Reduced Representation Bisulfite Sequencing for Methylation Analysis

Preparing Samples for the Illumina Sequencing Platform

Introduction

This protocol explains how to prepare libraries of genomic DNA (gDNA) for reduced representation bisulfite sequencing (RRBS), a sequencing-based methylation analysis application. The libraries are prepared for subsequent cluster generation starting from sample DNA digestion through adapter ligation, library purification, and quantification.

Input gDNA (2–5 µg) is digested at the CCGG sites with MspI. The fragments are blunt-ended and phosphorylated, and a single 'A' nucleotide is added to the 3' ends of the fragments in preparation for ligation to a methylated adapter that has a single-base 'T' overhang. The ligation products are purified and accurately size-selected by agarose gel electrophoresis. Size-selected DNA is bisulfite-treated and purified. The treated DNA is PCR-amplified to enrich for fragments that have adapters on both ends. The final purified product is then quantitated prior to cluster generation.

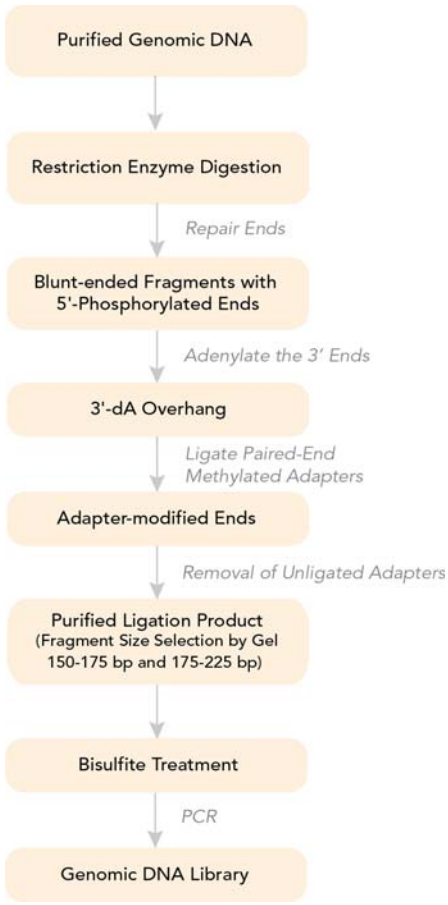
NOTE
It is essential to run a control lane on the same flowcell as the RRBS sample libraries. Because bisulfite treatment greatly decreases the C base in the genomic sequence, the resulting three based libraries are not optimal for base calls. Any genomic library containing balanced A,G,T, and C components (e.g., genomic resequencing, RNA-seq, ChIP-seq) could be served as a control lane for RRBS sample libraries.

NOTE
It is highly recommended that you limit the cluster number for RRBS libraries to a 75–80% optimized cluster number for other genomic libraries, because of the unique CGG starting sequence in the first three bases.

Sample Prep Workflow

The following figure illustrates the steps in the RRBS for Methylation Analysis protocol.

Figure 1 RRBS for Methylation Analysis Workflow



Best Practices

When preparing genomic DNA libraries for sequencing, you should always adhere to good molecular biology practices.

Liquid Handling

Good liquid handling measures are essential, particularly when quantifying libraries or diluting concentrated libraries for making clusters. Small differences in volumes (±0.5 µl) can sometimes give rise to very large differences in cluster numbers (~100,000). Small volume pipetting can also be a source of potential error in protocols that require generation of standard curves, such as PicoGreen assays or qPCR, or those that require small but precise volumes, such as the Agilent BioAnalyzer. If small volumes are unavoidable, then due diligence should be taken to ensure that pipettes are correctly calibrated. Also, care should be taken, because solutions of high molecular weight dsDNA can be viscous and not evenly dispersed, resulting in aliquot measurements that are not representative of the true concentration of the solution.

To minimize pipetting errors, especially with small volume enzyme additions, prepare the reagents for multiple samples simultaneously. As a result, you pipette once from the reagent tubes with a larger volume, rather than many times with 1 µl volumes. Prepare a master mix of enzymes, water, buffer, etc. and aliquot this in a single pipetting movement to individual samples to standardize across multiple samples.

