**Patents, Licensing and Trademarks**

© 2016 NuGEN Technologies, Inc. All rights reserved. The Encore®, Ovation® and Applause® families of products and methods of their use are covered by several issued U.S. and International patents and pending applications (www.nugen.com). NuGEN, Ovation, SPIA, Ribo-SPIA, Applause, Encore, Prelude, Mondrian and Imagine More From Less are trademarks or registered trademarks of NuGEN Technologies, Inc. Other marks appearing in these materials are marks of their respective owners.

The purchase of this product conveys to the buyer the limited, non-exclusive, non-transferable right (without the right to modify, reverse engineer, resell, repackage or further sublicense) under these patent applications and any patents issuing from these patent applications to use this product and methods, accompanying this user guide, for research and development purposes solely in accordance with the intended use described and the written instructions provided in this user guide. No license to make or sell products by use of this product is granted to the buyer whether expressly, by implication, by estoppels or otherwise. In particular, the purchase of this product does not include or carry any right or license to use, develop or otherwise exploit this product commercially and no rights are conveyed to the buyer to use the product or components of the product for purposes including commercial services or clinical diagnostics.

For information on purchasing a license to the NuGEN patents for uses other than in conjunction with this product or to use this product for purposes other than research, please contact NuGEN Technologies, Inc., 201 Industrial Road, Suite 310, San Carlos, CA 94070. Phone 888-654-6544 or 650-590-3600; FAX 888-296-6544 or 650-590-3630.

**Warranty**

NuGEN warrants that this product meets the performance standards described in the Company's product and technical literature for a period of six months from the date of purchase, provided that the product is handled and stored according to published instructions, and that the product is not altered or misused. If the product fails to meet these performance standards, NuGEN will replace the product free of charge or issue a credit for the purchase price. NuGEN's liability under this warranty shall not exceed the purchase price of the product. NuGEN shall assume no liability for direct, indirect, consequential or incidental damages arising from the use, results of use or inability to use its products. NuGEN reserves the right to change, alter or modify any product to enhance its performance and design.

NuGEN's products are developed, designed and sold FOR RESEARCH USE ONLY. This product is not to be used for diagnostic or therapeutic purposes, nor is it to be administered to humans or animals.

Except as expressly set forth herein, no right to modify, reverse engineer, distribute, offer to sell or sell NuGEN's product is conveyed or implied by buyer's purchase of this NuGEN product. The buyer agrees to use NuGEN products accompanying the product insert in accordance with the intended use and the written instructions provided.
## Contents

### I. Introduction
- A. Background .................................................................................................. 1
- B. Performance Specifications ........................................................................ 5
- C. Quality Control .......................................................................................... 5
- D. Storage and Stability .................................................................................. 6
- E. Safety Data Sheet (SDS) .............................................................................. 6

### II. Kit Components
- A. Reagents Provided .................................................................................... 7
- B. Additional Equipment, Reagents and Labware .......................................... 8

### III. Planning the Experiment
- A. Input DNA Requirements .......................................................................... 9
- B. Amplified Library Storage .......................................................................... 9
- C. Using the Ovation RRBS Methyl-Seq System 1–16 on Illumina NGS Platforms .......................................................................................... 9
- D. Index Structure and Index Read Recommendations .................................. 11
- E. Data Analysis and Parsing Multiplex Libraries .......................................... 12

### IV. Protocol
- A. Overview ................................................................................................... 14
- B. Protocol Notes ............................................................................................ 14
- C. Agencourt® Purification Beads ................................................................. 15
- D. Programming the Thermal Cycler ............................................................... 16
- E. MspI Digestion ........................................................................................... 17
- F. Ligation ........................................................................................................ 18
- G. Final Repair ................................................................................................. 19
- H. Bisulfite Conversion .................................................................................... 20
- I. Library Amplification ................................................................................... 21
- J. Amplified Library Purification ..................................................................... 22
- K. Quantitative and Qualitative Assessment of the Library ......................... 23

---

**Table of Contents**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>A. Background</td>
<td>1</td>
</tr>
<tr>
<td>B. Performance Specifications</td>
<td>5</td>
</tr>
<tr>
<td>C. Quality Control</td>
<td>5</td>
</tr>
<tr>
<td>D. Storage and Stability</td>
<td>6</td>
</tr>
<tr>
<td>E. Safety Data Sheet (SDS)</td>
<td>6</td>
</tr>
<tr>
<td>Kit Components</td>
<td>7</td>
</tr>
<tr>
<td>A. Reagents Provided</td>
<td>7</td>
</tr>
<tr>
<td>B. Additional Equipment, Reagents and Labware</td>
<td>8</td>
</tr>
<tr>
<td>Planning the Experiment</td>
<td>9</td>
</tr>
<tr>
<td>A. Input DNA Requirements</td>
<td>9</td>
</tr>
<tr>
<td>B. Amplified Library Storage</td>
<td>9</td>
</tr>
<tr>
<td>C. Using the Ovation RRBS Methyl-Seq System 1–16 on Illumina NGS Platforms</td>
<td>9</td>
</tr>
<tr>
<td>D. Index Structure and Index Read Recommendations</td>
<td>11</td>
</tr>
<tr>
<td>E. Data Analysis and Parsing Multiplex Libraries</td>
<td>12</td>
</tr>
<tr>
<td>Protocol</td>
<td>14</td>
</tr>
<tr>
<td>A. Overview</td>
<td>14</td>
</tr>
<tr>
<td>B. Protocol Notes</td>
<td>14</td>
</tr>
<tr>
<td>C. Agencourt® Purification Beads</td>
<td>15</td>
</tr>
<tr>
<td>D. Programming the Thermal Cycler</td>
<td>16</td>
</tr>
<tr>
<td>E. MspI Digestion</td>
<td>17</td>
</tr>
<tr>
<td>F. Ligation</td>
<td>18</td>
</tr>
<tr>
<td>G. Final Repair</td>
<td>19</td>
</tr>
<tr>
<td>H. Bisulfite Conversion</td>
<td>20</td>
</tr>
<tr>
<td>I. Library Amplification</td>
<td>21</td>
</tr>
<tr>
<td>J. Amplified Library Purification</td>
<td>22</td>
</tr>
<tr>
<td>K. Quantitative and Qualitative Assessment of the Library</td>
<td>23</td>
</tr>
</tbody>
</table>
V. Technical Support .................................................................................................................. 24

VI. Appendix .................................................................................................................................. 25
   A. Barcode Sequences and Guidelines for Multiplex Experiments ............................. 25
   B. PCR Amplification When Using Degraded DNA or Less Than 100 ng of
      High-Quality DNA .............................................................................................................. 26
   C. Data Analysis of Ovation RRBS Methyl-Seq Libraries ........................................... 29
   D. Diversity Trimming Examples ....................................................................................... 32
   E. Effects of Read Length on Mapping Rate ...................................................................... 34
   F. Frequently Asked Questions (FAQs) ............................................................................ 36
   G. Update History .................................................................................................................... 38
A. Background

Reduced representation bisulfite sequencing (RRBS) is a technique used to generate single base resolution DNA methylation (5-methylC) information across a genomic sample. By analyzing a reduced representation of the genome, the amount of sequencing required is greatly reduced relative to whole genome bisulfite sequencing. Since the RRBS technique was first described (Meissner, et al. (2005) Nucleic Acids Res 33(18):5868), there have been many enhancements and improvements (see for example, Boyle, et al. (2012) Genome Biol 13:R92). The current approach utilizes the methylation insensitive restriction enzyme MspI, which recognizes CCGG. As a result of partial fragmentation during bisulfite conversion, PCR, and efficiency of cluster generation, only a subset of these fragments, typically under 300 bp in length, are sequenced. Figure 1 is a Bioanalyzer trace of MspI digested human genomic DNA. Note how only a small amount of the digested DNA is represented in the under 300 bp fraction. However these smaller fragments are derived from genomic DNA that has a high frequency of MspI sites and therefore a high frequency of potential CpG methylation sites. This is how RRBS achieves reduced representation.

Figure 1. Bioanalyzer trace of MspI digested human genomic DNA.

Figure 2 follows a hypothetical MspI fragment as it is processed with the Ovation RRBS Methyl-Seq System. The digested fragments are directly ligated to cohesive-ended adaptors, without the requirement for blunting or A-tailing. This both simplifies the workflow and increases efficiency. MspI enzyme is still active during ligation. Therefore, any genomic fragments that ligate to each other will be re-cleaved. Since adaptor ligation does not recreate an MspI site, these desired ligation products are stable.
I. Introduction

Figure 2. MspI fragment during processing with the Ovation RRBS Methyl-Seq System 1–16.

5’ …CCGG(OT)CCGG… 3’
3’ …GGCC(OB)GGCC… 5’

MspI digestion

5’ pCGG(OT)C 3’
3’ C(OB)GGp 5’

Adaptor ligation

5’ (Ad)CGG(OT)CGG(Ad) 3’
3’ (Ad)GCC(OB)GCC(Ad) 5’

DNA sequence derived from library construction is shown in blue. Ad = Adaptor

Bisulfite conversion

5’ (Ad)YGG(BCOT)TCG(Ad) 3’
and
3’ (Ad)GCT(BCOB)GGY(Ad) 5’

BCOT = Bisulfite Converted Original Top genomic strand
BCOB = Bisulfite Converted Original Bottom genomic strand
Y = C or T, depending on methylation status

CG motifs that may contain 5-methyl C are underlined.

