



**LGN 5839:**  
**Construção de Bibliotecas GBS (Genotyping By Sequencing) visando a descoberta de SNPs**

**05 a 16 de dezembro de 2022**

**Professores Responsáveis LGN 5839:**

Profa. Dra. Maria Imaculada Zucchi (APTA/PG-UNICAMP), [mizucchi@sp.gov.br](mailto:mizucchi@sp.gov.br)

Profa. Dra. Maria Lúcia Carneiro Vieira (ESALQ/USP), [mlcvieir@usp.br](mailto:mlcvieir@usp.br)

**Colaboradores**

Caroline Bertocco	<a href="mailto:carolinebertocco0@gmail.com">carolinebertocco0@gmail.com</a>
Jonathan M. Marroquin	<a href="mailto:jmorales1089@gmail.com">jmorales1089@gmail.com</a>
Matheus Scaketti	<a href="mailto:mts.scaketti@gmail.com">mts.scaketti@gmail.com</a>
Joao Rabello	<a href="mailto:joaorabelooufrj@gmail.com">joaorabelooufrj@gmail.com</a>
Roberto Portella	<a href="mailto:roberto.portella.bio@gmail.com">roberto.portella.bio@gmail.com</a>
André Augusto Stella	<a href="mailto:andre.augusto.stella@usp.br">andre.augusto.stella@usp.br</a>
Kauanne Martins	<a href="mailto:kauanne.karolline@gmail.com">kauanne.karolline@gmail.com</a>
Matheus S. Moro	<a href="mailto:m.sartorimoro@gmail.com">m.sartorimoro@gmail.com</a>

**Local: Aulas práticas -laboratório Profa. Maria Lucia Carneiro Vieira** –Depto de Genética -ESALQ USP e **aulas teóricas- sala dos alunos do GGGC-** Depto de Genética -ESALQ-USP (frente ao quiosque).

**Ementa da disciplina:****LGN 5839-1: Construção de bibliotecas GBS (*Genotyping by Sequencing*) visando a descoberta de SNPs****Período:** S2/2022

T:60 T:30 P:20 E:10

**Programa:**

The goal of this workshop is to learn step-by-step how to prepare libraries for Genotyping-By-Sequencing (GBS). By the end of the course, the students should be able to 1) Describe the necessary steps to prepare libraries for GBS; 2) Assess the quality of DNA libraries digested with different restriction enzymes; 3) Prepare a GBS library to be sent for Illumina sequencing. We will also discuss some of the possible applications of GBS generated data for studies for population genomics; 4) We will show approaches for data filtering and identification of SNPs markers.

**Workshop structure:**

This workshop will cover theoretical and practical aspects of the development of GBS genomic libraries aiming the identification of SNP markers, data filtering using the pipeline of Stacks and Ipyrad, and introducing population genomic analyses.

**References:**

Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. 2011. *A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species*. PLoS One, 6(5), e19379.

Poland, J.; Brown, P.J.; Sorrells, M.E.; Jannink, J.C. 2012. *Development of high-density genetic maps for Barley and Wheat using a novel two-enzyme genotyping-by-sequencing approach*. PLoS One, 7(2): e32253.

Programa:

As aulas terão início às 09 horas durante as manhãs e às 14 horas durante as tardes.

No decorrer do curso ajustes nos horários poderão ser necessários e serão divulgados durante as aulas.

