação dos marcadores de interesse através da reação em cadeia da polimerase (PCR, do inglês *Polymerase Chain Reaction*), que permite a análise de DNA a partir de quantidades pequenas de amostras biológicas, como as provenientes de cabelo, manchas de sangue e outras que podem ser coletadas na cena de um crime. Em seguida à obtenção e amplificação do DNA, é feita a análise de tamanhos dos fragmentos amplificados. Uma vez que tais regiões são hipervariáveis é possível se obter um perfil característico e único de cada indivíduo.

Em laboratórios forenses do mundo inteiro são usados os mesmos marcadores STR para construção dos perfis com o propósito de que estes possam ser incluídos na plataforma mundial, cedida pelo FBI, o Sistema Combinado de Índices de DNA (CODIS, do inglês, *Combined DNA Index System*). O FBI padronizou e selecionou 13 marcadores STRs no início, porém desde janeiro de 2017, foram incluídos mais 7 marcadores, D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433 e D22S1045 sendo hoje portanto, 20 marcadores no total. No Brasil, existe o banco de dados nacional, o RIBPG (Rede Integrada de Bancos de Perfis Genéticos) que utiliza os mesmos STRs para a inclusão dos perfis. Esta padronização permite que as informações relativas aos perfis genéticos sejam trocadas em âmbito mundial.

FIGURA 1. Treze locos STRs que inicialmente compunham o CODIS juntamente com suas posições cromossômicas.



Fonte: Butler, 2006.

2.2 STRS DO CROMOSSOMO Y

Os STRs do cromossomo Y (Y-STRs) são os marcadores mais comumente usados no campo da genética forense devido a sua praticidade quando ao nível de diversidade e especificidade. Isso remete a uma diversidade de marcadores como opção para o emprego da PCR em casos de amostras comprometidas quantitativamente ou qualitativamente.

De todos os Y-STRs utilizados em genética forense ou em genética de populações, todos são referentes à região não recombinante do cromossomo Y. A região recombinante do cromossomo, regiões PAR, do inglês *pseudoautosomal regions* são regiões pseudo autossômicas homólogas ao cromossomo X da espécie humana e não possuem relevância para fins forenses, uma vez que a mesma sofre recombinações com o cromossomo X, a identidade de um perfil masculino estará comprometida quanto a sua origem (Snustad, 2013).

Sendo assim, as regiões utilizadas como marcadores STR são exclusivamente as regiões fora do espectro de recombinação com o cromossomo X, é a chamada região MSY a qual não sofre recombinação e é exclusiva do cromossomo Y, contém genes como o SRY, fundamental para a

diferenciação sexual e outros genes relevantes para a fertilidade da espécie humana e onde estão localizadas as sequências de STRs relevantes à genética forense (Snustad, 2013).

Alguns possuem mais dados conhecidos sobre suas particularidades, sendo eles DYS19, DYS385, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 e YCAII que foram os pioneiros no campo das análises forenses, bem como populacionais. Mais tarde foram adicionados mais sequencias STR ao campo, como o DYS434, DYS435, DYS436, DYS437, DYS438, DYS439, DYS460 (figura 2) entre outros que ganharam popularidade em estudos forenses e também, populacionais devido a praticidade em amplificar esses alelos com iniciadores específicos (Gusmão et. al., 2003).

O DNA do cromossomo Y apresenta uma herança paterna, na forma haplóide, de forma a não apresentar uma cópia homóloga como acontece com o restante dos cromossomos. A falta de recombinação na maior parte do cromossomo Y faz com que ele se comporte como um fragmento único de DNA, que se transmite intacto por via paterna à descendência masculina. A única fonte de variação nessa região são as mutações que, com o tempo, se acumulam gradativamente. Ao conjunto da informação relativa a vários marcadores STRs em um mesmo cromossomo dá-se o nome de haplótipo. Um haplótipo de Y-STR pode, portanto ser compartilhado por um indivíduo com seus irmãos, tios, primos, avô e demais antepassados masculinos. Portanto, estes últimos, não individualizam uma pessoa, mas individualizam um grupo de pessoas do sexo masculino em uma linhagem parental (KAYSER, 2007).

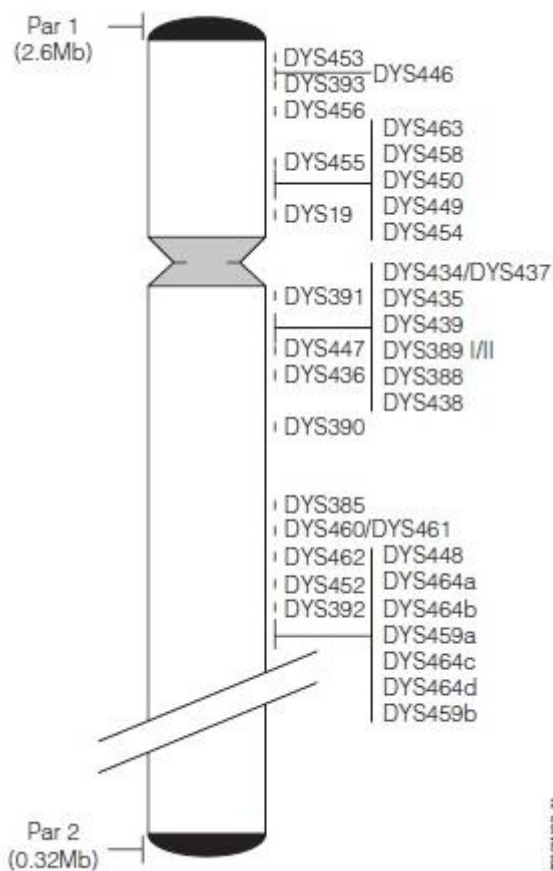
Desde a descoberta dos primeiros polimorfismos no cromossomo Y, o número de marcadores vem aumentando significativamente. A região não recombinante do cromossomo Y possui diferentes tipos de polimorfismos que apresentam taxas de mutação diferentes, permitindo assim selecionar o tipo de marcador mais adequado para os diferentes tipos de estudo e abordagens (KAYSER, 2007).