Two library molecules from each fragment are possible, one from each strand. Note that after bisulfite conversion, they are non-complementary. Also note that for each MspI fragment, the methylation information for the CpG in the MspI site to the left is preserved in the library molecule derived from the top strand, whereas the methylation information for the CpG in the MspI site to the right is preserved in the library molecule derived from the bottom strand.

Traditionally, PCR duplicates are identified in libraries made from randomly fragmented inserts by mapping inserts to the genome and discarding any paired end reads that share the same genomic coordinates. This approach doesn’t work for restriction digested samples, such as RRBS, because all fragments mapping to a genomic location will share the same ends. NuGEN provides a way to identify unique molecules in RRBS libraries by adding an additional 6 bases of random sequence after the 6 base barcode. See Figure 3 for a schematic of the Ovation RRBS Methyl-Seq System library molecule. If you wish to utilize this PCR duplicate marking feature, increase the index read from 6 to 12 nucleotides, then use the NuGEN-provided Duplicate Marking tool, NuDup, to identify and discard any PCR duplicates found. The tool works by looking at the random N6 sequence associated with all reads that map to the same genomic location. Any reads that both map to the same genomic location and contain the same N6 sequence are considered duplicates, and all but one copy are removed.
During sequencing on Illumina instruments, software identifies clusters over the first several cycles of sequencing. During sequencing of normal, high diversity libraries, overlapping clusters can be distinguished because they are different colors. If overlapping clusters contain the same sequence during the first few cycles, they may be mistaken as a single cluster. When the two sequences eventually diverge in a mixed cluster, low quality and/or incorrect base calls are the result. In the RRBS strategy, all reads begin with either CGG or TGG, depending on methylation state. This makes sequencing traditional RRBS libraries on Illumina instruments challenging. The Ovation RRBS Methyl-Seq System 1–16 overcomes this challenge by inserting a variable number of random bases at the beginning of both forward and reverse reads. The variable number of bases dephases the clusters across the image tile such that all four bases are present in every cycle. The random nature of the sequences reduces the chance of overlapping clusters containing the same initial sequence and being miscalled as a single cluster. As shown in Figure 4, the added diversity is not completely random. Instead, it is carefully chosen such that it can be identified and removed from the sequencing read using the NuGEN diversity trimming script. This preserves the first base of the Mspl fragment, which contains a CpG methylation measurement. The RRBS approach results in a very efficient use of sequencing capacity, with more than 1 CpG measurement per read. The enhancements provided by the Ovation RRBS Methyl-Seq System 1–16 deliver that same efficiency with improved workflow, higher quality base calls at higher cluster density, and the ability to identify unique molecules.
I. Introduction

Figure 4. Types of sequence diversity added to the beginning of forward reads.

Blue = DNA sequence derived from library construction  
Y = C or T, depending on methylation status  
D = A, G, or T  
R = A or G

The Ovation RRBS Methyl-Seq System 1–16 provides a simple, fast and scalable solution for producing directional reduced representation bisulfite-converted libraries for next generation sequencing.

As shown in Figure 5 the streamlined workflow consists of five main steps: MspI digestion, adaptor ligation, final repair, bisulfite conversion, and PCR amplification to produce the final library. The entire workflow can be completed in a single day, and yields DNA libraries ready for cluster formation and either single read or paired-end sequencing. The kit provides sufficient reagents to generate 32 libraries, two libraries each from 16 different barcodes. The barcodes have been carefully selected so that all 16 can be sequenced in eight color balanced duplexes. In addition, the barcodes are edit distance 3 from each other, which means they can be parsed with one mismatch, further increasing data yield per run.
I. Introduction

Figure 5. Schematic of the Ovation RRBS Methyl-Seq System 1–16 workflow.

Genomic DNA

MspI digestion

Add adaptors and ligate

Final repair

Bisulfite conversion

PCR amplification

Cluster formation and sequencing

B. Performance Specifications

The Ovation RRBS Methyl-Seq System 1–16 is designed to produce RRBS libraries from human genomic DNA suitable for either single read or paired-end sequencing on Illumina NGS platforms without gel-based size selection, using 100 ng of high-quality human genomic DNA as input. This system generates libraries suitable for cluster generation in a single day.

C. Quality Control

Every lot of the Ovation RRBS Methyl-Seq System 1–16 undergoes functional testing to meet specifications for library generation performance.
I. Introduction

D. Storage and Stability
The Ovation RRBS Methyl-Seq System 1–16 is shipped on dry ice and should be unpacked immediately upon receipt.

Note: This product contains components with multiple storage temperature requirements.

The vials labeled Agencourt® Beads (clear cap) should be removed from the top of the shipping carton upon delivery and stored at 4°C.

All other components should be stored at –20°C on internal shelves of a freezer without a defrost cycle. The kit has been tested to perform to specifications after as many as six freeze/thaw cycles. Kits handled and stored according to the above guidelines will perform to specifications for at least six months.

Important: Do not warm Ligation Adaptor Mixes above room temperature. Heating will severely degrade performance.

E. Safety Data Sheet (SDS)
If appropriate, an SDS for this product is available on the NuGEN website at www.nugen.com/products/ngs/ovation-rrbs-methyl-seq-system.
## A. Reagents Provided

### Table 1. Ovation RRBS Methyl-Seq System 1–16 reagents (Part No. 0353)

<table>
<thead>
<tr>
<th>COMPONENT</th>
<th>0353 PART NUMBER</th>
<th>VIAL LABEL</th>
<th>VIAL NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>MspI Buffer Mix</td>
<td>S02137</td>
<td>Blue</td>
<td>—</td>
</tr>
<tr>
<td>MspI Enzyme Mix</td>
<td>S02138</td>
<td>Blue</td>
<td>—</td>
</tr>
<tr>
<td>Ligation Buffer Mix</td>
<td>S01534</td>
<td>Yellow</td>
<td>L1 VER 4</td>
</tr>
<tr>
<td>Ligation Adaptor Mixes</td>
<td></td>
<td>Yellow</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S02140</td>
<td></td>
<td>L2V20DR-BC1</td>
</tr>
<tr>
<td></td>
<td>S02141</td>
<td></td>
<td>L2V20DR-BC2</td>
</tr>
<tr>
<td></td>
<td>S02142</td>
<td></td>
<td>L2V20DR-BC3</td>
</tr>
<tr>
<td></td>
<td>S02143</td>
<td></td>
<td>L2V20DR-BC4</td>
</tr>
<tr>
<td></td>
<td>S02144</td>
<td></td>
<td>L2V20DR-BC5</td>
</tr>
<tr>
<td></td>
<td>S02145</td>
<td></td>
<td>L2V20DR-BC6</td>
</tr>
<tr>
<td></td>
<td>S02146</td>
<td></td>
<td>L2V20DR-BC7</td>
</tr>
<tr>
<td></td>
<td>S02147</td>
<td></td>
<td>L2V20DR-BC8</td>
</tr>
<tr>
<td></td>
<td>S02148</td>
<td></td>
<td>L2V20DR-BC9</td>
</tr>
<tr>
<td></td>
<td>S02149</td>
<td></td>
<td>L2V20DR-BC10</td>
</tr>
<tr>
<td></td>
<td>S02150</td>
<td></td>
<td>L2V20DR-BC11</td>
</tr>
<tr>
<td></td>
<td>S02151</td>
<td></td>
<td>L2V20DR-BC12</td>
</tr>
<tr>
<td></td>
<td>S02152</td>
<td></td>
<td>L2V20DR-BC13</td>
</tr>
<tr>
<td></td>
<td>S02153</td>
<td></td>
<td>L2V20DR-BC14</td>
</tr>
<tr>
<td></td>
<td>S02154</td>
<td></td>
<td>L2V20DR-BC15</td>
</tr>
<tr>
<td></td>
<td>S02155</td>
<td></td>
<td>L2V20DR-BC16</td>
</tr>
<tr>
<td>Ligation Enzyme Mix</td>
<td>S01535</td>
<td>Yellow</td>
<td>L3 VER 4</td>
</tr>
<tr>
<td>Final Repair Buffer Mix</td>
<td>S02139</td>
<td>Purple</td>
<td>FR1 VER 4</td>
</tr>
<tr>
<td>Final Repair Enzyme Mix</td>
<td>S01569</td>
<td>Purple</td>
<td>FR2</td>
</tr>
<tr>
<td>Amplification Primer Mix</td>
<td>S01739</td>
<td>Red</td>
<td>P2 VER 8</td>
</tr>
<tr>
<td>Amplification Enzyme Mix</td>
<td>S01740</td>
<td>Red</td>
<td>P3 VER 3</td>
</tr>
<tr>
<td>25 μM Sequencing Primer</td>
<td>S01757</td>
<td>Orange</td>
<td>MetSeq Primer 1</td>
</tr>
<tr>
<td>Nuclease-free Water</td>
<td>S01001</td>
<td>Green</td>
<td>D1</td>
</tr>
<tr>
<td>Agencourt Beads</td>
<td>S01502</td>
<td>Clear</td>
<td>—</td>
</tr>
<tr>
<td>DNA Resuspension Buffer</td>
<td>S02156</td>
<td>Clear</td>
<td>DR1</td>
</tr>
</tbody>
</table>
II. Kit Components

