

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

Giovanni Vitral Pinto

IMPLEMENTAÇÃO DO SEQUENCIAMENTO GENÉTICO MASSIVO EM PARALELO NA
SEÇÃO TÉCNICA DE BIOLOGIA E BACTERIOLOGIA LEGAL DA POLÍCIA CIVIL DO
ESTADO DE MINAS GERAIS

BELO HORIZONTE

2022

Giovanni Vitral Pinto

IMPLEMENTAÇÃO DO SEQUENCIAMENTO GENÉTICO MASSIVO EM PARALELO NA
SEÇÃO TÉCNICA DE BIOLOGIA E BACTERIOLOGIA LEGAL DA POLÍCIA CIVIL DO
ESTADO DE MINAS GERAIS.

Dissertação apresentada ao Programa de Pós-Graduação
em Genética do Instituto de Ciências Biológicas da
Universidade Federal de Minas Gerais como parte das
exigências para obtenção do título de Mestre em Genética.

Orientador: Prof. Dr. Eduardo Martin Tarazona Santos

Coorientador: Dr. Rennan Garcias Moreira

Belo Horizonte

2022

043

Pinto, Giovanni Vitral.

Implementação do sequenciamento genético massivo em paralelo na Seção Técnica de Biologia e Bacteriologia Legal da Polícia Civil do Estado de Minas Gerais [manuscrito] / Giovanni Vitral Pinto. - 2022.

79 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. Eduardo Martins Tarazona Santos.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Genética. 2. Genética forense. 3. Marcadores moleculares. 4. Eletroforese. I. Tarazona Santos, Eduardo Martín. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 575



UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Genética

ATA DE DEFESA DE DISSERTAÇÃO / TESE

ATA DA DEFESA DE DISSERTAÇÃO	321/2022 entrada
Giovanni Vitral Pinto	1º/2020 CPF: 036.749.096-00

Às nove horas do dia 11 de fevereiro de 2022, reuniu-se remotamente (rede mundial de computadores) a Comissão Examinadora de Dissertação, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: "IMPLEMENTAÇÃO DO SEQUENCIAMENTO GENÉTICO MASSIVO EM PARALELO NA SEÇÃO TÉCNICA DE BIOLOGIA E BACTERIOLOGIA LEGAL DA POLÍCIA CIVIL DO ESTADO DE MINAS GERAIS", requisito para obtenção do grau de Mestre em Genética. Abrindo a sessão, o Presidente da Comissão, Eduardo Tarazona Santos, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	Instituição	CPF	Indicação
Eduardo Martin Tarazona Santos	UFMG	012.494.056-02	Aprovado
Rennan Garcias Moreira	UFMG	059.401.386-08	Aprovado
Evanguedes Kalapothakis	UFMG	494.307.426-04	Aprovado
Rinaldo Wellerson Pereira	Universidade Católica de Brasília	844.342.796-53	Aprovado

Pelas indicações, o candidato foi considerado: **APROVADO**

O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 11 de fevereiro de 2022.

Eduardo Tarazona Santos - UFMG

Rennan Garcias Moreira - UFMG

Evanguedes Kalapothakis - UFMG

Rinaldo Wellerson Pereira



Documento assinado eletronicamente por **Evanguedes Kalapothakis**, Presidente, em 11/02/2022, às 12:53, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rennan Garcias Moreira**, Biólogo, em 11/02/2022, às 15:38, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eduardo Martin Tarazona Santos**, Professor do Magistério Superior, em 14/02/2022, às 23:10, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rinaldo Wellerson Pereira**, Usuário Externo, em 11/03/2022, às 09:49, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1242624** e o código CRC **35E8E404**.

Na busca incessante pela Justiça, devemos lançar a luz da ciência para iluminar os caminhos da verdade.

Dedico este trabalho aos meus filhos Lucas e Luiza.

AGRADECIMENTOS

Agradeço a minha família, pelo apoio, paciência e compreensão em todas os momentos da minha formação, em especial aos meus pais, Lélío Augusto Pinto e Maria Clarice Vitral Pinto, e aos meus avós Alfen dos Santos Vitral (*in memoriam*) e Lea de Paula Vitral (*in memoriam*). Agradeço à minha irmã, Isabella Vitral Pinto, pelo incentivo de sempre continuar no caminho do aprendizado.

À minha esposa, Thamara Vieira de Oliveira Rocha, pela dedicação aos nossos filhos, Lucas Vitral e Luiza Vitral, nas horas mais difíceis do meu trabalho de Mestrado.

Agradeço ao Professor Dr. Eduardo Tarazona, meu orientador, pela confiança, parceria, disponibilidade e por todo conhecimento compartilhado.

Agradeço ao Dr. Rennan Moreira, meu coorientador, pelo apoio em todos os momentos que precisei, estando sempre disponível para orientar e com seus apontamentos, abrilhantar o meu trabalho.

Ao LDGH (Laboratório de Diversidade e Genética Humana), em especial à Camila, Júlia, Carol e Ricardo, por terem me recebido de braços abertos.

Aos meus amigos, Peritos Criminais, do Laboratório de DNA da Polícia Civil do Estado de Minas Gerais (STBBL) que me apoiaram, incentivaram e contribuíram com os resultados apresentados neste trabalho: Reinaldo Moreira, Carlos Danilo, Elza Cristina e Higgor Dornelas.

Ao amigo Bruno Diniz, Perito Criminal, que vibrou em todas as etapas contribuindo com ótimas sugestões no desenho experimental deste trabalho.

Aos demais amigos e colegas do Laboratório de DNA da Polícia Civil do Estado de Minas Gerais (STBBL) pelos ótimos momentos de convivência que sempre tivemos.

Ao amigo Rodrigo Alves, Perito Criminal, por transformar o árduo trabalho pericial em momentos mais “leves”, além de sempre me incentivar a me tornar um profissional melhor.

Ao grande amigo Jorge Freitas, Perito Criminal Federal, que foi um parceiro dentro da Polícia Federal e que esteve disponível sempre que precisei.

À Superintendência de Polícia Técnico-Científica da Polícia Civil do Estado de Minas Gerais pelo apoio incondicional.

1. Resumo

A análise de amostras biológicas em genética forense é baseada especialmente na obtenção de perfis genéticos pelo processo de PCR (reação em cadeia da polimerase) seguida de eletroforese capilar (EC). Devido às características e limitações deste processo, a Seção Técnica de Biologia e Bacteriologia Legal do Instituto de Criminalística da Polícia Civil do Estado de Minas Gerais (STBBL - PCMG), trabalha para implementar a tecnologia de sequenciamento massivo em paralelo (SMP), uma das tecnologias conhecidas como *Next Generation Sequencing* (NGS).

Com o objetivo de implementar a metodologia de SMP para fins forenses em amostras de violência sexual na STBBL – PCMG, foram selecionadas 30 amostras referentes a casos concretos de violência sexual, um controle positivo (CP) e um controle negativo (CN), para serem sequenciadas usando o *ForenSeq™ DNA Signature Prep Kit* da Verogen®. Essas amostras foram submetidas à metodologia de PCR - EC e posteriormente sequenciadas pelo método de SMP. As métricas da qualidade da corrida de implementação do SMP foram analisadas e evidenciaram o correto funcionamento do equipamento *MiSeq™ FGX da Illumina*. Os dados gerados pelo sequenciamento foram analisados e os resultados obtidos puderam ser comparados com os resultados fornecidos pela metodologia PCR - EC. Para a detecção de locos autossômicos, a metodologia de SMP não gerou perfil genético em 9,4% das amostras, enquanto a PCR - EC não gerou perfil genético em 21,9% das amostras. Quando utilizada a metodologia de SMP, em 81,3% das amostras foram obtidos perfis genéticos completos, enquanto a metodologia de PCR - EC forneceu perfis genéticos completos em 50% das amostras. A análise dos locos Y STRs, a partir da metodologia de PCR - EC, gerou perfis de haplótipo completos e/ou incompletos em 58,8% das amostras, enquanto, a metodologia de SMP gerou perfis de haplótipo completos e/ou incompletos em 43,7%. Soma-se ao melhor resultado apresentado pelo SMP para os STRs autossômicos, um conjunto de informações, tais como, X STRs e SNPs de identidade, os quais, a metodologia de PCR - EC, atualmente padronizada na STBBL, não é capaz de fornecer. Portanto, considerando os resultados deste estudo, concluímos que a metodologia de SMP foi implementada com sucesso no STBBL - PCMG gerando como resultado um Procedimento Operacional Padrão (POP) o qual poderá ser executado pelos Peritos da Seção. A partir do conjunto de dados obtidos pelo SMP é possível agregar informações relevantes à análise forense, tais como, aumentar o poder de discriminação, auxiliar em casos de violência sexual, em casos de paternidade resultante de estupro e em análise de amostras contendo misturas de material genético.

Palavras-chave: Genética forense, marcadores moleculares, eletroforese capilar, *next generation sequencing*.

2. Abstract

The analysis of biological samples in forensic genetics is based especially on obtaining genetic profiles by the process of PCR (polymerase chain reaction) followed by capillary electrophoresis (EC). Due to the characteristics and limitations of this process, the Technical Section of Biology and Legal Bacteriology of the Criminalistics Institute of the Civil Police of the State of Minas Gerais (STBBL - PCMG), works to implement the massive parallel sequencing (SMP) technology, one of the technologies known as Next Generation Sequencing (NGS).

In order to implement the SMP methodology for forensic purposes in samples of sexual violence at STBBL - PCMG, 30 samples referring to specific cases of sexual violence, one positive control (PC) plus one negative control (NC), to be sequenced using Verogen[®]'s ForenSeq[™] DNA Signature Prep Kit. These samples were submitted to the PCR - EC methodology and subsequently sequenced by the SMP method. The quality metrics of the SMP implementation run were analyzed and evidenced the correct functioning of the Illumina MiSeq[™] FGX equipment. The data generated by the sequencing were analyzed and the results obtained could be compared with the results provided by the PCR - EC methodology. For the detection of autosomal loci, the SMP methodology did not generate a genetic profile in 9.4% of the samples, while PCR - EC did not generate a genetic profile in 21.9% of the samples. When using the SMP methodology, in 81.3% of the samples complete genetic profiles were obtained, while the PCR - EC methodology provided complete genetic profiles in 50% of the samples.

The analysis of Y STRs loci, using the PCR - EC methodology, generated complete and/or incomplete haplotype profiles in 58.8% of the samples, while the SMP methodology generated complete and/or incomplete haplotype profiles in 43.7%. In addition to the best result presented by the SMP for autosomal STRs, a set of information, such as X STRs and identity SNPs, which the PCR - EC methodology, currently standardized in STBBL, is not able to provide. Therefore, considering the results of this study, we conclude that the SMP methodology was successfully implemented in STBBL/PCMG generating as a result a Standard Operating Procedure (SOP) which can be executed by the Section Experts. From the data set obtained by the SMP, it is possible to add relevant information to forensic analysis, such as increasing the power of discrimination, assisting in cases of sexual violence, in cases of paternity resulting from rape and in the analysis of samples containing mixtures of material genetic.

Keywords: Forensic genetics, molecular markers, capillary electrophoresis, next generation sequencing.

Lista de figuras

Figura 1.	Representação esquemática da eletroforese capilar em equipamento analisador genético	14
Figura 2.	Figura esquemática das etapas envolvidas no sequenciamento massivo paralelo até a captação do sinal emitido pela incorporação sequencial das bases	15
Figura 3.	<i>Stutters</i> observados nos locos D8S1179, D23S11 e D18S51	17
Figura 4.	Eletroferograma mostrando os locos presentes no <i>Kit PowerPlex® Fusion 6C</i>	26
Figura 5.	Diagrama esquemático dos locos presentes no <i>Kit PowerPlex® Y23</i>	27
Figura 6.	Etapas que envolvem o sequenciamento de amostras utilizando o equipamento <i>MiSeq FGx™</i> da <i>Illumina</i>	29
Figura 7.	Quantidade de locos autossômicos genotipados nas amostras analisadas obtidos com <i>Powerplex® Fusion 6C</i>	32
Figura 8.	Quantidade de locos Y STRs genotipados por amostra analisada	34
Figura 9.	Parte de um eletroferograma (representação gráfica do perfil genético) com material genético de apenas um contribuinte	35
Figura 10A.	Parte de um eletroferograma com material genético de dois indivíduos	36
Figura 10B.	Representação esquemática da forma do pico da amelogenina quando há mistura de material genético masculino e feminino na proporção 1:1	36
Figura 11.	Resultados observados (perfis genéticos completos, perfis genéticos incompletos e ausência de perfil genético) na análise de locos autossômicos	37
Figura 12.	Resultados observados (perfis genéticos completos, perfis genéticos incompletos e ausência de perfil genético) na análise de locos Y STRs	37
Figura 13.	Parte de um gráfico com a distribuição do número de <i>reads</i> e os alelos detectados nos locos analisados	39
Figura 14.	Parte de um gráfico com a distribuição do número de <i>reads</i> e os alelos detectados nos locos analisados (mistura de material genético)	40
Figura 15.	Representação do número de locos obtidos por amostra analisada, utilizando as metodologias de PCR – EC e SMP.	41
Figura 16.	Mapa de calor que representa a comparação entre os resultados obtidos, para os locos STRs autossômicos, nas metodologias de PCR - EC e SMP	42
Figura S.1.	Painel <i>Quality Metrics</i>	55
Figura S.2.	Painel mostrando o número de <i>reads</i> por amostra	56

Lista de tabelas

Tabela 1. Categorias considerando o índice de degradação (ID)	21
Tabela 2. Critérios estabelecidos para escolha dos casos a serem analisados	22
Tabela 3. Combinações dos critérios estabelecidos para escolha dos casos	23
Tabela 4. Apresenta a descrição das 32 amostras analisadas neste estudo	23
Tabela 5. Apresenta a descrição das amostras analisadas e a quantidade de cada tipo de amostra	24
Tabela 6. Resultados obtidos a partir da quantificação dos extratos de DNA das amostras analisadas	31
Tabela 7. Amostras que apresentaram perfis genéticos incompletos, para os locos autossômicos (PCR – EC) e apresentaram perfis genéticos completos (27 locos) com a utilização da metodologia de SMP	38
Tabela 8. Comparação entre as metodologias de PCR - EC e SMP quanto aos perfis genéticos autossômicos obtidos	41
Tabela 9. Comparação entre os resultados obtidos nas metodologias utilizadas, considerando a complexidade do caso e a complexidade da análise	42
Tabela 10. Comparação entre os resultados obtidos nas metodologias utilizadas, considerando a complexidade do caso e a complexidade da análise	43
Tabela 11. Comparação entre os resultados obtidos nas metodologias utilizadas, considerando a complexidade do caso e a complexidade da análise	43
Tabela 12. Comparação entre os resultados obtidos nas metodologias utilizadas, considerando a complexidade do caso e a complexidade da análise	44
Tabela 13. Comparação entre os resultados obtidos, para os locos Y STRs, nas amostras testadas em ambas metodologias de PCR - EC e SMP	45
Tabela 14. Comparação entre a quantidade de perfis de haplótipo Y STRs, completos e incompletos, obtidos a partir das metodologias de PCR - EC e SMP	45
Tabela 15. Quantidade de locos X STRs genotipados nas amostras analisadas, com a utilização da metodologia SMP	46
Tabela 16. Quantidade de SNPs de identidade genotipados nas amostras analisadas, com a utilização da metodologia SMP	47
Tabela S.1. Parâmetros da qualidade de corrida da metodologia de SMP	54

Sumário

Introdução	12
Justificativa	18
Objetivo Geral	19
Objetivos específicos	19
Metodologia	20
Metodologias de extração e quantificação do DNA	24
Extração de DNA	25
Quantificação de DNA	25
Testagem das amostras provenientes de violência sexual utilizando PCR – EC	25
PCR (reação em cadeia da polimerase)	25
PCR utilizando o Kit PowerPlex® Fusion 6C (Promega® Corporation, Madison, WI)	25
PCR utilizando o Kit PowerPlex® Y23, (Promega® Corporation, Madison, WI)	27
Eletroforese Capilar	28
Análise dos resultados - <i>softwares</i> utilizados	28
Testagem das amostras provenientes de violência sexual utilizando SMP	29
Resultados	30
Extração e quantificação	30
Genotipagem utilizando a metodologia de PCR – EC	32
PCR utilizando o Kit <i>PowerPlex® Fusion 6C (Promega® Corporation, Madison, WI)</i> e resultados obtidos na eletroforese capilar	32
Relação entre o número de locos autossômicos obtidos e o total de locos analisados na EC	32
PCR utilizando o Kit <i>PowerPlex® Y23 (Promega® Corporation, Madison, WI)</i> e resultados obtidos na Eletroforese Capilar	32
Resultados obtidos a partir da eletroforese capilar considerando a presença ou ausência de mistura de material genético nas amostras	33
Genotipagem usando a metodologia de Sequenciamento Massivo em Paralelo	36
Resultados obtidos na etapa de sequenciamento para os locos autossômicos	37
Resultados obtidos a partir do SMP considerando o número de contribuintes com material genético nas amostras	38
STRs autossômicos: ganhos obtidos com a utilização da metodologia de SMP	40
Análise comparativa entre os resultados obtidos a partir das metodologias de PCR – EC e SMP, considerando os locos STRs autossômicos.	40
Y STRs obtidos com a utilização da metodologia de SMP	44
Locos X STRs obtidos com a utilização da metodologia de SMP	46
SNPs de identidade obtidos com o ForenSeq™ DNA Signature Prep Kit da Verogen®	46
Bancos de dados congelados de produtos intermediários gerados a partir do SMP	47

Discussões	47
Referências	51
APÊNDICE 1: Resultados das métricas da qualidade obtidas na corrida de implementação do SMP	54
APÊNDICE 2 - Procedimento Operacional Padrão – Construção de Bibliotecas para Sequenciamento Massivo em Paralelo pela Tecnologia Illumina	58

Introdução

A ciência forense é uma área multidisciplinar que envolve a troca de conhecimentos entre diferentes especialidades [1] tais como conhecimentos em química, física, matemática, biologia, além de áreas relacionadas ao direito, contabilidade, informática, entre outras. O objetivo de se utilizar tais conhecimentos é dar suporte técnico e científico a processos criminais [2].

Em investigações criminais, o foco principal do profissional forense, é registrar a materialidade do evento criminal e determinar e/ou confirmar a autoria ou descartar o envolvimento de possível (eis) suspeito (s). As técnicas empregadas permitem identificar se uma pessoa esteve ou não na cena do crime a partir da análise de impressões papilares, ou então amostras biológicas encontradas neste local. No Estado de Minas Gerais a Perícia Criminal é atividade vinculada à estrutura da Polícia Civil e em Belo Horizonte, a Perícia Criminal funciona no Instituto de Criminalística (IC) onde se encontra instalado o Laboratório de Genética Forense - Seção Técnica de Biologia e Bacteriologia Legal (STBBL) da Polícia Civil do Estado de Minas Gerais (PCMG). Considerando o que é determinado pelo Código de Processo Penal, no seu artigo 158: “Quando a infração deixar vestígios, será indispensável o exame de corpo de delito, direto ou indireto, não podendo supri-lo a confissão do acusado” [3]. Quando este vestígio for de natureza biológica, alguns exames podem ser realizados na STBBL, tais como, pesquisa de sangue humano, pesquisa de espermatozoides ou análise de DNA.

A Genética Forense é a área da perícia criminal que utiliza os conhecimentos em biologia molecular, a partir da utilização de técnicas envolvendo a análise de DNA. Dentre suas competências destaca-se a análise de vestígios biológicos provenientes de locais de crime, tais como, crimes contra a vida (homicídios), crimes contra o patrimônio (roubos, furtos, arrombamentos), crimes contra a dignidade sexual (estupro), paternidade criminal (gravidez fruto de estupro) e identificação de cadáveres. Um dos aspectos centrais da genética forense é o uso de marcadores genéticos polimórficos, os quais podem ser usados para identificar determinado indivíduo através do seu genótipo. Existem muitas sequências repetitivas ao longo do genoma humano que possuem alto poder de discriminação. Estas regiões são chamadas de microssatélites ou STRs (*Short Tandem Repeats*) e podem variar em tipos e tamanhos (motivo de repetição entre 2 e 7 pares de base) [4]. Desde meados dos anos de 1980, regiões repetitivas do DNA são usadas em casos forenses, devido ao elevado poder de discriminação e sua capacidade em incluir indivíduos em cenas de crimes e também em excluir suspeitos de terem participado de eventos criminais [5]. No início da década de 90, os STRs foram usados pela primeira vez como marcadores moleculares em testes de paternidade humana e posteriormente foram amplamente utilizados para elucidação de casos forenses.

Na PCMG, atualmente, a metodologia padronizada para a genotipagem de amostras forenses é a PCR (reação em cadeia da polimerase) seguida de EC (eletroforese capilar). A análise de DNA utilizando PCR - EC tem apresentado bastante sucesso para uma grande variedade de vestígios biológicos, tais como, pelos, manchas de saliva, manchas de sangue, manchas de sêmen, células epiteliais e diversos objetos que eventualmente servem de suporte para amostras biológicas [6]. Alguns exemplos de amostras possíveis de serem analisadas seguem abaixo [4]:

- 1- DNA do agressor depositado no corpo ou na roupa da vítima;
- 2- DNA do agressor depositado em um objeto;
- 3- DNA do agressor depositado no local do crime;
- 4- DNA da vítima depositado no corpo ou na roupa do agressor;
- 5- DNA da vítima depositado em um objeto;
- 6- DNA da vítima depositado no local do crime.

Os geneticistas forenses frequentemente enfrentam desafios diante de amostras como as citadas acima, pois, geralmente contêm DNA em baixa quantidade e/ou qualidade e podem conter um ou mais indivíduos contribuindo com material genético [7]. A quantidade de DNA presente nas amostras, a qualidade do DNA extraído e a presença/ausência de mistura de material genético é determinada na etapa de quantificação a partir da utilização do Kit Quantifiler® Trio DNA Quantification Kit (Thermo Fisher Scientific, Oyster Point, CA), vide tópico 7.1.2-Quantificação de DNA. A genotipagem de amostras utilizando eletroforese capilar baseia-se principalmente na detecção de fragmentos (amplicons) de DNA de diferentes tamanhos e que são marcados com fluorescência. Estes amplicons são diferenciados pelo tamanho e, portanto, um dos fatores limitantes da eletroforese capilar é a impossibilidade de se ultrapassar um número máximo de locos relevantes em uma mesma reação de PCR [8,9]. Isto acontece porque amplicons diferentes em sequência, mas que possuem o mesmo tamanho, migram lado a lado na malha do polímero (utilizado para preencher os capilares do equipamento de eletroforese), não podendo ser diferenciados nesta metodologia de análise (figura 1). Para superar essas limitações, o laboratório de genética forense necessita desenvolver diferentes fluxos de trabalho para solucionar um único caso, ou até mesmo, recorrer a outros laboratórios para expandir os STRs utilizados [9].

Nos últimos anos o *Next Generation Sequencing (NGS)* tem surgido como uma variedade de técnicas capazes de superar as limitações das análises baseadas em EC para genotipagem de amostras forenses. A STBBL adquiriu o equipamento *MiSeq™ FGx da Illumina* o qual baseia-se na tecnologia de sequenciamento massivo em paralelo por síntese (SMP), amplamente adotada em múltiplas áreas da ciência, inclusive nas ciências forenses.