O baixo índice de taxa de mutação do cromossomo Y, faz dessa região não recombinante pouco diferente e sem grandes alterações em uma mesma linhagem de paternidade, o que faz desses marcadores úteis para estudos de

paternidade cível ou criminal, estudos de parentesco e na identificação de pessoas desaparecidas devido a sua transmissão dentro de uma linhagem paterna (PRINZ, M. et al. 1997, 2001, SWGDAM 2014).

Além das aplicações mencionadas acima, outra de suma importância em termos de utilização dos Y-STRs acontece em casos de violência sexual. As amostras de mistura provenientes de casos de agressões sexuais são, predominantemente, constituídas por quantidades ínfimas de DNA masculino mascaradas em grandes quantidades de DNA feminino, o que dificulta tecnicamente a análise relativa ao contribuinte minoritário. Sendo assim, nestes casos muitas vezes não é possível obter um perfil genético de STRs autossômicos que permita a identificação única do agressor. Nestas situações, é possível analisar apenas os Y - STRs, o que aliado a outras provas periciais podem auxiliar na identificação do agressor. (GUSMÃO et al, 2006, 2003).

FIGURA 2. STRs totais presentes no cromossomo Y da espécie humana identificados até o momento.



2.3 SISTEMA COMERCIAL PROMEGA POWER PLEX Y23

Os sistemas comerciais de amplificação de marcadores de regiões STRs para o cromossomo Y são comercializados pela Promega, uma das maiores empresas fornecedoras de sistemas comerciais para fins forenses de amplificação em marcadores STRs.

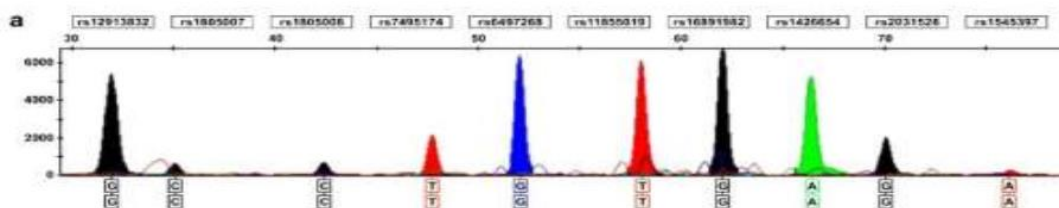
Diversos sistemas comerciais já foram validados pela Promega e anteriormente ao sistema Power Plex, no laboratório de Genética Molecular Forense do Instituto de Criminalística do Paraná, era utilizado o sistema comercial Yfiler (Life Technologies) e este foi substituído pelo sistema Power Plex, devido ao aumento de sensibilidade do sistema em relação ao Y filer quanto à amplificação eficaz de DNA masculino em relação a abundância de material feminino presente na amostra (PROMEGA CORPORATION, 2017).

O sistema comercial Power Plex Y23 permite à amplificação de 23 locos de STRs do cromossomo Y do genoma humano, os quais foram selecionados como locos principais aceitos no sistema do CODIS, sendo eles DYS576, DYS389I, DYS448, DYS389II, DYS19, DYS391, DYS481, DYS549, DYS533, DYS438, DYS411, DYS570, DYS635, DYS438, DYS437, DYS570, DYS635, DYS390, DYS439, DYS392, DYS643, DYS393, DYS458, DYFS385a / b, DYS456 e Y-GATA-H4.

A genotipagem se dá por análise de fragmentos através de eletroforese capilar. Essa técnica baseia-se na determinação do tamanho de fragmentos amplificados por PCR, utilizando iniciadores (*primers*) marcados com fluorescência, por comparação com um marcador de tamanho conhecido (*size standard*). O uso de diferentes fluorescências permite a análise de vários fragmentos (locos) ao mesmo tempo. Após a PCR, os fragmentos são submetidos a uma eletroforese capilar, nos sequenciadores automáticos de DNA. Diferentes fluorocromos são empregados e uma vez excitados por um feixe de laser emitem luz em diferentes comprimentos de onda. Os produtos da reação de sequenciamento, marcados com os fluorocromos, ao serem

submetidos a eletroforese passam pelo feixe de laser, que promove a excitação dos fluorocromos. A luz emitida pelos fluorocromos é detectada por um fotomultiplicador e a informação é processada através de um software específico que traduz a informação em uma sequência de picos mostrados na forma de eletroferograma (Figura 3).