B. Additional Equipment, Reagents and Labware

Required Materials

- **Equipment**
  - Agilent 2100 Bioanalyzer and High Sensitivity DNA Kit or materials and equipment for electrophoretic analysis of nucleic acids
  - Qubit® 2.0 Fluorometer and dsDNA HS Assay Kit (Life Technologies)
  - Microcentrifuge for individual 1.5 mL and 0.5 mL tubes
  - 0.5–10 µL pipette, 2–20 µL pipette, 20–200 µL pipette, 200–1000 µL pipette
  - Vortexer
  - Thermal cycler with 0.2 mL tube heat block, heated lid, and 100 µL reaction capacity

- **Reagents**
  - EpiTect Fast DNA Bisulfite Kit (QIAGEN Cat. #59824 for 50 preps or Cat. #59826 for 200 preps)
  - Ethanol (Sigma-Aldrich, Cat. #E7023), for purification steps
  - OPTIONAL: EvaGreen® Dye, 20X in water (Biotium, Cat. #31000)

- **Supplies and Labware**
  - Nuclease-free pipette tips
  - 1.5 mL and 0.5 mL RNase-free microcentrifuge tubes
  - 8 X 0.2 mL strip PCR tubes or 0.2 mL thin-wall PCR plates (recommended) or 0.2 mL individual thin-wall PCR tubes
  - Magnetic separation device for the amplified library purification
  - Disposable gloves
  - Kimwipes
  - Ice bucket
  - Cleaning solutions such as DNA-OFF™ (MP Biomedicals, Cat. #QD0500)
  - OPTIONAL: PhiX Control (Illumina, Cat. #FC-110-3001)
  - OPTIONAL: Real-time PCR system

To Order:

- Affymetrix, www.affymetrix.com
- Biotium, www.biotium.com
- Illumina, www.illumina.com
- Life Technologies, www.lifetechnologies.com
- MP Biomedicals, www.mpbio.com
- QIAGEN, www.qiagen.com
- Sigma-Aldrich, Inc., www.sigmaaldrich.com
III. Planning the Experiment

A. Input DNA Requirements

The Ovation RRBS Methyl-Seq System 1–16 is designed to work with inputs of 100 ng of intact human genomic DNA. Quantitation of gDNA by a dsDNA assay, such as PicoGreen® or the Qubit System, is highly recommended. Do not rely solely on OD260 readings for quantitation of input material. DNA samples must be free of contaminating proteins, RNA, organic solvents (including phenol and ethanol) and salts. We recommend using a commercially available system for gDNA isolation. The A260:A280 ratio for DNA samples should be in excess of 1.8. Using DNA samples with lower ratios may compromise your results.

Although not recommended, it is possible to generate RRBS libraries from less than 100 ng of gDNA, or from degraded gDNA, such as DNA extracted from formalin fixed, paraffin embedded (FFPE) specimens. If you choose to proceed with such samples, Library Amplification will require additional cycles of PCR. As a general guideline, use one additional cycle of PCR for each two fold decrease in starting material. To accurately determine the number of PCR cycles to perform, follow the optional real-time PCR protocol in Appendix B.

You may observe an increase in adaptor artifacts, as well as decreased library complexity, in libraries made from less than 100 ng or degraded samples that require additional cycles of PCR amplification. Adaptor artifacts typically appear as a peak of approximately 145 bp. These artifacts can be reduced by performing a second bead purification. Add 1 volume of beads to the purified library, and follow the amplified library purification protocol.

Signs of reduced complexity should be monitored by sequencing the unique 6 random bases appended to the index on each reverse adaptor molecule, then using the Duplicate Marking tool (NuDup; see Appendix C) to identify unique molecules. After removing true PCR duplicates you will have a reliable measure of library complexity. Therefore, sequencing the additional 6 bases after the 6-base barcode is especially useful when attempting RRBS on rare or degraded samples. See Section III. D. for information on index structure and sequencing, and Appendix C for information on duplicate marking and removal.

B. Amplified Library Storage

Amplified libraries may be stored at –20°C.

C. Using the Ovation RRBS Methyl-Seq System 1–16 on Illumina NGS Platforms

The Ovation RRBS Methyl-Seq System 1–16 uses the same approach to multiplexing found in the standard Illumina method. These libraries should be sequenced using the Illumina protocol for multiplex sequencing. The DR barcode sequences, found in Appendix A, must be entered into the Illumina software prior to analysis.
III. Planning the Experiment

**Important Note:** The design of the Ovation RRBS Methyl-Seq System 1–16 requires the use of a custom Read 1 sequencing primer, MetSeq Primer 1, included in this kit at a concentration of 25 µM. Sequencing with custom primers may not be supported on all Illumina platforms. Please follow the custom primer recommendations for your specific sequencer. The Standard Read 1 Primer is also required when multiplexing with PhiX or other libraries. The standard Illumina primers are used for the index, and if desired, reverse reads.

The barcode sequences used in this kit were carefully chosen for their ability to parse properly and for color balancing. Therefore, strict requirements exist for choosing barcoded libraries for a multiplexed sequencing run. Refer to Appendix A, Table 7 for multiplexing guidelines.

If you wish to use the Duplicate Marking feature built into the Ovation RRBS Methyl-Seq System, add an additional 6 nt to the index read. For more information on this feature, see Appendix C., Data Analysis.

The Ovation RRBS Methyl-Seq System 1–16 produces directional bisulfite-converted libraries. This means Read 1 will sequence the C-to-T converted strand and the overall nucleotide balance for Read 1 will show a low proportion of C bases. Illumina has recommendations for how to obtain high-quality base calls from libraries containing unbalanced nucleotide ratios. For the highest quality results, follow the low-diversity sequencing recommendations for your specific instrument. If you are sequencing on a HiSeq or MiSeq, make sure your instrument is running the following software versions or later:

- HiSeq — HCS v2.2.38 (includes RTA v1.18.61)
- MiSeq — MSR 2.6.2 or higher (RTA v1.17.28)

These sequencing recommendations, in combination with the sequence diversity incorporated in the Ovation RRBS Methyl-Seq System adaptors, can produce high-quality RRBS reads on compatible platforms without the need to spike in balanced library such as PhiX. However, spiking in 5% PhiX or another, previously characterized high-quality balanced library can be useful for troubleshooting purposes in the event of a failed run, while only reducing the data output by 5%. Such a control can be used to tell if the sequencing run failed in general, or if there is a problem specific to the particular RRBS library being sequenced. Please refer to the Illumina technical support for your instrument for more information.

One of the principle benefits of the RRBS method is that the first base of the read contains a CpG methylation measurement. In principle, reads need only be long enough to accurately map them to the genome. However, mapping rates are directly affected by sequence read length, and use of longer reads can lead to more uniquely mapping reads and coverage of a greater number of CpG loci. Please refer to Appendix E. Effects of Read Length on Mapping Rate, for sample data demonstrating the differences in alignment by read length.
III. Planning the Experiment

D. Index Structure and Index Read Recommendations

The RRBS System uses 6-base barcodes for sample multiplexing. The 6-base barcode sequences can be found in Appendix A. In addition to the 6-base barcode, the adapter contains 6 random bases immediately following the 6-base barcode, for a total of 12 bases. The additional 6 bases are used for duplicate read determination using the Duplicate Marking Tool (see Optional Duplicate Determination in Appendix C). To take advantage of this feature, the libraries should be sequenced using 12 cycles for the index read.

Parsing and Generating an index.fastq File For De-duplication

Illumina does not provide a simple way to obtain the sequence information contained in the 12-base pair index read including the 6 random bases that are necessary for duplicate read determination. Several recommended methods to generate the necessary index fastq file are provided below.

MiSeq Instrument

Parsing multiplex runs using the MiSeq built-in Illumina software replaces the barcode sequence from each library with a numerical substitute, which removes the duplicate information provided by the N6 sequence present after the barcode. To retrieve this information using the MiSeq instrument use one of the options given below.

Option 1: (Recommended) contact Illumina Technical Support and request a modification of the MiSeq config file to allow generation of an index fastq file during data analysis.

Option 2: Modify the MiSeq config file to allow generation of an index fastq file during data analysis:

1. Stop the MiSeq Reporter process.
2. Locate the “MiSeq Reporter.exe.config” file located in C:/Illumina/MiSeq Reporter
3. Open config file and search for a line that reads:
   
   `<add key="CreateFastqForIndexReads" value="0"/>
   
   • If this line is present, change the value from “0” to “1”.
   • If this line is not present, add the line to the config file using the add keys function under the appSettings tab with the value set as “1”.
4. Restart the MiSeq reporter process.
5. Re-queue the run for data analysis if required. The 6-base barcodes followed by NNNNNNN should be entered into the sample sheet to enable proper multiplex library parsing.
III. Planning the Experiment

**HiSeq 2000/2500 and Next-Seq Instruments**

**Important Note:** If your HiSeq sequencer is using a software version prior to HCS 2.2.58, there may be incompatibilities with a 12 base index read. It is recommended that the HiSeq software be updated to the most recent version.