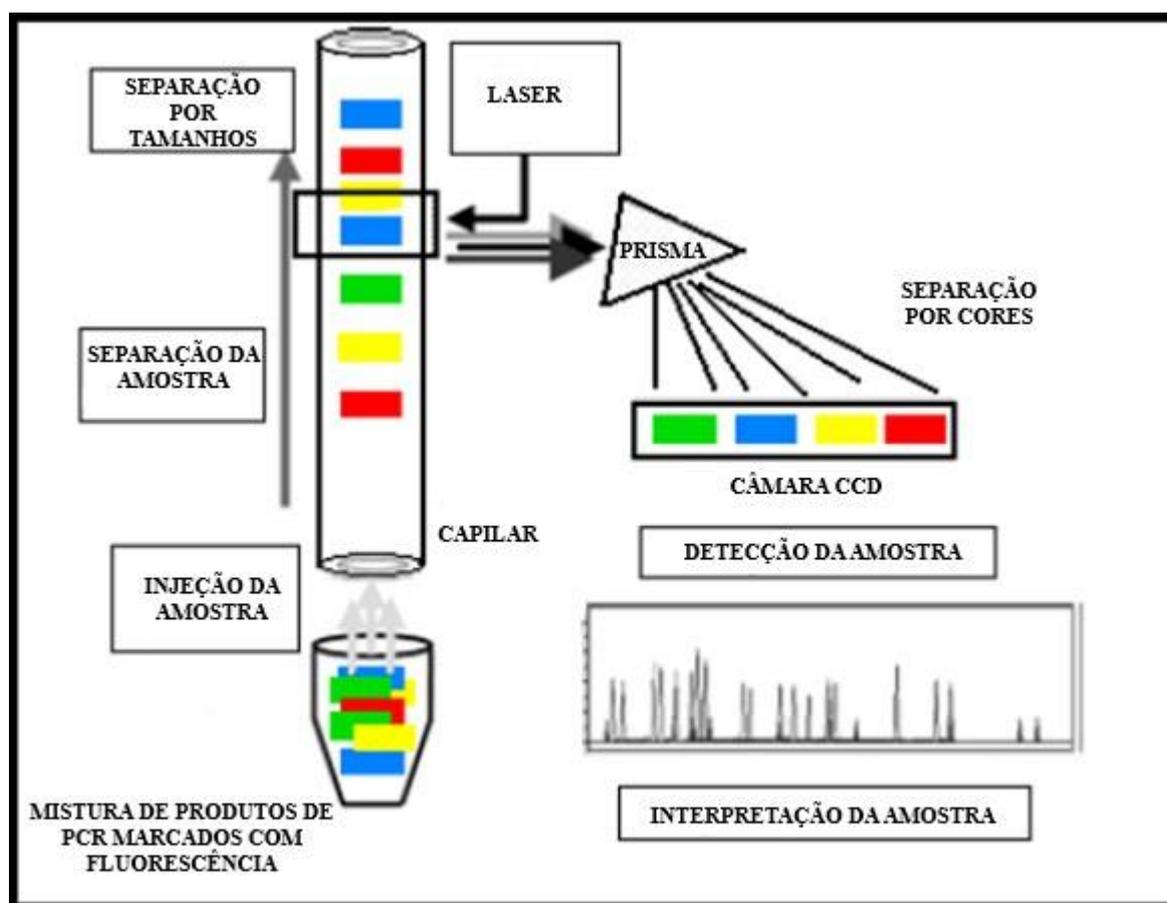


Figura 1: Representação esquemática da eletroforese capilar em equipamento analisador genético. Fonte: Adaptado de Crispim, Bruno do Amaral (2012) [37]

A corrida de implementação do SMP na STBBL gerou métricas da qualidade que foram analisadas e evidenciaram o correto funcionamento do equipamento *MiSeqTM FGX da Illumina* (apêndice 1). A tecnologia de SMP do *MiSeqTM FGx da Illumina* usa uma câmera de alta resolução para gravar as imagens dos dNTPs marcados com fluorescências que são incorporados sequencialmente durante a síntese da fita complementar, permitindo o sequenciamento massivo em paralelo de milhões de amplicons alvo [9]. O SMP se inicia com a adição de adaptadores nas extremidades das sequências de DNA alvo, a partir de duas reações de PCR. Após a PCR1 e a PCR 2, os fragmentos de DNA se ligam à superfície da *flow cell* e passam por uma nova etapa de PCR chamada amplificação em ponte. A amplificação em ponte gera milhares de *clusters* dos fragmentos de DNA os quais passarão pelo processo de sequenciamento. Este processo permite distinguir alelos que têm o mesmo comprimento, porém, possuem sequências diferentes (figura 2).

O *MiSeqTM ForenSeq Genomics*, ou *MiSeqTM FGx da Illumina*, foi desenvolvido para análise de DNA humano envolvendo casos forenses. Este sistema é composto por quatro componentes: o *ForenSeqTM DNA Signature Prep Kit* (relacionado à amplificação do DNA alvo e pela individualização das amostras com a utilização de diferentes *indexes* ou *barcodes*), o *MiSeqTM*

FGx Reagent Kit (composto pela *flow cell* e pelo cartucho de reagentes que serão inseridos no equipamento), o *MiSeqTM FGx sequencing instrument* (equipamento de sequenciamento com seu servidor acoplado) e o *ForenSeqTM Universal Analysis Software* (*software* responsável por realizar as análises dos resultados obtidos e gerar os perfis genéticos das amostras) [9].

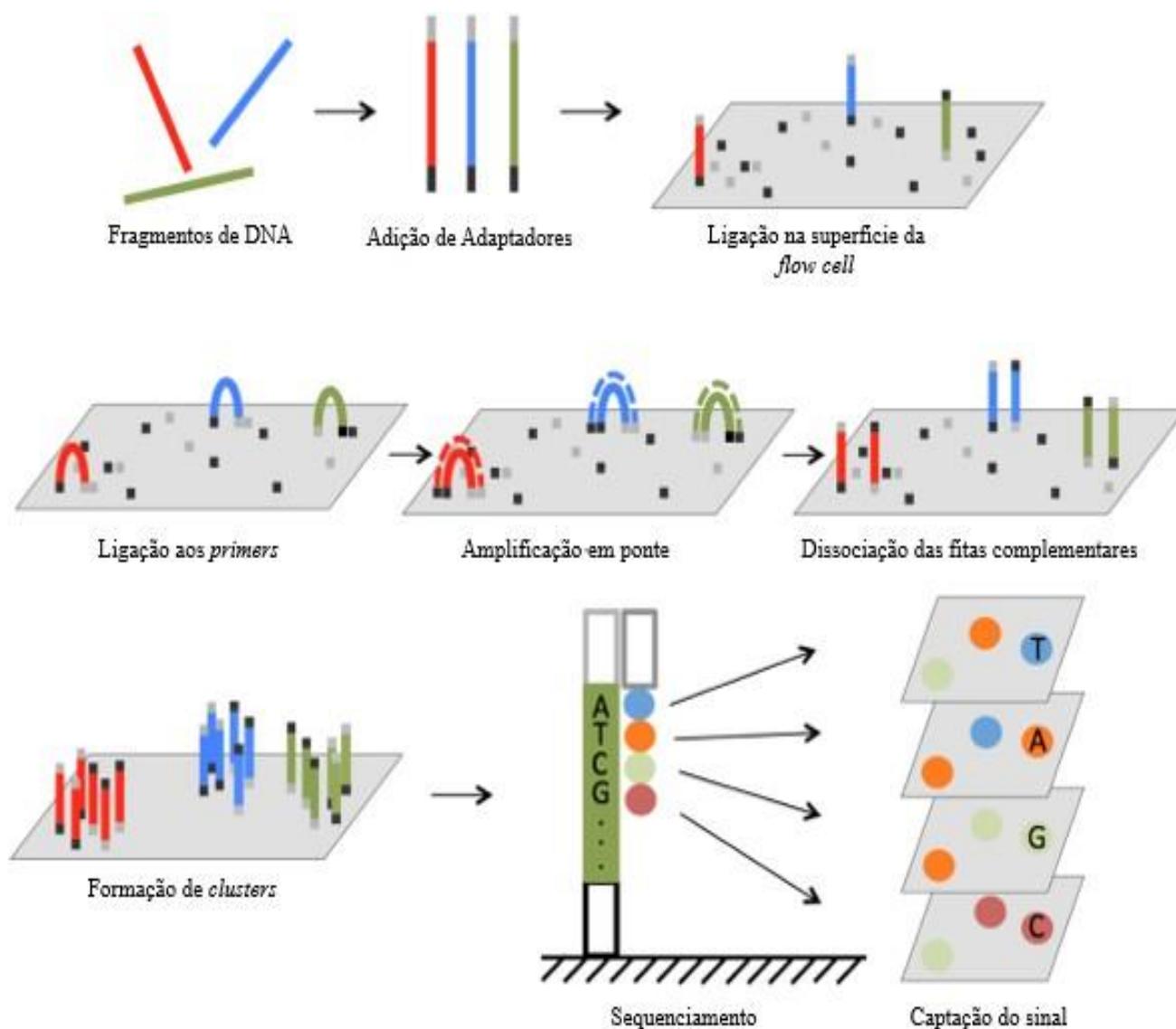


Figura 2: Figura esquemática das etapas envolvidas no sequenciamento massivo em paralelo até a captação do sinal emitido pela incorporação sequencial das bases. Adaptado de <https://edisciplinas.usp.br/> (Ciências Biomédicas - FMRP/USP - Disciplina RCB0300 - Biotecnologia III - Princípios de Sequenciamento em escala genômica [39].

O *MiSeqTM FGx System* foi submetido a processos de validação pelo *SWGDM* (*Scientific Working Group on DNA Analysis Method*) [10], que é um grupo de cientistas internacionais que discutem, avaliam e compartilham protocolos e pesquisas relacionadas à biologia forense e

fornece parâmetros de qualidade relacionadas à genética forense para organizações policiais, como o *FBI (Federal Bureau Investigation)*, por exemplo.

O uso da tecnologia de sequenciamento massivo em paralelo tem possibilitado a obtenção simultânea de dados de STRs autossômicos e sexuais (X e Y STRs), além de diversos SNPs (*Single Nucleotide Polymorphism*) resultando em um volume muito maior de informações por experimento quando comparada com a técnica rotineiramente utilizada nos laboratórios de genética forense, que é a PCR seguida de eletroforese capilar [11]. Os SNPs de identidade podem ser usados na genética forense de forma complementar aos STRs com o objetivo de aumentar o poder de discriminação do exame de DNA. A detecção de SNPs de identidade pode ser determinante nos casos de violência sexual, onde há grande chance de se encontrar DNA degradado e a ocorrência de misturas de material genético é frequente.

A capacidade da metodologia de SMP, utilizando o *MiSeq™ FGx* da *Illumina*, em analisar muitos locos STRs e SNPs simultaneamente, a torna mais vantajosa quando comparada à eletroforese capilar, pois, aumenta a possibilidade de identificação genética das amostras analisadas. Com uso do SMP é possível a solução de casos em que as amostras apresentem menor número de locos autossômicos genotipados, resultado comum de ser observado em amostras com pouco DNA ou com DNA degradado, pois, o volume final de informação (obtenção de SNPs de identidade, X STRs e Y STRs) pode ser suficiente para conclusão da análise [7].

O poder de discriminação também é aprimorado devido ao aumento do número de alelos observados com o SMP, onde a informação das sequências obtidas pode ser usada para diferenciar alelos do mesmo tamanho, porém, com sequências diferentes, o que pode ser de grande importância na análise de parentesco ou na análise de misturas de materiais biológicos [11]. Em estudos populacionais, o *RMP (random match probability)*, que se traduz na possibilidade de se encontrar o mesmo perfil genético sequenciado na população em geral, diminuiu cerca de 20 (vinte) vezes, já o índice de paternidade aumentou cerca de 7 vezes quando três STRs D3S1358, D12S391 e D21S11 foram genotipados por SMP em determinada população, quando comparados com resultados obtidos por PCR-EC [12].

Esses três locos são altamente polimórficos com mais de uma sub-repetição (isoalelos), sendo que o número de alelos detectados é quase o triplo em relação aos obtidos utilizando-se a metodologia de PCR-EC [7].

Conforme dados obtidos por *Chiara Gelardi et. al* (2014), aproximadamente 30% dos genótipos homocigotos detectados por PCR - EC tornam-se heterocigotos quando sequenciados por SMP [12].

Somados aos locos STRs autossômicos, o *ForenSeq™ DNA Signature Prep Kit da Verogen®* possui locos STRs para o cromossomo Y e para o cromossomo X, além de SNPs de

identidade. Os locos Y STRs são específicos para indivíduos do sexo masculino e tem a capacidade de informar se dois indivíduos pertencem à mesma linhagem patrilinea e em casos de violência sexual, são uma ferramenta complementar aos STRs autossômicos. Os locos X STRs são específicos para o cromossomo X e fornecem um conjunto de dados complementares aos STRs autossômicos, podendo ser usados em casos de violência sexual, com o objetivo de auxiliar na identificação do agressor ou até mesmo na solução de casos de paternidade criminal (gravidez resultante de estupro). Os SNPs (*single nucleotide polymorphisms*) de identidade também são usados de forma complementar aos marcadores STRs autossômicos. Eles têm a capacidade de aumentar o poder de discriminação dos exames relacionados à genética forense e tem como características alta estabilidade genética e maior abundância no genoma humano quando comparados aos STRs [13].

A utilização do SMP também pode ser muito útil na análise de misturas, por possuir a capacidade de diferenciar alelos do agressor que estejam em posição de *stutter* em relação aos alelos da vítima. *Stutters* são artefatos de PCR resultantes da cópia de locos STR pela DNA polimerase. Estes artefatos de PCR possuem uma unidade de repetição a menos que o pico do alelo principal. Produtos *stutter* com uma unidade de repetição maior que o pico do alelo principal são raros quando trabalhamos com locos STR de tetranucleotídeos. [4]. A figura 3 mostra artefatos *stutter* observados nos locos D8S1179, D21S11 e D18S51.

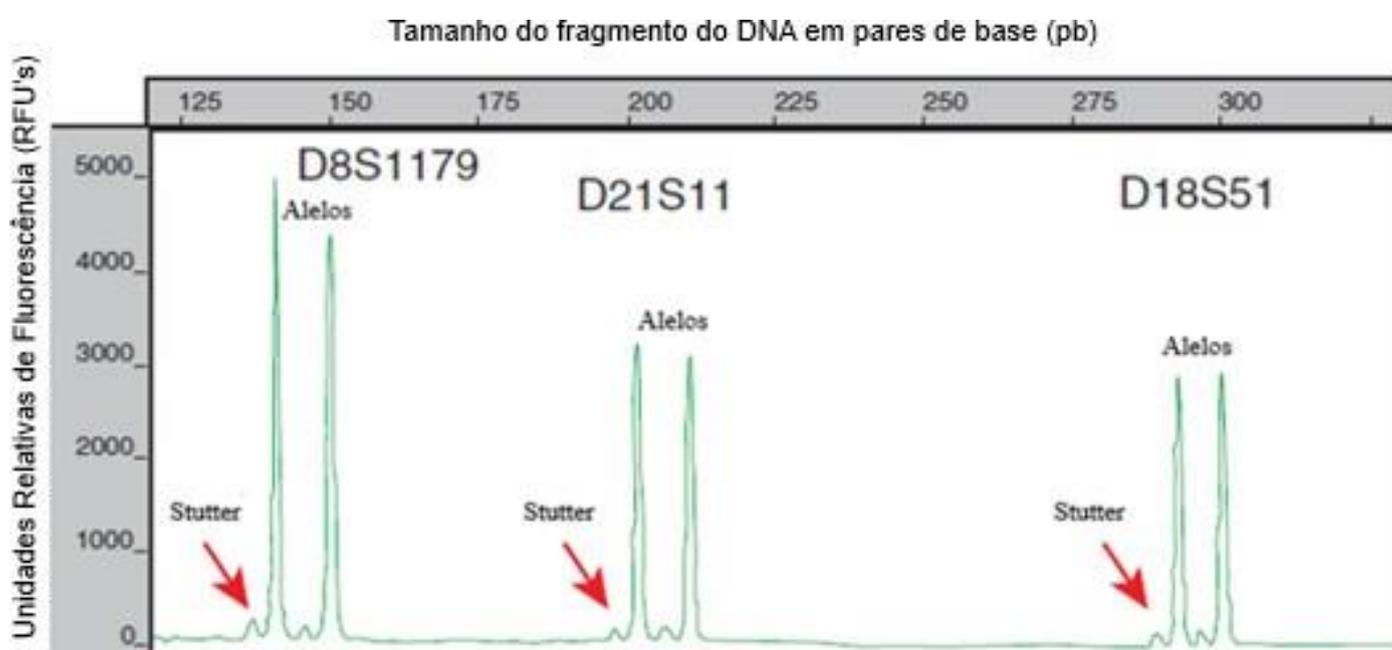


Figura 3: *Stutters* observados nos locos D8S1179, D21S11 e D18S51 [4].

O estudo de caso “*How Next Generation Sequencing Resolved a Difficult Case, Leading to the First Criminal Conviction of its Kind*” [14] publicado pela Verogen® em novembro de 2019, apresentou a primeira condenação criminal baseada em resultados fornecidos pela metodologia de NGS e que se baseou na identificação dos alelos que estavam em posição de *stutter*, apontando que estes alelos pertenciam ao investigado e não à vítima. Quando há uma mistura de material genético, entre vítima de violência sexual e seu agressor, em que o agressor é o contribuinte minoritário, os alelos deste contribuinte minoritário podem se encontrar na posição de *stutter* em relação ao alelo da vítima. Neste caso a PCR - EC não é capaz de distinguir a origem desse alelo, pois, ambos (alelo do agressor e *stutter*) possuem o mesmo tamanho. O SMP é capaz de diferenciar o *stutter* do alelo do agressor, quando, apesar de terem o mesmo tamanho, apresentam sequências de bases diferentes.

Justificativa

A STBBL - PCMG enfrenta diferentes desafios na rotina de análise das amostras biológicas recebidas na instituição. São limitações relacionadas à natureza das amostras e às técnicas utilizadas:

1-Natureza das amostras forenses: As amostras forenses possuem normalmente quantidade de DNA abaixo de 1,0 ng e para amplificação utilizando o kit multiplex comercial disponível na STBBL, *PowerPlex® Fusion 6C (Promega® Corporation, Madison, WI)*, o valor recomendado é de 1,0 ng de DNA [17]. Além da baixa concentração de DNA, muitas vezes, só há um único material disponível para a realização do exame, como, por exemplo, um suabe vaginal ou um fragmento de tecido impregnado com material biológico, que uma vez consumido, não há possibilidade de realização de nova análise [15]. Portanto, a utilização do SMP aumenta a chance de se obter dados mesmo em amostras com baixa quantidade de DNA devido à possibilidade de se acessar um número maior de locos em relação à PCR - EC.

2-Amostras com o DNA degradado: O material biológico pode ter sido exposto às intempéries ou a microrganismos que degradam as moléculas de DNA em pequenos fragmentos. Para que ocorra a amplificação do DNA usando a técnica de PCR, a molécula original de DNA deve estar intacta nas regiões de ligação dos *primers* e na região de extensão. Portanto, quanto mais degradado o DNA, menos moléculas terão o comprimento necessário para a amplificação por PCR [15]. O SMP estuda uma variedade de SNPs que possibilitam a obtenção de dados mesmo em amostras degradadas.

3- Inibidores da reação de PCR: Amostras forenses podem estar sujeitas à contaminação por substâncias que inibem a reação de PCR, tais como, material terroso, areia, madeira, corantes têxteis, couros e outros [15,16]. Como resultado da presença de inibidores na reação de PCR podemos observar a perda de alelos dos STRs de tamanhos maiores ou até mesmo ausência de alelos em todos os locos [15]. Considerando que o SMP estuda uma grande variedade de STRs e SNPs, a chance de obtenção de informações referentes aos locos estudados é aumentada.

4- Grande volume de amostras relacionadas a casos de violência sexual custodiadas na STBBL: A implementação da metodologia de SMP permitiria a análise de grande número de amostras e a obtenção de grande volume de dados a partir de uma única reação.

5- Amostras relacionadas a violência sexual podem conter misturas de materiais genéticos (vítima e agressor): A análise dessas amostras é bastante desafiadora quando se utiliza o sistema PCR – Eletroforese Capilar, pois, alelos com comprimentos semelhantes não podem ser diferenciados.

6- Relevância social: Segundo a OMS (Organização Mundial de Saúde) pode envolver aspectos de saúde, sociais e financeiros.

7- Aquisição do *MiSeqTM FGX da Illumina*: A STBBL - PCMG é o primeiro laboratório de genética forense Estadual do país a adquirir o equipamento *MiSeqTM FGX da Illumina*. Considerando-se a complexidade desta tecnologia é imprescindível a capacitação de pessoal para este tipo de análise.

Objetivo Geral

Implementar a metodologia de Sequenciamento Massivo em Paralelo (SMP) para fins forenses em amostras de violência sexual no Laboratório de Genética Forense – STBBL – PCMG por meio do estabelecimento de um procedimento operacional padrão e avaliar sua performance em comparação com a metodologia de PCR – EC já estabelecida.

Objetivos específicos

- 6.1- Selecionar 30 amostras relacionadas a casos reais de violência sexual custodiados na STBBL;
- 6.2- Genotipar 30 amostras utilizando a técnica de PCR – EC;
- 6.3- Construir bibliotecas para o SMP das mesmas 30 amostras do item 6.2;
- 6.4- Sequenciar o *pool* de bibliotecas;

- 6.5- Analisar os resultados gerados na corrida usando o *ForenSeq™ Universal Analysis Software* da *Verogen®*;
- 6.6- Estabelecer estudo comparativo entre as metodologias PCR – EC e SMP;
- 6.7- Criação de banco de dados congelados dos produtos intermediários obtidos do SMP;
- 6.8 - Elaborar um POP (procedimento operacional padrão) para a metodologia de SMP considerando a tecnologia *Illumina – Verogen®*.

Metodologia

Para a realização do estudo foram selecionadas 30 amostras de casos concretos relacionados à crimes de violência sexual (quadro 3) considerando as variáveis: qualidade e quantidade do DNA (avaliadas na etapa de quantificação) e complexidade de análise, definida pela presença ou ausência de misturas de DNA (amostras com material genético de mais de uma pessoa). Foram adicionados ao estudo um controle positivo (CP – DNA genômico humano masculino 2800 M da *Promega® Corporation, Madison, WI*) e um controle negativo (CN – água ultrapura). As amostras forenses podem variar quanto à quantidade e qualidade do DNA analisado. Foram consideradas amostras com quantidade de DNA suficiente para serem submetidas à amplificação utilizando o *Powerplex® Fusion 6C Kit* [17] e o *ForenSeq™ DNA Signature Prep Kit* [18] as amostras que continham pelo menos 1 ng de DNA. Aquelas que possuíam quantidade inferior a 1 ng foram consideradas como amostras contendo baixa quantidade de DNA. A qualidade da amostra é medida pelo índice de degradação (ID) determinado na etapa de quantificação utilizando o *Quantifiler® Trio DNA Quantification Kit (Thermo Fisher Scientific, Oyster Point, CA)*. O ID é medido pela razão entre a eficiência de amplificação dos produtos de PCR de baixo peso molecular (80 pares de base) e a eficiência de amplificação dos produtos de PCR de alto peso molecular (214 pares de base). A partir dos valores obtidos verifica-se o grau de degradação da amostra. O valor do ID para que uma amostra fosse considerada degradada foi > 1,5. Os valores de ID os quais basearam esse estudo foram obtidos a partir da publicação de *Stefano Vernarecci et. al* [19] e estão descritos na tabela 1. Portanto, apenas as amostras que possuíam ID < 1,5 foram consideradas como não degradadas.

A etapa de quantificação utilizando o *Quantifiler® Trio DNA Quantification Kit (Thermo Fisher Scientific, Oyster Point, CA)* indica se a amostra contém mistura de material biológico entre indivíduos do sexo feminino e masculino. O número de contribuintes de determinada amostra biológica determina o grau de complexidade da análise a ser realizada.

Quando a amostra biológica possui apenas um contribuinte os cálculos estatísticos utilizados são baseados no *RMP (random match probabilities)*, que é a probabilidade de uma

coincidência entre perfis genéticos ocorrer aleatoriamente em determinada população. Na STBBL o *software* utilizado para calcular o *RMP* é o *LRmix* [20]. Quando a amostra biológica apresenta mistura de material genético a análise é mais complexa, pois, deve-se verificar se há presença de menor e maior contribuintes ou se os indivíduos apresentam contribuição igualitária para a produção da amostra. A mistura deve ser submetida ao cálculo estatístico utilizando o *software LRmix* [20] que possibilita verificar a relação entre duas hipóteses: Hipótese 1 (H1) / Hipótese 2 (H2), onde H1 considera a evidência genética dada a hipótese da acusação (a mistura é composta pela amostra da vítima e do investigado) e H2 considera a evidência genética dada a hipótese da defesa (a mistura é composta pela amostra da vítima e de outro homem da população diferente do investigado).

Tabela 1: Categorias considerando o índice de degradação (ID)

Quatro categorias separadas de acordo com o índice de degradação	
Índice de Degradação	Categoria da degradação
0 - 1,5	sem degradação
1,5 – 4	levemente degradada
4 – 10	degradada
> 10	muito degradada

O perfil de mistura também deve ser submetido ao cálculo de *MRE* (*match rarity estimate*) [21] o qual determina o quanto a mistura obtida é rara em determinada população. A ferramenta utilizada para a realização do cálculo do *MRE* é o *software CODIS* (*Combined DNA Index System*) de propriedade da Polícia Federal Norte Americana (*FBI - Federal Bureau Investigation*) e cedido à Secretaria de Estado de Justiça e Segurança Pública de Minas Gerais (SEJUSP/MG) para ser utilizado na STBBL.