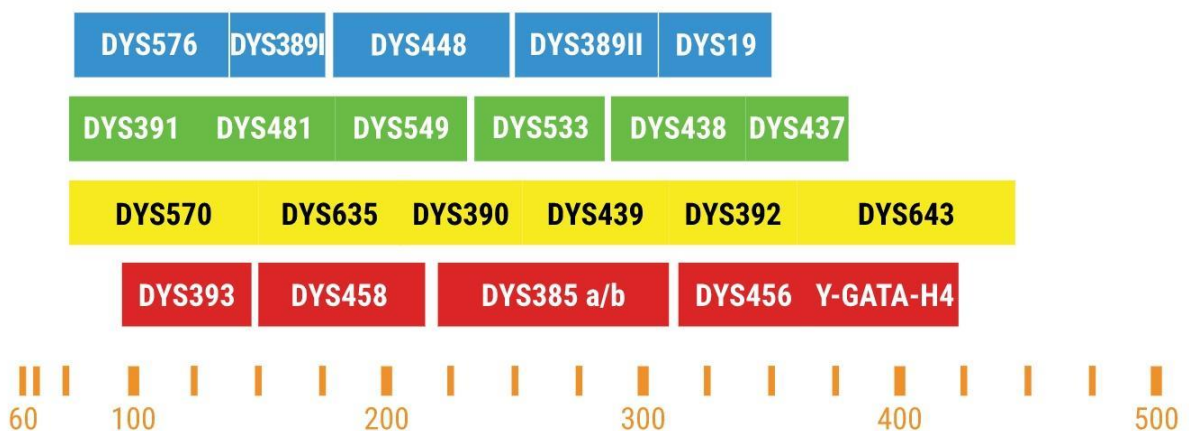
Figura 3. Exemplo de eletroferograma gerado por análise de tamanho de fragmentos em sequenciador automático de DNA.



Fonte: O autor

No Power Plex Y23 os iniciadores utilizados na PCR são marcados com quatro fluoróforos diferentes. Produtos de amplificação de tamanhos similares ficarão marcados com o mesmo fluoróforo, enquanto que o de tamanhos diferentes ficam com o mesmo (Figura 4) (Promega, 2017).

Figura 4. Fluoróforos presentes nos locos correspondentes com o tamanho do fragmento para 23 locos.



Fonte: Promega, 2017.

2.4 VALIDAÇÃO INTERNA

A validação interna consiste em acumular dados obtidos em testes de laboratório a fim de demonstrar a eficiência e funcionamento dos métodos e procedimentos estabelecidos. Também auxilia na definição de uma margem das condições para se obter resultados confiáveis em um laboratório. Diferentemente da validação de desenvolvimento na qual se aplicam testes no funcionamento direto de equipamentos e reagentes, a validação interna se fundamenta na garantia da qualidade dos dados obtidos através do equipamento ou do reagente em questão. Conforme os dados forem documentados e resumidos, o responsável deve aprovar os estudos internos de validação (Butler 2015, FBI 2011, QAS 2011).

2.5 THRESHOLDS

Os *thresholds* são limiares de uma curva gaussiana ou curva padrão que nos possibilitam inferir a veracidade de picos STRs obtidos na genotipagem de DNA, diferenciando de ruídos ou artefatos presentes na amostra em análise e necessitam ser determinados previamente por cálculos estatísticos (BUTLER, 2015). Esses picos que serão analisados são gerados por fluorescência presente nas amostras amplificadas por PCR e detectadas pelo sequenciador em unidade relativa de fluorescência (RFU) do inglês *relative fluorescence units*. Existem vários tipos de limiares ou *thresholds*, porém o que vamos abordar no presente trabalho será o analítico.

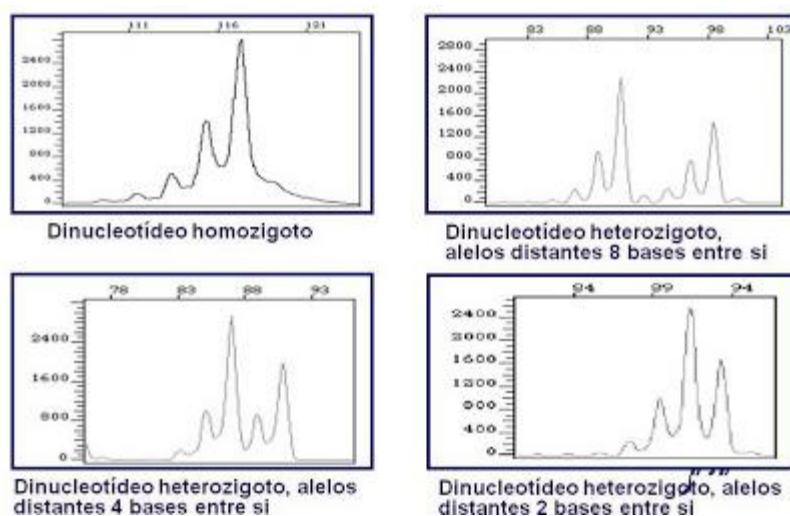
A quantidade de RFU captada pelo aparelho na determinação de um pico real varia de aparelho para aparelho, porém nesse trabalho todos os dados para a análise serão gerados a partir da genotipagem de um único aparelho sequenciador, de modo a padronizar RFUs com a metodologia empregada.

Os principais problemas associados a análise de fragmentos por eletroforese capilar e que precisam ser identificados, corrigidos e serão alvo da validação são os picos resultantes de artefatos de PCR (stutter), perda alélica (*allele drop-out*) e sinal muito intenso (*offscale*).