When setting up a sequencing run, specify 12 bases of index sequencing (no sample sheet is required). However, if you wish to include a sample sheet, specify only the 6 bases of the actual barcode. Do not include the N6 in your sample sheet. Use the method described below to parse and generate the N6 index fastq files for HiSeq2000/2500 and NextSeq instruments using bcl2fastq2 version 2.17.

1. Browse to the location of the run folder (called "RunFolder" in this example).
2. Run bcl2fastq2. Use the "--use-bases-mask Y*,I6Y*" option to generate an Index fastq file along with the forward read (for paired end reads use "--use-bases-mask Y*,I6Y*, Y*"). For example:

```
/usr/local/bin/bcl2fastq --runfolder-dir . --output-dir ./Data/Intensities/BaseCalls/ --no-lane-splitting --sample-sheet SampleSheet.csv --use-bases-mask Y*,I6Y* --minimum-trimmed-read-length 0 --mask-short-adapter-reads 0
```

**Note:** To generate the read and index fastq files without parsing, modify the `--use-bases-mask` option to "--use-bases-mask Y*,Y*". The generated fastq files can then be parsed using alternative software. In this case, the index fastq file will be labeled "R2" rather than "I1".

For information on sequencing with other sequencers, please contact NuGEN Technical Support.

**E. Data Analysis and Parsing Multiplex Libraries**

For the Ovation RRBS Methyl-Seq System 1–16, follow the recommendations in the Illumina technical support documentation on parsing barcodes. The sequences of the Ovation RRBS Methyl-Seq System 1–16 barcodes must be entered prior to parsing (the sequences are listed in Appendix A). With bisulfite-converted libraries, we notice a slightly higher rate of unmatched barcodes relative to non-bisulfite-converted libraries. However, because the Ovation RRBS Methyl-Seq System uses edit distance 3 barcodes, you can allow one mismatch during parsing. This will significantly reduce the fraction of unmatched barcode reads.

Once the data have been parsed according to sample index, the reads must be trimmed before attempting alignment. Trimming is done in two steps. First, any low-quality bases and adaptor sequences are removed from the 3’ end, then the sequence diversity provided by the Ovation RRBS Methyl-Seq System adaptor is removed. At
III. Planning the Experiment

This point reads are ready for downstream analysis, such as mapping to the genome and determining methylation status. Please see Appendix C for detailed recommendations on trimming, alignment and optional duplicate removal for unique molecule identification.
A. Overview

The library preparation process used in the Ovation RRBS Methyl-Seq System 1–16 is performed in five stages:

1. DNA digestion with MspI 1.0 hour
2. Adaptor ligation 0.75 hours
3. Final repair 0.25 hours
4. Bisulfite conversion 3.0 hours
5. Amplification and purification 1.5 hours

Total time to prepare amplified library 6.5 hours

Components in the Ovation RRBS Methyl-Seq System 1–16 are color coded, with each color linked to a specific stage of the process. Performing each stage requires making a master mix, then adding it to the reaction, followed by incubation. Master mixes are prepared by mixing components provided for that stage.

It is important when performing multiplex sequencing to produce each library independently, and not to mix adaptors during the actual library construction protocol. Multiplexing is achieved by mixing the amplified libraries prior to cluster formation. The barcode sequences in this kit were carefully chosen for their ability to parse properly and for color balancing. Therefore, strict requirements exist for multiplexing. Refer to Appendix A, Table 7 for multiplexing guidelines.

B. Protocol Notes

- Use the water provided with the kit (green: D1) or an alternate source of nuclease-free water. We do not recommend the use of DEPC-treated water with this protocol.
- Set up a minimum of four reactions at a time to ensure that you are not pipetting very small volumes.
- Thaw components used in each step and immediately place them on ice. Do not thaw all reagents at once.
- Always keep thawed reagents and reaction tubes on ice unless otherwise instructed.
- After thawing and mixing buffer mixes, if any precipitate is observed, re-dissolve it completely prior to use. You may gently warm the buffer mix for 2 minutes at room temperature followed by brief vortexing. Do not warm any enzyme, primer, or adaptor mixes.
- When placing small amounts of reagents into the reaction mix, pipet up and down several times to ensure complete transfer from the pipet tip into the reaction mix.
- When instructed to pipet mix, gently aspirate and dispense a volume that is at least half of the total volume of the reaction mix.
IV. Protocol

- Always allow the thermal cycler to reach the initial incubation temperature prior to placing the tubes or plates in the block.
- When preparing master mixes, use a minimal amount of extra material to ensure that you are able to run the maximum number of reactions using the components provided in the kit.
- Components and reagents from other NuGEN kits should not be used with the Ovation RRBS Methyl-Seq System 1–16.
- Use only fresh ethanol stocks to make 70% ethanol for the purification protocols.
- Make the ethanol mixes fresh, carefully measuring both the ethanol and water with pipettes. Lower concentrations of ethanol in wash solutions will result in loss of yield as the higher aqueous content will dissolve the DNA and wash it off the beads or column.

C. Agencourt® Purification Beads

Tips and Notes

The protocol for the Ovation RRBS Methyl-Seq System 1–16 requires significant modifications to the Agencourt beads standard procedure. You must follow the protocols outlined in this user guide for the use of these beads. However, you may review the Beckman Coulter user guide to become familiar with the manufacturer’s recommendations.

The bead purification processes used in this kit consist of the following steps:

1. Binding of DNA to beads
2. Magnetic separation of beads from supernatant
3. Ethanol wash of bound beads to remove salts, etc.
4. Elution of bound DNA from beads

Figure 6. Agencourt® Beads process overview

Reproduced from original picture from Agencourt/Beckman Coulter Genomics
IV. Protocol

**Additional Tips and Notes**

- Remove beads from 4°C and leave at room temperature for at least 15 minutes before use. Ensure that they have completely reached room temperature before using. Cold beads reduce recovery.
- Fully resuspend beads by vortexing before adding to sample.
- Note that ratio of Agencourt bead volume to sample volume varies among different steps in the protocol. Furthermore, the bead:sample ratios used differ from the standard Agencourt protocol.
- It is critical to let the beads separate on the magnet for a full 5 minutes. Removing binding buffer before the beads have completely separated will impact DNA yields.
- After completing the binding step, it is important to minimize bead loss when removing the binding buffer. With the samples placed on the magnet, remove only the amount of the binding buffer specified at each of the individual purification steps. Some liquid will remain at the bottom of the tube, helping to minimize bead loss at this step.
- Be certain to minimize bead loss throughout the procedure to maximize DNA yield.
- Ensure that the ethanol wash is freshly prepared from fresh ethanol stocks at the indicated concentration. Lower percent ethanol mixes will reduce recovery.
- During the ethanol washes, it is critical that all residual ethanol be removed prior to continuing with the next step. Therefore, when removing the final ethanol wash, first remove most of the ethanol, then allow excess ethanol to collect at the bottom of the tube before removing the remaining ethanol. This process reduces the required bead air drying time.
- After drying the beads for the time specified in the protocol, inspect each tube carefully and make certain that all the ethanol has evaporated before proceeding.
- We strongly recommend firmly placing strip tubes or partial plates within the magnetic plate. We do not advise the use of individual tubes as they are difficult to position stably on the magnetic plates.

**D. Programming the Thermal Cycler**

Use a thermal cycler with a heat block designed for 0.2 mL tubes, equipped with a heated lid, and with a capacity of 100 μL reaction volume. Prepare the programs shown in Table 2, following the operating instructions provided by the thermal cycler manufacturer. For thermal cyclers with an adjustable heated lid, set the lid temperature to 100°C only when sample temperature reaches above 30°C. For thermal cyclers with a fixed temperature heated lid (e.g., ABI GeneAmp® PCR 9600 and 9700 models), use the default settings (typically 100 to 105°C).
IV. Protocol

Table 2. Thermal cycler programming

<table>
<thead>
<tr>
<th>Mspl DIGESTION</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Program 1</td>
<td>Mspl Digestion</td>
</tr>
<tr>
<td></td>
<td>37°C – 60 min, hold at 4°C</td>
</tr>
<tr>
<td>LIGATION</td>
<td></td>
</tr>
<tr>
<td>Program 2</td>
<td>Ligation</td>
</tr>
<tr>
<td></td>
<td>25°C – 30 min, 70°C – 10 min, hold at 4°C</td>
</tr>
<tr>
<td>FINAL REPAIR</td>
<td></td>
</tr>
<tr>
<td>Program 3</td>
<td>Final Repair</td>
</tr>
<tr>
<td></td>
<td>60°C – 10 min, 70°C – 10 min, hold at 4°C</td>
</tr>
<tr>
<td>BISULFITE CONVERSION</td>
<td></td>
</tr>
<tr>
<td>Program 4</td>
<td>Bisulfite Conversion</td>
</tr>
<tr>
<td></td>
<td>2 cycles of [95°C – 5 min, 60°C – 20 min], hold at 20°C</td>
</tr>
<tr>
<td>AMPLIFICATION</td>
<td></td>
</tr>
<tr>
<td>Program 5</td>
<td>Library Amplification*</td>
</tr>
<tr>
<td></td>
<td>95°C – 2 min, 12 cycles of (95°C – 15 sec, 60°C – 1 min, 72°C – 30 sec, hold at 10°C</td>
</tr>
</tbody>
</table>

*Important Note: If you choose to use less than 100 ng of intact gDNA, or if your DNA is severely degraded (such as with FFPE), you may need more than 12 cycles of PCR. As an estimate, perform one additional cycle of PCR for each two-fold reduction of input. For example, for 25 ng of intact gDNA, use 14 cycles of PCR.