Para a realização deste estudo foram analisados, preferencialmente, 4 tipos de casos (tabelas 2 e 3):

(i) casos com DNA de qualidade ($ID < 1,5$) e quantidade adequadas (1 ng de DNA), sem resultados de misturas de DNA; (ii) casos com DNA degradado ($ID > 1,5$) e quantidade de DNA insuficiente (< 1 ng de DNA), sem resultados de misturas de DNA.

(iii) casos com DNA de qualidade ($ID < 1,5$) e em quantidade adequadas (1 ng de DNA), com misturas de DNA; (iv) casos com DNA degradado ($ID > 1,5$) e quantidade de DNA insuficiente (< 1 ng de DNA) e com misturas de DNA.

A partir da combinação de critérios estabelecidos na tabela 3, as amostras selecionadas foram agrupadas de acordo com as tabelas 4 e 5.

Tabela 2: Critérios estabelecidos para escolha dos casos a serem analisados.

Complexidade do caso baixa	Caso com amostras contendo DNA em quantidade e qualidade adequadas (suabes orais, suabes genitais, perianais e vestes).
Complexidade do caso alta	Caso com amostras contendo DNA em quantidade e qualidade insuficientes (suabes orais, suabes genitais, perianais e vestes).
Análise simples	Amostras as quais foram obtidos perfis genéticos únicos. Exemplos: perfil genético masculino único obtido das amostras coletadas da vítima ou de vestes. Neste caso, o cálculo estatístico é realizado calculando-se a frequência de incidência do perfil genético encontrado.
Análise complexa	Amostras as quais foram obtidas misturas de perfis genéticos. Exemplos: perfil genético masculino e feminino obtido das amostras coletadas da vítima ou de vestes. Neste caso, o cálculo estatístico é realizado utilizando <i>software</i> validado para amostras envolvendo misturas.

Tabela 3: Combinações dos critérios estabelecidos para escolha dos casos a serem analisados.

Complexidade do caso x complexidade da análise	
Categoria A	Categoria B
-Complexidade do caso baixa	-Complexidade do caso alta
-Análise simples	-Análise simples
Categoria C	Categoria D
-Complexidade do caso baixa	-Complexidade do caso alta
-Análise complexa	-Análise complexa

A descrição de cada amostra é acompanhada pelo número do caso, o qual é sequencial e único dentro da STBBL. Este código numérico é utilizado desde a entrada da amostra no Laboratório de Genética Forense da Polícia Civil e tem como objetivo preservar o anonimato das amostras analisadas. Portanto, cada caso, é composto por todas as amostras relacionadas ao mesmo caso e geralmente são compostas por amostras de referência (suabes orais) coletadas da vítima e do investigado e de amostras questionadas (coletadas de vestes, de objetos, do corpo da vítima etc).

Tabela 4: Apresenta a descrição das 32 amostras analisadas neste estudo. A categoria de cada amostra foi determinada de acordo com a tabela 3.

Amostra	Nº do caso	Descrição das amostras	Categoria
1	79-20	Suabe da mucosa oral coletado da vítima	D
2	79-20	Suabe da mucosa oral coletado do investigado	B
3	79-20	Fração espermática do suabe vaginal coletado da vítima	D
4	79-20	Fração não espermática do suabe vaginal coletado da vítima	D
5	79-20	Fração espermática do suabe perianal coletado da vítima	D
6	79-20	Fração não espermática do suabe perianal coletado da vítima	D
7	79-20	Suabe subungueal coletado da mão direita da vítima	D
8	79-20	Suabe subungueal coletado da mão direita da vítima	D
9	79-20	Fragmento de tecido coletado da calcinha da vítima	D
10	996-20	Suabe da mucosa oral coletado do investigado	B
11	996-20	Fração espermática do suabe vaginal coletado da vítima	B
12	996-20	Fração não espermática do suabe vaginal coletado da vítima	B
13	448-20	Suabe da mucosa oral coletado do investigado	B
14	448-20	Fração espermática do suabe vaginal coletado da vítima	C
15	448-20	Fração não espermática do suabe vaginal coletado da vítima	D
16	448-20	Fração espermática do suabe perianal coletado da vítima	D
17	448-20	Fração não espermática do suabe perianal coletado da vítima	D

18	650-20	Suabe da mucosa oral coletado da vítima	A
19	650-20	Suabe da mucosa oral coletado do investigado	A
20	650-20	Fração espermática do suabe vaginal coletado da vítima	C
21	650-20	Fração não espermática do suabe vaginal coletado da vítima	C
22	650-20	Fração espermática do suabe perianal coletado da vítima	Indeterminada
23	650-20	Fração não espermática do suabe perianal coletado da vítima	D
24	650-20	Fragmento de tecido coletado da calcinha da vítima	D
25	3468-19	Fragmento de músculo da vítima	C
26	3468-19	Suabe da mucosa oral coletado do investigado	B
27	3468-19	Fração espermática do suabe vaginal coletado da vítima	B
28	3468-19	Fração não espermática do suabe vaginal coletado da vítima	B
29	3468-19	Leito ungueal coletado da vítima	D
30	229-20	Fração espermática do suabe vaginal coletado da vítima	A
31	CP	Controle Positivo	B
32	CN	Controle Negativo	D

*Continuação da tabela 4.

Tabela 5: Apresenta a descrição das amostras analisadas e a quantidade de cada tipo de amostra.

Natureza das amostras analisadas	Número de amostras
Suabes da mucosa oral	7
Fragmento de músculo	1
Frações espermáticas dos suabes vaginais	6
Frações não espermáticas dos suabes vaginais	5
Frações espermáticas dos suabes perianais	3
Frações não espermáticas dos suabes perianais	3
Suabes subungueais	3
Fragmentos de tecidos coletados de vestes	2
Controle positivo	1
Controle negativo	1
Total	32

Metodologias de extração e quantificação do DNA

As metodologias de extração e quantificação do DNA utilizadas nas 30 amostras foram comuns para ambos os métodos utilizados neste estudo (PCR – EC e SMP).

Extração de DNA

A extração do DNA consiste no tratamento químico e físico (temperatura) com o objetivo de romper as membranas celulares e posterior purificação do DNA obtido.

A extração das amostras de referência (suabes de mucosa oral) foi realizada com resina *Chelex 100*[®] (que consiste em uma resina quelante) associada à proteinase K [22,23,24].

Para a extração de amostras provenientes de violência sexual foi utilizado o equipamento *Maxwell*[®] 16 *Tissue DNA Purification Kit* (*Promega*[®] Corporation, Madison, WI) com protocolo de extração diferenciada [25], separando células epiteliais de células espermáticas. Este método permite isolar o DNA do agressor do DNA da vítima, visto que, na maioria das vezes, há mistura destes materiais [4].

Quantificação de DNA

Neste estudo foi utilizado o equipamento *QuantStudio Real Time PCR*[®] da *Thermo Fisher Scientific, Oyster Point, CA* [26] associado ao Kit de quantificação *Quantifiler*[®] *Trio DNA Thermo Fisher Scientific, Oyster Point, CA* [27]. Este processo visa quantificar o DNA total da amostra extraída e o DNA humano masculino, além de mensurar o grau de degradação da amostra biológica analisada [28].

A amostras que apresentaram quantidade de DNA acima de 1 ng, tiveram suas amostras diluídas para a quantidade final de 1 ng. Aquelas amostras que apresentaram quantidade de DNA inferior a 1 ng, foram concentradas utilizando o equipamento *SpeedVac Concentrator* da (*Thermo Fisher Scientific, Oyster Point, CA*).

Testagem das amostras provenientes de violência sexual utilizando PCR – EC

A testagem das amostras provenientes de violência sexual foi realizada utilizando inicialmente a metodologia de PCR - EC que é a metodologia padronizada na STBBL - PCMG.

PCR (reação em cadeia da polimerase)

PCR utilizando o Kit *PowerPlex*[®] *Fusion 6C* (*Promega*[®] Corporation, Madison, WI)

Após as etapas de extração e quantificação, as amostras foram submetidas à amplificação utilizando o *PowerPlex*[®] *Fusion 6C* (*Promega*[®] Corporation, Madison, WI), o qual possui um

conjunto de *primers* para amplificar 23 locos autossômicos, 3 locos específicos para o cromossomo Y e a amelogenina (marcador genético que permite a determinação do sexo biológico). A figura 4 mostra um eletroferograma com os locos presentes no *Kit PowerPlex® Fusion 6C (Promega® Corporation, Madison, WI)* [29] onde os valores plotados no eixo vertical (variando de 0 a 8.000) representam as Unidades Relativas de Fluorescência (*RFUs*), ou seja, a intensidade de fluorescência emitida pelos fluoróforos. Os valores plotados no eixo horizontal (variando de 75 a 475) representam os tamanhos dos fragmentos em pares de base (pb). Os locos são apresentados em cinco cores diferentes as quais são emitidas pelos fluoróforos FL-6C (azul), JOE-6C (verde), TMR-6C (amarelo, representado em preto para melhor visualização), CXR-6C (vermelho) e TOM-6C (roxo) ligados aos diferentes produtos de PCR.

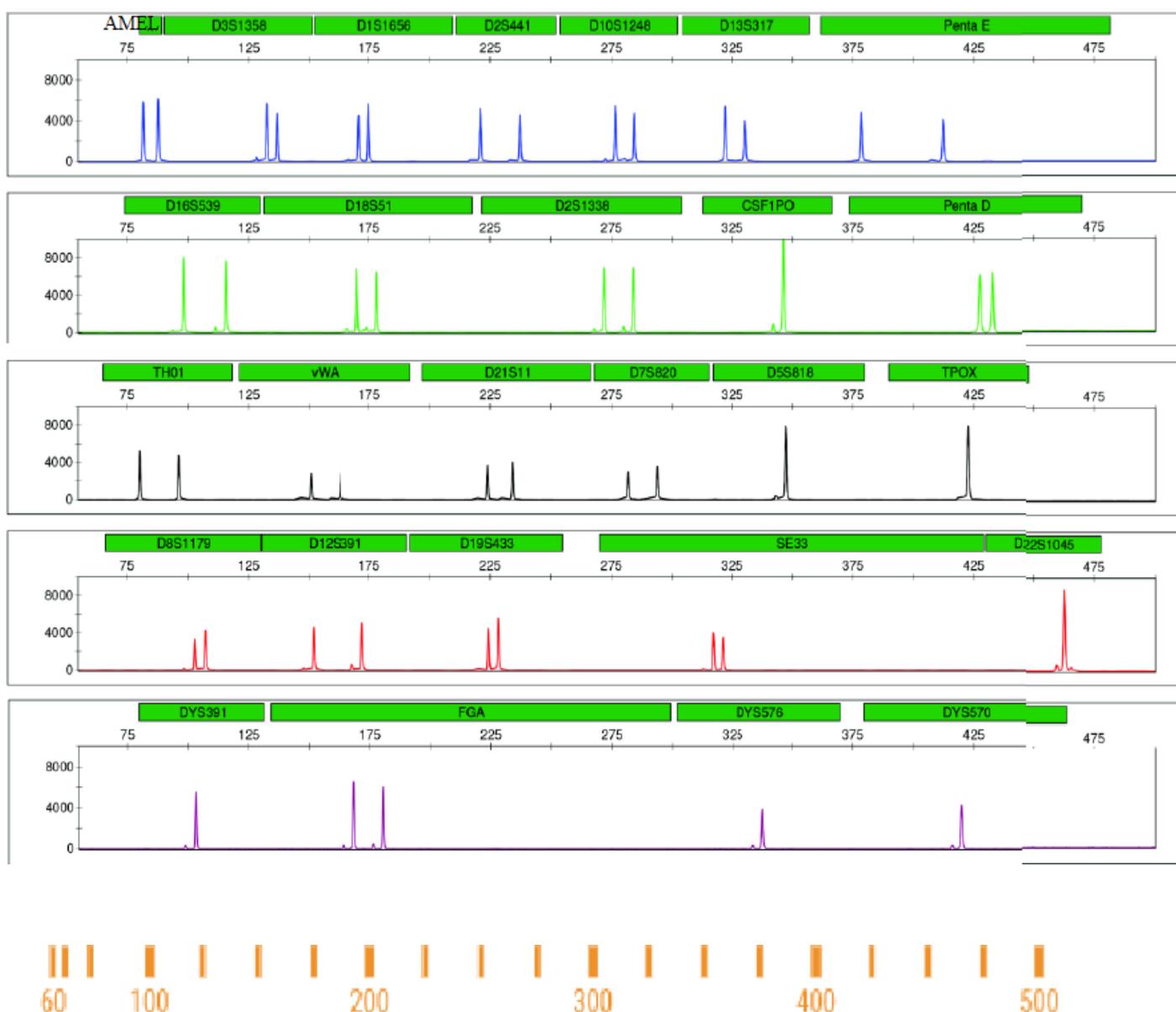


Figura 4: Eletroferograma mostrando os locos presentes no *Kit PowerPlex® Fusion 6C* (Promega® Corporation, Madison, WI) a partir da amplificação do controle positivo 2800 M. Esta figura apresenta 23 locos autossômicos, 3 locos específicos para cromossomo Y e a AMEL (amelogenina). O padrão interno de corrida é detectado no canal alaranjado e é marcado com o WEN (*WEN Internal Lane Standard 500*) Fonte: *Developmental Validation of the PowerPlex® Fusion 6C System* [30].

PCR utilizando o Kit PowerPlex® Y23, (Promega® Corporation, Madison, WI)

As amostras relacionadas à violência sexual devem ser submetidas à amplificação utilizando kits específicos para locos do cromossomo Y (Y STRs), de acordo com protocolos estabelecidos na STBBL [31], pois, em muitas amostras desta natureza que envolvem vítima do sexo feminino, o material masculino (do agressor) pode estar em proporções muito inferiores quando comparado ao material da vítima o que impossibilita a identificação do perfil genético do estuprador devido à amplificação preferencial do material genético da vítima. Para enfrentar a desproporção de materiais biológicos (feminino e masculino) é utilizado o *kit PowerPlex® Y23* (Promega® Corporation, Madison, WI), o qual possui um conjunto de primers para amplificar 23 locos Y STRs. Para o estudo de Y STRs, das 30 amostras, 17 amostras foram selecionadas para serem testadas, respeitando os protocolos da STBBL, os quais determinam que para a realização de PCR para detecção de material masculino nas amostras relacionadas à violência sexual devemos selecionar o mínimo de amostras por caso, considerando a quantidade de DNA determinada na etapa de quantificação [31, 32].

A figura 5 mostra um diagrama esquemático dos fluoróforos e dos tamanhos dos produtos de PCR dos locos presentes no *Kit PowerPlex® Y23* (Promega® Corporation, Madison, WI) [33].

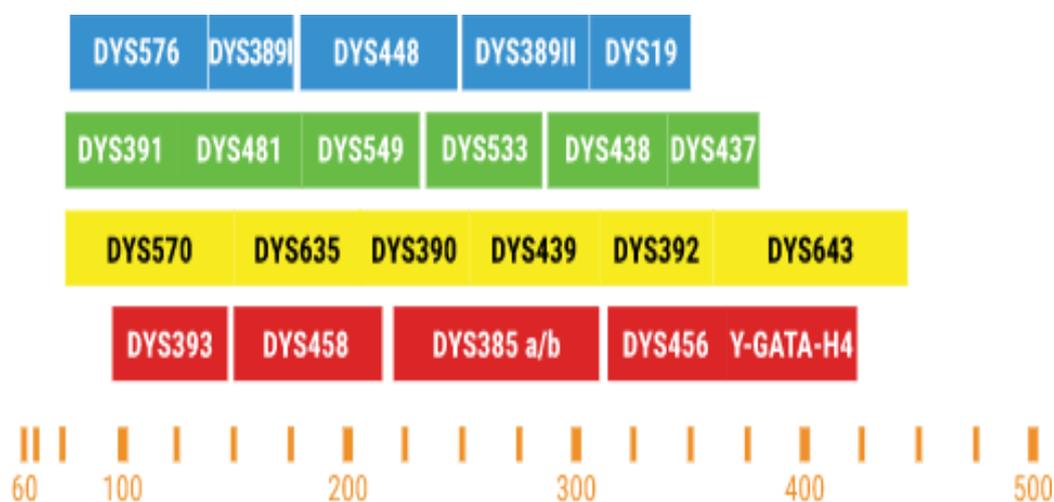


Figura 5: Diagrama esquemático dos locos presentes no *Kit PowerPlex® Y23* e seus respectivos tamanhos em pb. Os locos são apresentados em quatro cores diferentes as quais representam os fluoróforos: Fluoresceína (azul), JOE (verde), TMR-ET (amarelo) e CXR-ET (vermelho). O padrão interno de corrida é detectado no canal alaranjado e é marcado com o WEN (*WEN Internal Lane Standard 500*) [34].

Eletroforese Capilar

A eletroforese capilar foi realizada utilizando o analisador genético *3500 XL Genetic Analyser* (*Thermo Fisher Scientific, Oyster Point, CA*) com 24 capilares de 36 cm de comprimento e matriz de separação (polímero POP-4™ da *Thermo Fisher Scientific, Oyster Point, CA*). O software utilizado para obtenção dos dados brutos da corrida eletroforética foi o *3500 Data Collection Software*.

Análise dos resultados - softwares utilizados

Os dados obtidos a partir da eletroforese capilar, foram analisados no *GeneMapper™ ID-X Software* (*Thermo Fisher Scientific, Oyster Point, CA*). Estes dados “brutos” são convertidos em uma sequência numérica em forma de picos representados em um eletroferograma (perfil genético). Na STBBL, os perfis genéticos devem atender parâmetros de qualidade para serem validados como resultados aceitáveis. Esses parâmetros são os limiares analíticos e o *PHR* (*peak height ratio*). O limiar analítico indica que as fluorescências emitidas acima do valor estabelecido são, de fato, picos válidos e não ruídos de base do equipamento, já o *PHR* é a máxima diferença permitida entre as alturas dos picos dos alelos para um loco heterozigoto. Os parâmetros de qualidade estabelecidos na STBBL são: (i) limiar analítico o valor de 200 *RFUs* (unidades relativas de fluorescência) e (ii) o *PHR* de 60%. Os perfis genéticos obtidos nos permitem, em alguns casos, identificar um agressor sexual e, em outros, excluir indivíduo (s) suspeito (s) como autor (es) do crime em questão. Quando os resultados indicam a inclusão do investigado na amostra coletada da vítima, utilizamos cálculos estatísticos para agregar valor matemático aos resultados obtidos [31]. Para a realização dos cálculos estatísticos que envolvem as evidências genéticas, o *software* disponível na STBBL é o *LRmix Studio* [20].

Testagem das amostras provenientes de violência sexual utilizando SMP

A STBBL - PCMG possui, para realização do SMP, o equipamento *MiSeq[®] FGx* da *Illumina* que utiliza o *ForenSeq[™] DNA Signature Prep Kit* da *Verogen[®]* o qual possui dois *Mixes* de *primers* (*Mix A* e *Mix B*). O *Mix A* possui *primers* para 27 STRs autossômicos, 7 X STRs, 24 Y STRs e 94 SNPs de identidade. O *Mix B* possui *primers* para os mesmos *locos* do *Mix A* somados a *primers* para 54 SNPs de ancestralidade, 22 SNPs de informação fenotípica e 2 SNPs com ambas características (ancestralidade e fenotípica) [32]. Para o sequenciamento das amostras deste estudo foi escolhido o *Mix B* por possibilitar a obtenção de maior quantidade de dados, porém, os dados de ancestralidade e fenotípicos não farão parte do escopo deste trabalho. O sequenciamento das amostras foi realizado no *MiSeq FGx[™] sequencing instrument* com o *MiSeq FGx[™] Reagent Kit*. A figura 6 mostra as etapas do processo de análise usando o *MiSeq FGx[™] Forensic Genomic Solution*.

Após a extração e quantificação do DNA das amostras, elas foram submetidas à primeira PCR com o objetivo de acrescentar “etiquetas” ou *tags* (sequências que contém entre 4 e 8 nucleotídeos) em regiões específicas do DNA, sendo que a quantidade de DNA inicial recomendada é de 1 ng [18]. A segunda reação de PCR consiste na adição de *indexes* ou adaptadores (i5 e i7 *Illumina* adaptadores), com comprimento de 120 pb, aos amplicons gerados na primeira PCR. Esses adaptadores são ligados aos *tags* que haviam sido adicionados na etapa anterior e permitem que as amostras que compõem a biblioteca se liguem às sequências complementares (oligonucleotídeos) na *flowcell* durante a geração dos *clusters* [18,35].

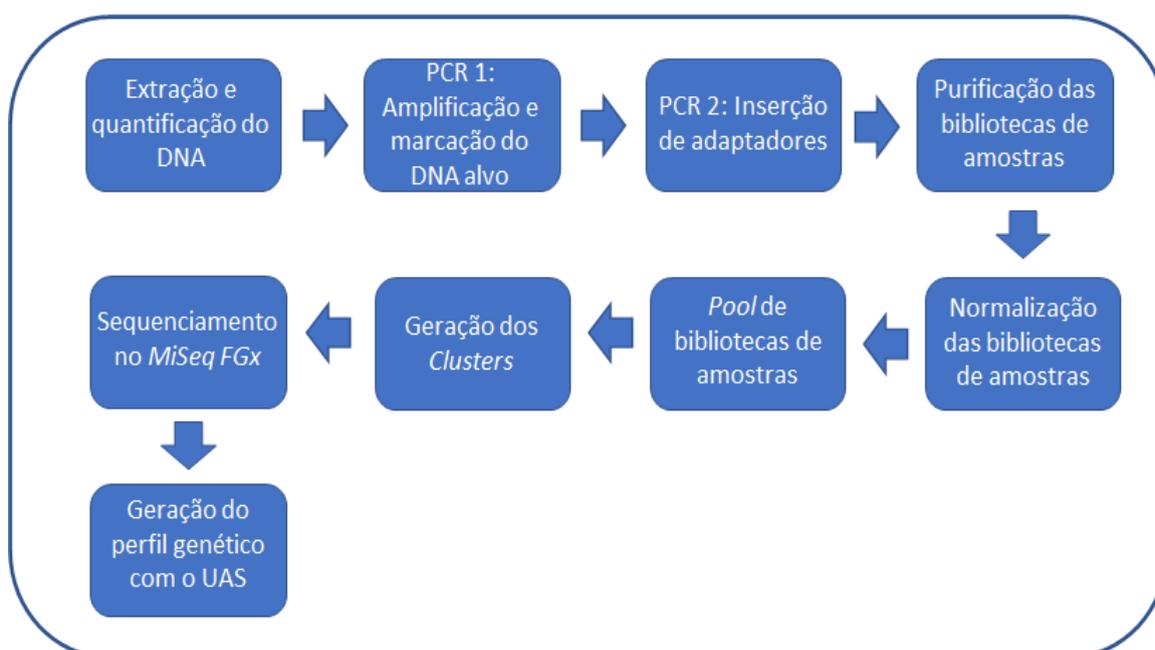


Figura 6: Etapas que envolvem o sequenciamento de amostras utilizando o equipamento *MiSeq FGx[™]* da *Illumina*.

As bibliotecas de amostras formadas a partir das duas PCRs são submetidas a sucessivas lavagens para remover componentes da reação, tais como, dímeros de *indexes* e pequenos fragmentos de DNA, usando *SPB (Sample Purification Beads)* que são *beads* (esferas) magnéticas [18]. Para o processo de normalização das bibliotecas, foram utilizadas *beads* magnéticas que integram o kit *ForenSeq Signature Prep*. A etapa de normalização tem como objetivo fazer com que as bibliotecas estejam igualmente representadas na solução ao serem submetidas à reação de sequenciamento [18,35]. Após o processo de normalização é feito o preparo do *pool* de bibliotecas que utiliza volumes iguais da biblioteca normalizada para criar um *pool* que será sequenciado ao mesmo tempo na *flowcell* [18]. O processo na *flowcell* se inicia com a formação dos *clusters* por amplificação em “ponte” e, posteriormente, ocorre o sequenciamento massivo em paralelo de cada *cluster*. O sequenciamento é realizado no *MiSeq Desktop Sequencer (Illumina)* com o *MiSeq FGx Reagent Kit (Illumina)* seguindo os protocolos recomendados pelo fabricante [18]. No sequenciamento por síntese (método utilizado pelo *MiSeq Desktop Sequencer – Illumina*), todos os quatro nucleotídeos marcados com fluoróforos estão presentes na reação e apenas um nucleotídeo é adicionado na fita crescente de DNA em todos os *clusters*. Os nucleotídeos são reversivelmente bloqueados, o que impede a incorporação de mais de um nucleotídeo ao mesmo tempo. Quando um nucleotídeo é incorporado, um sinal fluorescente é emitido, é realizada a leitura do sinal e em seguida a terminação que bloqueia o nucleotídeo é clivada, permitindo a ligação da próxima base. Após a reação de sequenciamento, a análise dos dados é realizada com o *ForenSeq Universal Analysis Software (UAS) – Verogen®*. O *software* gera os resultados em forma de planilhas e gráficos de *reads* que deverão passar por um processo de checagem. Esta etapa consiste em verificar quais resultados podem ser validados, se o número de *reads* e verificando os *thresholds* analíticos e de interpretação.