Potential DNA Contaminants

Incorrect DNA quantification may result from DNA contamination, for example, by interference from superfluous nucleic acids in a sample (e.g., RNA, small nucleic acid fragments, nucleotides, single-stranded DNA), excess proteins, or other contaminating materials. DNA quality may also affect the quantity of usable DNA in a sample. For example, if the DNA is damaged (e.g., heavily nicked or containing extensive apurinic/apyrimidinic sites), then many of these fragments may fail during library preparation. High molecular weight dsDNA derived from host genomes can also interfere with accurate quantification. For example, bacterial artificial chromosomes (BACs) and other bacterially-derived plasmids usually contain a few percent of the chromosomal DNA from the host cells, despite the best purification efforts. These sequences may ultimately give rise to unwanted clusters on a flow cell lane. However, this contamination can be accurately quantified by analyzing aligned reads generated during sequencing against known bacterial sequences and subtracting these out. High molecular weight contamination may also be estimated prior to library preparation using qPCR assays designed to target unique chromosomal markers.

Temperature Considerations

Temperature is another important consideration for making genomic DNA libraries. Elevated temperatures should be particularly avoided in the steps preceding the adapter ligation. DNA fragments that have a high AT content are more likely to denature into single strands than GC-rich fragments, which can result in an increased probability of creating a bias in the sequencing coverage. As a general rule, libraries should be kept at temperatures below ≤37°C. Temperature is less of an issue after the adapters have been ligated onto the ends of the DNA, although care should be taken not to denature the library prior to the agarose gel electrophoresis process, because single-stranded DNA has a different migration rate.

DNA Input Recommendations

Input DNA Quantification

The ultimate success or failure of a library preparation strongly depends on using an accurately quantified amount of input DNA, particularly when starting quantities are <1 µg. Thus, correct quantification of gDNA is essential. This protocol is optimized for 2-5 µg input DNA.

Assessing DNA Quantity and Quality

Absorbance measurements at 260 nm are commonly used to quantify DNA. The ratio of absorbance at 260 nm to absorbance at 280 nm is used as an indication of sample purity, and values of 1.8–2.0 are considered indicative of relatively pure DNA. However, both measurements can be compromised by the presence of RNA or small nucleic acid fragments such as nucleotides. Thus, gDNA samples should be carefully collected to ensure that they are free of contaminants, and the most accurate spectrophotometric method available should be used to quantify the input gDNA.

DNA quantification methods that rely on intercalating fluorescent dyes measure only double-stranded DNA and are less subject to excess nucleic acids. However, these methods require the preparation of calibration curves and are highly sensitive to pipetting error. Ensure that pipettes are correctly calibrated and are not used at the volume extremes of their performance specifications.

Gel electrophoresis is a powerful means for revealing the condition (even the presence or absence) of DNA in a sample. Impurities, such as detergents or proteins, can be revealed by smearing of DNA bands. RNA, which interferes with 260 nm readings, is often visible at the bottom of a gel. A ladder or smear below a band of interest may indicate nicking or other damage to DNA. Where possible, or necessary, a gel should be run to assess the condition of the DNA sample.

Consistency of Results

Given that fluorescent methods only measure double stranded DNA and 260 nm determinations measure both single and double stranded DNA, fluorescence analysis can be performed to assess the sample denaturation state (i.e. single or double stranded). Sample preparation can begin with a fixed amount of double-stranded DNA prior to fragmentation and cleanup, and thereafter, can be measured by 260 nm determination. If the DNA amount is less than expected, the amount of input DNA can be adjusted before proceeding to the end-repair steps. A further validation step can be performed by analyzing an aliquot for the presence of contaminants by electrophoresis or using an automated instrument, such as the Agilent Bioanalyzer.

Consumables and Equipment

Check to ensure that you have all of the necessary user-supplied consumables and equipment before proceeding to sample preparation.