E. Mspl Digestion

1. Remove the Mspl Buffer Mix (blue), Mspl Enzyme Mix (blue) and Nuclease-free Water (green: D1) from –20°C storage.
2. Spin down contents of Mspl Enzyme Mix and place on ice.
3. Thaw Mspl Buffer Mix and nuclease-free water at room temperature. Mix Mspl Buffer Mix by vortexing, spin and place on ice.
4. Adjust each 100 ng gDNA sample to 8.5 µL with nuclease-free water.
IV. Protocol

5. Prepare a master mix by combining MspI Buffer Mix and MspI Enzyme Mix in a 0.5 mL capped tube, according to the volumes shown in Table 3.

Table 3. MspI Master Mix (volumes listed are for a single reaction)

<table>
<thead>
<tr>
<th>Mspl BUFFER MIX (BLUE)</th>
<th>Mspl ENZYME MIX (BLUE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 µL</td>
<td>0.5 µL</td>
</tr>
</tbody>
</table>

6. Add 1.5 μL of the MspI Master Mix to each sample tube.

7. Mix by pipetting, cap and spin tubes and place on ice.

8. Place tubes in a thermal cycler programmed to run Program 1 (MspI Digestion; see Table 2):

   37°C – 60 min, hold at 4°C

9. Remove tubes from the thermal cycler, spin to collect condensation and place on ice.

10. Continue with the Ligation protocol.

F. Ligation

1. Obtain the Ligation Buffer Mix (yellow: L1), Ligation Adaptor Mixes (yellow: L2V20DR-BC1 through L2V20DR-BC16), and Ligation Enzyme Mix (yellow: L3) from –20°C storage.

2. Spin down L3 and place on ice.

3. Thaw L1 and L2 at room temperature. Mix by vortexing, spin and place on ice. The Ligation Buffer Mix (L1) is extremely viscous. Care should be taken to ensure it is well mixed after thawing (it is helpful to alternately vortex the tube right-side up and upside down).

   Important: Do not warm Ligation Adaptor Mixes above room temperature. Heating will severely degrade performance.

4. Add 3 µL of the appropriate L2 Ligation Adaptor Mix to each sample. Mix thoroughly by pipetting.

5. Just prior to use, make a master mix by combining the Nuclease-free Water (green: D1), L1 and L3 in a 0.5 mL capped tube, according to the volumes shown.
IV. Protocol

in Table 4. Mix by pipetting slowly, without introducing bubbles, spin and place on ice. Use mix immediately.

**Note:** The L1 Ligation Buffer Mix is very viscous. Please be sure to pipet this reagent slowly, and take care to ensure that the Ligation Master Mix and the ligation reactions are well mixed (visually observe that the solutions become homogeneous).

<table>
<thead>
<tr>
<th>Water (Green: D1)</th>
<th>Ligation Buffer Mix (Yellow: L1 ver 4)</th>
<th>Ligation Enzyme Mix (Yellow: L3 ver 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0 µL</td>
<td>4.0 µL</td>
<td>1.0 µL</td>
</tr>
</tbody>
</table>

6. Add 7 µL Ligation Master Mix to each reaction tube. Mix thoroughly by pipetting slowly and gently, spin and place on ice. Proceed immediately with the incubation.

7. Place tubes in a pre-warmed thermal cycler programmed to run Program 2 (Ligation; see Table 2):
   
   25°C – 30 min, 70°C – 10 min, hold at 4°C

8. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice.

9. Continue with the Final Repair protocol.

G. Final Repair

1. Remove the Final Repair Buffer Mix (Purple: FR1 ver 4) and Final Repair Enzyme Mix (Purple: FR2) from –20°C storage.

2. Spin down contents of FR2 and place on ice.

3. Thaw FR1 at room temperature. Mix by vortexing, spin and place on ice.

4. Prepare a master mix by combining FR1 and FR2 in a 0.5 mL capped tube, according to the volumes shown in Table 5.

<table>
<thead>
<tr>
<th>Final Repair Buffer Mix (Purple: FR1 ver 4)</th>
<th>Final Repair Enzyme Mix (Purple: FR2)</th>
<th>Water (Green: D1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.0 µL</td>
<td>0.5 µL</td>
<td>13.5 µL</td>
</tr>
</tbody>
</table>

Mix by pipetting slowly, without introducing bubbles, spin for 2 seconds and place on ice. Use mix immediately.
IV. Protocol

5. Add 20 μL of the Final Repair Master Mix to each of the 20 μL ligation reactions.
6. Mix by pipetting, cap and spin tubes and place on ice.
7. Place the tubes in a thermal cycler pre-heated to 60°C and programmed to run Program 3 (Final Repair; see Table 2):
   60°C – 10 min, 70°C – 10 min, hold at 4°C
8. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice.
9. Continue with the Bisulfite Conversion protocol.

**Important Note:** Bisulfite conversion is critical for successful library amplification.

H. Bisulfite Conversion

Notes for conversion:

1. We recommend using the QIAGEN EpiTect Fast DNA Bisulfite Kit (QIAGEN Cat. #59824 for 50 preps or Cat. #59826 for 200 preps). Other commercial bisulfite conversion kits may be suitable as well, but these have not been validated by NuGEN scientists.

2. The 40 μL product of the Final Repair reaction can be input directly into the bisulfite conversion kit. Use the QIAGEN EpiTect Fast protocol “Bisulfite Conversion of Unmethylated Cytosines in DNA”, on page 19, following the guidelines for 1–500 ng in a maximum volume of 40 μL (low concentration). Use 40 μL of input and 15 μL of DNA Protect Buffer. Use the following thermal cycler conditions for best conversion rates:
   95°C – 5 min, 60°C – 20 min, 95°C – 5 min, 60°C – 20 min, hold at 20°C.
   After thermal cycling, you may observe a blue ring at the meniscus. This is normal and will not affect the results.

3. Continue with the rest of the QIAGEN protocol. Use carrier RNA in the BL Buffer.

4. Elute the purified, bisulfite-converted DNA in 23 μL of EB. This should yield 20 μL of bisulfite-converted DNA ready for amplification. If necessary, adjust final eluted volume of bisulfite-converted DNA to 20 μL with nuclease-free water.
IV. Protocol

I. Library Amplification

**Note:** in cases in which you are unsure how many cycles of PCR to perform (such as when starting with degraded DNA or significantly less than 100 ng of high-quality DNA), please consult the real-time protocol in Appendix B.

1. Remove the Amplification Primer Mix (red: P2 ver 8) and Amplification Enzyme Mix (red: P3) from −20°C storage. Also remove the Agencourt beads from 4°C storage. Place at room temperature for use after Library Amplification.

2. Spin P3 and place on ice.

3. Thaw P2 at room temperature. Mix by vortexing, spin and place on ice.

4. Obtain the 20 μL product from the Bisulfite Conversion protocol.

5. Make a master mix by combining P2 and P3 in an appropriately sized capped tube according to the volumes shown in Table 6. Mix well by pipetting taking care to avoid bubbles, spin and place on ice.

<table>
<thead>
<tr>
<th>AMP PRIMER MIX (RED: P2 ver 8)</th>
<th>AMP ENZYME MIX (RED: P3 ver 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 μL</td>
<td>25 μL</td>
</tr>
</tbody>
</table>

6. On ice, add 30 μL of the Amplification Master Mix to each sample.

7. Place tubes in a pre-warmed thermal cycler programmed to run Program 5 (Library Amplification; see Table 2):

   95°C – 2 min, 12 cycles of (95°C – 15 sec, 60°C – 1 min, 72°C – 30 sec, hold at 10°C

**Important Note:** If you choose to use less than 100 ng of intact gDNA, or if your DNA is severely degraded (such as with FFPE), you may need more than 12 cycles of PCR. Please see Appendix B for guidance on how to proceed in such situations.

8. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice.

9. Continue with the Amplified Library Purification protocol.
IV. Protocol

J. Amplified Library Purification

1. Retrieve the Agencourt beads and ensure they are at room temperature prior to use.

2. Resuspend the beads by vortexing the tube. Ensure the beads are fully resuspended before adding to the sample. After resuspending, do not spin the beads.

3. At room temperature, add 50 μL (1 volume) of the bead suspension to each reaction. The bead suspension is quite viscous, therefore pipet slowly to ensure accuracy.