Resultados

Extração e quantificação

Todas as 30 amostras foram submetidas à extração de DNA, variando a metodologia de extração de acordo com a natureza da amostra. Todos os extratos obtidos foram quantificados com o equipamento *QuantStudio Real Time PCR® (Thermo Fisher Scientific, Oyster Point, CA)* [13] associado ao *Kit* de quantificação *Quantifiler® Trio DNA Quantification Kit (Thermo Fisher Scientific, Oyster Point, CA)* [14]. A tabela 6 apresenta os dados relativos à quantidade de DNA autossômico (DNA total), de DNA masculino (obtidos a partir da amplificação de locos do cromossomo Y) e forneceu a relação de DNA masculino / feminino, o qual é um importante

parâmetro nas análises de casos onde ocorrem misturas de material genético (vítima do sexo feminino e agressor do sexo masculino). O processo de quantificação não identificou a presença de inibidores de reação de PCR nas amostras testadas.

Tabela 6: Resultados obtidos a partir da quantificação dos extratos de DNA das amostras analisadas.

Amostra	Quantidade de DNA autossômico (ng/ul)	Quantidade de DNA Y (ng/ul)	Razão DNA Masculino/Feminino	Índice de Degradação
1	51,03	0,002	1:20.029	7,47
2	17,9	11,23	indeterminado	2,06
3	4,36	0,006	1:725	1,64
4	41,08	0,0002	1:224.571	1,54
5	0,4	0,001	1:399	1,08
6	2	0,002	1:999	1,86
7	0,29	0,02	1:14	2,78
8	0,86	0,02	1:46	2,92
9	0,26	0,14	indeterminado	2,59
10	0,02	0,01	indeterminado	3,62
11	0,003	indeterminado	indeterminado	3,57
12	11,86	indeterminado	indeterminado	4,58
13	29,05	21,3	indeterminado	2,54
14	9,84	4,55	1:1	1,49
15	91,62	0,28	1:326	1,56
16	0,14	0,02	1:4	1,24
17	16,88	0,01	1:1.122	2,1
18	7,81	indeterminado	indeterminado	1,37
19	9,3	7,24	indeterminado	1,36
20	2,8	0,01	1:279	1,42
21	83,88	0,05	1:1.626	1,46
22	indeterminado	indeterminado	indeterminado	-
23	0,03	0,001	1:21	5,27
24	67,18	0,01	1:4.466	2,02
25	79,61	0,04	1:1.955	1,06
26	33,25	38,3	indeterminado	1,77
27	115,95	indeterminado	indeterminado	3,78
28	0,86	indeterminado	indeterminado	2,18
29	5,88	0,07	1:83	2,63
30	1,79	2,2	indeterminado	0,8
CP	0,2	0,16	indeterminado	1,47
CN	0,002	0,0003	1:4.388	-

Genotipagem utilizando a metodologia de PCR – EC

PCR utilizando o Kit *PowerPlex® Fusion 6C* (*Promega® Corporation, Madison, WI*) e resultados obtidos na eletroforese capilar

Na análise dos marcadores autossômicos presentes no *kit Powerplex® Fusion 6C*, foram obtidos 25 perfis genéticos, os quais foram divididos em completos e incompletos. Foram considerados perfis genéticos completos aqueles que geraram 23 locos autossômicos (número total de locos autossômicos presentes no kit de amplificação utilizado) e foram considerados perfis genéticos incompletos aqueles que geraram até 22 locos autossômicos ($0 < n^{\circ} \text{ de locos} \leq 22$). A figura 7 apresenta a quantidade de locos autossômicos genotipados em cada amostra analisada.

A análise das amostras utilizando PCR – EC, apresentou como resultado, 16 (dezesesseis) amostras com perfis genéticos completos, 9 (nove) amostras com perfis genéticos incompletos e 7 (sete) amostras as quais não geraram perfis genéticos.

Relação entre o número de locos autossômicos obtidos e o total de locos analisados na EC

Após análise dos resultados obtidos para marcadores autossômicos, verificou-se que em 7 amostras (21,9%) não foram obtidos perfis genéticos, em 9 amostras (28,1%) foram obtidos perfis genéticos que continham entre 1 e 22 locos genotipados e em 16 amostras (50%) foram obtidos perfis genéticos completos, ou seja, 23 locos autossômicos em 23 locos autossômicos analisados.

PCR utilizando o Kit *PowerPlex® Y23* (*Promega® Corporation, Madison, WI*) e resultados obtidos na Eletroforese Capilar

Foram selecionadas 17 amostras para serem submetidas à PCR utilizando o Kit *PowerPlex® Y23* (*Promega® Corporation, Madison, WI*), respeitando os protocolos da STBBL, os quais determinam que para a realização de PCR para detecção de material masculino nas amostras relacionadas à violência sexual devemos selecionar o mínimo de amostras por caso, considerando a quantidade de DNA determinada na etapa de quantificação [31, 32].

Locos genotipados por amostra analisada

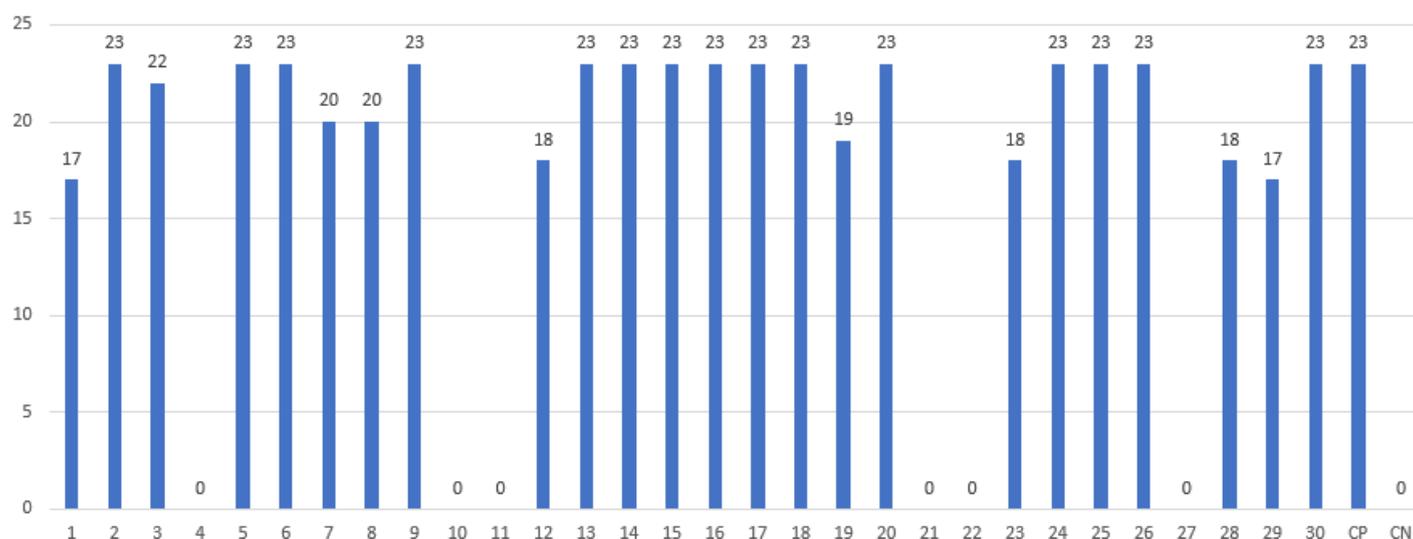


Figura 7: Quantidade de locos autossômicos genotipados nas amostras analisadas obtidos com *Powerplex® Fusion 6C (Promega® Corporation, Madison, WI)*.

A análise das 17 amostras selecionadas (2,3,7,8,9,10,11,13,14,16,19,21,23,24,27,28, e CP) utilizando o kit *PowerPlex® Y23 (Promega® Corporation, Madison, WI)*, seguida de eletroforese capilar, resultou na obtenção de 10 (dez) perfis de haplótipo, ou seja, perfis contendo locos Y STRs. Estes perfis foram separados em completos e incompletos, onde os perfis de haplótipo completos apresentaram 23 locos genotipados (número total de locos analisados no kit *PowerPlex® Y23* da *Promega® Corporation, Madison, WI*) e os perfis de haplótipo incompletos apresentaram até 22 locos Y STRs ($0 < n^{\circ} \text{ de locos} \leq 22$). Portanto, a análise das 17 (dezessete) amostras selecionadas utilizando o Kit *PowerPlex® Y23 (Promega® Corporation, Madison, WI)* resultou na obtenção de 6 (35,2%) perfis de haplótipos completos, 4 (23,6%) perfis de haplótipo incompletos ($0 < n^{\circ} \text{ de locos} \leq 22$) e em sete 7 (41,2%) amostras não foram genotipados nenhum loco. A figura 8 apresenta a quantidade de locos Y STRs genotipados em cada amostra analisada.

Resultados obtidos a partir da eletroforese capilar considerando a presença ou ausência de mistura de material genético nas amostras

A eletroforese capilar detectou, em 21 amostras, perfis genéticos únicos (material genético de apenas um contribuinte), em 4 amostras foram detectados perfis de mistura de material genético de mais de um contribuinte (indivíduo) e em 7 amostras não foram detectados perfis genéticos.



Figura 8: Quantidade de locos Y STRs genotipados por amostra analisada.

As amostras que apresentaram perfis genéticos de apenas um contribuinte são representadas por locos com um pico (indivíduo homocigoto) ou dois picos (indivíduo heterocigoto). Podem ser amostras de referência de vítimas, de investigados ou amostras questionadas, como, por exemplo, vestes e suaves vaginais/perianais. As amostras questionadas também podem apresentar perfis de mistura de material genético com mais de um contribuinte (vítima e agressor sexual). A figura 9 apresenta um exemplo de perfil genético com um contribuinte. Nesta figura é possível observar o padrão de balanceamento dos picos heterocigotos (alturas proporcionais entre dois picos de um mesmo loco), a amelogenina (AMEL) XX indicando tratar-se de um indivíduo do sexo feminino e a qualidade geral do eletroferograma. O resultado mostrado foi obtido com a utilização do *software GeneMapper™ ID-X Software (Thermo Fisher Scientific, Oyster Point, CA)* a partir da metodologia PCR - EC.

O uso da metodologia de PCR – EC detectou resultados com perfil de mistura de material genético de mais de um contribuinte em 4 (quatro) amostras. A figura 10 apresenta um exemplo de perfil de mistura de material genético de mais de um contribuinte.

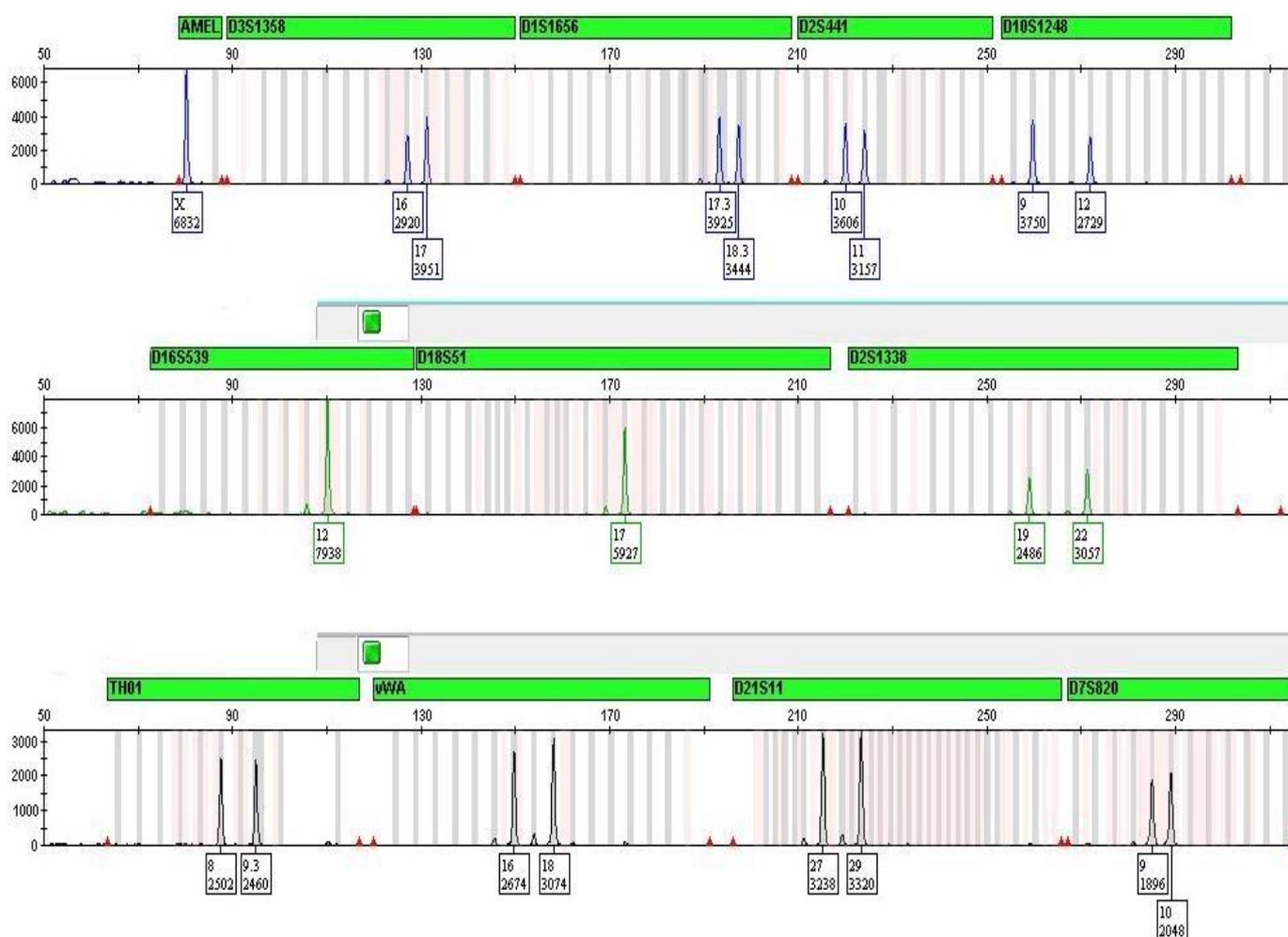


Figura 9: Parte de um eletroferograma (representação gráfica do perfil genético) com material genético de apenas um contribuinte. Nesta figura, os picos são representados em três cores diferentes representados por diferentes fluoróforos presentes na reação de PCR (azul, verde e preto).

Figura 10: A*- Parte de um eletroferograma com material genético de dois indivíduos. Observar a presença de três ou quatro picos, indicando mais de um contribuinte na amostra (setas). B** - Representação esquemática da forma do pico da amelogenina quando há mistura de material genético masculino e feminino na proporção 1:1. A amelogenina XY indica a presença de indivíduo do sexo masculino, porém, a desproporção entre o X e o Y, onde se observa 3 X em relação a 1 Y, indica que a quantidade de material feminino (XX) é aproximadamente a mesma quantidade de material masculino (XY). Fonte*STBBL, 2020; ***Forensic DNA Typing, John M. Butler, 2ª edição, pag: 114.*

Figura 10 A:

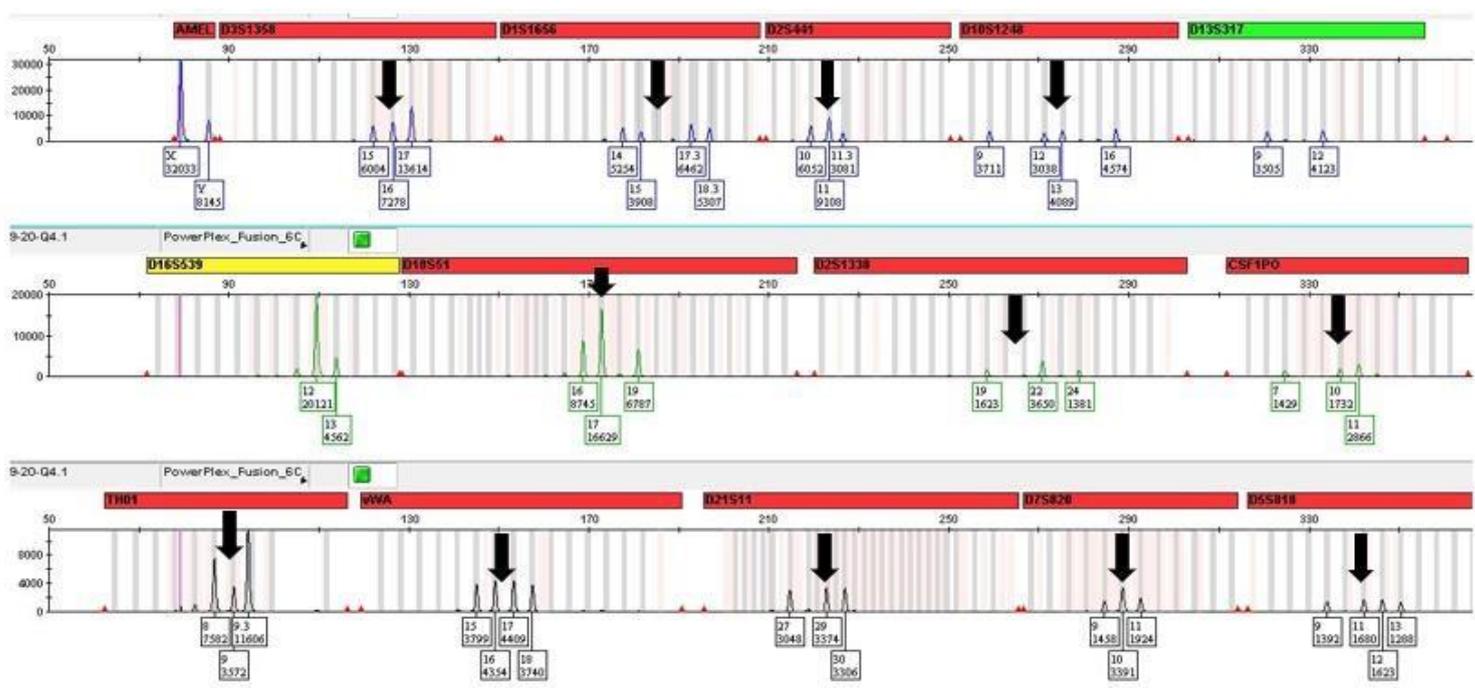
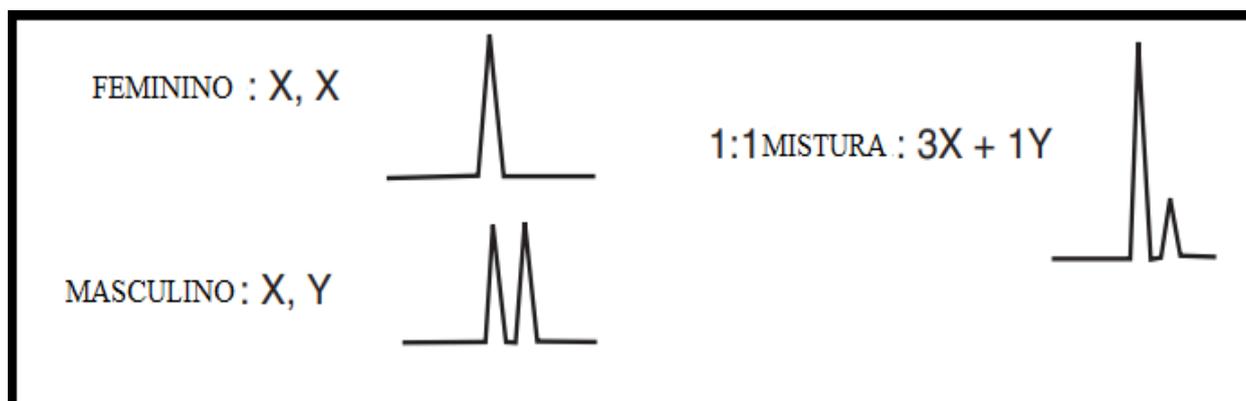


Figura 10 B:



Além dos resultados apresentados (material genético de apenas um contribuinte e mistura de material genético de mais de um contribuinte) também foram observadas amostras as quais não geraram perfis genéticos. As figuras 11 e 12 apresentam uma síntese dos resultados obtidos na análise de STRs autossômicos e Y STRs na metodologia de PCR – EC.

Genotipagem usando a metodologia de Sequenciamento Massivo em Paralelo

Após a análise com a metodologia PCR – EC, as amostras foram submetidas à metodologia de SMP utilizando o *ForenSeq™ DNA Signature Prep Kit* da Verogen®, o *MiSeq™ FGx Reagent Kit*, o *MiSeq™ FGx Sequencing Instrument* e o *ForenSeq™ Universal Analysis Software*.

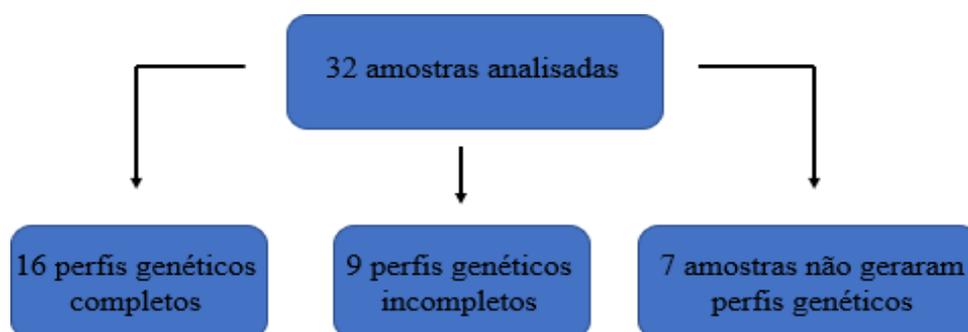


Figura 11: Resultados observados (perfis genéticos completos, perfis genéticos incompletos e ausência de perfil genético) na análise de locos autossômicos (30 amostras e 2 controles).

Resultados obtidos na etapa de sequenciamento para os locos autossômicos

Para os locos autossômicos, utilizando o *ForenSeq™ Universal Analysis Software*, verificou-se que, em 3 amostras (9,4%) não foram obtidos perfis genéticos, em 2 amostras (6,2%) foram obtidos perfis genéticos que continham entre 5 e 10 locos tipados, em 1 amostra (3,1%) foi obtido perfil genético que continham 23 locos genotipados e em 26 amostras (81,3%) foram obtidos perfis genéticos completos, ou seja, 27 locos autossômicos em 27 locos autossômicos analisados. A tabela 7 mostra aquelas amostras que apresentaram perfis genéticos incompletos na PCR – EC e quando submetidas ao SMP apresentaram perfis genéticos completos.

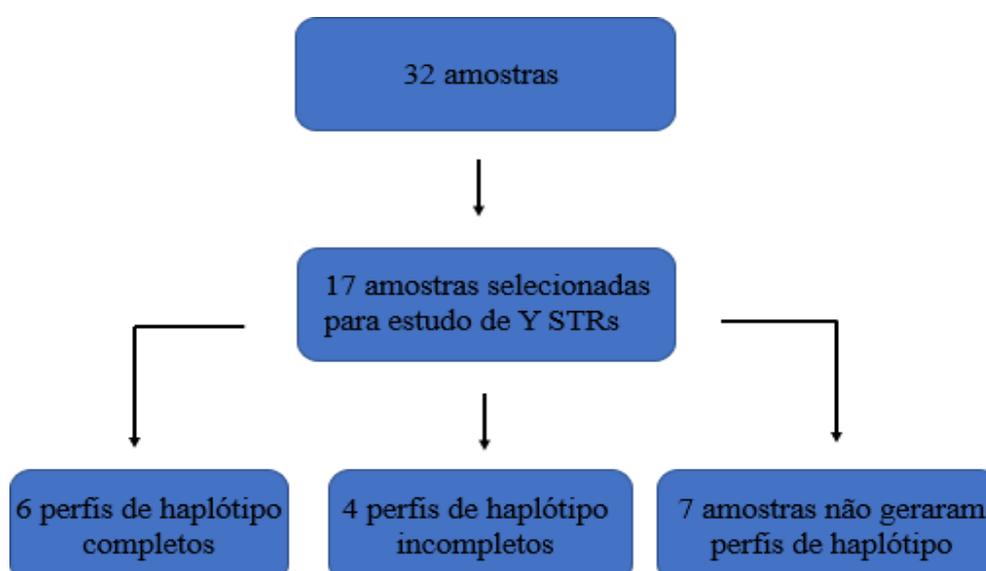


Figura 12: Resultados observados (perfis de haplótipo completos, perfis de haplótipo incompletos e ausência de perfil de haplótipo) na análise de locos Y STRs (30 amostras e 2 controles).