05/12	Manhã	TEÓRICO	<b>Apresentação do curso</b> - Introdução ao curso - <b>Palestra: Introdução ao GBS visando Conservação e Genômica de Populações</b> Profa. Dra. Maria Imaculada Zucchi
		TEÓRICO	<b>Apresentação dos alunos (2 slides cada sobre a pesquisa)</b>
	Tarde	TEÓRICO	<b>Apresentação do protocolo</b> Profa. Maria I. Zucchi, Ms. Jonathan Morales, Ms. Caroline Bertocco
		PRÁTICO /LABORATORIAL	<b>REAÇÃO DE DIGESTÃO</b>
06/12	Manhã	PRÁTICO /LABORATORIAL	<b>LIGAÇÃO DOS ADAPTADORES</b>
		TEÓRICO	- <b>Palestra: Genômica e estrutura populacional da Macaúba (<i>Acrocomia aculeata</i>), visando subsidiar a domesticação da espécie</b> Dra. Brenda Diaz Hernandez - <b>Palestra: Genômica e estrutura populacional da macaúba (<i>Acrocomia aculeata</i>), utilizando SNPs outliers putativos candidatos a seleção positiva.</b> Ms. Jonatham Morales Marroquin, - <b>Palestra: Genômica de populações no <i>Euterpe oleracea</i>, o acai do Para.</b> Ms Ana Flavia Francisconi
	Tarde	PRÁTICO /LABORATORIAL	<b>POOL DE AMOSTRAS E PURIFICAÇÃO DO POOL</b>
		TEÓRICO	- <b>Palestra: Genômica populacional e metagenômica parcial do complexo de percevejos da soja: <i>Euschistus heros</i> e <i>Piezodorus guildinii</i> das Américas</b> Dr. Matheus Sartori Moro - <b>Genômica populacional em <i>Spodoptera frugiperda</i>, visando subsidiar o manejo da praga</b> Ms Tamylin K. Shizuka
07/12	Manhã	TEÓRICO	- <b>Palestra: Diversidade química e genômica de populações naturais de <i>Lychnophora pinaster</i></b> <b>MART</b> Dr. Roberto Portella - <b>Palestra: GENÔMICA POPULACIONAL DE <i>Digitaria insularis</i> NAS PRINCIPAIS REGIÕES PRODUTORAS DE SOJA NO BRASIL</b> Dr. Acácio Goncalves
		PRÁTICO /LABORATORIAL	<b>AMPLIFICAÇÃO POR PCR E PURIFICAÇÃO</b>
	PRÁTICO /LABORATORIAL	<b>QUANTIFICAÇÃO DAS BIBLIOTECAS POR qPCR</b>	

	Tarde	<b>PRÁTICO /LABORATORIAL</b>	<b>AVALIAÇÃO DOS FRAGMENTOS AMPLIFICADOS POR GEL DE AGAROSE</b>
08/12	Manhã e Tarde	<b>PRÁTICO-COMPUTACIONAL</b>	-Palestra: <b>Pré-análise dos dados de sequenciamento I</b> <b>Introdução ao Linux</b> Ms. Matheus Scaketti
		<b>PRÁTICO-COMPUTACIONAL</b>	- <b>Pré-análise dos dados de sequenciamento II</b> - <b>Controle de qualidade e identificação de SNPs em espécies não-modelo usando o STACKS</b> Matheus Scaketti
09/12	Manhã	<b>PRÁTICO-COMPUTACIONAL</b>	- <b>Palestra: Pré-análise dos dados de sequenciamento III</b> <b>Pacote IpyRad</b> Dr. Matheus Sartori Moro
	Tarde	<b>PRÁTICO-COMPUTACIONAL</b>	- <b>Palestra: Pré-análise dos dados de sequenciamento IV</b> <b>Introdução Identificação de SNPs OUTILERS</b> Dr. Roberto Portela
		<b>PRÁTICO-COMPUTACIONAL</b>	- <b>Palestra: Amostragem genética V</b> <b>Pacote Sassi</b> Dr. Matheus Scaketti
12/12	Manhã e Tarde	<b>TEORICO - PRÁTICO</b>	- <b>Palestra – Introdução a tecnologia NANOPORE: Uma abordagem teórica -prática.</b> Ms Igor Carvalho
		<b>TEORICO - PRÁTICO</b>	- <b>Palestra: Interprise X Nanopore: Desafios e Vantagens.</b> Ms. Patricia Favoreto – Interprise
13/12	Manhã e Tarde	<b>TEORICO - PRÁTICO COMPUTACIONAL</b>	- <b>Palestra: Introdução a genética e genômica de populações. Uma breve revisão sobre as estimativas populacionais.</b> Profa Maria Zucchi
		<b>PRÁTICO-COMPUTACIONAL</b>	- <b>Palestra: Pré-análise dos dados de sequenciamento VI</b> <b>Uso de NGS em análises populacionais, pacotes populacionais em R</b> Ms. Ana Flavia Francisconi
14/12	Manhã e Tarde	<b>TEORICO</b>	- <b>SEMINARIOS</b>
			- <b>ENCERRAMENTO DO CURSO</b>