Table 1 User-Supplied Consumables

Consumable	Supplier
0.2 ml nuclease-free thin-wall PCR tubes	USA Scientific, catalog # 1402-4700
2.0 ml DNA LoBind tubes	Eppendorf, catalog # 022431048
50 bp Redi-Load DNA ladder	NEB, catalog # N3231L
50X TAE Buffer	BIO-RAD, catalog # 161-0743
100 bp Redi-Load DNA ladder	NEB, catalog # N3236L
Certified low range ultra agarose	BIO-RAD, catalog # 161-3106
DNeasy Kit (or equivalent)	QIAGEN, catalog # : 69581 (4 samples) 69582 (12 samples) 69504 (50 samples) 69506 (250 samples)
EpiTect Bisulfite Kit	QIAGEN, catalog # 59104
Ethanol	General lab supplier
Ethidium Bromide	Sigma, catalog # E1510
GeneCatchers	Gel Company, catalog # PKB4.0 or PKB6.5
Loading buffer (50 mM Tris pH 8.0, 40 mM EDTA, 40% (w/v) sucrose)	General lab supplier
Methylation Adapter Oligo Kit	Illumina, catalog # ME-100-0010
MinElute PCR Purification Kit	QIAGEN, catalog # 28004
MspI (100U/µl)	NEB, catalog # R0106M
Paired-End Sample Prep Kit	Illumina, catalog # PE-102-1001 (10 samples) or PE-102-1002 (40 samples)
PfuTurbo Cx Hotstart DNA Polymerase	Stratagene, catalog # 600410
Purified DNA (2–5 µg, 2 µg recommended) DNA should be as intact as possible, with an OD260/280 ratio of 1.8–2.0	General lab supplier
QIAquick Gel Extraction Kit	QIAGEN, catalog # 28704
QIAquick PCR Purification Kit	QIAGEN, catalog # 28104
Ultra Pure Water	General lab supplier

Table 2 User-Supplied Equipment

Equipment	Supplier
[Optional] 2100 Bioanalyzer Desktop System	Agilent, part # G2940CA
[Optional] Agilent DNA 1000 Kit	Agilent, part # 5067-1504
Benchtop microcentrifuge	General lab supplier
Benchtop centrifuge with swing-out rotor (e.g., Sorvall Legend RT)	General lab supplier
Dark Reader transilluminator or a UV transilluminator	Clare Chemical Research, catalog # D195M
Electrophoresis unit	General lab supplier
Gel trays and tank	General lab supplier
Thermal cycler	General lab supplier

Purify DNA

This procedure uses the methyl insensitive enzyme MspI to digest gDNA at the CCGG site.

User-Supplied Consumables

- DNA Sample (2–5 µg)
- MspI
- NEB4
- QIAGEN DNeasy Kit (or equivalent)
- QIAquick PCR Purification Kit (QIAGEN, catalog # 28104)
- Ultra Pure Water

Procedure

1. Prepare 2–5 µg of gDNA using a QIAGEN DNeasy or equivalent kit.
2. Prepare the reaction mix in a thermal cycler or water bath:

Reagent	Volume
NEB4	10 µl
MspI	4 µl
DNA	2–5 µg
Ultra Pure Water	to increase total volume to 100 µl
Total Volume	**100 µl**

3. Mix gently, but thoroughly, and centrifuge briefly.
4. Incubate on the thermal cycler for at least 18 hours at 37°C.
5. Follow the instructions in the QIAquick PCR Purification Kit to purify on one QIAquick column, eluting in 30 µl of QIAGEN EB.

SAFE STOPPING POINT
If you do not plan to proceed to *Perform End Repair* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

Perform End Repair

This procedure converts the overhangs resulting from restriction digestion into blunt ends using T4 DNA polymerase and Klenow enzyme. In addition, T4 PNK in this reaction phosphorylates the 5' ends of the DNA fragments.

User-Supplied Consumables

- Paired-End Sample Prep Kit content:
  - 10 mM dNTP Mix
  - Klenow Enzyme
  - T4 DNA Ligase Buffer with 10 mM ATP
  - T4 DNA Polymerase
  - T4 PNK
  - Ultra-pure Water
- QIAquick PCR Purification Kit (QIAGEN, catalog # 28104)

Procedure

1. Prepare the reaction mix on ice in the following order:

Reagent	Volume (µl)
Water	45
DNA Sample	30
T4 DNA Ligase Buffer with 10 mM ATP	10
10 mM dNTP Mix	4
T4 DNA Polymerase	5
Klenow Enzyme	1
T4 PNK	5
Total Volume	**100**

2. Mix gently, but thoroughly, and centrifuge briefly.
3. Incubate on the thermal cycler for 30 minutes at 20°C.
4. Follow the instructions in the QIAquick PCR Purification Kit to purify on one QIAquick column, eluting in 32 µl of QIAGEN EB.