Tabela 7: Amostras que não geraram perfis genéticos ou que apresentaram perfis genéticos incompletos, para os locos autossômicos, com a metodologia PCR – EC e apresentaram perfis genéticos completos (27 locos gênicos) a partir da utilização da metodologia de SMP.

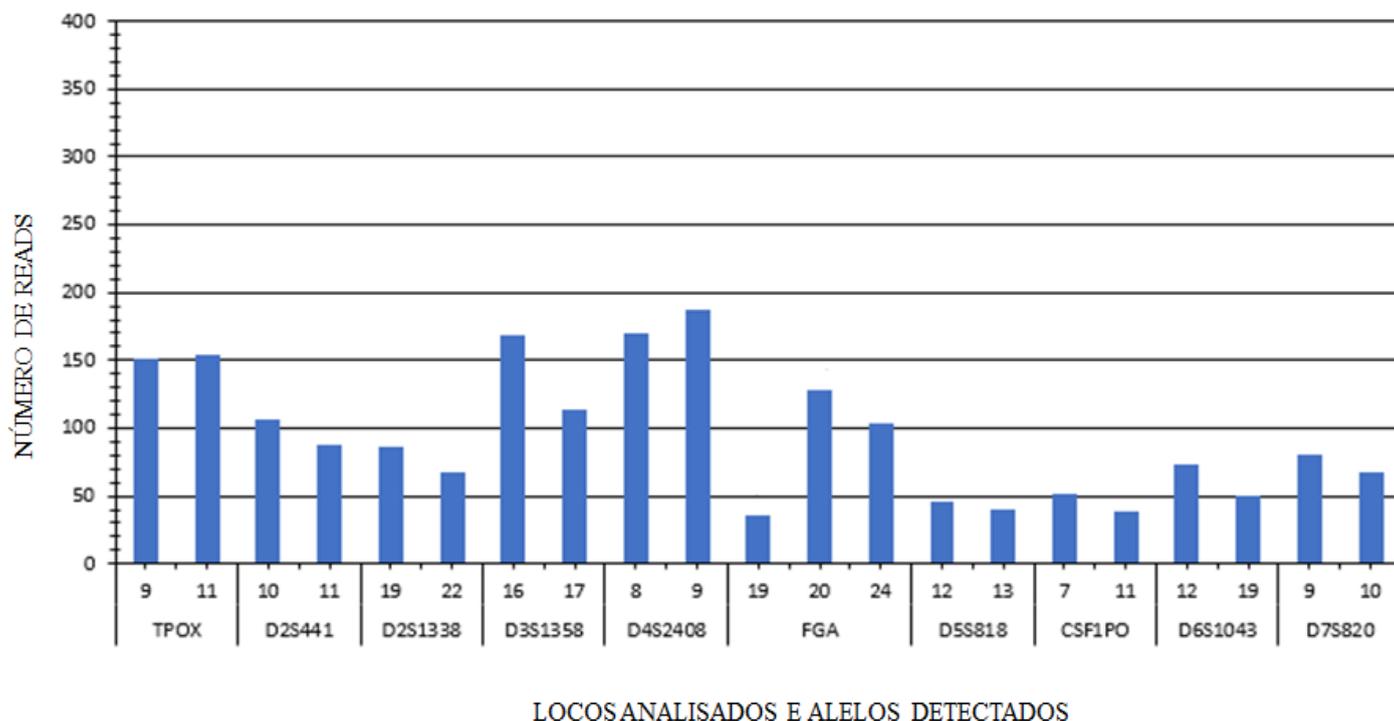
Amostra	locos autossômicos obtidos	Amostra	locos autossômicos obtidos
	PCR – EC		SMP
3	22	3	27
4	0	4	27
7	20	7	27
8	20	8	27
12	18	12	27
19	19	19	27
21	0	21	27
27	0	27	27
28	18	28	27
29	17	29	27

Resultados obtidos a partir do SMP considerando o número de contribuintes com material genético nas amostras

O sequenciamento massivo em paralelo detectou em 25 amostras perfis genéticos únicos (material genético de apenas um indivíduo), em 4 amostras foram detectados perfis de mistura de material genético de mais de um indivíduo e em 3 amostras não foram detectados perfis genéticos. Esses resultados foram compatíveis (compatibilidade em relação ao número de indivíduos nas amostras) com os resultados observados na PCR – EC.

Assim como observado na análise utilizando PCR – EC, a metodologia de SMP também apresentou amostras com perfis genéticos de apenas um contribuinte, as quais eram compostas por amostras de referência de vítimas e/ou investigados e por amostras questionadas. Também foram observadas amostras questionadas com perfis de mistura de material genético, com pelo menos dois contribuintes (vítima e agressor sexual). A figura 13 apresenta um exemplo de perfil genético com um contribuinte e a detecção de *stutter* no loco FGA.

Figura 13: Parte de um gráfico com a distribuição do número de *reads* e os alelos detectados nos locos analisados. Este gráfico (representação do perfil genético) foi gerado pelo *UAS* e mostra resultado quando há material genético de apenas um contribuinte. Aqui estão representados os seguintes locos STRs: TPOX, D2S441, D2S1338, D3S1358, D4S2408, FGA, D5S818, CSF1PO, D6S1043 e D7S820.



Além das amostras com contribuintes únicos, existem aquelas em que podemos encontrar mais de um indivíduo contribuindo com seu material genético. Neste estudo utilizando SMP foram encontrados resultados com perfil de mistura de material genético de mais de um contribuinte em 4 (quatro) amostras. A figura 14 apresenta um exemplo de perfil genético com mistura de material genético de mais de um contribuinte. Somam-se a estes resultados apresentados (material genético de apenas um contribuinte e mistura de material genético de mais de um contribuinte), resultados em que não foram observados perfis genéticos, e nestes casos, o *ForenSeq™ Universal Analysis Software* não gera o gráfico com a distribuição do número de *reads*.

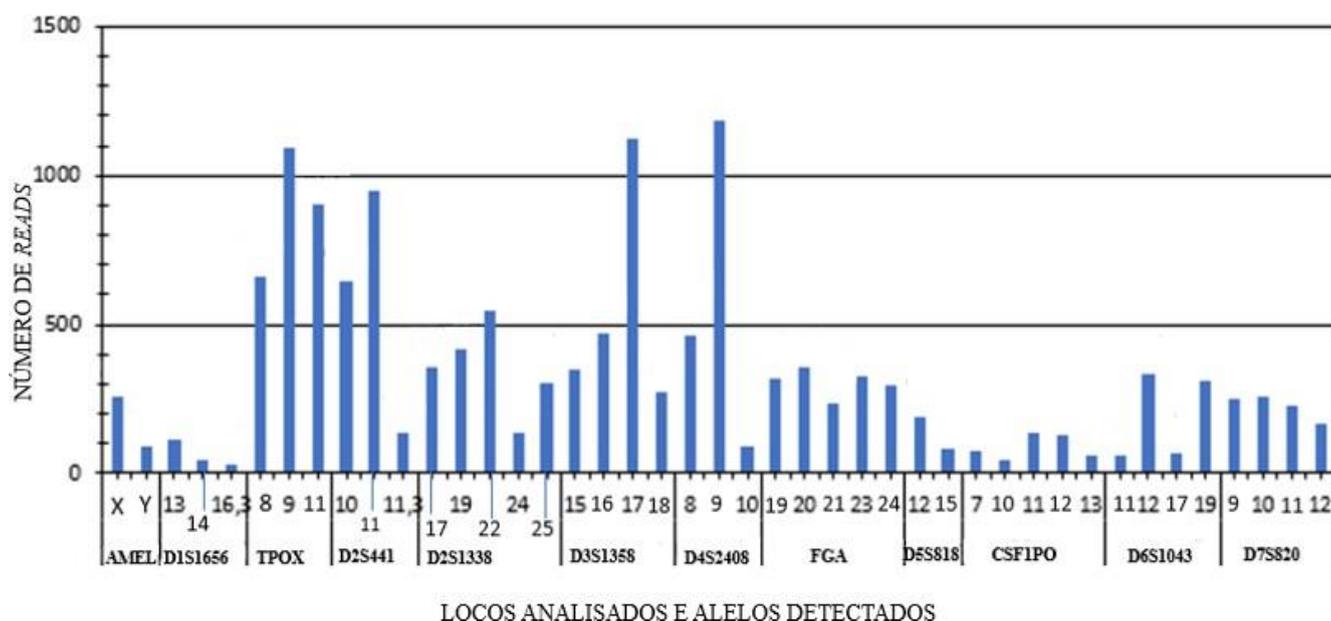


Figura 14: Parte de um gráfico com a distribuição do número de *reads* e os alelos detectados nos locos analisados. Este gráfico (representação do perfil genético) foi gerado pelo *UAS* e mostra resultado quando há material genético de mais de um contribuinte.

STRs autossômicos: ganhos obtidos com a utilização da metodologia de SMP

A partir da utilização da metodologia de SMP, ao se analisar os STRs autossômicos, foram obtidos ganhos em relação à metodologia de PCR – EC. Das amostras analisadas, apenas 3 foram consideradas sem sucesso na obtenção de perfis genéticos, 2 amostras apresentaram de 1 a 22 locos genotipados, 1 amostra apresentou 23 locos genotipados e 26 amostras apresentaram 27 locos genotipados, conforme apresentado na figura 15. A metodologia de PCR – EC genotipou em média 16 locos gênicos enquanto a metodologia de SMP genotipou 23 locos gênicos, em média. Quanto maior o número de locos STRs genotipados maior será o poder de discriminação da análise efetuada. O aumento dos locos genotipados em amostras forenses geram *RMP* (*random match probabilities*) e/ou *LRs* (*likelihood ratios*) mais robustos. Na estatística forense, o *RMP* (*random match probabilities*) é a probabilidade de um *match* ocorrer por coincidência enquanto o *LR* (*likelihood ratio*) representa a força da evidência extraída do DNA [36].

Análise comparativa entre os resultados obtidos a partir das metodologias de PCR – EC e SMP, considerando os locos STRs autossômicos.

Os dados gerados nesse estudo permitiram a comparação entre os métodos de PCR - EC e SMP. A tabela 8 mostra o número de perfis genéticos obtidos em ambas metodologias, considerando tratarem-se de perfis genéticos completos, incompletos e ausência de perfil genético.

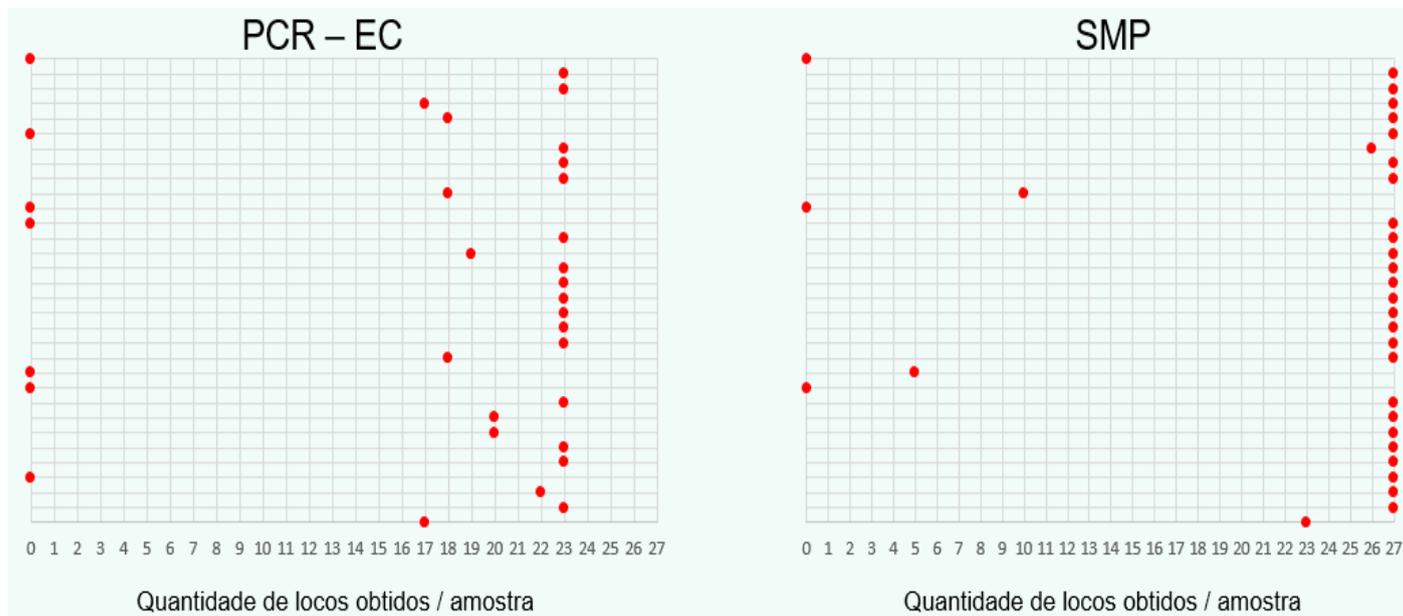


Figura 15: Representação do número de locos obtidos por amostra analisada, utilizando as metodologias de PCR – EC e SMP. Cada ponto vermelho representa uma amostra testada.

Os resultados comparativos entre as metodologias, que consideram a complexidade do caso (alta ou baixa) e a complexidade da análise (simples ou complexa), estão descritos nas tabelas 9, 10, 11 e 12. A amostra 22, apresentou resultado “indeterminado” para todos os parâmetros analisados na etapa da quantificação e não gerou perfil genético em ambas as metodologias. Aquelas amostras em que a razão entre os materiais masculino e feminino na quantificação foi “indeterminado” foram consideradas como amostras de perfil genético de um contribuinte.

Tabela 8: Comparação entre as metodologias de PCR - EC e SMP quanto aos perfis genéticos autossômicos obtidos (completos, incompletos e/ou ausência de perfil genético).

	Perfis Genéticos Completos	Perfis Genéticos incompletos	Ausência de perfil genético	Total
Metodologia PCR – EC	16	9	7	32
Metodologia SMP	26	3	3	32

Os resultados obtidos para os locos STRs autossômicos nas metodologias de EC e SMP estão compilados, de forma comparativa, na figura 16.

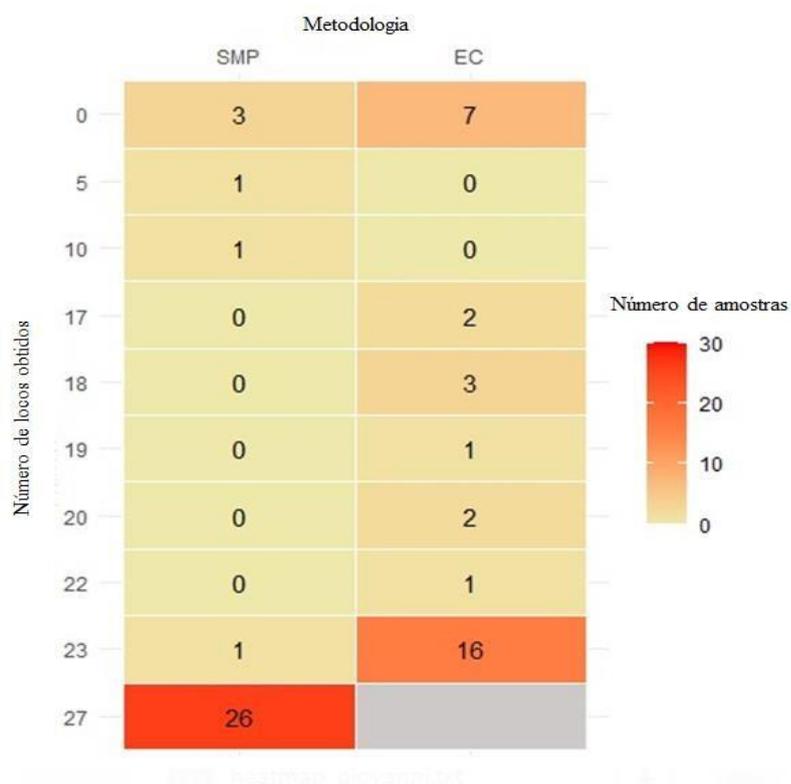


Figura 16: Mapa de calor que representa a comparação entre os resultados obtidos, para os locos STRs autossômicos, nas metodologias de PCR - EC e SMP. *27 é o máximo de locos autossômicos possíveis para o *ForenSeq™ DNA Signature Prep Kit* da Verogen®. **23 é o máximo de locos autossômicos possíveis para o *PowerPlex® Fusion 6C System* (Promega® Corporation, Madison, WI).

Tabela 9: Comparação entre os resultados obtidos nas metodologias utilizadas, para os locos autossômicos estudados, considerando a complexidade do caso (baixa) e a complexidade da análise (simples).

Complexidade baixa / Análise Simples

Amostra	PCR - EC	SMP
	Resultado	Resultado
9	Mistura de material genético	Mistura de material genético
18	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
19	Perfil genético incompleto**, sem mistura	Perfil genético completo*, sem mistura
30	Mistura de material genético	Mistura de material genético

*Completo: obtenção de perfil genético contendo todos os locos; **Incompleto: obtenção de perfil genético sem a genotipagem de todos os locos.

Tabela 10: Comparação entre os resultados obtidos nas metodologias utilizadas, para os locos autossômicos estudados, considerando a complexidade do caso (alta) e a complexidade da análise (simples).

Complexidade alta / Análise Simples		
EC		SMP
Amostra	Resultado	Resultado
2	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
10	Ausência de perfil genético	Ausência de perfil genético
11	Ausência de perfil genético	Perfil genético incompleto*, sem mistura
12	Perfil genético incompleto*, sem mistura	Perfil genético completo*, sem mistura
13	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
26	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
27	Ausência de perfil genético	Perfil genético completo*, sem mistura
28	Perfil genético incompleto*, sem mistura	Perfil genético completo*, sem mistura
31 (CP)	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura

*Completo: obtenção de perfil genético contendo todos os locos; **Incompleto: obtenção de perfil genético sem a genotipagem de todos os locos.

Tabela 11: Comparação entre os resultados obtidos nas metodologias utilizadas, para os locos autossômicos estudados, considerando a complexidade do caso (baixa) e a complexidade da análise (alta).

Complexidade baixa / Análise complexa		
EC		SMP
Amostra	Resultado	Resultado
14	Mistura de material genético	Mistura de material genético
20	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
21	Ausência de perfil genético	Perfil genético completo*, sem mistura
25	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura

*Completo: obtenção de perfil genético contendo todos os locos; **Incompleto: obtenção de perfil genético sem a genotipagem de todos os locos.

Tabela 12: Comparação entre os resultados obtidos nas metodologias utilizadas, para os locos autossômicos estudados, considerando a complexidade do caso (alta) e a complexidade da análise (alta).

Complexidade alta / Análise complexa		
	EC	SMP
Amostra	Resultado	Resultado
1	Perfil genético incompleto**, sem mistura	Perfil genético incompleto*, sem mistura
3	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
4	Ausência de perfil genético	Perfil genético completo*, sem mistura
5	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
6	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
7	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
8	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
15	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
16	Mistura de material genético	Mistura de material genético
17	Perfil genético completo*, sem mistura	Perfil genético completo*, sem mistura
23	Perfil genético incompleto**, sem mistura	Perfil genético incompleto**, sem mistura
24	Ausência de perfil genético	Perfil genético completo*, sem mistura
29	Perfil genético incompleto**, sem mistura	Perfil genético incompleto**, sem mistura
32 (CN)	Ausência de perfil genético	Ausência de perfil genético

*Completo: obtenção de perfil genético contendo todos os locos; **Incompleto: obtenção de perfil genético sem a genotipagem de todos os locos.

A quantidade de locos obtidos por amostra na análise dos Y STRs, a partir das metodologias de PCR - EC e de SMP são apresentados na tabela 13. Para a metodologia SMP todas as amostras foram submetidas ao sequenciamento e para a metodologia de PCR - EC apenas as amostras 2,3,7,8,9,10,11,13,14,16,19,21,23,24,27,28 e 31 foram testadas.

Y STRs obtidos com a utilização da metodologia de SMP

Os resultados obtidos para os locos Y STRs na reação de sequenciamento foram: em 18 amostras (56,3%) não foram obtidos perfis de haplótipo, em 13 amostras (40,6%) foram obtidos perfis de haplótipo que continham entre 1 e 23 locos (perfis de haplótipo incompletos) e em 1 amostra (3,1%) foi obtido perfil de haplótipo completo, ou seja, 24 locos Y STRs em 24 locos analisados. Estes dados são apresentados na tabela 14 de modo que se possa comparar os resultados obtidos a partir da metodologia de PCR - EC e SMP.

Tabela 13: Comparação entre os resultados obtidos, para os locos Y STRs, nas amostras testadas em ambas metodologias de PCR - EC e SMP.

Amostras analisadas	locos Y STRs obtidos	locos Y STRs obtidos
	PCR – EC	SMP
2	23	22
3	0	0
7	21	7
8	21	4
9	23	18
10	16	0
11	0	0
13	23	23
14	23	23
16	23	19
19	12	23
21	0	0
23	0	0
24	0	0
27	0	0
28	0	0
31	23	24

Tabela 14: Comparação entre a quantidade de perfis de haplótipo Y STRs, completos e incompletos, obtidos a partir das metodologias de PCR - EC e SMP. Também são apresentadas aquelas amostras as quais não foram obtidos perfis de haplótipo. Das 18 amostras, as quais, os perfis de haplótipo não foram obtidos (PHNO), 3 eram pertencentes à vítima do sexo feminino e 1 era o controle negativo, portanto, essas amostras não foram contabilizadas para o cálculo apresentado nesta tabela 12.

Metodologia	Amostras testadas	PHC*	PHI*	PHNO*	% Amostras PHC	% Amostras PHI	% Amostras PHNO
PCR - EC	17	6	4	7	35,3	23,5	41,2
SMP	32	1	13	18	3,6	46,4	50,0

*PHC: Perfis de haplótipo completos; PHI: Perfis de haplótipo incompletos; PHNO: perfis de haplótipo não obtidos.

Locos X STRs obtidos com a utilização da metodologia de SMP

O *ForenSeq™ DNA Signature Prep Kit* da *Verogen®* possui um conjunto de primers específicos para amplificação de 7 locos X STRs e os resultados obtidos são descritos a seguir. Das amostras analisadas 1 amostra (3,1%) apresentou um loco genotipado, 1 amostra (3,1%) apresentou cinco locos genotipados, 7 amostras (21,9%) apresentaram seis locos genotipados, 20 amostras (62,5%) apresentaram sete locos genotipados e 3 amostras (9,4%) não apresentaram nenhum loco genotipado, sendo que uma delas era o controle negativo. Os resultados para os locos referentes ao cromossomo X estão representados na tabela 15.

Tabela 15: Quantidade de locos X STRs genotipados nas amostras analisadas, com a utilização da metodologia SMP.

Amostras	locos X STRs obtidos	Amostras	locos X STRs obtidos
1	6	17	7
2	7	18	7
3	7	19	7
4	7	20	7
5	6	21	7
6	7	22	0
7	7	23	5
8	7	24	7
9	6	25	6
10	0	26	7
11	1	27	6
12	7	28	7
13	6	29	7
14	7	30	6
15	7	31	7
16	7	32	0*

*Amostra 32: controle negativo. Continuação da tabela 15.

SNPs de identidade obtidos com o *ForenSeq™ DNA Signature Prep Kit* da *Verogen®*

A partir da implementação da metodologia de *SMP* com o *ForenSeq™ DNA Signature Prep Kit* da *Verogen®* que possui um conjunto de primers específicos para amplificação de 94 SNPs de identidade foram obtidos os seguintes resultados: Das amostras analisadas, 1 amostra (3,1%) apresentou cinco SNPs genotipados, 1 amostra (3,1%) apresentou 21 SNPs genotipados, 5 amostras (15,6%) apresentaram entre 73 SNPs e 90 SNPs genotipados, 21 amostras (65,7%) apresentaram entre 91 SNPs e 94 SNPs genotipados e 4 amostras (12,5%) não apresentaram

nenhum SNP genotipado, sendo que uma delas era o controle negativo. Os resultados para os SNPs de identidade estão representados na tabela 16.