4. Mix thoroughly by pipetting 10 times. It may be helpful to use a multichannel pipettor to ensure the incubation times are uniform.

5. Incubate at room temperature for 10 minutes.

6. Transfer the tubes to the magnet and let stand 5 minutes to completely clear the solution of beads.

7. Carefully remove 90 μL of the binding buffer and discard it. Leaving some of the volume behind minimizes bead loss at this step.

   Note: The beads should not disperse; instead, they will stay on the walls of the tubes. Significant loss of beads at this stage will impact the amount of purified DNA, so ensure beads are not removed with the binding buffer or the wash.

8. Remove the samples from the magnet, add 150 μL of freshly prepared 70% ethanol and pipette to fully resuspend the beads.

9. Transfer the tubes to the magnet and let stand 5 minutes to fully clear the beads. Remove the 70% ethanol wash using a pipette.

10. Repeat steps 8 and 9 for a total of 2 washes.

   Note: With the final wash, it is critical to remove as much of the ethanol as possible. Use at least 2 pipetting steps and allow excess ethanol to collect at the bottom of the tubes after removing most of the ethanol in the first pipetting step.

11. Air dry the beads on the magnet for a minimum of 10 minutes. Inspect each tube carefully to ensure that all the ethanol has evaporated. It is critical that all residual ethanol be removed prior to continuing.

12. Remove the tubes from the magnet.

13. Add 20 μL DNA Resuspension Buffer (Clear: DR1) to the dried beads. Mix thoroughly to ensure all the beads are resuspended.

14. Transfer the tubes to the magnet and let stand for 5 minutes.

15. Carefully remove 18 μL of the eluate, ensuring as few beads as possible are carried over, and transfer to a fresh set of tubes. When pipetting any portion of this eluted library downstream, be sure to let stand briefly on a magnet to minimize bead carryover.

Ensure that all residual ethanol is removed prior to continuing.
16. Proceed to Quantitative and Qualitative Assessment of the Library.

**Important Note:** After quantitation, barcoded libraries that will be run in the same flow cell should be mixed in equimolar ratios prior to processing on the flow cell, cBot or Cluster Station.

### K. Quantitative and Qualitative Assessment of the Library

1. Measure library concentration using 2 µL of each library with the Qubit® 2.0 Fluorometer and dsDNA HS Assay Kit (Life Technologies). Dilute an aliquot to 5 ng/µL using DR1 buffer, load 1 µL on the Bioanalyzer High Sensitivity DNA Chip. Fragment distribution should be as shown in Figure 7. The three peaks at 200 bp, 265 bp, and 330 bp are due to MspI-containing micro-satellite repeats, and are characteristic of RRBS libraries made from human DNA.

**Figure 7. Library size distribution on Bioanalyzer High Sensitivity DNA Chip**

2. If using qPCR for quantification (recommended), use 250 bp as library size for calculations.
V. Technical Support

For help with any of our products, please contact NuGEN Technical Support at 650.590.3674 (direct) or 888.654.6544, option 2 (toll-free, U.S. only). Send faxes to 888.296.6544 (toll-free) or email techserv@nugen.com.

In Europe contact NuGEN at +31(0)135780215 (Phone) or +31(0)135780216 (Fax) or email europe@nugen.com.

In all other locations, contact your NuGEN distributor for technical support.
A. Barcode Sequences and Guidelines for Multiplex Experiments

Barcode sequences and multiplex guidelines for adaptors used in Ovation RRBS Methyl-Seq System 1–16 can be found in Table 7. These 6-nucleotide barcode sequences must be input into the Illumina Sequencing System prior to parsing of the data.

You may combine anywhere from 2–16 barcoded libraries to allow for a range of multiplex sequencing. However, the barcodes were carefully chosen for their ability to parse properly and for color balancing and therefore have strict pairing requirements when performing 2-plex multiplexing. Users wishing to perform greater than a 2-plex multiplexing must choose a Duplex Set (as defined in Table 7), combined with any of the remaining barcoded libraries.

All barcode sequences are separated by an edit distance of three. For further details on the barcode design strategy, please refer to Faircloth BC, Glenn TC (2012). Not All Sequence Tags Are Created Equal: Designing and Validating Sequence Identification Tags Robust to Indels. PLoS ONE 7(8): e42543. doi:10.1371/journal.pone.0042543.

If you choose to mix Ovation Methyl-Seq System libraries with other libraries, check to ensure that the barcodes are compatible (i.e. can be parsed). If you intend to parse with one mismatch allowed, make sure that the spike in library barcodes are an edit distance of 3 or greater from the RRBS barcodes used in that lane. The PhiX library from Illumina does not contain an index. As a result, the sequencer will produce a low-quality index read from PhiX clusters. To remove PhiX reads prior to parsing, filter by index read quality and remove reads with quality less than 20.
Table 7. Barcode sequences for dedicated read (DR) adaptors used in Ovation RRBS Methyl-Seq System 1–16 (Part No. 0353)

<table>
<thead>
<tr>
<th>LIGATION ADAPTOR MIX</th>
<th>BARCODE SEQUENCE</th>
<th>BARCODE PAIRING (2-PLEX)</th>
<th>BARCODE PAIRING (&gt; 2-PLEX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2V20DR-BC1</td>
<td>AACCAG</td>
<td></td>
<td>Duplex Set 1</td>
</tr>
<tr>
<td>L2V20DR-BC2</td>
<td>TGGTGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2V20DR-BC3</td>
<td>AGTGGAG</td>
<td></td>
<td>Duplex Set 2</td>
</tr>
<tr>
<td>L2V20DR-BC4</td>
<td>GCACGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2V20DR-BC5</td>
<td>ACCTCA</td>
<td></td>
<td>Duplex Set 3</td>
</tr>
<tr>
<td>L2V20DR-BC6</td>
<td>GTGCTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2V20DR-BC7</td>
<td>AAGCCT</td>
<td></td>
<td>Duplex Set 4</td>
</tr>
<tr>
<td>L2V20DR-BC8</td>
<td>GTCGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2V20DR-BC9</td>
<td>AAGAGG</td>
<td></td>
<td>Duplex Set 5</td>
</tr>
<tr>
<td>L2V20DR-BC10</td>
<td>GAGAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2V20DR-BC11</td>
<td>AGCATG</td>
<td></td>
<td>Duplex Set 6</td>
</tr>
<tr>
<td>L2V20DR-BC12</td>
<td>GAGTCGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2V20DR-BC13</td>
<td>CGTAGA</td>
<td></td>
<td>Duplex Set 7</td>
</tr>
<tr>
<td>L2V20DR-BC14</td>
<td>TCAGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2V20DR-BC15</td>
<td>CACAGT</td>
<td></td>
<td>Duplex Set 8</td>
</tr>
<tr>
<td>L2V20DR-BC16</td>
<td>TTGGCA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

One of the duplex sets from the column to the left must be used in combination with any of the other remaining 14 individual barcodes.

B. PCR Amplification When Using Degraded DNA or Less Than 100 ng of High-Quality DNA

In the standard Ovation RRBS Library System protocol, libraries are generated using 100 ng of high-quality, MspI-digested genomic DNA, and Library Amplification is performed by preparing an Amplification Master Mix and adding 30 µL of this to 20 µL of bisulfite-converted sample for a total PCR volume of 50 µL. In the event you are using degraded DNA or less than 100 ng of high-quality DNA, please use one of the following two options:
VI. Appendix

A) If you are using high-quality DNA and you know the amount of DNA, you can adjust the number of PCR cycles to obtain sufficient material for analysis and sequencing. NuGEN recommends that you perform one additional cycle of PCR for each two-fold reduction of input. For example, for 25 ng of intact gDNA, use 14 cycles of PCR.

B) If you are using degraded DNA, or you are unsure of the amount and/or quality of the DNA, NuGEN recommends you perform a 1/5 scale real-time PCR with an aliquot of your sample as follows:

1. Make a master mix by combining the Amplification Primer Mix (red: P2) and the Amplification Enzyme Mix (red: P3) in an appropriately sized capped tube according to the volumes shown in Table 8. Add P3 at the last moment and mix well by pipetting taking care to avoid bubbles, spin and place on ice.

<table>
<thead>
<tr>
<th>AMP PRIMER MIX (RED: P2 VER 8)</th>
<th>20X EVAGREEN</th>
<th>AMP ENZYME MIX (RED: P3 VER 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 µL</td>
<td>0.5 µL</td>
<td>4.5 µL</td>
</tr>
</tbody>
</table>

**Note:** EvaGreen (Biotium, Cat. #31000) is less inhibitory to PCR than SYBR Green®, however SYBR Green can also be used at a 1X final concentration in the real-time PCR reaction.

2. For each library, aliquot 6 µL of the Amplification Master Mix into a well of on a real-time PCR plate.

3. Add 4 µL of sample to each well (after bisulfite conversion but before PCR amplification) for a total real-time PCR volume of 10 µL.