SAFE STOPPING POINT
If you do not plan to proceed to *Adenylate 3' Ends* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

Adenylate 3' Ends

A single ‘A’ nucleotide is added to the 3’ ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction. A corresponding single ‘T’ nucleotide on the 3’ end of the adapter provides a complementary overhang for ligating the methylated adapter to the fragment. This strategy ensures a low rate of chimera (concatenated template) formation.

User-Supplied Consumables

- MinElute PCR Purification Kit (QIAGEN, catalog # 28004)
- Paired-End Sample Prep Kit content:
  - Klenow Buffer
  - 1 mM dATP
  - Klenow Exo -

NOTE

This process requires a QIAquick MinElute column rather than a normal QIAquick column.

Procedure

1. Prepare the reaction mix on ice in the following order:

Reagent	Volume (µl)
DNA Sample	32
Klenow Buffer	5
1 mM dATP	10
Klenow Exo -	3
Total Volume	**50**

2. Incubate on the thermal cycler for 30 minutes at 37°C.
3. Follow the instructions in the MinElute PCR Purification Kit to purify on one QIAquick MinElute column, eluting in 10 µl of QIAGEN EB.

Ligate Adapters

This process ligates adapters to the ends of the DNA fragments. The reaction adds distinct sequences to the 5′ and 3′ ends of each strand in the genomic fragment. Later in the workflow, additional sequences are added by tailed primers during PCR. These additional sequences are necessary for library amplification on the flow cell during cluster formation.

User-Supplied Consumables

- Paired-End Sample Prep Kit content:
  - DNA Ligase Buffer, 2X
  - DNA Ligase
- Methylation Adapter Oligo Kit content:
  - Methylated Adapter Oligo
- MinElute PCR Purification Kit (QIAGEN, catalog # 28004)

Procedure

A molar excess of adapter to fragments is used to increase the yield of adapter ligation to both ends of the DNA fragments. If the ratio is too high, then the yield of adapter dimers also increases, for example, if 2 µg or less of input DNA is used instead of 5 µg. The volume of adapter oligo mix added in the following procedure is recommended for an initial input DNA quantity of 2 µg.

1 Prepare the reaction mix on ice in the following order:

Reagent	Volume (µl)
DNA Sample	10
Water	6
DNA Ligase Buffer 2X	25
Methylation Adapter Oligo	4
DNA Ligase	5
Total Volume	**50**

2 Incubate on the thermal cycler for 30 minutes at 20°C.

3 Follow the instructions in the MinElute PCR Purification Kit to purify on one QIAquick MinElute column, eluting in 10 µl of QIAGEN EB.

SAFE STOPPING POINT

If you do not plan to proceed to *Purify Ligation Products* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

Purify Ligation Products

This process purifies the products of the ligation reaction on a gel and removes unligated adapters, removes any adapters that may have ligated to one another, and selects a narrow 150–175 bp and 175–225 bp size-range of DNA fragments for bisulfite treatment.

CAUTION
Follow the exact 150–175 bp and 175–225 bp size selection range and methods stated in this procedure to ensure reproducibility.

NOTE
Test your electrophoresis unit in advance to ensure that you can readily resolve DNA in the range below 1000 base pairs. The DNA smear should be sufficiently resolved to enable you to excise a narrow band of a chosen size with a standard deviation as low as 5% of the median (i.e., a gel slice at 400 bp, where +/- one standard deviation is equivalent to a size range of 380–420 bp). The conditions described are typical gel electrophoresis conditions which have been validated by Illumina.

Size Selection

Perform gel electrophoresis and band excision after adapter ligation to remove excess adapter and adapter dimers and to tighten the range of fragment sizes. Ligation reaction products are separated on an agarose gel and a ~2 mm wide gel slice containing DNA of the desired size is excised.