Tabela 16: Quantidade de SNPs de identidade genotipados nas amostras analisadas, com a utilização da metodologia SMP.

Amostras analisadas	iSNPs obtidos	Amostras analisadas	iSNPs obtidos
1	73	17	94
2	94	18	93
3	94	19	93
4	94	20	94
5	94	21	0
6	94	22	0
7	93	23	21
8	94	24	94
9	93	25	89
10	0	26	94
11	5	27	90
12	94	28	94
13	92	29	94
14	94	30	90
15	94	31	89
16	94	32	0*

*Amostra 32: controle negativo

Bancos de dados congelados de produtos intermediários gerados a partir do SMP

As informações geradas a partir do SMP das amostras deste estudo, foram armazenadas na STBBL, na forma de um banco de dados, para que possam ser analisadas em pesquisas futuras. Este banco de dados é composto por arquivos no formato FASTQ, gerado pelo *software* UAS da Verogen®.

Discussões

A implementação de novas tecnologias no laboratório de genética forense da PCMG é de grande importância para auxiliar no combate aos crimes violentos. As amostras originadas de vítimas e de locais de crimes de natureza sexual compõem aproximadamente 70% da casuística da STBBL. A natureza destas amostras, que frequentemente possuem DNA em quantidade e

qualidade insuficientes, desafia a equipe de geneticistas forenses da STBBL. Além disso, a morosidade do processo de análise destas amostras resulta em uma demanda reprimida, a qual torna-se inviável de ser sanada, com a tecnologia tradicional.

Com o intuito de analisar amostras desafiadoras, aumentar a sua produtividade e resolutividade, a Seção Técnica de Biologia Legal, adquiriu o equipamento *MiSeq FGxTM* da *Illumina*. Este equipamento baseia-se na metodologia de Sequenciamento Massivo em Paralelo e a sua implementação na rotina do Laboratório de Genética Forense foi o escopo deste trabalho.

Uma das questões desafiadoras da aquisição e manutenção do equipamento *MiSeq[®] FGx* da *Illumina* (utilizado especificamente para a área forense) e insumos para seu funcionamento, é de cunho logístico. Todos os serviços relacionados a esse equipamento foram terceirizados pela *Illumina[®]* e até o início de 2021 eram de responsabilidade da empresa *Verogen[®]*. A *Verogen[®]* não possuía uma estrutura de manutenção e suporte, no Brasil, que possibilitasse o mesmo suporte fornecido pela *Illumina[®]*, pois, a demanda do *MiSeq[®] FGx*, no Brasil, ainda é exígua, existindo, atualmente apenas 4 equipamentos em todo território nacional. Portanto, no atendimento de manutenções corretivas, onde há a necessidade de reposição de peças, a demora deste processo era um fator limitante no uso dessa tecnologia nos laboratórios de genética forense no Brasil. Em meados do ano de 2021 houve uma parceria entre a *Verogen[®]* e a *Qiagen[®]* com o objetivo de dar suporte ao portfólio de identificação humana, incluindo às soluções de *Next Generation Sequencing*; outro ponto complicador que é bastante relevante é o fato de que dados gerados em versões de *software* descontinuadas não podem ser analisados em versões mais atuais; os custos operacionais também devem ser considerados, pois, de acordo com o artigo conduzido por Jennifer D. Churchill *et. al* (2016) [37], o custo por amostra utilizando o *ForenSeq DNA Signature Prep Kit* da *Verogen[®]* é de aproximadamente USD 84 (oitenta e quatro dólares) quando são sequenciadas 32 amostras simultaneamente, e USD 45 (quarenta e cinco dólares) quando são sequenciadas 96 amostras simultaneamente, gerando dados de até 230 marcadores. Os kits utilizados para corrida com o método de eletroforese capilar, tais como, *Globalfiler Express[®]* (*Thermo Fisher Scientific, Waltham, MA USA*) e *Yfiler Plus* (*Thermo Fisher Scientific*) custam aproximadamente USD 20 (vinte dólares) e USD 32 (trinta e dois dólares) por amostra, respectivamente, sendo obtidos dados de 24 marcadores com o *Globalfiler Express[®]* e 27 marcadores com o *Yfiler[®]*. Levantamentos realizados na STBBL em outubro de 2021 mostraram que o sequenciamento simultâneo de 96 amostras utilizando o *ForenSeq DNA Signature Prep Kit* da *Verogen[®]* tem o custo aproximado de R\$600,00 (seiscentos reais) por amostra. Enquanto que a genotipagem de 96 amostras utilizando o *PowerPlex[®] Fusion 6C System* (*Promega[®] Corporation, Madison, WI*) tem o custo aproximado de R\$120,00 (cento e vinte reais) por amostra. Apesar de apresentar um custo financeiro de análise maior do que a PCR – EC, o uso do SMP justifica-se

pelo grande volume de dados produzidos e pelo aumento da possibilidade de resolução de casos criminais.

Além disso, o equipamento *MiSeq FGx™* da *Illumina®* utilizando o *ForenSeq™ DNA Signature Prep* da *Verogen*, possui a capacidade de analisar centenas de marcadores forenses em uma única reação, em até 96 amostras simultaneamente, o que confere a este equipamento uma enorme vantagem quando comparado à metodologia de PCR – EC, atualmente padronizada na STBBL.

Nas análises de genética forense, principalmente nas relacionadas a crimes de natureza sexual, as amostras com mistura de material genético são desafiadoras, principalmente quando utilizamos a metodologia de PCR – EC, a qual os STRs amplificados (alelos) migram numa matriz (polímero) e são separados de acordo com o tamanho (os STRs menores migram na malha do polímero com maior velocidade, enquanto os STRs maiores migram mais lentamente). Quando esses STRs de indivíduos diferentes (contribuintes diferentes na mistura) possuem o mesmo tamanho, não é possível distinguir a quem pertence o material genético analisado utilizando a PCR – EC. Outra situação desafiadora é quando o alelo do agressor sexual (sendo contribuinte minoritário de uma mistura) se encontra na posição de *stutter* em relação ao alelo da vítima, sendo que a PCR – EC não é capaz de distinguir a origem desse alelo (se é o alelo do agressor ou se é *stutter* do alelo da vítima), pois, ambos possuem o mesmo tamanho.

Em algumas situações pode ocorrer a não obtenção de perfil genético devido à falta de material biológico na amostra, ao seu grau de degradação ou à presença de contaminantes. Amostras degradadas submetidas à PCR - EC podem não ser amplificadas devido ao tamanho dos *primers* que compõem os kits de amplificação utilizados na rotina laboratorial. Neste caso o DNA molde está degradado a ponto do *primer* não se ligar a ele não sendo possível o processo de amplificação e detecção de perfil genético pretendido com a metodologia PCR – EC.

Considerando a presença ou ausência de mistura de material genético de mais de um indivíduo, o SMP apresentou resultados compatíveis (com relação à presença/ausência de mistura de material genético) com os resultados observados na PCR – EC, porém, o SMP oferece a oportunidade de serem analisadas as sequências dos alelos obtidos.

Os resultados obtidos neste trabalho mostraram que a metodologia de SMP foi superior em eficiência, na qual, podemos observar maior quantidade de locos obtidos e, portanto, maior possibilidade de se identificar indivíduos pela análise de perfis genéticos. Este ganho em quantidade de locos genotipados obtidos com a metodologia de SMP, reflete no *RMP (random match probability)*, com o qual podemos incluir ou excluir investigados de agressões sexuais. Também foram observados melhores resultados com o SMP quando foram analisadas as amostras de acordo com os parâmetros complexidade do caso e complexidade da análise. Neste contexto,

os resultados de ambas metodologias foram compatíveis em 22 amostras, porém, em 9 amostras os resultados obtidos com o SMP foram superiores, tendo sido obtidos maior quantidade de locos genotipados ou, até mesmo, perfis genéticos que não haviam sido obtidos na PCR – EC. As amostras 4, 21 e 27, apesar de apresentarem quantidade e qualidade de DNA adequadas, não geraram perfis genéticos na PCR – EC e quando sequenciadas geraram perfis genéticos completos. Esta divergência observada, pode ter sido por erro do analista no momento da manipulação das amostras quando do preparo da corrida eletroforética.

Para os locos relacionados aos Y STRs, a utilização da metodologia de SMP proporcionou o sequenciamento de todas as amostras enquanto para a metodologia de PCR – EC apenas as amostras 2,3,7,8,9,10,11,13,14,16,19,21,23,24,27,28 e 31 foram testadas, respeitando os protocolos internos de análise da STBBL. Considerando as amostras testadas em ambas as metodologias (Y STRs e SMP), apenas duas amostras (19 e 31) apresentaram resultado mais informativo no SMP do que na PCR – EC. Para o restante das amostras (2,3,7,8,9,10,11,13,14,16,21,23,24,27 e 28) a metodologia de PCR - EC apresentou melhores resultados quando comparada com SMP. Uma possível explicação para a obtenção destes resultados, pode estar nas diferenças dos desenhos dos *primers* envolvendo as metodologias, o que poderia afetar os resultados finais. Porém, esta análise envolveria um trabalho de bioinformática mais aprofundado o qual não é o objetivo deste trabalho.

É possível tentar otimizar os resultados para os marcadores do cromossomo Y, colocando menos amostras no cartucho de sequenciamento, o que possibilitaria a obtenção de maior número de *reads* por amostra. Outra possibilidade, seria diminuir a quantidade de amostras no cartucho de sequenciamento e aumentar o *input* de DNA daquelas amostras que não apresentaram bons resultados para os marcadores do cromossomo Y. Porém, é preciso avaliar a necessidade real da obtenção destes resultados de Y STRs para os casos em questão uma vez que os custos financeiros (somando-se os valores dos reagentes e horas de trabalho do Perito Analista) são muito elevados. Cabe destacar que, o estudo de haplótipos de cromossomo Y é complementar à análise de STRs autossômicos, pois, o resultado obtido a partir da análise destes marcadores não permite incluir um único indivíduo em uma cena de crime visto que todos indivíduos do sexo masculino da mesma linhagem patrilínea apresentarão o mesmo perfil de haplótipo. Para que a análise seja conclusiva no sentido de incluir indivíduos como “produtores” de vestígios deixados em cenas de crime, a análise de marcadores autossômicos ou de SNPs de identidade são as ferramentas a serem consideradas. Apesar destes resultados observados para os locos Y STRs, o SMP oferece a possibilidade de geração de maior quantidade de dados na análise das amostras forenses. Após a realização de todas as análises, os dados gerados na corrida de sequenciamento foram armazenados, de forma segura na STBBL, compondo um banco de dados dos produtos

intermediários para que possam ser acessados em caso de estudos futuros. A implementação do SMP resultou na elaboração de um POP (apêndice 2) o qual possibilitará aos peritos criminais da STBBL a execução da metodologia de sequenciamento nas amostras relacionadas à violência sexual custodiadas na PCMG. Este estudo foi direcionado para a implementação da metodologia de SMP na STBBL, não estando focado na análise comparativa e verificação de concordância entre os genótipos obtidos considerando-se as duas metodologias abordadas (PCR – EC e SMP). Todos os resultados obtidos farão parte de um estudo mais aprofundado ao qual pretendo trabalhar no meu doutoramento onde os dados serão analisados estatisticamente.

Referências

- 1-Peer-Reviewed Forensic Consultation in Practice: Multidisciplinary Oversight in Common Expertise. Michael Welner,1 M.D.; Emily E. Davey, M.A.; and Adam Bernstein, B.A.
- 2-A utilização da Ciência Forense e da Investigação Criminal Como Estratégia Didática na Compreensão de Conceitos Científicos. Ana Paula Sebastiany,1 Michelle Camara Pizzato, José Cláudio Del Pino & Tania Denise Miskinis Salgado.
- 3-http://www.planalto.gov.br/ccivil_03/decreto-lei/del3689.htm.
- 4-Fundamentals of Forensic DNA Typing. John M. Butler (2010).
- 5-A preliminary assessment of the ForenSeq™ FGx System: next generation sequencing of an STR and SNP multiplex. Ashley L. Silvia & Nathan Shugarts & Jenifer Smith.
- 6-Nuclear and Mitochondrial DNA Quantification of Various Forensic Materials. H Andréasson, M Nilsson, B Budowle, H Lundberg, M Allen. DOI: 10.1016/j.forsciint.2005.11.024.
- 7-Next generation sequencing and its applications in forensic genetics; Claus Børsting*, Niels Morling. Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.
- 8-Laurence Devesse, David Ballard, Lucinda Davenport, Immy Riethorst, Gabriella Mason-Buck, Denise Syndercombe Court, Concordance of the ForenSeq™ system and characterisation of sequence-specific autosomal STR alleles across two major population groups, Forensic Science International: Genetics <https://doi.org/10.1016/j.fsigen.2017.10.012>.
- 9-Developmental validation of the MiSeq FGx Forensic Genomics System for Targeted Next Generation Sequencing in Forensic DNA Casework and Database Laboratories. <http://dx.doi.org/10.1016/j.fsigen.2017.01.011>.
- 10- SWGDAM - Scientific Working Group on DNA Analysis Methods Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories - APPROVED 01/12/2017. - <https://www.swgdam.org/>.

- 11- Population and performance analyses of four major populations with Illumina's FGx Forensic Genomics System, JenniferD.Churchill <http://dx.doi.org/10.1016/j.fsigen.2017.06.004>.
- 12- Chiara Gelardi, Eszter Rockenbauer, Sigrun Dalsgaard, Claus Børsting, Niels Morling, Second generation sequencing of three STRs D3S1358, D12S391 and D21S11 in Danes and a new nomenclature for sequenced STR alleles, *Forensic Science International: Genetics*, Volume 12, 2014, Pages 38-41, ISSN 1872-4973, <https://doi.org/10.1016/j.fsigen.2014.04.016>.
- 13- Chang L, Yu H, Miao X, Wen S, Zhang B and Li S (2021) Evaluation of a Custom SNP Panel of Identifying and Rectifying of Misjudged Paternity in Deficiency Cases. *Front. Genet.* 12:602429. doi: 10.3389/fgene.2021.602429
- 14- How Next Generation Sequencing Resolved a Difficult Case, Leading to the First Criminal Conviction of its Kind. Peter de Knijff.
- 15- Forensic DNA Typing. Biology, Technology, and Genetics of STR Markers. John M. Butler (2005).
- 16- PCR inhibitors – occurrence, properties and removal C. Schrader, A. Schielke, L. Ellerbroek and R. Johne.
- 17- Use of the PowerPlex® Fusion 6C System to Amplify Extracted DNA. Instructions for Use of Products DC2705, DC2720 and DC2780. <https://www.promega.com.br/products/forensic-dna-analysis-ce/str-amplification/powerplex-fusion-6c-system/?catNum=DC2705#protocols>.
- 18- ForenSeq™ DNA Signature Prep Reference Guide. VEROGEN PROPRIETARY document # VD2018005 Rev. A June 2018.
- 19- Quantifiler1 Trio Kit and forensic samples management: A matter of degradation. S. Vernarecci et al. / *Forensic Science International: Genetics* 16 (2015) 77–85. <http://dx.doi.org/10.1016/j.fsigen.2014.12.005>
- 20- LRMIX STUDIO - <https://www.isfg.org/Software>.
- 21- https://www.justica.gov.br/sua-seguranca/seguranca-publica/ribpg/manual/resolucao_14-2019_aprova_o_manual.pdf/view.
- 22- Chelex 100 as a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material P. Sean Walsh, David A. Metzger, and Russell Higuchi Cetus Corporation and Illinois State Police.
- 23- DNA Extraction Methods in Forensic Analysis Steven B. Lee San Jose State University, San Jose, CA, USA Jaiprakash G. Shewale Rowpar Pharmaceuticals Inc., Scottsdale, AZ, USA.
- 24- MET-STBBL-5.6 – Extração de DNA com resina Chelex para suabes bucais. Pop padronizado na STBBL – PCMG.
- 25- MET-STBBL-5.3 – Realização da extração de DNA utilizando o equipamento Maxwell®16 casework para manchas e suportes sólidos. Pop padronizado na STBBL – PCMG.

26- <https://www.thermofisher.com/br/en/home/life-science/pcr/real-time-pcr/real-time-pcr-instruments/quantstudio-3-5-real-time-pcr-system.html>.

27- <https://www.thermofisher.com/order/catalog/product/4482910#/4482910>.

28- MET-AGQSTBBL-2.10 – Quantificação do DNA humano utilizando o equipamento QuantStudio™ 5 Real-Time PCR system kit QuantFiler™ Trio DNA Quantification. Pop padronizado na STBBL – PCMG.

29- A selection guide for the new generation 6-dye DNA profiling systems Sze-wah Lin, Christina Li, Stephen C.Y. Ip. doi.org/10.1016/j.fsigen.2017.05.010.

30- Ensenberger MG, Lenz KA, Matthies LK, Hadinoto GM, Schienman JE, Przech AJ, Morganti MW, Renstrom DT, Baker VM, Gawrys KM, Hoogendoorn M, Steffen CR, Martín P, Alonso A, Olson HR, Sprecher CJ, Storts DR. Developmental validation of the PowerPlex® Fusion 6C System. *Forensic Sci Int Genet.* 2016 Mar; 21:134-44. doi: 10.1016/j.fsigen.2015.12.011. Epub 2015 Dec 22. PMID: 26774099.

31- Fluxograma *PowerPlex Y23 Kit* da *Promega Corporation*. Fluxograma padronizado na STBBL – PCMG.

32- POP de Metodologia para análise de amostras referentes à violência sexual da STBBL.

33- <https://www.promega.com.br/>.

34- Thompson JM, Ewing MM, Frank WE, Pogemiller JJ, Nolde CA, Koehler DJ, Shaffer AM, Rabbach DR, Fulmer PM, Sprecher CJ, Storts DR. Developmental validation of the PowerPlex® Y23 System: a single multiplex Y-STR analysis system for casework and database samples. *Forensic Sci Int Genet.* 2013 Feb;7(2):240-50. doi: 10.1016/j.fsigen.2012.10.013. Epub 2013 Jan 20. PMID: 23337322.

35- A review of the method and validation of the MiSeq FGx™ Forensic Genomics Solution. Ryan England, SallyAnn Harbison. DOI: 10.1002/wfs2.1351.

36- Ng, J., Oldt, R.F., Kanthaswamy, S., Assessing the FBI's Native American STR database for random match probability calculations, *Legal Medicine* (2017), doi: <https://doi.org/10.1016/j.legalmed.2017.10.012>.

37- Evaluation of the Illumina1 Beta Version ForenSeq™ DNA Signature Prep Kit for use in genetic profiling. Jennifer D. Churchill, Sarah E. Schmedes, Jonathan L. King, Bruce Budowle.

38- Discriminação alélica em ovinos naturalizados do Pantanal Sul-Matogrossense por meio de marcadores microssatélites. Article in *Journal of the Selva Andina Research Society* · August 2012. DOI: 10.36610/j.jsars.2012.030100003.

39- <https://edisciplinas.usp.br/> (Ciências Biomédicas - FMRP/USP - Disciplina RCB0300 - Biotecnologia III - Princípios de Sequenciamento em escala genômica).

APÊNDICE 1: Resultados das métricas da qualidade obtidas na corrida de implementação do SMP

Após a etapa de sequenciamento foram geradas métricas que reportam a qualidade da corrida de sequenciamento de acordo com os parâmetros estabelecidos no *ForenSeq™ DNA Signature Prep Reference Guide. VEROGEN PROPRIETARY document # VD2018005 Rev. A June 2018* (tabela S.1). Por se tratar de implementação de nova metodologia, foram analisadas essas métricas da qualidade da corrida de implementação do SMP que evidenciam o sucesso da corrida de implementação a partir do correto funcionamento do equipamento *MiSeq® FGX da Illumina*. Para todos os parâmetros da qualidade (*Cluster Density, Cluster passing filter, Phasing, Pre-Phasing, Sample Representation, Read and Index Quality Metrics, ForenSeq Human Sequencing control, Positive Control Metrics e Negative Control Metrics*) foram obtidos resultados adequados sugerindo que a corrida de implementação da metodologia de SMP foi bem sucedida.

1-*Cluster Density* (k/mm²): Expressa o número de *clusters* por mm². Para o *ForenSeq™ DNA Signature Prep Kit* da *Verogen* o intervalo recomendado é de 400 a 1.650 k/mm² [18]. O resultado obtido neste estudo foi de 1.037 k/mm² indicando a produção de dados/resultados adequados para análise (figura S.1).

Tabela S.1: Parâmetros da qualidade de corrida da metodologia de SMP.

PARÂMETRO DA CORRIDA	VALOR RECOMENDADO*	VALOR OBTIDO NA CORRIDA
<i>CLUSTER DENSITY K/MM²</i>	400 - 1.650 K/MM ²	1.037 K/MM ²
<i>CLUSTER PASSING FILTER (%)</i>	≥ 80 %	92,90 %
<i>PHASING (%)</i>	≤ 0,25 %	0,231 %
<i>PRE-PHASING (%)</i>	≤ 0,15 %	0,086 %
<i>SAMPLE REPRESENTATION</i>	≥ 85.000 <i>READS</i>	23 AMOSTRAS COM RESULTADOS ACIMA DE 85.000 <i>READS</i>

Fonte: *ForenSeq™ DNA Signature Prep Reference Guide. Verogen Proprietary document # VD2018005 Ver. A June 2018.*

2-*Cluster passing filter (%)*: Porcentagem de *clusters* que passaram no “*Illumina chastity filter*”. Este filtro dá a medida da qualidade do sinal das bases incorporadas e o recomendado é que seja maior ou igual a 80 % [18]. O resultado obtido neste estudo foi de 92,90 % indicando a produção de dados/resultados adequados para análise (figura S.1).

3-*Phasing (%)*: Porcentagem de moléculas que estão um ciclo atrás do ciclo atual (este parâmetro é aplicado tanto para a *Read 1* quanto para a *Read 2*). Recomenda-se valores menores ou iguais a 0,25 % [18]. O resultado obtido neste estudo foi de 0,231 % indicando a produção de dados/resultados adequados para análise (figura S.1).

4-*Pre-Phasing (%)*: Porcentagem de moléculas que estão um ciclo à frente do ciclo atual (este parâmetro é aplicado tanto para a *Read 1* quanto para a *Read 2*). Recomenda-se valores menores ou iguais a 0,15 % [18]. O resultado obtido neste estudo foi de 0,086 % indicando a produção de dados/resultados adequados para análise (figura S.1).

Figura S. 1: Painel *Quality Metrics* mostrando os resultados obtidos para os parâmetros: *cluster density*, *cluster passing filter*, *phasing* e *pre-phasing*.