4. Perform real-time PCR with the following cycling conditions:

95°C – 2 min, 30 cycles of (95°C – 15 sec, 60°C – 1 min, 72°C – 30 sec

5. Monitor the SYBR Green channel, and do not use ROX normalization.

6. After thermal cycling is complete, examine the amplification plot for each sample to determine the number of PCR cycles to perform during the Library Amplification step. Select a cycle in the late exponential phase. For example, in the amplification plot shown in Figure 8, cycle 9 was chosen. This provides sufficient amplification without entering the plateau phase.
Figure 8. Determining the number of PCR cycles to perform during the library amplification

Note: If using an Applied Biosystems real-time PCR instrument, select Plot > Rn vs. Cycle, not Plot > deltaRn vs. Cycle. Selecting Plot > deltaRn vs. Cycle may give unexpected results similar to those shown in Figure 9.

Figure 9. Results from selecting Plot > deltaRn vs. Cycle

After determining the optimal number of PCR cycles to perform, the remaining 16 µL of bisulfite-converted library can be amplified as follows:

1. Make a master mix by combining the Amplification Primer Mix (red: P2) and the Amplification Enzyme Mix (red: P3) in an appropriately sized capped tube
VI. Appendix

according to the volumes shown in Table 9. Add P3 at the last moment and mix well by pipetting, taking care to avoid bubbles. Spin and place on ice.

**Table 9. Amplification Master Mix (volumes listed are for a single reaction)**

<table>
<thead>
<tr>
<th>AMP PRIMER MIX</th>
<th>AMP ENZYME MIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>(RED: P2 ver 8)</td>
<td>(RED: P3 ver 3)</td>
</tr>
<tr>
<td>4.0 µL</td>
<td>20 µL</td>
</tr>
</tbody>
</table>

2. For each library, aliquot 24 µL of the Amplification Master Mix into a well of on a real-time PCR plate.

3. Add 16 µL of sample to each well.

4. Perform PCR with the following cycling conditions, where N = number of cycles determined from the above real-time PCR assay:

   - 95°C – 2 min
   - N cycles of (95°C – 15 sec, 60°C – 1 min, 72°C – 30 sec), hold at 10°C

5. Proceed with the Amplified Library Purification protocol (Section V. J.) by adding 1 volume (40 µL) of Agencourt RNAClean XP Beads to each amplified sample and following the remaining steps of the purification protocol.

**C. Data Analysis of Ovation RRBS Methyl-Seq Libraries**

The Ovation RRBS Methyl-Seq System generates libraries compatible with Illumina sequencing platforms. After parsing the data by sample index, libraries must be trimmed prior to alignment as described below to remove adaptor sequence, low quality reads, and diversity bases. Ensure you have installed the most current version of Trim Galore, Bismark, Bowtie2, and Samtools prior to data analysis. Additional scripts used for data analysis are available through NuGEN Technical Support (techserv@nugen.com). Optional de-duplication can be performed after alignment to the reference genome following the instructions below.

**Adaptor and Quality Trimming**

To accurately identify the diversity sequence and MspI (C^CGG) site it is important to first trim any adaptor sequence that may be present on the 3’ end of your reads. Trim Galore (www.bioinformatics.babraham.ac.uk/projects/trim_galore/) works well for this purpose, but there may be other equivalent options available. Trim Galore will also trim some or all of a read due to low quality. Run the program with default parameters and do not use the --RRBS option.
Trim single end reads with the following command:

```bash
trim_galore -a AGATCGGAAGAGC R1.FQ
```

If you have paired-end reads, use this command instead:

```bash
trim_galore --paired -a AGATCGGAAGAGC -a2 AAATCAAAAAAAC R1.FQ R2.FQ
```

### Diversity Trimming and Filtering

Following adaptor and quality trimming and prior to alignment, the additional sequence added by the diversity adaptors must be removed from the data. This trimming is performed by a custom python script provided by NuGEN. To obtain this script, contact NuGEN Technical Support at techserv@nugen.com. The script removes any reads that do not contain an MspI site signature (YGG) at the 5’ end. For paired end data an MspI site signature is required at the 5’ end of both sequences.

The script accepts as input one or two fastq file strings, given either as complete filenames or as a pattern in quotes. When a pattern is given, the script will find all the filenames matching a specified pattern according to the rules used by the Unix shell (*,?). You may access the help option of this script for more details (-h).

Example usage for single end reads after adaptor and quality trimming with a complete filename:

```bash
python trimRRBSdiversityAdaptCustomers.py -1 sample_R1.fq
```

with a pattern:

```bash
python trimRRBSdiversityAdaptCustomers.py -1 '*R1.fq'
```

Example usage for paired-end reads after adaptor and quality trimming with a complete filename:

```bash
python trimRRBSdiversityAdaptCustomers.py -1 sample_R1.fq -2 sample_R2.fq
```

with a pattern:

```bash
python trimRRBSdiversityAdaptCustomers.py -1 '*R1.fq' -2 '*R2.fq'
```

The script will generate new file(s) with "_trimmed.fq" appended to the filename. The reads will have been trimmed at the 5’ end to remove the diversity sequence (0–3 bases), and all reads should begin with YGG, where Y is C or T. On the 3’ end, 5 bases are trimmed from every read (6 bases are trimmed for paired-end to prevent alignment issues).

The trimmed fastq file should be used for downstream analysis including Bismark.
VI. Appendix

Alignment to Genome

After trimming, the data can be aligned to the genome of interest. Bismark (http://www.bioinformatics.babraham.ac.uk/projects/bismark/) is a tool that aligns bisulfite converted sequencing reads to the genome and also performs methylation calls in the same step. The program supports single and paired-end reads and both ungapped and gapped alignments. Other equivalent options may be available.

To align single end reads:

\[
\text{bismark --bowtie2 /location/bismark/genome/ R1\_trimmed.FQ}
\]

For paired-end reads:

\[
\text{bismark --bowtie2 /location/bismark/genome/ -1 R1\_trimmed.FQ -2 R2\_trimmed.FQ}
\]

Note: Recent versions of Bismark automatically generate a BAM file instead of a SAM file. In order to perform the optional duplicate determination step, the resulting BAM file must be converted to a SAM file, or else run Bismark with option "--sam"

Continue with downstream data analysis or to unique molecule identification as described in the following section.

Duplicate Determination with NuDup (Optional):

The N6 molecular tag is a novel approach to the unambiguous identification of unique molecules. Traditionally, PCR duplicates are identified in libraries made from randomly fragmented inserts by mapping inserts to the genome and discarding any paired end reads that share the same genomic coordinates. This approach doesn’t work for restriction digested samples, such as RRBS, because all fragments mapping to a genomic location will share the same ends.

The Duplicate Marking tool utilizes information provided by the unique N6 sequence to discriminate between true PCR duplicates and independent adaptor ligation events to fragments with the same start site resulting in the recovery of more usable data.

First, Bismark output files must be modified for input into NuDup using the following command:

\[
\text{strip\_bismark\_sam.sh bismarkout\_stripped.sam}
\]

Note: Recent versions of Bismark automatically generate a BAM file instead of a SAM file. In order to use the stripping tool, the resulting BAM file must be converted to a SAM file, or else run Bismark with option "--sam"

Next, run NuDup using the modified SAM files as input:

For single end reads:

\[
\text{python nudup.py -f index.fq -o outputname bismarkout\_stripped.sam}
\]
VI. Appendix

For paired-end reads:

```
python nudup.py -2 -f index.fq -o outputname bismarkout_stripped.sam
```

Continue with downstream data analysis.

**Note:** These commands assume that a 12-base index read was generated. If longer index reads were generated, contact NuGEN Technical Support.

### D. Diversity Trimming Examples

#### Examples of Trimming the 5' Ends of Forward Reads

Bases in **blue** denote sequence derived from the adaptor. In this example, the fragment was derived from the genomic sequence, starting and ending with MspI sites:

5' CCAGAGTT...AAGGGCCGG 3'
3' GCCCTCAA...TTCCCGGCC 5'

After MspI digestion:

5' CGGAGTT...AAGGGC 3'
3' CTCAA...TTCCCGGC 5'

After ligation to adaptors, both with three bases of diversity:

5' RDDCGGAGTT...AAGGGCCGHY 3'
3' YHGCTCAA...TTCCCGGCDR 5'

After bisulfite conversion and PCR amplification of the top strand:

5' RDDYGGAGTT...AAGGGTCGHY 3'
3' YHHRCTCAA...TTCCAGCDDR 5'

Assuming the insert is smaller than the read length, the forward read after Trim Galore is used to trim the adaptor from the 3' end will be:

5' RDDYGGAGTT...AAGGGTCGHY 3'

The result of the NuGEN diversity trim of the forward read (if it’s a single-end read) will be:

5' YGGAGTT...AAGGGT 3'

The reverse read after Trim Galore is used will be:

5' RDDCGACCCCTT...AACTCCRHYH 3'

The result of the NuGEN diversity trim of the reverse read:

5' ACCCTT...AACT 3'

The adaptor can contain between 0 and 3 bases of diversity.
Tables 12–15 show how the script trims all types of adaptor variation.