## PREPARO DE BIBLIOTECAS GENÔMICAS DE FRAGMENTOS DE DNA PARA GENOTIPAGEM POR SEQUENCIAMENTO (GBS)

### Introdução

Este é um protocolo de preparo de bibliotecas genômicas de fragmentos de DNA para o Sequenciamento de Nova Geração (em inglês, *Next-Generation Sequencing*, NGS) nas plataformas Illumina. O protocolo que será utilizado durante este curso foi descrito por Poland *et al.* (2011), esquematizado na Figura 1. Primeiramente, o DNA genômico de cada indivíduo é digerido com duas enzimas de restrição (que podem variar de acordo com espécie) escolhidas levando em conta o tamanho do genoma e a quantidade de sequências repetitivas. Em seguida, adaptadores específicos para sequenciadores Illumina são ligados aos fragmentos digeridos. Nesta etapa cada amostra recebe um adaptador que contém sequências *barcodes* específicas. Os fragmentos são então reunidos em um único *pool* (multiplex) para cada espécie. Cada multiplex é purificado com kits de purificação (alternativamente com *beads* magnéticas) e, em seguida, é feita uma amplificação por meio de PCR para o enriquecimento dos fragmentos no multiplex. Após essa etapa, as bibliotecas são uma vez mais purificadas e podem ser validadas e quantificadas para então serem submetidas ao sequenciamento.