2.5.1 PICOS STUTTER

Os picos *stutter* são artefatos resultantes da amplificação dos STRs, caracterizados pela presença de uma unidade de repetição mais curta em relação ao alelo principal (frequência de 3-12%) (Figura 5). Para um dado loco STR, a probabilidade de stutter geralmente aumenta com o tamanho dos alelos (maior número de unidades de repetição) e com o tamanho da unidade de repetição (em pares de bases). A quantificação dos picos pode facilitar a interpretação dos resultados quando ocorre sobreposição dos picos com *stutters*.

Figura 5. Perfis de picos stutter.



Fonte: O autor

2.5.2 PERDA ALÉLICA

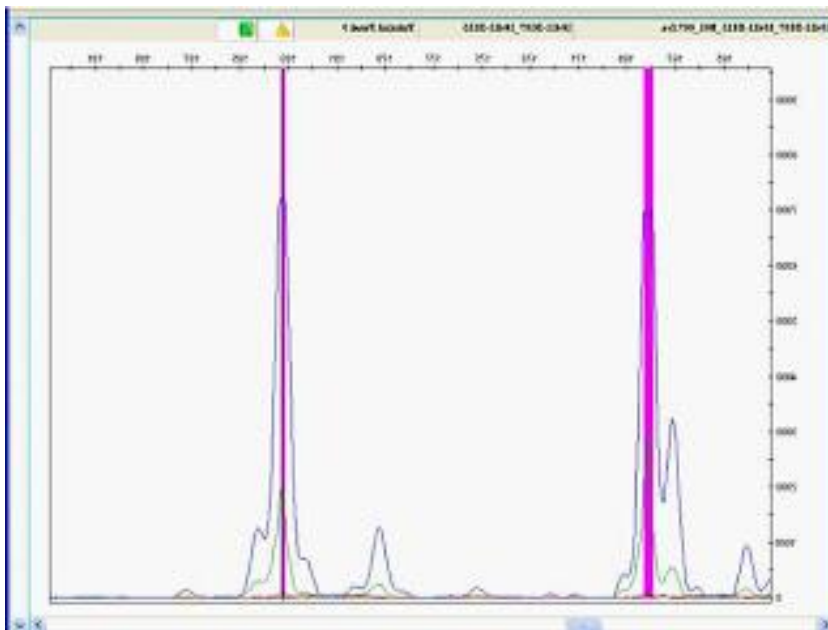
A perda alélica (*drop-out*) refere-se a falha de amplificação de locos/alelos específicos. A possibilidade de ocorrer *drop-out* aumenta quando se utilizam-se pequenas quantidades de DNA, quando existe algum contaminante inibidor na amostra (ex: hematina, ácido húmico, fenol), ou quando as amostras estão degradadas ou comprometidas. Esta também

costuma ser mais frequente quando há duas cópias para o mesmo gene, isto é, no lote diploide cromossômico, sendo o Y composto de apenas um lote, a perda alélica e seus riscos aqui são menores quando comparados aos cromossomos autossômicos.

2.5.3 SINAL MUITO INTENSO (OFFSCALE)

A ocorrência de um sinal de fluorescência muito intenso ou do padrão de peso molecular (size standard) pode derrubar os valores de qualidade dos genótipos atribuídos às amostras. Picos com sinais muito intensos são marcados pelo software GeneMapper com uma faixa lilás (Figura 6).

Figura 6. Exemplo de amostra com sinal muito intenso, ultrapassando o limiar de detecção.



Fonte: O autor

3 OBJETIVOS

De acordo com o exposto, percebe-se que a determinação de perfis genéticos confiáveis é de suma importância nas análises forenses, uma vez que estes são utilizados para incriminar ou inocentar indivíduos e

também são inclusos em bancos de dados de perfis genéticos. Sendo assim, o objetivos deste trabalho são:

- Analisar e determinar o *threshold* analítico para amplificação de locos específicos do cromossomo Y.
- Contribuir para a qualidade dos resultados gerados pelo Laboratório de Genética Molecular Forense do Instituto de Criminalística do Paraná.

4 MATERIAIS E MÉTODOS

4.1 AMOSTRAS

Foram utilizados como amostras o controle com perfil masculino do sistema comercial da Promega Y23, o controle 2800M com concentração de 10ng/μL, lote 0000209890, o controle com perfil masculino Power Quant System Male (PQSM) com concentração de 50ng/μL, lote 0000150482 e o controle interno masculino do laboratório (LGMF) concentrado a aproximadamente 3ng/μL e por fim o controle negativo com água ultrapura autoclavada.

As amplificações foram feitas em triplicatas juntamente com o controle negativo, que acompanhou todas as etapas do processo auxiliando na detecção de casos de contaminação assim como também em casos de alterações indesejadas nas amostras com DNA.

Para tanto, foram realizadas diluições das amostras em água ultrapura na concentração de 1, 0.5, 0.25, 0.125, 0.0625, 0.03125 e 0.01563ng/uL de DNA de cada amostra seguindo-se a amplificação com o sistema comercial Power Plex Y23. Para tanto foi necessária a realização da quantificação de DNA nas amostras para então ser realizado o teste de sensibilidade e estocasticidade.

a. TESTE DE SENSIBILIDADE E ESTOCASTICIDADE

Os estudos de sensibilidade são necessários durante a validação interna nos laboratórios para determinar a quantidade mínima de material genético ao qual o teste é sensível além da faixa recomendada pelo fabricante.