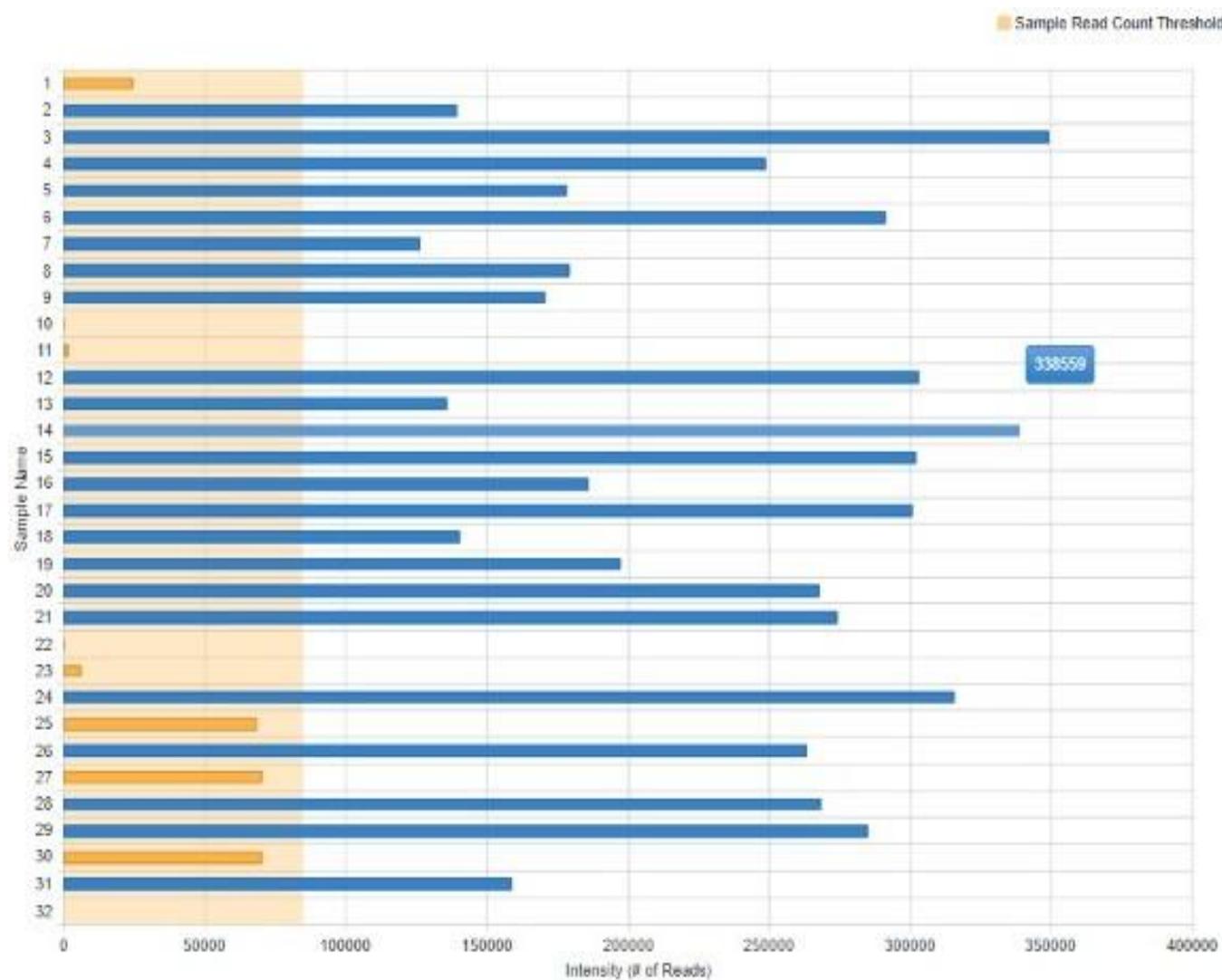


Fonte: *ForenSeq™ Universal Analysis Software (UAS) – Verogen® - Corrida de implementação.*

5- *Sample Representation (reads)*: Mostra a intensidade (número de *reads*) de cada amostra. Recomenda-se valores acima de 85.000 *reads* [18]. O resultado obtido neste estudo está representado na figura S.2. Encontraram-se dentro dos valores recomendados (barras azuis), ou seja, maior/igual a 85.000 *reads*, 23 amostras analisadas (22 de casos reais mais o controle positivo). Um total de oito amostras apresentaram resultados abaixo de 85.000 *reads* (barras laranjas), sendo três entre 50.000 e 85.000 *reads*, uma entre 20.000 e 25.000 *reads* e duas entre

5.000 e 10.000 *reads*. Duas amostras apresentaram sinal abaixo de 5.000 e uma (controle negativo) não apresentou sinal na corrida.

Figura S. 2: Painel mostrando o número de *reads* por amostra.



Fonte: *ForenSeq™ Universal Analysis Software (UAS) – Verogen® - Corrida de implementação.*

6- *READ AND INDEX QUALITY METRICS*: A qualidade das *Reads* (1 e 2) e dos *indexes* (i5 e i7) é representada por cores, sendo que, a cor verde obtida nesta corrida de implementação, indica que as *Reads* (1 e 2) e os *indexes* (i7 e i5) foram completamente sequenciados e a corrida completou 398 ciclos de um total de 398 [18].

7- *ForenSeq Human Sequencing control*: O resultado obtido para este parâmetro foi satisfatório indicando que o *Human Sequencing control* preencheu os critérios de intensidade e concordância de genótipo com o banco de dados do equipamento de sequenciamento.

8-*Positive Control Metrics*: Foi gerado alerta para análise do resultado do controle positivo (cor laranja da aba relativa ao resultado do controle positivo). Foi constatado que apesar de o número total de *reads* desta amostra ter sido adequado (acima de 85.000 *reads*) 8 SNPs não foram sequenciados nesta amostra. Porém, todos os STRs do controle positivo foram sequenciados indicando sucesso da reação de sequenciamento.

9-*Negative Control Metrics*: Foi obtido resultado condizente com o controle negativo (cor verde da aba relativa ao resultado do controle negativo), ou seja, não foi detectada a presença de SNPs ou STRs nesta amostra.

APÊNDICE 2 - Procedimento Operacional Padrão – Construção de Bibliotecas para Sequenciamento Massivo em Paralelo pela Tecnologia Illumina

1. Objetivo

Estabelecer procedimentos para construção de bibliotecas para Sequenciamento Massivo em Paralelo.

2. Aplicação

Aplica-se à sala de genotipagem e sequenciamento de DNA da Seção Técnica de Biologia e Bacteriologia Legal (STBBL). O equipamento utilizado neste processo é o sequenciador *MiSeq™ FGx* da *Illumina*.

3. Definições

O equipamento *MiSeq™ FGx* da *Illumina* baseia-se na tecnologia de sequenciamento massivo em paralelo por síntese (SMP), amplamente adotada em múltiplas áreas da ciência, inclusive nas ciências forenses. A tecnologia de SMP do *MiSeq™ FGx* da *Illumina* usa uma câmera de alta resolução para gravar as imagens dos dNTPs marcados com fluorescências que são incorporados sequencialmente durante a síntese da fita complementar, permitindo o sequenciamento massivo em paralelo de milhões de fragmentos de DNA alvo.

4. Amostras

Amostras de DNA extraídas e quantificadas de acordo com POP padronizado na STBBL.

5. Materiais

5.1. Materiais gerais

Sala reservada, com ventilação adequada, freezer e geladeira para armazenamento de reagentes, capelas de fluxo laminar, equipamentos termocicladores, lençol de papel descartável, caneta esferográfica, caneta marcadora permanente, lápis, pipetas automáticas, pipetas multicanais automáticas, vórtex, micro centrífuga, equipamento de proteção individual (avental, gorro, máscara, luvas, pró-pé, óculos de proteção).

5.2. Metodologia, procedimentos técnicos e materiais usados nas etapas do sequenciamento

5.2.1. PCR1 – Ligação de “etiquetas” ou “tags” no DNA alvo

5.2.1.1. Materiais:

DNA controle 2800M, DNA *Primer Mix A* (DPMA), reagente FEM (*mix* de enzima), PCR1 (*mix* para reação de PCR1), tampão TBE, microtubos de 1,5 ml, placa (*skirted* ou *semiskirted*) para PCR de 96 poços com 0,3 ml de profundidade, filme adesivo selante para placa, água livre de nuclease, tubos *strip* com 8 (livres de RNase/DNase) e suas respectivas tampas, 1 ng de DNA genômico humano.

Nota: Usar o filme adesivo selante para selar a placa antes de colocá-la no termociclador. Usar o filme adesivo selante para selar a placa nas etapas onde houver necessidade.

5.2.1.2. Metodologia:

1. Preparar os seguintes consumíveis:

Item	Armazenamento	Instruções
DPMA	- 25°C a -15°C	Descongelar à temperatura da sala
FEM	- 25°C a -15°C	Descongelar à temperatura da sala
PCR1	- 25°C a -15°C	Descongelar à temperatura da sala
TBE 1 X	- 25°C a -15°C	Descongelar à temperatura da sala
2800M	2°C a 8°C	Esperar 30 minutos para colocar à temperatura da sala

2. Criar uma lista de amostras contendo a posição de cada amostra e os seus respectivos *indexes*.

3. Salvar no termociclador o programa específico para a PCR1.

3.1. Programa para a PCR1

- ▶ 98°C por 3 minutos
- ▶ 8 ciclos de:
 - ▶ 96°C por 45 segundos
 - ▶ 80°C por 30 segundos
 - ▶ 54°C por 2 minutos com modo "rampa"
 - ▶ 68°C por 2 minutos com modo "rampa"
- ▶ 10 ciclos de:
 - ▶ 96°C por 30 segundos
 - ▶ 68°C por 3 minutos com modo "rampa"
- ▶ 68°C por 10 minutos

4. Rotular a placa de PCR como abaixo:

► *FSP (ForenSeq Sample Plate)*

5. Quantificar o DNA extraído das amostras utilizando o método de quantificação (qPCR) padronizado na STBBL, conforme POP específico.

6. Diluir 1ng de DNA purificado (que será o *input* de DNA) para 0,2 ng/ul com água livre de nuclease.

7. Criar o *master mix* com quantidade suficiente para o total de amostras, usando os seguintes reagentes:

► PCR1 – 4,7 ul

► FEM – 0,3 ul

► DPMA – 5,0 ul

Nota: As quantidades dos reagentes do *master mix* citadas acima são para uma amostra. Deve-se multiplicar esses valores pela quantidade de amostras.

8. Homogeneizar o *mix* e depois centrifugá-lo rapidamente (apenas um *spin* para retirar as gotículas da tampa do tubo).

9. Se forem processadas mais de 8 amostras, deve-se usar os tubos *strip* e distribuir o *master mix* utilizando uma pipeta multicanal.

10. Distribuir 10ul de *master mix* em cada poço da placa *FSP*.

11. Diluir 2ul do controle positivo (2800M) com 98ul de água livre de nuclease em um microtubo de 1,5 ml. Homogeneizar levemente e centrifugar rapidamente (apenas um *spin* para retirar as gotículas da tampa do tubo).

12. Adicionar 5ul do controle positivo da reação de PCR (2800M) no poço apropriado de acordo com a lista de amostras.

13. Adicionar 5ul de água livre de nuclease como controle negativo da reação de PCR no poço apropriado de acordo com a lista de amostras.

14. Adicionar 5ul de DNA purificado (0,2 ng/ul) de amostra em cada poço de acordo com a lista de amostras. Homogeneizar com pipeta.

15. Selar a placa com filme adesivo selante e centrifugar a 1000 x g por 30 segundos em centrífuga de placa.

16. Colocar a placa no termociclador e correr o programa da PCR1.

Nota: Se o analista pretender parar nesta etapa deve-se selar a placa com filme adesivo selante e armazenar em geladeira (entre 2°C e 8°C) por até 2 dias. Alternativamente a placa pode ser mantida em termociclador overnight.

5.2.2. PCR2 – Ligação dos *indexes* (adaptadores) i5 e i7 ao DNA alvo

5.2.2.1. Materiais:

Placa para fixação dos *indexes* do *ForenSeq Kit*, tubos com tampa de cor alaranjada contendo o *Index 1* (i7), tubos com tampa de cor branca contendo o *Index 2* (i5), PCR2 (*mix* para reação de PCR2), micro centrífuga e filme adesivo selante para placa e tubos *strip* com 8 (livres de RNase/DNase) e suas respectivas tampas.

Notas:

1. Usar o filme adesivo selante para selar a placa antes de colocá-la no termociclador. Usar o filme adesivo selante para selar a placa nas etapas onde houver necessidade.
2. Se forem processadas mais de 8 bibliotecas, distribuir o *mix* da PCR2 a partir dos tubos *strip* usando uma pipeta multicanal.
3. Adicionar o *mix* da PCR2 lentamente para cada poço da placa de forma a evitar bolhas de ar.

5.2.2.2. Metodologia:

1. Preparar os seguintes consumíveis:

Item	Armazenamento	Instruções
Adaptadores (i5 e i7)	- 25°C a -15°C	Somente remova os adaptadores que estiverem sendo utilizados. Descongelar à temperatura da sala por 20 minutos. Vórtexos tubos para homogeneizar. Centrifugar rapidamente (<i>spin</i>).
<i>Mix</i> PCR2	- 25°C a -15°C	Descongelar à temperatura da sala

2. Salvar o programa da PCR2 no termociclador:

- ▶ Escolher no pré-aquecimento da tampa e programar a 100°C
- ▶ 98°C por 30 segundos
- ▶ 15 ciclos de:
 - ▶ 98°C por 20 segundos
 - ▶ 66°C por 30 segundos
 - ▶ 68°C por 90 segundos
- ▶ 68°C por 10 minutos
- ▶ manter a 10°C

3. Centrifugar a *FSP* a 1000 x g por 30 segundos.

4. Organizar os adaptadores 1 (i7) nas colunas horizontais, de 1 – 12, da placa para fixação dos *indexes* do *ForenSeq* (figura 1).

5. Organizar os adaptadores 2 (i5) nas filas verticais, de A – H, da placa para fixação dos *indexes* do *ForenSeq* (figura 1).

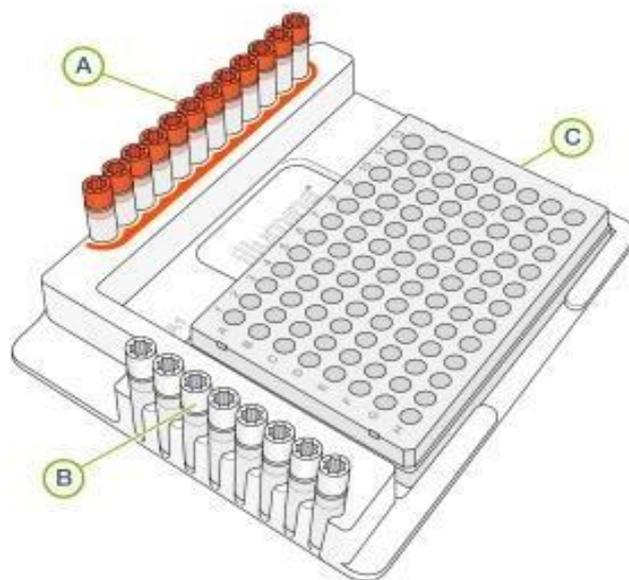


Figura 1: Placa para fixação dos *indexes* do *ForenSeq*. (A) Colunas horizontais de 1 – 12: adaptadores 1 (i7) com tampa de cor alaranjada; (B) filas verticais de A – H: adaptadores 2 (i5) com tampa de cor branca; (C) placa *FSP*. Fonte da imagem: *ForenSeq DNA Signature Prep Reference Guide*.

6. Colocar a placa *FSP* na placa para fixação dos *indexes* do *ForenSeq*.

7. Usando uma pipeta multicanal, adicionar 4 ul do adaptador 1 (i7) em cada coluna horizontal. Tampar os tubos dos adaptadores com novas tampas alaranjadas.

8. Usando uma pipeta multicanal, adicionar 4 ul do adaptador 2 (i5) em cada fila vertical. Tampar os tubos dos adaptadores com novas tampas de cor branca.

9. Vórtex o mix da PCR2 e depois centrifugar rapidamente (*spin*).

10. Adicionar 27ul do *mix* da PCR2 em cada poço da placa.

11. Centrifugar a 1000 x g por 30 segundos.

12. Colocar a placa no termociclador e “correr” o programa da PCR2.

Nota: Se o analista pretender parar nesta etapa deve-se selar a placa com filme adesivo selante e armazenar em geladeira (entre 2°C e 8°C) por até 7 dias. Alternativamente a placa pode ser mantida em termociclador overnight.

5.2.3. Purificação das bibliotecas

Este processo usa *SPB* (*Sample Purification Beads*) para purificar as bibliotecas, retirando todo resto dos outros componentes.

5.2.3.1. Materiais:

RSB (*Resuspension Buffer*), *SPB* (*Sample Purification Beads*), placa de 96 poços de 0,3 ml (*skirte* ou *semiskirted*), placa de 96 poços tipo *midi plate* (placa com os poços maiores), etanol 80% preparado no momento da purificação das bibliotecas, filme adesivo selante, suporte magnético para placa, reservatório para reagentes livre de RNase e de DNase.

Notas:

1. Vórtex o *SPB* antes de cada uso.
2. Vórtex o *SPB* frequentemente para se ter certeza de que as *beads* estarão distribuídas de forma homogenia.
3. Aspirar e dispensar o *SPB* vagarosamente devido a viscosidade da solução.

5.2.3.2. Metodologia:

1. Preparar os seguintes consumíveis:

Item	Armazenamento	Instruções
<i>RSB</i>	2°C a 8°C	Deixar em repouso por 30 minutos para trazer para a temperatura da sala
<i>SPB</i>	2°C a 8°C	Deixar em repouso por 30 minutos para trazer para a temperatura da sala

Nota: Para garantir uma ótima performance e rendimento das bibliotecas, tenha certeza de que as *SPB beads* estejam em sua totalidade na temperatura da sala.

2. Rotular as placas da forma descrita abaixo:

- ▶ *PBP* (*Purification Bead Plate*) – placa *midi*
- ▶ *PLP* (*Purified Library Plate*) – placa de PCR

3. Aliquotar o *SPB* de acordo com o número de amostras: 50ul de *SPB* x quantidade de amostras analisadas. Exemplo: para o total de 30 amostras, multiplicar 50 x 30 = 1.500ul e aliquotar em reservatório de reagente para pipeta multicanal.

Nota: O reagente *SPB* deve ser constantemente homogeneizado para evitar a precipitação das *beads* magnéticas.

4. Pipetar da alíquota realizada do passo anterior 45ul do *SPB* para cada poço da placa *PBP* de acordo com o mapa de amostras.

5. Centrifugar a placa *FSP* a 1000 x g por 30 segundos.
 6. Transferir 45ul da placa *FSP* para o poço correspondente na placa *PBP*, de acordo com o mapa de amostras.
 7. Selar a placa com o filme selante e agitar no agitador de placas a 1800 rpm por 2 minutos.
 8. Incubar a temperatura da sala por 5 minutos.
 9. Colocar a placa no suporte magnético e aguardar até a solução clarear (aproximadamente 2 minutos).
 10. Remover e descartar todo o sobrenadante de cada poço.
 11. Lavar duas vezes como descrito abaixo:
 - a. Adicionar 200ul de etanol 80% recém preparado, em cada poço.
 - b. Incubar no suporte magnético por 30 segundos.
 - c. Remover e descartar todo o sobrenadante de cada poço.
 12. Centrifugar a 1000 x g por 30 segundos.
 13. Colocar a placa no suporte magnético.
 14. Usar uma pipeta de 20ul para remover o etanol residual de cada poço.
 15. Remover a placa do suporte magnético.
 16. Adicionar 52,5ul de *RBS* a cada poço.
 17. Selar a placa com filme selante e agitar a 1800 rpm por 2 minutos. Se as *beads* não se ressuspenderem, homogeneizar com pipeta ou repetir a agitação a 1800 rpm por 2 minutos.
 18. Incubar à temperatura da sala por 2 minutos.
 19. Colocar a placa no suporte magnético e aguardar até o líquido clarear (aproximadamente 2 minutos).
 20. Transferir 50ul para o poço correspondente na placa *PLP*.
 21. Centrifugar a 1000 x g por 30 segundos.
- Nota: Se o analista pretender parar nesta etapa deve-se selar a placa com filme adesivo selante e armazenar em freezer (entre -25°C e -15°C) por até 1 ano.

5.2.4. Normalização das bibliotecas

Este processo prepara as bibliotecas de DNA para a geração dos clusters, garantindo que estas bibliotecas estejam igualmente representadas na corrida de sequenciamento.

5.2.4.1. Materiais:

HP3 (2N-NaOH), *LNA1* (*Library Normalization Additives 1*), *LNB1* (*Library Normalization Beads 1*), *LNS2* (*Library Normalization Storage Buffer 2*), *LNW1* (*Library Normalization Wash 1*), microtubos de 1,5 ml, tubo falcon de 15 ml, placa para PCR com 96 poços e 0,3 ml (*skirted* ou

semiskirted), placa *midi* com 96 poços, filme adesivo selante para placa, água livre de nuclease, reservatório de reagente para pipeta multicanal (livre de RNase e DNase).

Nota: 1. **Cuidado** – Este conjunto de reagentes possui compostos tóxicos, tais como, formamida e β -mercaptoetanol. O procedimento deve ser usado em área ventilada ou utilizando-se máscara de proteção. Deve-se utilizar frascos específicos e devidamente rotulados, para o descarte de sobrenadantes (restos destas soluções tóxicas).

2. Sobre os reagentes:

2.1. LNA1 – após vórtex, segurar o LNA1 em frente à luz e certificar-se que não possuem cristais presentes e que todo o precipitado foi dissolvido.

2.2. LNB1 – após vórtex, certificar-se que as *beads* do LNB1 estão ressuspensas e não há pellets formados no fundo do tubo. Pode-se utilizar pipeta de 1000ul para ajudar na homogeneização das *beads*. Ressuspender as *beads* é essencial para alcançar a formação de uma densidade de *clusters* consistente e otimizar a resolução das bibliotecas individuais quando sequenciadas simultaneamente.

2.3. As bibliotecas restantes na placa *PLP* pode ser estocada. Selar a placa com filme adesivo selante para placa e estocar entre -25°C e -15°C por até 1 ano.

5.2.4.2. Metodologia:

1. Preparar os seguintes consumíveis:

Item	Armazenamento	Instruções
HP3	-25°C a -15°C	Descongelar à temperatura da sala.
LNA1	-25°C a -15°C	Descongelar à temperatura da sala. Vórtex com inversão intermitente.
LNB1	2°C a 8°C	Aguardar 30 minutos para trazer para temperatura da sala. Vórtex por pelo menos 1 minuto, invertendo 5 vezes a cada 15 segundos. Homogeneizar com pipeta até não haver pellet de <i>beads</i> , ou seja, até ocorrer ressuspensão completa.
LNW1	2°C a 8°C	Aguardar 30 minutos para trazer para temperatura da sala.
LNS2	15°C a 30°C	Remova da área de estoque.

2. Rotular os tubos e placas de acordo com o descrito abaixo:

- ▶ LNA1/LNB1 *master mix* – microtubo de 1,5 ml ou tubo Falcon de 15 ml
- ▶ *NWP (Normalization Working Plate)* – placa *midi*

- ▶ *NLP (Normalization Library Plate)* – placa de PCR
- 3. Disponibilizar um frasco para descarte de resíduos tóxicos para líquidos e sólidos.
- 4. Criar um *master mix* no tubo rotulado LNA1/LNB1 *master mix*.
 - ▶ LNA1 – 46,8ul por amostra
 - ▶ LNB1 – 8,5ul por amostra
- 5. Vórtex e inversão do tubo várias vezes com o objetivo de homogeneizar.
- 6. Verter em um reservatório de reagentes para pipeta multicanal.
- 7. Transferir 45ul para cada poço da placa *NWP* que irá conter as bibliotecas de acordo com o mapa de amostras.
- 8. Para se evitar a aspiração de *beads*, colocar a placa *PLP* no suporte magnético e aguardar até que o líquido fique límpido (aproximadamente 2 minutos).
- 9. Transferir 20ul de cada poço da placa *PLP* para o poço correspondente na placa *NWP*.
- 10. Selar a placa com filme selante e agitar a 1800 rpm por 30 minutos.
- 11. Enquanto a placa estiver agitando, preparar os seguintes passos:
 - 11.1. Preparar HP3 0,1 N em um novo microtubo de 1,5 ml como se segue:
 - ▶ 33,3ul de água livre de nuclease, por amostra.
 - ▶ 1,8ul de HP3, por amostra.
 - ▶ Inverter o tubo várias vezes com o objetivo de homogeneizar.
 - ▶ Deixar na bancada para usá-lo no momento que o protocolo indicar.
 - 11.2. Adicionar 30ul de *LNS2* a cada poço da placa *NLP* que irá conter as bibliotecas de acordo com o mapa de amostras.
- 12. Imediatamente após o processo de agitação da placa *NWP* terminar, colocar a placa *NWP* no suporte magnético e aguardar o líquido ficar límpido (aproximadamente 2 minutos).
- 13. Remover e descartar o sobrenadante de cada poço.
- 14. Remover a placa do suporte magnético.
- 15. Lavar duas vezes com 45ul de *LNWI* como se segue:
 - 15.1. adicionar 45ul de *LNW1* a cada poço.
 - 15.2. selar a placa com filme adesivo selante e agitar a 1800 rpm por 5 minutos.
 - 15.3. colocar a placa no suporte magnético e aguardar até que o líquido esteja límpido (aproximadamente 2 minutos).
 - 15.4. remover e descartar o sobrenadante de cada poço.
- 16. Remover a placa do suporte magnético.
- 17. Centrifugar a 1000 x g por 30 segundos.
- 18. Colocar a placa no suporte magnético e aguardar até que o líquido fique límpido (aproximadamente 2 minutos).

19. Usar uma pipeta de 20ul para remover o sobrenadante residual de cada poço.
 20. Remover a placa do suporte magnético.
 21. Adicionar 32ul de HP3 0,1 N recém preparado em cada poço.
 22. Selar a placa com filme adesivo selante e agitar a 1800 rpm por 5 minutos. Se as *beads* não tiverem sido ressuspensas, homogeneizar com pipeta ou repetir a agitação a 1800 rpm por 5 minutos.
 23. Colocar a placa no suporte magnético e aguardar até que o líquido fique límpido (aproximadamente 2 minutos).
 24. Transferir 30ul para o poço correspondente da placa *NLP*. Homogeneizar com pipeta.
 25. Centrifugar a 1000 x g por 30 segundos.
- Nota: Se o analista pretender parar nesta etapa deve-se selar a placa com filme adesivo selante e armazenar em freezer (entre -25°C e -15°C) por até 30 dias.