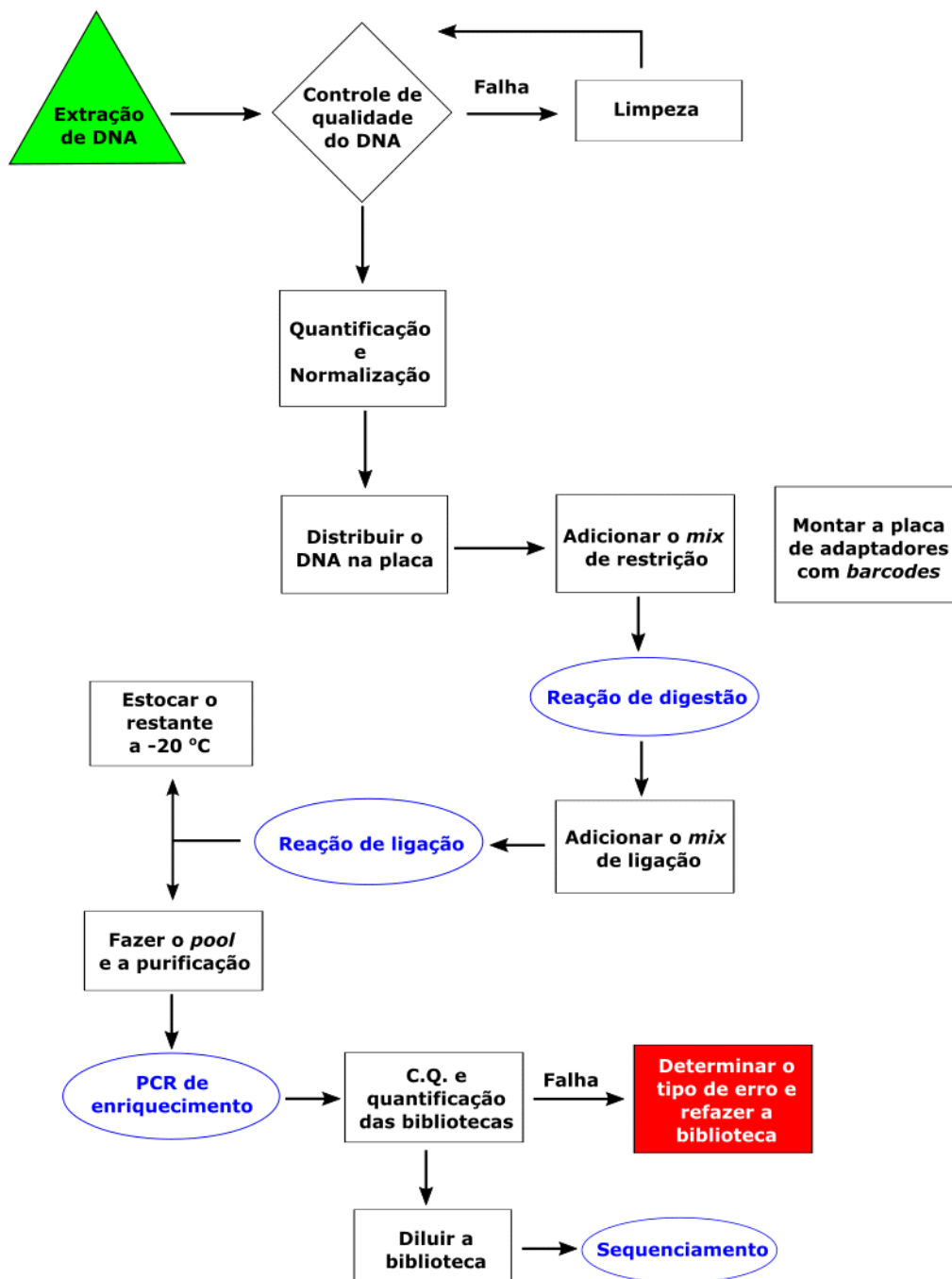
### Reagentes e Equipamentos

A tabela 1 lista os equipamentos, os reagentes e as soluções necessárias para o desenvolvimento do protocolo descrito a seguir.

**Tabela 1:** Equipamentos, reagentes e soluções.

Item	Fabricante	N de catálogo
2100 Bioanalyzer Desktop Bundle	Agilent	G2940CA
Adaptadores barcodes + adaptadores comuns	IDT	-
Adhesive PCR Film	Thermo Fisher	AB2558
Agilent DNA 1000 Kit	Agilent	5067-1504
Centrífuga de placas com refrigeração	Eppendorff	5804-R
CFX384 Touch™ Real-Time PCR Detection System	BioRad	CFX384
Enzima <i>MspI</i>	NEB	R0127L
Enzima <i>PstI</i>	NEB	R0140L
Microplaca para PCR de 96 poços	Axygen	PCR-AB-C
QIAquick PCR Purification Kit	Qiagen	28106
Qubit v3.0 Fluoremeter	Thermo	Q33216
T4 DNA Ligase, 500un,	Promega	M1804
Termociclador com bloco para placas	Biorad	C1000
Tris-EDTA Buffer Solution	Sigma	93283-500ML
Phusion® High-Fidelity PCR Master Mix	NEB	M0531S
UltraPure™ DNase/RNase-Free Distilled Water	Thermo Fisher	10977-015

### Workflow



**Figura 1:** Esquema do preparo das bibliotecas genômicas do tipo ddGBS para o sequenciamento de nova geração, como descrito por Poland *et al.* (2012).

### **Preparação de bibliotecas ddGBS para plataforma Illumina – *PstI* + *MspI***

Jesse Poland (*USDA-ARS / Kansas State University*)

Patrick Brown (*University of Illinois*)

Fonte: Poland, J.; Brown, P.J.; Sorrells, M.E.; Jannink, J.C. 2012. *Development of high-density genetic maps for Barley and Wheat using a novel two-enzyme genotyping-by-sequencing approach*. PLoS One, 7(2): e32253.