A empresa Promega recomenda o procedimento de validação interna nos laboratórios para delimitar um padrão mínimo de amostra a ser usado no sistema comercial, além da faixa já testada nos procedimentos de validação do sistema pela empresa. Para os procedimentos de cada laboratório é recomendado determinar através do teste de sensibilidade, acompanhado aos métodos de quantificação do laboratório que sejam aplicáveis no restante das etapas seguintes.

Os picos STRs obtidos pelo sequenciamento, são influenciados diretamente pela quantidade de DNA dos ensaios a fim de se obter os resultados desejados e filtrar resultados de natureza aleatória como picos não verdadeiros comumente associados a amostras de baixa qualidade devido a baixa quantidade de material genético na amostra (SWGAM 2014, NIST, 2015).

4.2.1 QUANTIFICAÇÃO

As amostras que foram diluídas para a realização do teste de sensibilidade foram quantificadas, pois o sistema comercial necessita de uma quantidade ideal de DNA para amplificar com precisão a fim de se obter resultados satisfatórios, reduzindo a probabilidade da ocorrência de artefatos de PCR ou sinais intensos como *off scale*, ou perda de alelos como também em casos de amostras criminais, excluir eventuais possibilidades de se obter na amostra, DNA não humano. (Butler, 2012)

A quantificação foi realizada através de PCR em tempo real no equipamento 7500 da Applied Biosystems, na qual se detecta intensidade de fluorescência à medida que os produtos são amplificados. As amostras são quantificadas por comparação com uma curva padrão onde as concentrações já estão determinadas pelo programa responsável pela leitura da fluorescência

e esta é usada como controle para as absorvâncias geradas pelas amostras. O sistema de amplificação por fluoróforos é específico para DNA humano, portanto toda a quantidade de DNA quantificado será de total relevância para a PCR posterior e desprezado a quantidade de DNA não humano caso houver, na amostra.

Os resultados das quantificações são mostrados na Tabela 1.

Tabela 1. QUANTIFICAÇÃO DOS CONTROLES

CONTROLES PERFIL Y23	CONCENTRAÇÃO (ng/ µL)
2800	6,32
CONTROLE LGMF	2,64
PQSM	44,15
BRANCO	0

3.3 PCR E SEQUENCIAMENTO POR GENOTIPAGEM

O DNA quantificado foi submetido à amplificação por PCR utilizando-se o sistema comercial Power Plex Y23 (Promega).

As amplificações foram feitas seguindo-se as recomendações do fabricante nas diluições determinadas e em triplicatas como ilustrado nas tabelas 2, 3 e 4. O produto amplificado foi submetido à análise de fragmentos no sequenciador 3500 Genetic Analyzer Applied Biosystems, juntamente com a escala alélica (*allelic ladder*), o tamanho padrão ILS600 (*size standard*) e formamida Hi-Di e posteriormente submetido à análise pelo software GeneMapper IDX.

TABELA 2. AMOSTRAS CONTROLE 2800

	1	2	3
A	1	CONTROLE POSITIVO	1
B	0,5	0,01563	0,5
C	0,25	0,03125	0,25
D	0,125	0,0625	0,125
E	0,0625	0,125	0,0625
F	0,03125	0,25	0,03125
G	0,01563	0,5	0,01563
H	LADDER	1	CONTROLE NEGATIVO

TABELA 3. AMOSTRAS CONTROLE LGMF

	1	2	3
A	CONTROLE NEGATIVO	1	LADDER
B	0,01563	0,5	0,01563
C	0,03125	0,25	0,03125
D	0,0625	0,125	0,0625
E	0,125	0,0625	0,125
F	0,25	0,03125	0,25
G	0,5	0,01563	0,5
H	1	CONTROLE POSITIVO	1

TABELA 4. AMOSTRAS CONTROLE PQSM

	1	2	3
A	1	LADDER	1
B	0,5	0,01563	0,5
C	0,25	0,03125	0,25
D	0,125	0,0625	0,125

E	0,0625	0,125	0,0625
F	0,03125	0,25	0,03125
G	0,01563	0,5	0,01563
H	CONTROLE NEGATIVO	1	CONTROLE POSITIVO

Após as amostras serem submetidas ao sequenciamento por genotipagem de acordo com as Tabelas 2, 3 e 4 obtemos os perfis gerados de acordo com os controles amplificados nas regiões de interesse do cromossomo Y, onde as amostras biológicas são convertidas a dados de fluorescência em comprimentos de onda específicos aos quais são essencialmente separados por cores de acordo com o tamanho em pares de bases (pb) onde serão interpretados através do programa GeneMapperIDX 1.4.

3.3. ANÁLISE ESTATÍSTICA

A análise estatística se consistiu em cálculos de média e desvio padrão e resultou em um intervalo de confiança de unidades de fluorescência obtida para a caracterização do *threshold* analítico, ou seja, na veracidade daquele pico obtido, se ele é realmente um alelo ou se são ruídos ou artefatos de PCR.

5 RESULTADOS E DISCUSSÃO

4.1 ANÁLISE DE PICOS NO PROGRAMA GeneMapperIDX 1.4

Como já foi explicado previamente, na análise de tamanho de fragmentos os alelos são estabelecidos através dos picos de floorencência gerados nos eletroferogramas. Na figura 7, há picos verdadeiros e não verdadeiros.