NOTE
Cutting a band of 150–175 bp and 175–225 bp on a 2% agarose gel will result in an insert size of approximately 30–100 bp and 60–160 bp size distribution of two libraries respectively and empirically, accounting for the influence of the adapters on the gel mobility.

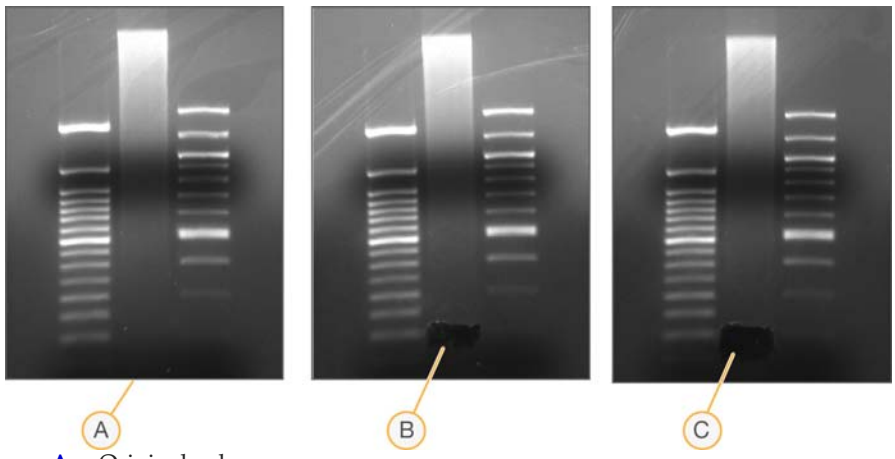
User-Supplied Consumables

- 2.0 ml DNA LoBind Tubes (2)
- Certified Low Range Ultra Agarose
- 50x TAE Buffer
- Distilled Water
- Ethidium Bromide
- Gene Catchers (2)
- Loading Buffer
- 100 bp Redi-Load DNA Ladder
- 50 bp Redi-Load DNA Ladder
- QIAquick Gel Extraction Kit (QIAGEN, catalog # 28704)

Procedure

Referenced the following figure while performing these procedures to extract gel slices:

Figure 2 Excised Gel Slices



A Original gel
B First gel slice at 175–225 bp
C Second gel slice at 150–175 bp

CAUTION
Do not purify multiple samples on a single gel due to the risk of cross-contamination between libraries.

NOTE
It is important to excise as accurate a band as possible from the gel during gel purification. Use a Dark Reader to visualize DNA on agarose gels.

1. Prepare a 50 ml, 2% agarose gel with distilled water and TAE according to the manufacturer's instructions. The final concentration of TAE should be 1X.
2. Add Ethidium Bromide (EtBr) after the TAE-agarose has cooled. The final concentration of EtBr should be 400 ng/ml (i.e., add 20 µg EtBr to 50 ml of 1X TAE-agarose).
3. Cast the gel using a comb that can accommodate 56 µl in each well.
   Recommended well size: 1 mm (length) x 8 mm (width) x 7 mm (height).
4. Add 5 µl of Loading Buffer to 250 µl of the 50 bp Redi-Load Ladder.
5. Add 5 µl of Loading Buffer to 250 µl of the 100 bp Redi-Load Ladder.
6. Add 2 µl of Loading Buffer to 10 µl of the DNA from the purified ligation reaction.
7. Load 10 µl of the ladders onto one lane of the gel, as shown in Figure 2.
8. Load the entire sample onto another lane of the gel, as shown in Figure 2.

NOTE
Flanking the library on both sides with ladders may make the library excision easier.

9. Run gel at 120 V for 10 minutes, then 60 V for 120 minutes (6 V/cm).

10 View the gel on a Dark Reader transilluminator or a UV transilluminator.

11 Photograph the gel before a slice is excised.

12 Place a GeneCatcher vertically above the sample in the gel at the desired size of the template.

13 Excise a 2 mm slice of the sample lane at exactly 175–225 bp using the markers as a guide. Use two Gene Catchers for this band range if needed.