## **Adaptadores**

Os adaptadores são oligonucleotídeos comuns (e.g. IDT, com dessalinização padrão, escala de síntese 25 M). Para cada adaptador, um par de oligos complementares é comprado e devem ser anelados para formar o adaptador dupla-fita antes de iniciar o protocolo. Após o anelamento, estes adaptadores são bastante estáveis e podem ser estocados a  $-20^{\circ}\text{C}$  indefinidamente. Para o anelamento os adaptadores são submetidos a  $95^{\circ}\text{C}$  e resfriados a  $30^{\circ}\text{C}$  lentamente ( $-1^{\circ}\text{C}$  por minuto). Este procedimento foi realizado em um termociclador comum com o seguinte programa de PCR: 1 ciclo de  $95^{\circ}\text{C}$  por 1 minuto; e 65 ciclos diminuindo a temperatura em  $1^{\circ}\text{C}$  por 1 minuto, iniciando em  $95^{\circ}\text{C}$  e terminando em  $30^{\circ}\text{C}$ . Os adaptadores contendo *barcodes* podem ser encomendados em placa e todo o procedimento de anelamento pode ser realizado em formato de placa. Os adaptadores (*barcodes* e comuns) devem ser quantificados após o anelamento para garantir que DNA dupla-fita tenha sido formado adequadamente e para ajustar a concentração de trabalho. Uma concentração uniforme de adaptadores é essencial para que o número de *reads* obtidos para cada indivíduo após o sequenciamento da biblioteca multiplex seja uniforme.

## **Anelamento dos adaptadores**

Tampão de eluição (EB) 1X – 10 mM Tris-HCl, pH 8,0-8,5

Tampão do Adaptador (AB) 10X – 500 mM NaCl, 100 mM Tris-HCl

Para os Adaptadores1 (com *barcodes*), o procedimento é realizado em placas:

- Suspende com EB 1X os oligos simples-fita para a concentração de 100 M
- Fazer alíquotas de 100 L do Adaptador1 dupla-fita com concentração de 10 M:
  - 10 L de cada adaptador simples-fita (com concentração de 100 M)
  - 10 L de AB 10X
  - 70 L de água Milli-Q
- Aquecer a  $95^{\circ}\text{C}$  e resfriar a  $30^{\circ}\text{C}$  (diminuindo  $1^{\circ}\text{C}$  por minuto), *hold*  $4^{\circ}\text{C}$ .
- Diluir os adaptadores para aproximadamente 3 M (3 adaptador: 10 água Milli-Q) e quantificar usando Qubit (espera-se obter  $\sim 50\text{ ng/L}$ ).
- Normalizar a  $1,6\text{ ng/L}$  (= 0,1 M). Obs.: Esta é a concentração em ng/L da versão “curta” do Adaptador1. Adaptadores na versão “completa” deverão ter uma concentração em ng/L maior para atingirem a mesma molaridade.

Para os adaptadores comuns *MspI* (Adaptador2), o procedimento é realizado em microtubos individuais, seguindo os mesmos passos, mas deixando as alíquotas de trabalho do Adaptador2 com concentração de 10 M.

## **Estoque de adaptadores:**

Cada poço na placa de trabalho contendo os adaptadores terá 0,02 M de Adaptador1 e 3 M de Adaptador2.

Para fazer a placa de adaptadores adicionar em placas de 96 poços:

20 L de Adaptador1 (0,1 M)

30 L de Adaptador2 (10 M)

50 L de AB 1X

Misturar bem, aplicar um *spin* na placa e armazenar a  $-20^{\circ}\text{C}$ .

## **Quantificação e normalização do DNA**

A concentração do DNA é crítica para produzir um número uniforme de *tags* e *reads* para cada amostra. É recomendado que o DNA seja quantificado usando um método baseado em fluorescência, tal como Qubit ou PicoGreen.

## **Normalização do DNA**

Normalizar a concentração das amostras de DNA para  $30\text{ ng/L}$ , e distribuir 7 L em placas de 96 poços.