A base do threshold no programa GeneMapperIDX 1.4 foi delimitada para 1 RFU sendo essa a linha de base ou ruído de fundo, de forma a todos os picos aparecem no painel dos locos, assim todos os picos ficam evidenciados. Foi delimitada uma faixa que servirá como faixa de espectro do threshold analítico (AT, do inglês, analythical threshold) que será determinado. Usamos essa faixa

para fins de exclusão de picos, em torno de 300 RFU (Figura 8), assim tudo que estiver acima dessa marca, será excluído e o que estiver abaixo permanecerá para as análises. A escolha do valor de referência é devido a padrões internos laboratoriais onde denomina o que estiver acima de 300 RFUs seguramente será um pico alélico, enquanto valores abaixo não há a segurança de haver um pico correspondendo à um alelo ou ao ruído do aparelho.

Figura 7. Flouróforo azul com os picos totais.

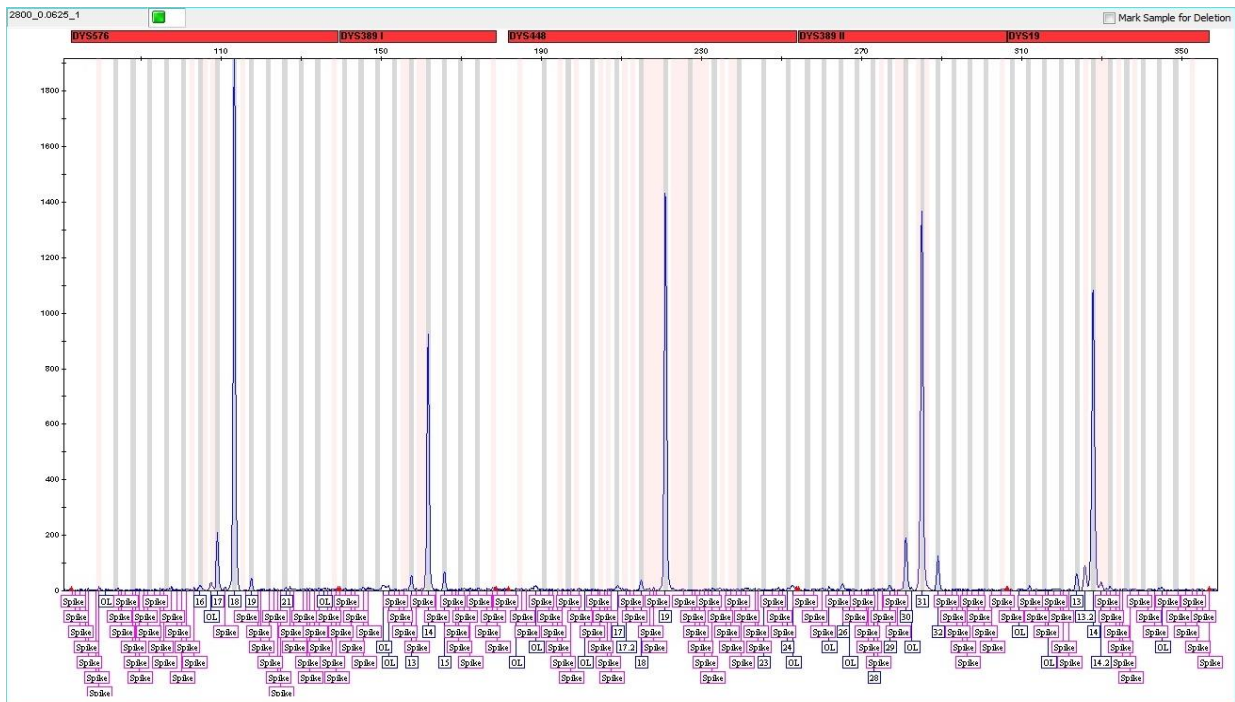
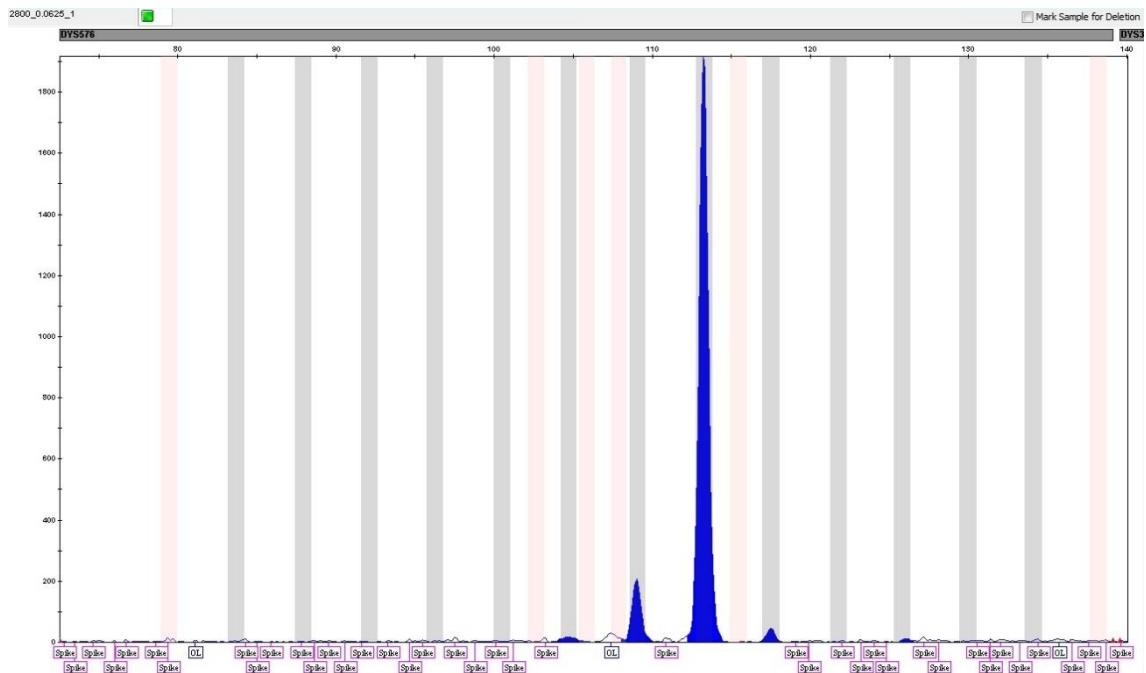