14 Place the gel slice in a new 2.0 ml DNA LoBind tube.

15 Discard the GeneCatcher to avoid sample cross-contamination.

16 Photograph the gel after the slice was excised.

17 Place a new GeneCatcher vertically above the sample in the gel at the desired size of the template.

18 Excise a 2 mm slice of the sample lane at exactly 150–175 bp using the markers as a guide.

19 Place the gel slice in a new 2.0 ml DNA LoBind tube.

CAUTION
It is important to place the second fragment into a separate collection tube and keep the two size ranges separated in the subsequent bisulfite treatment, PCR, and cluster generation processes.

20 Photograph the gel after the second slice was excised.

21 Follow the instructions in the QIAquick Gel Extraction Kit to purify on one QIAquick column.Weigh the gel slice as indicated in the instructions, in order to calculate how much initial buffer to add, eluting in 20 µl of QIAGEN EB twice and combine them into one tube.

The final volume should be 40 µl for each band region.

SAFE STOPPING POINT
If you do not plan to proceed to *Enrich DNA Fragments* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

Enrich DNA Fragments

This process uses PCR to selectively enrich those DNA fragments that have methylated adapter molecules on both ends, add sequences necessary for the adapters to bind to the flowcell, and to amplify the amount of DNA in the library for accurate quantification. The PCR is performed with two primers that anneal to the ends of the adapters. The number of PCR cycles should be minimized to avoid skewing the representation of the library and the recommended numbers are as follows:

Input of DNA to Library Prep	Number of PCR Cycles
2 µg	12–15
5 µg	8–10

User-Supplied Consumables

- MinElute PCR Purification Kit (QIAGEN, catalog # 28004)
- Paired-End Sample Prep Kit content:
  - 10 mM dNTP Mix
  - PCR Primer PE 1.0
  - PCR Primer PE 2.0
  - Ultra Pure Water
- PfuTurbo Cx Hotstart DNA Polymerase Kit content:
  - PfuTurbo Cx Hotstart DNA Polymerase
  - Pfu Turbo Cx Reaction Buffer

Procedure

The volume of methyl-adapter ligated fragments added to the PCR reaction below are based on an initial input DNA quantity of 2 µg.

CAUTION
To avoid sample cross-contamination, set up PCR reactions (all components except the template DNA) in a designated clean area, preferably a PCR hood with UV sterilization and positive air flow.

1 Prepare the reaction mix on ice in a 200 µl thin wall PCR tube in the following order:

Reagent	Volume (µl)
DNA	15
Ultra Pure Water	25.75
Pfu Turbo Cx Reaction Buffer	5
10 mM dNTP Mix	1.25
PCR primer PE 2.0	1
PCR primer PE 1.0	1
PfuTurbo Cx Hotstart DNA Polymerase	1
Total Volume	**50**

2 Mix gently, but thoroughly, and centrifuge briefly.

3 Amplify using the following PCR process with a heated lid:

- **a** 5 minutes at 95°C
- **b** 30 seconds at 98°C
- **c** 15 cycles of:
  10 seconds at 98°C
  30 seconds at 65°C
  30 seconds at 72°C
- **d** 5 minutes at 72°C
- **e** Hold at 4°C

4 Follow the instructions in the MinElute PCR Purification Kit to purify on one QIAquick MinElute column, eluting in 15 µl of QIAGEN EB.

SAFE STOPPING POINT

If you do not plan to proceed to *Validate Library* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

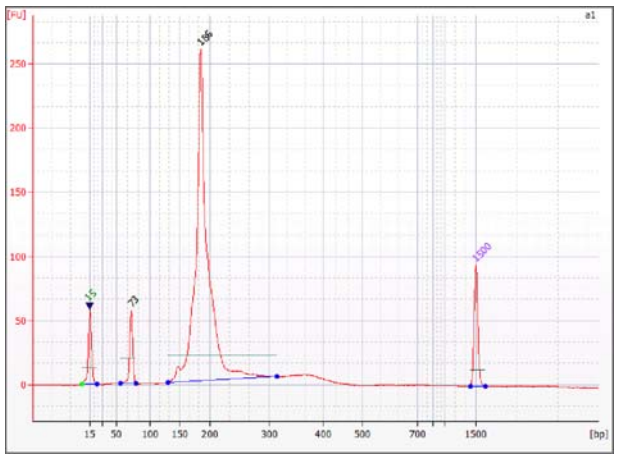
Validate Library

Perform the following quality control steps on your DNA library.