5.2.5. Conjunto (*pool*) das bibliotecas

Este processo combina iguais volumes de biblioteca normalizada para criar um conjunto de bibliotecas que serão sequenciadas ao mesmo tempo na mesma *flow cell*.

5.2.5.1. Materiais:

- ▶ microtubo de 1,5 ml.
- ▶ filme adesivo selante para placa.
- ▶ tubos strip de 8, livres de RNase e de DNase.

5.2.5.1. Metodologia:

1. Determinar quais bibliotecas serão sequenciadas. Para o *mixA* (PCR1) podem ser sequenciadas até 96 bibliotecas e para o *mixB* (PCR1) podem ser sequenciadas até 32 bibliotecas.
2. Rotular o tubo *PNL*, para indicar o *Pooled Normalized Libraries*.
3. Transferir 5ul de cada biblioteca para um novo tubo *strip* com 8.
4. Armazenar a placa na área de pós-PCR entre -25°C e -15°C por até 30 dias.
5. Transferir o conteúdo final do tubo *strip* com 8 para o tubo *PNL*.
6. Vórtex e centrifugar rapidamente (*spin*).

Nota: Se o analista pretender parar nesta etapa deve-se selar a placa com filme adesivo selante e armazenar em freezer (entre -25°C e -15°C) por até 30 dias.

5.2.6. Desnaturação e diluição das bibliotecas

Este processo dilui as bibliotecas em HT1 (*Hybridization Buffer*), adiciona HSC (*Human Sequencing Control*) e desnatura as bibliotecas, por aquecimento, para prepara-las para o sequenciamento.

Nota: Realizar este processo imediatamente antes de adicionar as bibliotecas no cartucho de reagentes para que seja garantida o eficiente carregamento da amostra na *flow cell*.

5.2.6.1. Materiais:

ForenSeq DNA Signature Prep Kit contendo o HP3 (2N – NaOH) e o HSC (*Human Sequencing Control*), microtubos (2) de 1,5 ml, *MiSeq Reagent Kit* contendo o HT1 (*Hybridization Buffer*) e o cartucho de reagentes e água livre de nuclease.

Nota sobre os reagentes: **Vide como preparar o cartucho de reagentes no item 5.2.7.**

5.2.6.2. Metodologia:

1. Preparar os seguintes consumíveis:

Item	Armazenamento	Instruções
HP3	- 25°C a -15°C	Descongelar à temperatura da sala.
HSC	- 25°C a -15°C	Descongelar à temperatura da sala.
HT1	- 25°C a -15°C	Descongelar à temperatura da sala.
Cartucho de reagentes	- 25°C a -15°C	Descongelar à temperatura da sala.

2. Pré-aquecer o termo bloco a 96°C.

3. Preparar o que se segue abaixo:

- ▶ Remover o suporte para resfriamento de tubos do freezer ou utilizar recipiente com gelo.
- ▶ alternativamente ao suporte para resfriamento pode-se preparar um recipiente com água e gelo combinando 3 partes de gelo e 1 parte de água livre de nuclease.

4. Rotular os tubos como descrito abaixo:

- ▶ HSC *mixture*.
- ▶ DNL para indicar *Denatured Normalized Libraries*.

5. Criar uma reação de desnaturação no tubo HSC *mixture*.

- ▶ HSC – 2ul
- ▶ HP3 – 2ul
- ▶ água livre de nuclease – 36ul

6. Vórtex e centrifugação rápida (*spin*).

7. Incubar à temperatura da sala por 5 minutos.
8. Adicionar 591ul de HT1 ao tubo *DNL*.
9. Transferir 7ul do tubo *PNL* para o tubo *DNL*. Pipetar para homogeneizar.
10. O tubo *PNL* pode ser armazenado entre -25°C e -15°C por até 30 dias. Após 30 dias há uma considerável redução na densidade de *clusters*.
11. Transferir 2ul da *HSC mixture* para o tubo *DNL*. Pipetar pra homogeneizar. Não armazenar o *HSC mixture* por muito tempo, para evitar a redução na densidade de *clusters*.
12. Vórtex e centrifugação rápida (*spin*).
13. Colocar o tubo *DNL* no termo bloco a 96°C por 2 minutos.
14. Inverter o tubo várias vezes para homogeneizar.
15. Colocar o tubo *DNL* imediatamente no suporte para resfriamento de tubos por 5 minutos.
16. Colocar todo o conteúdo imediatamente dentro do cartucho de reagentes.

5.2.7. Fluxo de trabalho do MiSeq FGX

5.2.7.1. Preparo do cartucho de reagentes

As instruções seguintes descrevem como se descongelar o cartucho de reagentes usando um recipiente com água à temperatura da sala.

Nota: O cartucho de reagentes do *MiSeq FGX* só pode ser utilizado com a corrida de sequenciamento do ForenSeq e não funciona para corrida *RUO (Research Use Only)*.

1. Remova o cartucho de reagentes do freezer.
2. Colocar o cartucho em um recipiente com água deionizada, à temperatura da sala, suficiente para submergir a base do cartucho (vide linha de marcação do cartucho da figura 2). Não deixar a quantidade de água ultrapassar a linha de marcação do cartucho.



Figura 2: Linha de marcação que determina o máximo de água que se deve mergulhar o cartucho. Fonte: *MiSeq FGX Instrument Reference Guide - Verogen*[®].

3. Aguardar o descongelamento total por aproximadamente 60 a 90 minutos.

4. Remover o cartucho do recipiente com água e suavemente bater sua base contra a bancada para deslocar a água que permanecer na sua base. Secar a base do cartucho e ter certeza que não tenha respingos na sua parte superior.

5.2.7.1.1. Inspeccionar o cartucho de reagentes

1. Inverter o cartucho de reagentes 10 vezes para homogeneizar os reagentes recém descongelados e efetuar uma inspeção visual para verificar se houve descongelamento total.

Nota: É crítico que os reagentes do cartucho estejam totalmente descongelados e homogeneizados para que ocorra um sequenciamento apropriado.

2. Inspeccionar visualmente os reagentes das posições 1, 2 e 4 para se ter certeza de que eles estão totalmente homogeneizados e livre de precipitados.

3. “Bater” o cartucho cuidadosamente contra a bancada para reduzir as bolhas de ar dentro dos reagentes. Importante: Certificar-se que o fundo do cartucho de reagentes esteja livre de bolhas, para que a aspiração dos reagentes pelo equipamento *MiSeq FGX* não seja prejudicada.

4. Colocar o cartucho de reagentes no gelo ou mantê-lo em geladeira (2°C e 8°C) por até 6 horas até que a configuração da corrida esteja pronta. Para resultados melhores, proceder diretamente com a adição das amostras e depois configurar a corrida.

5.2.7.1.2. Adicionar as amostras no cartucho de reagentes

Após o preparo do cartucho de reagentes do *MiSeq FGX*, as amostras podem ser adicionadas para a corrida de sequenciamento.

1. Usar um lenço *KimWipes* (lenços que não soltam fibras) para limpar o selo de alumínio que recobre a posição rotulada “*Load Samples*”.

2. Usar uma pipeta de 1 ml com a ponteira limpa para perfurar o selo de alumínio que recobre a posição rotulada “*Load Samples*”.

Nota importante: Não perfurar qualquer outra posição do cartucho. Outras posições do cartucho são perfuradas automaticamente durante a corrida de sequenciamento.

3. Pipetar 600ul de tubo *DNL* (biblioteca de amostras desnaturadas e normalizadas) no reservatório rotulado “*Load Samples*”. Tocar o orifício com a ponta da pipeta enquanto se dispensa a amostra.

Checar se há bolhas de ar dentro do reservatório após a dispensação da amostra. Se houver bolhas de ar “bater” levemente o cartucho contra a bancada para desfazer as bolhas.

4. Proceder diretamente aos passos da corrida usando a interface do *MiSeq FGx Control Software*.

5.2.7.1.3. *Login* e seleção do tipo de corrida

1. Na tela de *login*, entrar com usuário e senha do *ForenSeq Universal Analysis Software*, e então clicar em *Next*.

Nota: O usuário e senha encontra-se anexado na parte lateral do *MiSeq FGx*.

2. Na tela de abertura selecionar *Sequence*.

3. Quando a tela referente à seleção de tipo de corrida aparecer, deve-se selecionar *Forensic Genomics*.

4. Seguir os comandos para carregar a “*flow cell*” e reagentes e configurar a corrida (como descrito nos passos seguintes).

5.2.7.1.4. Preparando a *flow cell* – limpeza

A *flow cell* vem imersa em tampão de armazenagem em recipiente apropriado.

1. Calçar um par de luvas novas, livres de talco.

2. Usando uma pinça de preferência plástica, retirar a *flow cell* pela porção plástica do cartucho e removê-la do recipiente no qual está armazenada (figura 3).



Figura 3: Remoção da *flow cell*. Fonte: Fonte: *MiSeq FGX Instrument Reference Guide - Verogen®*.

3. Lavar suavemente a *flow cell* com água livre de nuclease certificando-se de que tanto o vidro quanto o cartucho plástico sejam completamente lavados do excesso de sal (figura 4). O excesso de sal pode afetar o encaixe da *flow cell* no instrumento. Se os sais cristalizarem na área a qual serão realizadas as imagens pelo equipamento, a formação destas imagens poderá ser afetada.



Figura 4: Lavando a flow cell. Fonte: *MiSeq FGX Instrument Reference Guide - Verogen*[®].

4. Secar completamente a *flow cell* e o cartucho plástico usando um lenço de papel (*KimWipes*) que não solte fiapos de tecido. Secar com cuidado ao redor da junta e da “porta” da *flow cell* (região demarcada de laranja na figura 5). Secar suavemente a área da junta e do vidro adjacente.

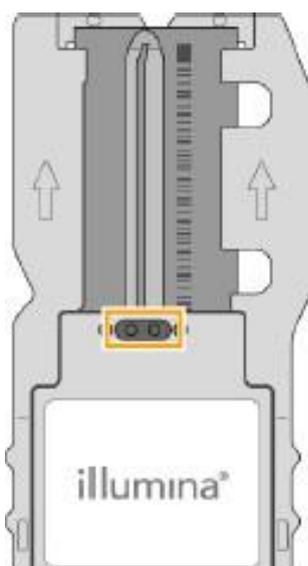


Figura 5: Porta e junta da *flow cell*. Fonte: *MiSeq FGX Instrument Reference Guide - Verogen*[®].

5. Usando um lenço com álcool, limpar o vidro da *flow cell*. Certificar-se de que o vidro esteja livre de sujidades, impressões digitais e fibras de tecido. **Evitar** o uso de lenço com álcool na porta e na junta da *flow cell*.

6. Secar qualquer excesso de álcool com lenço de limpeza de lentes sem fiapos (*KimWipes*). Fazer uma inspeção visual para garantir que a porta da *flow cell* esteja livre de obstruções e a junta esteja bem encaixada ao redor da porta da *flow cell*. Se a junta estiver deslocada da sua posição original, pressioná-la levemente para que ela retorne ao lugar correto e se encaixe perfeitamente ao redor da porta da *flow cell*.

5.2.7.1.5. Carregando a *flow cell* no equipamento *MiSeq FGx*

O carregamento da *flow cell* (figura 6) deve ocorrer quando o comando na tela do equipamento determinar.

1. Levante a porta do compartimento da *flow cell* e pressione o botão do lado direito do local de inserção da *flow cell*. Este compartimento interno se abrirá.

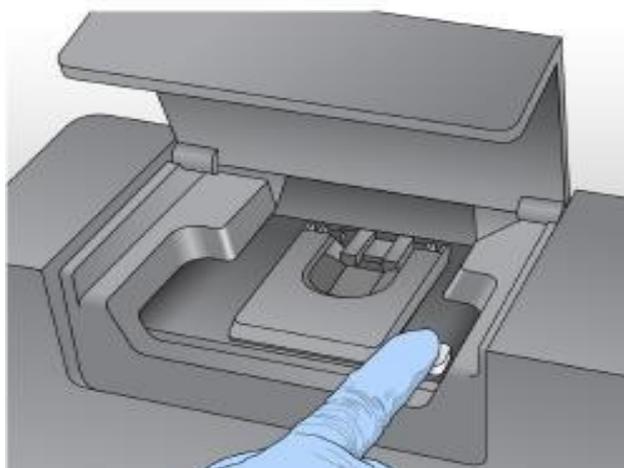


Figura 6: Botão de abertura do local de inserção da *flow cell*. Fonte: *MiSeq FGX Instrument Reference Guide - Verogen*[®].

2. Inspeccione visualmente a *flow cell* e seu compartimento para certificar-se de que não há fiapos de lenço. Caso exista fiapos de lenço ou outras impurezas, limpar o local com lenço (tipo *KimWipes*) umedecido em etanol ou isopropanol. Limpar cuidadosamente até que o local esteja limpo e seco.

3. Segurando a *flow cell* pelas extremidades do cartucho, colocá-la no compartimento como indicado na figura 7.

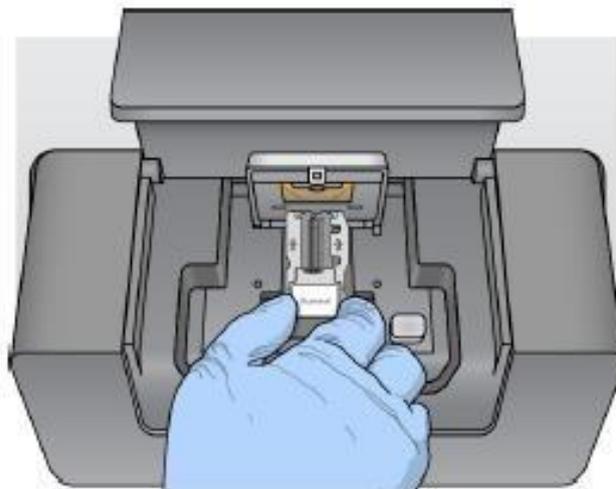


Figura 7: Colocando a *flow cell* no compartimento apropriado. Fonte: *MiSeq FGX Instrument Reference Guide - Verogen*[®].

4. Cuidadosamente pressione para baixo a porta do compartimento interno onde a *flow cell* se encaixa (figura 8).

Nota: Conforme a trava da *flow cell* é fechada, dois pinos de alinhamento perto da dobradiça possibilita o alinhamento e o posicionamento correto da *flow cell*. Um clique audível indica que a trava da *flow cell* está segura.

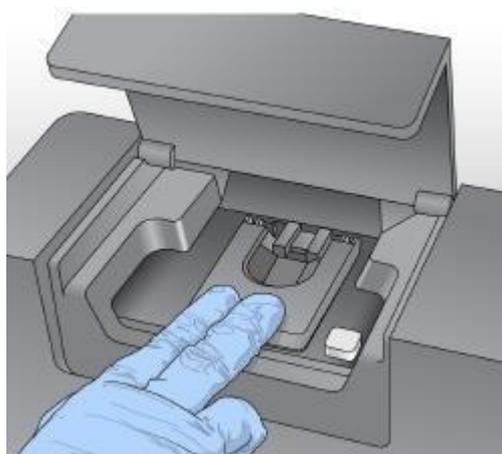


Figura 8: Fechando a trava da *flow cell*. Fonte: *MiSeq FGX Instrument Reference Guide - Verogen*[®].

5. Checar no canto inferior esquerdo da tela do equipamento *MiSeq FGx* a confirmação de que a identificação de rádio frequência (RFID) foi realizada com sucesso.

Nota: Caso ocorra erro na leitura do RFID, verificar o *MiSeq FGX Instrument Reference Guide - Verogen*[®], páginas 46 e 66 (*resolving RFID failure*).

6. Fechar a porta do compartimento da *flow cell*.

7. Selecionar em *Next* na tela de carregamento da *flow cell*. A tela de carregamento dos reagentes se abrirá.

5.2.7.1.6. Carregando os reagentes

Aqui existem dois passos para se carregar o equipamento com os reagentes. Primeiro, carregar a solução de SBS (PR2) e se certificar de que a garrada de descarte está vazia e depois disso, carregar o cartucho de reagentes.

■ Carregar a solução de SBS (PR2) e checar a garrafa de descarte

1. Remover a garrafa de solução de SBS (PR2) da geladeira (2°C a 8°C). Inverta suavemente a garrafa de SBS (PR2) para homogeneizar a solução e depois remova a tampa.
2. Abrir a porta do compartimento de reagentes.
3. Levantar a alça da trava até que se possa liberar as garrafas.
4. Colocar o frasco da solução SBS (PR2) no entalhe à direita do refrigerador de reagente (figura 9).
5. Certificar-se de que o recipiente de descarte está vazio. Caso não esteja vazio, esvaziá-lo em recipiente de descarte apropriado.

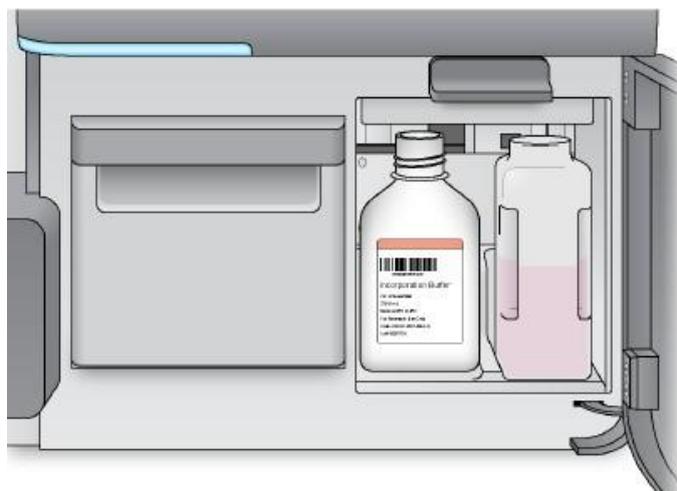


Figura 9: Carregamento do frasco da solução SBS (PR2). Fonte: *MiSeq FGX Instrument Reference Guide - Verogen*[®].

6. Abaixar lentamente a alça de travamento dos frascos (SBS e descarte) de acordo com a figura 10.
10. Certificar-se de que os ductos do equipamento estejam dentro dos frascos (SBS e descarte).

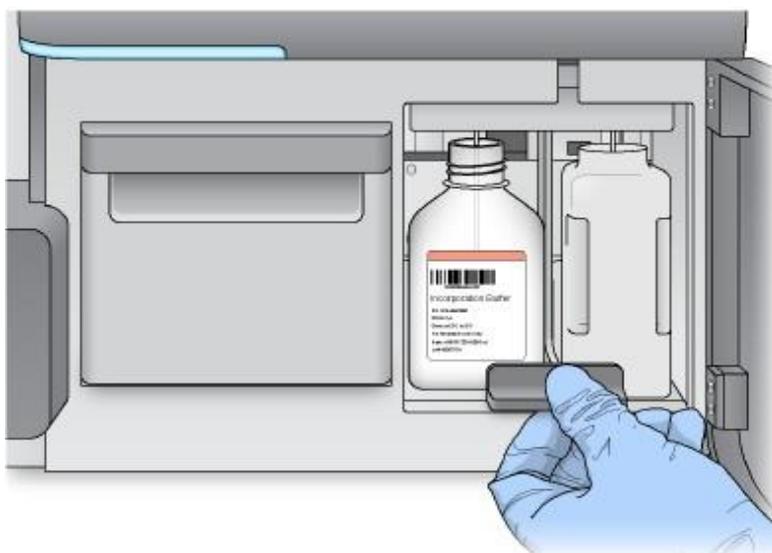


Figura 10: Abaixando a alça de travamento. Fonte: *MiSeq FGX Instrument Reference Guide - Verogen®*.

7. Checar no canto inferior esquerdo da tela do equipamento *MiSeq FGx* a confirmação de que a identificação de rádio frequência (RFID) foi realizada com sucesso.

Nota: Caso ocorra erro na leitura do RFID, verificar o *MiSeq FGX Instrument Reference Guide - Verogen®*, páginas 46 e 66 (*resolving RFID failure*).

8. Selecionar em *Next* na tela de carregamento dos reagentes.

■ Carregar o cartucho de reagentes

Nota: Não deixar a porta do refrigerador de reagentes aberta por muito tempo.

1. Abrir a porta do refrigerador de reagentes.

2. Segurar o cartucho de reagentes na sua porção final onde está rotulado *Illumina*, e deslizar o cartucho dentro do refrigerador de reagentes (figura 11).



Figura 11: Carregar o cartucho de reagentes. Fonte: *MiSeq FGX Instrument Reference Guide - Verogen®*.

3. Fechar a porta do refrigerador de reagentes.
4. Checar no canto inferior esquerdo da tela do equipamento *MiSeq FGx* a confirmação de que a identificação de rádio frequência (RFID) foi realizada com sucesso.

Nota: Caso ocorra erro na leitura do RFID, verificar o *MiSeq FGX Instrument Reference Guide - Verogen®*, páginas 46 e 66 (*resolving RFID failure*).

5. Fechar a porta do compartimento dos reagentes.
6. Selecionar em **Next** na tela de carregamento dos reagentes. Uma tela de revisão do processo se abrirá.

5.2.7.1.7. Iniciando a corrida

Após o carregamento da *flow cell* e dos reagentes, revisar os parâmetros de corrida e realizar uma checagem “pré-corrída” antes de iniciar a corrida propriamente dita.

■ Revisar os parâmetros de corrida

1. Revisar o nome da “Lista de Trabalho, o fluxo de trabalho da análise e o total de *Reads*.”
2. Selecionar em **Next** na tela do equipamento. A tela de checagem da pré-corrída se abrirá.

■ Revisar a checagem da pré-corrída

O sistema realiza uma checagem de todos os componentes da corrida, espaço em disco e conexões de rede antes do início da corrida.

Caso algum item não passe na checagem da pré-corrída, uma mensagem aparece na tela com instruções de como se corrigir o erro. Para mais informações, verificar a página 65 do *MiSeq FGX Instrument Reference Guide - Verogen® (Resolving Run Setup Errors)*.

Quando todos os itens passarem com sucesso pela checagem de pré-corrída, selecionar **Start Run**.

Nota importante: O equipamento MiSeq FGx é sensível à vibração. Tocar o equipamento após o início da corrida pode prejudicar os resultados do sequenciamento.

Após selecionar **Start Run**, não abra o compartimento da flow cell ou a porta do compartimento dos reagentes, não tocar o monitor do instrumento exceto se for para pausar a corrida. Para mais informações, verificar a página 30 do *MiSeq FGX Instrument Reference Guide - Verogen® (Pausing a Run)*.

5.2.7.1.8. Monitorando a corrida

1. Durante a corrida, monitore o progresso da corrida, intensidades e qualidade dos resultados que aparecem na tela de sequenciamento. A tela de sequenciamento é apenas para análise visual. Para mais informações visualizar a página 23 do *MiSeq FGX Instrument Reference Guide - Verogen® (Sequencing Screen)*.

2. Quando a corrida estiver completa, o botão **Next** aparece. Revisar os resultados na tela de sequenciamento antes de prosseguir.

Nota: A tela de visualização do sequenciamento permanece disponível até o botão **Next** ser selecionado. Após selecionar o botão **Next**, não será mais possível retornar à tela do sequenciamento.

3. Selecionar **Next** para sair da tela do sequenciamento e proceder à lavagem pós corrida.

5.2.7.1.9. Métricas da corrida

As métricas da corrida aparecem na tela de sequenciamento em diferentes pontos da corrida. Durante a formação dos *clusters* as métricas não são mostradas. Após o início da corrida as métricas aparecem nos ciclos indicados na tabela 1.

Tabela 1: Métricas da corrida de sequenciamento. Fonte: *MiSeq FGX Instrument Reference Guide - Verogen®*.

Métrica	Kit	Ciclo
Intensidade	<i>MiSeq FGx Reagent Kits</i>	Ciclo 1 – 7
Intensidade e densidade dos <i>Clusters</i>	<i>MiSeq FGx Reagent Kits</i>	Ciclo 8 – 25
Intensidade, densidade dos <i>Clusters</i> , % PF, produção e resultados da qualidade (<i>Q-scores</i>)	<i>MiSeq FGx Reagent Kits</i>	Ciclo 26 até o fim da corrida

6. Calibração e controle de qualidade

Controle da qualidade: aplicação adequada da metodologia.

7. Referências

1. *ForenSeq™ DNA Signature Prep Reference Guide*. Verogen Proprietary, document # VD2018005 Ver. A, June 2018.

2. *MiSeq FGX Instrument Reference Guide - Verogen®*.