Figura 8. Picos acima de 300 RFUs excluídos.



Com os picos acima da linha de base excluídos como mostrado na figura 8, foram utilizados dois métodos para calcular a determinação do *threshold* analítico para o sistema de amplificação Promega Power Plex Y 23.

a. MÉTODOS DA ANÁLISE ESTATÍSTICA

Para determinar a o limiar da curva ou *threshold* analítico de um pico verdadeiro que irá corresponder a um loco de um perfil genético, utilizamos de dois métodos empregados baseados em estudos relevantes na área. Utilizamos os métodos de Bregul et.al. (2012) de formas diferentes e o método de ISHI et. al. (2017).

4.2.1 MÉTODO 1

No primeiro método calculamos o AT para todos os picos de todos os fluoróforos, sem separação por cor, de forma a determinamos um AT geral, onde a fórmula utilizada foi:

$$AT = \text{Average Noise} + 10(\text{SD}) \text{ (ISHI et. al., 2017) ou } ATM1 = Ybl + ksYbl \text{ (Bregul, et. al., 2012)}$$

Onde AT corresponde a *Analytical Threshold*, *Average Noise* corresponde à média de picos abaixo da linha de 300 RFU e SD é o desvio padrão dos mesmos picos da média o mesmo vale para a fórmula de Bregul respectivamente. O número 10, corresponde ao k da fórmula de Bregul e está relacionado ao intervalo de confiança da curva normal ou gaussiana que corresponde um pico alélico.

Para tanto, obtemos o resultado de AT para 70,95 RFUs. Valor que está dentro do esperado para os padrões de picos verdadeiros já validados anteriormente pelo LGMF, porém possui baixa relevância quando fluoróforos de tamanho de pares de bases diferentes estão presentes, para tanto é necessário valores específicos para cada fluoróforo empregado.

4.2.2 MÉTODO 2

O método 2 é semelhante ao método 1, porém aqui separamos por fluoróforo e calculamos a média e desvio padrão para cada um dos marcadores para o cromossomo Y como mostrado a seguir na tabela 5.

Porém aqui nosso valor de k da fórmula será 3, correspondente ao intervalo de confiança de 99.87% determinado por Bregul et. al., 2012.

Tabela 5. VALORES DE THRESHOLD PARA CADA MARCADOR Y.

MARCADOR	MÉDIA	DESVIO PADRÃO(SD)	VALORES DE AT (EM RFUS)
DYS576	26,59	28,81	113,02
DYS389I	16,65	13,44	56,97
DYS448	22,52	24,03	94,65
DYS389II	22,09	35,96	129,68
DYS19	10,69	7,94	34,51
DYS391	42	44,39	175,18
DYS481	32,46	37,34	144,49
DYS549	24,30	22,15	90,76
DYS533	24,79	16,93	75,68

DYS438	21,79	16,34	70,7
DYS437	17,1	11,66	52,88
DYS570	60,62	54,18	223,17
DYS635	27,85	15,76	75,16
DYS390	39,13	37,28	151
DYS439	29,05	24,58	102,92
DYS392	24,66	19,78	84,01
DYS643	26,43	22,64	94,37
DYS393	59,53	58,11	233,88
DYS458	37,66	17,50	90,16
DYS385A/B	33,84	16,12	82,22
DYS456	30,78	17,51	83,33
Y-GATA-H4	25,12	9,93	54,93

Para diferentes marcadores obteve-se então com o método 2 diferentes valores de AT em um intervalo de confiança de $k = 99,87\%$, o que nos remete a valores diferentes para cada marcador, onde para cada peso observamos semelhanças nos valores obtidos. Como por exemplo, no marcador **DYS570**, o valor de AT é 223,17, enquanto no marcador **DYS393** o valor para AT é de 233,88. Valores parecidos para alelos de peso parecido, porém não igual como o método 1 nos oferece.

Essa semelhança também não é observada em todos os marcadores de peso parecidos, um exemplo disso é o marcador **DYS576** em relação ao marcador **DYS570**, possuem valores de AT calculados bem discriminados.

Esses valores individuais serão por fim aplicados à análises no Laboratório de Genética Molecular Forense, onde para cada marcador, valores acima da marca serão considerados picos referentes a alelos verdadeiros, enquanto valores abaixo serão considerados picos referentes ao ruído de sistema bem como artefatos de PCR, spikes e picos sttuter.

Portanto é ressaltada a importância do segundo método de análise para cada marcador, devido a possuírem pesos diferentes e mesmo quando há pesos semelhantes, os valores de AT podem não ser semelhantes e quando

padronizados como no primeiro método, se despreza a possibilidade de diferença entre eles e como consequência a validação de um pico não verdadeiro, comprometendo a veracidade de um perfil obtido. Em situações de mistura de perfis em casos de estupro onde há mais de um estuprador, a confiabilidade do laudo pode incriminar ou inocentar erroneamente um suspeito.