### **Digestão por restrição**

O protocolo descrito por Poland *et al.* (2012) emprega digestão dupla com as enzimas de restrição *Pst*I e *Msp*I. A extremidade coesiva gerada por *Pst*I se ligará ao Adaptador1 (que contém o *barcode*), enquanto a extremidade coesiva de *Msp*I se ligará ao adaptador comum Adaptador2 (Y-Adaptador). Na configuração *single-read*, a reação de sequenciamento acontece a partir do Adaptador1 e segue em sentido *downstream* incluindo a sequência *barcode*. As bibliotecas produzidas podem ser sequenciadas também a partir das duas extremidades na configuração *pair-end*.

Enzimas NEB são utilizadas neste protocolo uma vez que elas foram otimizadas para funcionarem com a mesma solução tampão (*Pst*I e *Msp*I podem ser utilizadas com o tampão “NEB Buffer3.1”).

### **Ligação**

O processo de ligação dos adaptadores é realizado no mesmo tubo (ou placa) no qual foi realizada a reação de digestão. A reação de ligação é realizada com a enzima NEB T4 DNA Ligase #M0202, com o tampão NEB Buffer4 e adição de ATP. Uma concentração muito elevada de enzima T4 é utilizada para garantir que a ligação de todos os fragmentos ocorra de maneira adequada. A concentração do Adaptador1 precisa ser ajustada dependendo da espécie. Por exemplo, para trigo e cevada, em torno de 0,1 pmol é adequado para ~200 ng de DNA genômico. O Adaptador2 pode ser adicionado em excesso, uma vez que ele não será amplificado a menos que a PCR tenha estendido o fragmento a partir do Adaptador1 na outra extremidade do fragmento. É importante que a enzima ligase seja inativada antes que seja feito o *pool* das amostras (65° C por 20 minutos após a reação de ligação).

Nota: Os autores do protocolo original sugerem a utilização de 50-100x mais Adaptador2 do que Adaptador1.

### **Multiplex**

Os produtos da reação de ligação são misturados e submetidos à amplificação por PCR em microtubos individuais. Este passo produz uma biblioteca contendo 96 amostras, que será sequenciada em uma linha (ou *flowcell*) única em equipamento Illumina.

### **Amplificação por PCR**

A biblioteca multiplex é amplificada por meio de PCR utilizando períodos de extensão curtos. Este procedimento visa o enriquecimento de fragmentos que contêm entre 200 e 500 bases, e que são mais apropriados à amplificação por ponte (*bridge PCR*) em sequenciadores Illumina. Apenas os fragmentos que contiverem um sítio de restrição *Pst*I e outro sítio de restrição *Msp*I serão amplificados. Os fragmentos com sítios de restrição *Msp*I-*Msp*I podem ser comuns, mas não serão amplificados devido aos AdaptadoresY. Os fragmentos com sítios de restrição *Pst*I-*Pst*I devem ser bastante raros.

PCRs múltiplas são realizadas para cada biblioteca na tentativa de reduzir a amplificação preferencial de fragmentos que podem ocorrer durante uma única reação de amplificação. As diferentes reações são então misturadas para compor a biblioteca final.

### **Protocolo**

- **Reação de digestão (volume final = 20 L)**

7 L de DNA (30 ng/L = ~210 ng de DNA total)

13 L de *mix* da reação de restrição



*Pst*I-HF – NEB #R3140 (20.000 U/mL)

*Msp*I – NEB #R0106 (20.000 U/mL)

Preparação do *mix* da reação de restrição:

	Por reação	1 placa (x100 + 10%)
NEB Buffer 4 (10X)	2,0 L	220 L
<i>Pst</i> I-HF [corte raro] (20 U/L)	0,4 L	44 L
<i>Msp</i> I [corte frequente] (20 U/L)	0,4 L	44 L
Água Milli-Q	10,2 L	1122 L

Normalizar as amostras de DNA para 30 ng/L e distribuir 7 L em uma placa de 96 poços;

Adicionar e misturar 13 L do *mix* da reação de restrição aos 7 L de DNA;

Realizar a reação de restrição seguindo os ciclos: 37° C por 12 h; 65° C por 20 min; 10° C indefinidamente.