1. Check the template size distribution by running an aliquot of the library on a gel or an Agilent Bioanalyzer.
2. If validating by gel, load 10% of the volume of the library on a gel and check that the size range is as expected.
3. If using an Agilent Bioanalyzer, load 1 µl of each library from the two libraries from the same sample in different wells on an Agilent DNA 1000 chip.

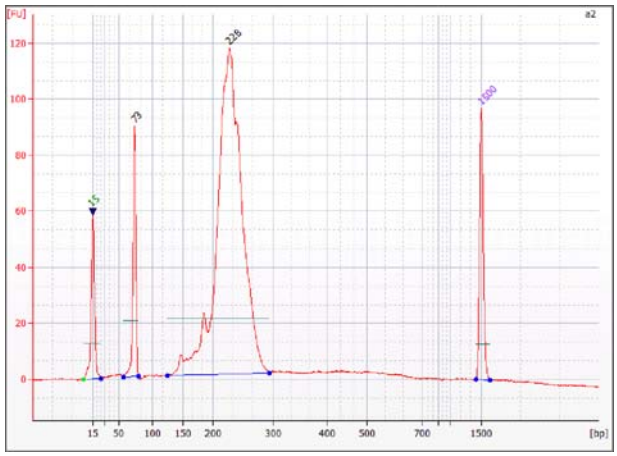
   Examples of each library prepared using this protocol are shown below. The small library has insert sizes ranging from 30–100 bp (i.e., a template size range of 150–220 bp).

Figure 3 150–175 bp RRBS Library Example



The large library has insert sizes ranging from 60–160 bp (i.e., a template size range of 180–280 bp).

Figure 4 175–225 bp RRBS Library Example



Library Quantification

In order to achieve the highest quality of data on Illumina sequencing platforms, it is important to create optimum cluster densities across every lane of every flow cell. This requires accurate quantification of DNA library templates. As described for sample input quantification, any method of measuring DNA concentration has certain advantages and potential drawbacks.

Optical Density or Intercalating Dye

Absorbance-based quantification can erroneously measure the presence of residual PCR primers or may not be sensitive enough to make an accurate measurement in circumstances where the final library yield is low. Measurements based on intercalating dyes such as PicoGreen may also be used, but care should be taken to minimize pipetting errors and to ensure that the library is not denatured to avoid inaccuracies.

Bioanalyzer

Automated analysis using an instrument like the Agilent Bioanalyzer can be used to quantify libraries and is an excellent alternative to traditional gel electrophoresis to determine the size of the library fragments. As with intercalating dye assays, single-stranded templates will not be detected nor contribute to the library quantification even though these templates will form clusters. Therefore, care must be taken to ensure that the library is not denatured and that pipetting errors are minimized.

qPCR

Quantitative real-time PCR (qPCR) is an alternative method of quantifying DNA that measures the relationship between the initial concentration of a template and how its concentration changes during progressing cycles of thermal amplification. qPCR quantification is not recommended for the RRBS for Methylation Analysis protocol due to complications in the library components.

Cluster Generation

It is essential to run a control lane on the same flowcell as the RRBS sample libraries. Because bisulfite treatment greatly decreases the C base in the genomic sequence, the resulting three based libraries are not optimal for base calls. Any genomic library containing balanced A,G,T, and C components (e.g., genomic resequencing, RNA-seq, ChIP-seq) could be served as a control lane for RRBS sample libraries.

It is highly recommended that you limit the cluster number for RRBS libraries to a 75–80% optimized cluster number for other genomic libraries, because of the unique CGG starting sequence in the first three bases.

DNA Template Storage

Store prepared DNA template at a concentration of 10 nM. Adjust the concentration for your prepared DNA samples (or pools of samples) to 10 nM using Tris-HCl 10 mM, pH 8.5. For long-term storage of DNA samples at a concentration of 10 nM, add Tween 20 to the sample to a final concentration of 0.1% Tween. This helps to prevent adsorption of the template to plastic tubes upon repeated freeze-thaw cycles, which would decrease the cluster numbers from a sample over time.