5 CONCLUSÕES

No primeiro método padronizou-se um valor em comum para todos os marcadores e no segundo, um valor específico para cada marcador. Concluimos que para fins de confiabilidade do método na caracterização de um perfil, o segundo método demonstrou maior intervalo de confiabilidade.

Diante do cenário atual onde diariamente casos de estupro são registrados nos mais diversos países, a importância de ferramentas na ciência que auxiliem a justiça para incriminar ou inocentar um suspeito são cada vez mais necessárias e o sistema comercial da Promega Power Plex Y23 possui um amplo espectro de marcadores que auxiliam na discriminação de um perfil masculino em amostras de mistura com células femininas. A determinação do AT por marcadores irá por sua vez, contribuir na veracidade de um perfil obtido, pelos peritos do Laboratório de Genética Forense do Paraná.

6 REFERÊNCIAS BIBLIOGRÁFICAS

BUTLER, J. M. **Advanced Topics in Forensic DNA Typing: Interpretation**. 1ed. National Institute Of Standards And Technology Gaithersburg, Maryland, USA: ELSEVIER, 2015. 585 p - 10,187,

BUTLER, J. M. **Advanced Topics in Forensic DNA Typing: Methodology**. 1ed. National Institute Of Standards And Technology Gaithersburg, Maryland, USA: ELSEVIER, 2012. 680 p. 37, 38, 39.

BUTLER, J. M. **Fundamentals of Forensic DNA Typing**. National Institute Of Standards And Technology Gaithersburg, Maryland, USA: ELSEVIER, 2010. 500 p.-100,

SNUSTAD, D.P. **Fundamentos de Genética**. 6ªed. Guanabara Koogan, 2013. 760p. 99, 100,101.

FBI, Federal Bureau Of Investigation et al. **QUALITY ASSURANCE STANDARDS FOR FORENSIC DNA TESTING LABORATORIES**. 2011. 29 p

GUSMÃO L., BUTLER J.M., CARRACEDO A., *et al.* (2006) **DNA commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis**, *International Journal Legal Medicine* **120**:191-200.

KAYSER, M.. (2007). **Y Chromosomal markers in forensic genetics**. Em Ralph Rapley and David Whitehouse. *Molecular Forensics*, Wiley.

MALAGHINI, M., ALONSO, C.A.M., DALL' STELLA, R, SCHNEIDER, V. J. **Análises de Material Genético na Investigação Criminal – um relato sobre a evolução dos processos de padronização**. Disponível em: http://www.labfa.com.br/texto_informatgencriminal.htm. Acesso em 26 de fevereiro de 2018.

PowerPlex® Y23 System. 1. ed. Madison: Promega Corporation, 2017. 85 p. v. 1. Disponível em: <<http://www.promega.com/protocols/>>. Acesso em: 29 jan. 2018

NIST – NATIONAL INSTITUTE OF STANDARTS AND TECHNOLOGY.
Workshop: Countdown to 2017: Internal Validation of the new CODIS Loci 26th annual International Symposium on Human Identification. 2015, Grapevine, Texas, United States.

PRINZ, M.; BOLL, K.; BAUM, H.; SHALER, B. **Multiplexing of Y chromosome specific STRs and performance for mixed samples**. Forensic Science International, [s.l.], v. 85, n. 3, p.209-218, mar. 1997. Elsevier BV.
[http://dx.doi.org/10.1016/s0379-0738\(96\)02096-8](http://dx.doi.org/10.1016/s0379-0738(96)02096-8).

RAMSDEN, JEREMY J (2009). **Bioinformatics: An Introduction**. New York: Springer. p. 191. 271 páginas

REECE, J. B., URRY, L. A., CAIN, M. L., WASSERMAN, S. A., MINORSKY, P. V., JACKSON, R. B. **Forensic evidence and genetic profiles**. 10th ed., pp. 430-431. San Francisco, CA: Pearson, 2011.

ROCHA, Andreia Possatti et al. Polimorfismos genéticos: implicações na patogênese do carcinoma medular de tireóide. **Arquivos Brasileiros de Endocrinologia & Metabologia**, [s.l.], v. 51, n. 5, p.723-730, jul. 2007. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/s0004-27302007000500009>.

STEFFEN, B. VALLONE, P. ROMSOS, E. COBLE, M. Countdown to 2017: Internal Validation of the New CODIS Loci. Workshop Countdown to 2017: Internal Validation of the New CODIS Loci 26th annual international Symposium on Human Identification, **NIST – NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY**, October 15th 2015, Grapevine.

SWGAM - SCIENTIFIC WORKING GROUP DNA ANALYSIS METHODS. (Org.). **SWGAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories**. United States, 2014. 20 p.

WALSH, P. Sean; METZGER, David A.; HIGUCHI, Russell. Chelex 100 as a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material. **Biotechniques**, Cetus Corporation And Illinois State Police, v. 54, n. 3, p.134-139, 2013.

WHATMAN PLC (United Kingdom). Ge Healthcare Life Sciences. **Cartões FTA® Whatman™**. 2011. Disponível em: <https://pt.vwr.com/assetsvc/asset/pt_PT/id/17887356/contents>. Acesso em: 06 fev. 2018.