Nota: Realizar a reação de ligação imediatamente.

• **Reação de ligação (volume final = 40 L)**

20 L da reação de digestão

5 L de adaptadores (Adaptador1 [0,02 M] = ~0,1 pmol + Adaptador2 [3 M] = ~15 pmol)

15 L de *mix* da reação de ligação

Preparação do *mix* da reação de ligação:

	Por reação	1 placa (x100 + 10%)
NEB Buffer 4 (10X)	2,0 L	220 L
ATP (10 mM)	1,0 L	110 L
T4 DNA Ligase (200 U/L)	0,5 L	55 L
Água Milli-Q	11,5 L	1265 L

Adicionar 5 L de adaptador aos 20 L da reação de restrição;

Adicionar e misturar 15 L do *mix* da reação de ligação;

Realizar a reação de ligação seguindo os ciclos: 22° C por 2 h; 65° C por 20 min; 10° C indefinidamente.

Nota: Após a reação de ligação a placa pode ser armazenada a -20° C.

• **Multiplex (*Pool* de amostras)**

Em um microtubo, misturar 5 L dos produtos de ligação de cada uma das amostras, totalizando 480 L. Adicionar 20 L de água ultrapura, e purificar utilizando colunas Qiagen (QIAquick PCR Purification Kit).

Obs.: O pool pode ser feito utilizando multicanal. Em uma *8-well strip* (ou 8 microtubos de PCR) misturar 6 L dos produtos de ligação de cada uma das colunas (1 a 12), utilizando a multicanal.

Misturar 60 L da mistura de reações de ligação de cada poço da *strip* em um único microtubo de 1,5 mL (previamente adicionar 20 L de água ao microtubo, volume total = 500 L);

Purificar o multiplex utilizando colunas Qiagen (QIAquick PCR Purification Kit).

### Observações:

Realizar duas purificações para cada biblioteca;  
Combinar 200 L do multiplex das reações de ligação e 1000 L do tampão PB (QIAquick PCR Purification Kit – protocolo abaixo) em um microtubo novo;  
Adicionar 600 L à coluna, centrifugar;  
Adicionar os outros 600 L, centrifugar;  
Seguir as recomendações do fabricante e **ressuspender em 60 L de água ultrapura**.  
Combinar as duas purificações em um único microtubo (volume final = 120 L).

### • **Amplificação das bibliotecas: 8 PCRs para cada biblioteca (volume final = 25 L)**

10 L da reação de ligação purificada  
12,5 L NEB Master Mix 2X  
2,0 L Iniciadores *forward* e *reverse* Illumina (10 M)  
0,5 L de água Milli-Q

Programa (16 ciclos): 95° C por 30 s; 16 ciclos de (95° C por 10 s, 62° C por 20 s, 72° C por 30 s); 72° C por 5 min; 4° C indefinidamente.

Para cada biblioteca, misturar os produtos das 8 PCRs (volume final ~200 L) e purificar utilizando colunas Qiagen (QIAquick PCR Purification Kit). Seguir as recomendações do fabricante e **ressuspender em 30 L de água ultrapura**.

\* o tempo de extensão na PCR pode ser aumentado ou diminuído para otimizar a distribuição dos tamanhos dos fragmentos amplificados.

Verificar a qualidade das bibliotecas em gel de Agarose 1% (p/v) ou em equipamento BioRad Experion.

### **Protocolo de purificação – QIAquick PCR Purification Kit**

1. Adicionar 5 volumes do Tampão PB a 1 volume de reação de PCR e misturar (se a cor da mistura for laranja ou violeta, adicionar 10 L de acetato de sódio 3M (pH 5,0) e misturar – a cor deve ser amarela);
2. Colocar uma coluna QIAquick em um tubo de 2 mL (*collection tube*);
3. Adicionar a mistura na coluna e centrifugar (13000 rpm) por 60 s, descartar o líquido e posicionar a coluna no mesmo *collection tube*;
4. Lavar adicionando 750 L de Tampão PE à coluna e centrifugar (13000 rpm) por 60 s, descartar o líquido e posicionar a coluna no mesmo *collection tube*;
5. Centrifugar novamente (13000 rpm) por 60 s para remover tampão residual;
6. Adicionar a coluna num microtubo de 1,5 mL novo;
7. Eluir o DNA com 30 L de Tampão EB (Tris-HCl 10 mM, pH 8,5), ou água (pH 7,0 – 8,5).  
Dispensar o tampão no centro da coluna, aguardar 1 min e centrifugar (13000 rpm) por 60 s.

### **Anexo I – Adaptadores (Poland et al, 2011)**

#### ADAPTADOR1

Adaptadores com *barcode* (Adaptador1)

Este é o adaptador com sequências *barcode* de comprimento variável. Um exemplo de Adaptador1 com a sequência *barcode* XXXXX:

```
5' CACGACGCTCTTCGATCTXXXXXTGCA GNNNNNNN 3'  
3' GTGCTGCGAGAAGGCTAGAXXXXX ACGTCNNNNNNN 5'
```

Esta é a versão "curta" dos adaptadores com sequência *barcode* (XXXXX)  
(Tm = 57)

```
_bot XXXXXagatcggagagcgtcgtg  
_top cacgacgctcttccgatctXXXXXtgca
```

#### ADAPTADOR2

Adaptadores comuns (Adaptador2)

Estes adaptadores são desenhados respectivamente para a segunda enzima usada no protocolo de digestão dupla. Eles são desenhados para sequenciamento *pair-end*. O Adaptador2 possui forma de Y, uma vez que o iniciador *reverse* (A2Rprimer) só poderá ser anelado caso o iniciador *forward* (A1Fprimer) tenha sido estendido a partir da outra extremidade do fragmento (um sítio de restrição *PstI* com o Adaptador1). O Adaptador2 possui uma extremidade coesiva 5'-CG para o sítio de restrição *MspI*. (Caso uma enzima deixar uma extremidade coesiva 3' (i.e. *PstI*) uma combinação diferente de adaptadores deve ser desenhada).

Adaptador2(PE) para *MspI* (C | CGG)

```
5' nnnnnnnnC CGAGATCGGAAGAGCGGGGACTTTAAGC  
3' nnnnnnnnGGC TCTAGCCTTCTCGCCAGTCTCCTTACGGCTCTGGCTAG
```

```
>Adpt2PE_top_cg  
cgAGATCGGAAGAGCGGGGACTTTAAGC
```

```
>Adpt2PE_bot  
GATCGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATCT
```

Adaptador2(PE) para *AvaII* (G | GWCC)

```
>Adpt2PE_top_gtc  
gtcAGATCGGAAGAGCGGGGACTTTAAGC
```

```
>Adpt2PE_top_gac  
gacAGATCGGAAGAGCGGGGACTTTAAGC
```

#### Iniciadores

Os iniciadores Illumina são idênticos aos oligos existentes nas *flowcells*, com a adição de sequências adaptadoras complementares ao Adaptador1 (*forward*) e, no caso de sequenciamento *pair-end* ao Adaptador2 (*reverse*).

```
>IlluminaF_PE (Tm=70) 58 pb (Tm=57)  
AATGATACGGCGACCAACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT
```

```
>IlluminaR_PE (Tm=70) 46 pb (Tm=62)  
CAAGCAGAAGACGGCATAACGAGATCGGTCTCGGCATTCTGCTGAA
```

```
Adaptador1 + iniciador + barcode = 62 - 68 pb  
Adaptador2 + iniciador = 61 pb  
Tamanho mínimo do fragmento = ~125 pb
```