Table 10. Trimming of the 5’ ends of forward reads

<table>
<thead>
<tr>
<th>READ BEFORE TRIMMING</th>
<th>READ AFTER TRIMMING</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ YGGAGTT...</td>
<td>5’ YGGAGTT...</td>
</tr>
<tr>
<td>5’ DYGGAGTT...</td>
<td>5’ YGGAGTT...</td>
</tr>
<tr>
<td>5’ DGYGGAGTT...</td>
<td>5’ YGGAGTT...</td>
</tr>
<tr>
<td>5’ RDDYGGAGTT...</td>
<td>5’ YGGAGTT...</td>
</tr>
</tbody>
</table>

Note: If YGG is not found in the first 6 bases, the read is discarded.

Table 11. Trimming of the 3’ ends of forward reads

<table>
<thead>
<tr>
<th>READ BEFORE TRIMMING</th>
<th>READ AFTER TRIMMING</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ ...AAGGGTCG</td>
<td>5’ ...AAG</td>
</tr>
<tr>
<td>5’ ...AAGGGTCGH</td>
<td>5’ ...AAGG</td>
</tr>
<tr>
<td>5’ ...AAGGGTCGHH</td>
<td>5’ ...AAGGG</td>
</tr>
<tr>
<td>5’ ...AAGGGTCGHHY</td>
<td>5’ ...AAGGGT</td>
</tr>
</tbody>
</table>

Table 12. Trimming of the 5’ ends of reverse reads

<table>
<thead>
<tr>
<th>READ BEFORE TRIMMING</th>
<th>READ AFTER TRIMMING</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ CGACCTT...</td>
<td>5’ ACCCTT...</td>
</tr>
<tr>
<td>5’ DCGACCTT...</td>
<td>5’ ACCCTT...</td>
</tr>
<tr>
<td>5’ DDCGACCTT...</td>
<td>5’ ACCCTT...</td>
</tr>
<tr>
<td>5’ RDDCGACCTT...</td>
<td>5’ ACCCTT...</td>
</tr>
</tbody>
</table>

Note: If CGR is not found in the first 6 bases, the read is discarded.
VI. Appendix

Table 13. Trimming of the 3’ ends of reverse reads

<table>
<thead>
<tr>
<th>READ BEFORE TRIMMING</th>
<th>READ AFTER TRIMMING</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ ...AACTCCRHHY</td>
<td>5’ ...AACTC</td>
</tr>
<tr>
<td>5’ ...AACTCCRHH</td>
<td>5’ ...AACT</td>
</tr>
<tr>
<td>5’ ...AACTCCRH</td>
<td>5’ ...AAC</td>
</tr>
<tr>
<td>5’ ...AACTCCR</td>
<td>5’ ...AA</td>
</tr>
</tbody>
</table>

See “Diversity Trimming and Filtering” on page 30 for file name examples.

E. Effects of Read Length on Mapping Rate

The following data illustrates how read length affects mapping rates. An Ovation RRBS Methyl-Seq System library was prepared from IMR90 cell line DNA and sequenced on a HiSeq2500 in Rapid Run mode using 2X 100 nt paired end reads. The raw data was used in full, or trimmed as indicated, before processing first with Trim Galore (www.bioinformatics.babraham.ac.uk/projects/trim_galore/) to remove adaptor sequence and low quality bases, and then with the NuGEN diversity trimming script. The resulting reads were then mapped to the hg19 human genome reference using Bismark (www.bioinformatics.bbsrc.ac.uk/projects/bismark/). Table 14 displays the percent of reads mapping uniquely and non-unically for single end reads of various lengths. Table 15 presents the same metrics for paired end reads. 29 nt and 36 nt reads are shown to enable comparison to published RRBS data (29 nt single end reads — Boyle, et al. (2012) Genome Biol 13:R92. 36nt single end reads — Varley, et al. (2013) Genome Res 23:555). While some reports use modified reference genomes to reflect only expected MspI fragments for mapping, for this analysis reads were mapped to the entire, unmodified human reference genome.

Table 14. Observed RRBS single end reads that are uniquely or non-uniquely mappable to the hg19 reference genome

<table>
<thead>
<tr>
<th>SINGLE END READ LENGTH</th>
<th>29nt</th>
<th>36nt</th>
<th>50nt</th>
<th>75nt</th>
<th>100nt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequences mapping uniquely</td>
<td>18.7%</td>
<td>51.2%</td>
<td>57.8%</td>
<td>64.8%</td>
<td>67.3%</td>
</tr>
<tr>
<td>Sequences mapping non-uniquely</td>
<td>81.3%</td>
<td>47.4%</td>
<td>37.8%</td>
<td>28.0%</td>
<td>23.5%</td>
</tr>
<tr>
<td>Total</td>
<td>100.0%</td>
<td>98.6%</td>
<td>95.6%</td>
<td>92.8%</td>
<td>90.8%</td>
</tr>
</tbody>
</table>

In addition to mappability, you may also want to consider how read length affects CpG loci coverage. Many MspI fragments contain internal CpG’s, so longer reads will sequence more CpGs. However, many MspI fragments are smaller than 100 bp, and
even smaller than 50 bp. For these fragments, long sequencing reads, or paired end reads, provide no additional CpG data.

### Table 15. Observed RRBS paired end reads that are uniquely or non-uniquely mappable to the hg19 reference genome

<table>
<thead>
<tr>
<th>SINGLE END READ LENGTH</th>
<th>2x29nt</th>
<th>2x36nt</th>
<th>2x50nt</th>
<th>2x75nt</th>
<th>2x100nt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequences mapping uniquely</td>
<td>57.5%</td>
<td>64.6%</td>
<td>66.7%</td>
<td>68.0%</td>
<td>67.6%</td>
</tr>
<tr>
<td>Sequences mapping non-uniquely</td>
<td>29.3%</td>
<td>25.5%</td>
<td>20.2%</td>
<td>17.8%</td>
<td>17.4%</td>
</tr>
<tr>
<td>Total</td>
<td>86.8%</td>
<td>90.1%</td>
<td>86.9%</td>
<td>85.8%</td>
<td>85.0%</td>
</tr>
</tbody>
</table>
VI. Appendix

F. Frequently Asked Questions (FAQs)

Q1. What kind of sequencing primers can I use with your library?
The design of the Ovation RRBS Methyl-Seq System 1–16 requires the use of a custom Read 1 sequencing primer, MetSeq Primer 1, which is included in this kit at a concentration of 25 µM. The standard primers provided in the Illumina sequencing kit are sufficient for Read 2 and for sequencing the barcodes (Index Read). The Standard Read 1 Primer is also required when using PhiX or other libraries to increase base complexity.

Q2. Does NuGEN provide reagents for performing the bisulfite conversion step of the protocol?
NuGEN does not provide the reagents used in the bisulfite conversion steps. We recommend using the QIAGEN EpiTect Fast DNA Bisulfite Kit (QIAGEN Cat. #59824). Other commercial bisulfite conversion kits may be suitable as well, but these have not been validated.

Q3. Can the Ovation RRBS Methyl-Seq System 1–16 be used with paired-end sequencing?
Yes, it can be used for both single end and paired-end sequencing.

Q4. How much material should I load into the sequencer or cBot System?
Please follow manufacturer’s recommendations for library QC, quantitation, balancing and loading of the amplified library on the sequencer or cBot System.

Q5. Does the Ovation RRBS Methyl-Seq System 1–16 work with the Illumina Cluster Station (predecessor of the cBot instrument)?
Yes, the system is compatible with the Illumina Cluster Station.

Q6. Is the Ovation RRBS Methyl-Seq System compatible with all Illumina sequencing platforms?
Illumina may not support the use of a custom sequencing primer or low diversity libraries on all platforms. Please follow the custom primer and low-diversity library recommendations for your specific sequencer.

Q7. How does your protocol improve the efficiency of ligation and avoid adaptor dimer formation?
The Ovation RRBS Methyl-Seq System 1–16 utilizes optimized chemistries to capitalize on the increased efficiency of sticky-end adaptor ligation while also minimizing the amount of adaptor dimer in the library.

Q8. How can gel purification be eliminated from the workflow and still prevent adaptor dimer formation?
The Ovation RRBS Methyl-Seq System 1–16 workflow uses bead-based purification and efficient primer design, thus eliminating the need for gel-based purification.
Q9. Can I modify the number of PCR amplification cycles recommended by the Ovation RRBS Methyl-Seq System 1–16 workflow when using different DNA input amounts? 
Real-time PCR can be used to determine the appropriate number of PCR cycles. Please see Appendix B for a detailed protocol. For more information, contact NuGEN Technical Support.

Q10. What kind of error correction is used to minimize the impact of sequencing errors in the barcodes?
Each of the DR barcode sequences shown in Table 7 are separated by an edit distance of three. This means that three events, such as insertion, deletion or substitution must occur before any barcode sequence is converted into another barcode sequence. A benefit of having an edit distance of three in the barcodes is that one error can be corrected without the chance of barcode misassignment. Parsing software can be adjusted to perform such error correction if there is a large proportion of unsegregated reads.

Q11. Can I combine the barcoded libraries prior to amplification?
The stoichiometry of barcoded libraries may be adversely affected by this modification to the workflow. We suggest that the libraries be amplified and quantitated independently before being balanced and pooled for use on the cBot, Cluster Station, or MiSeq.

Q12. What is the expected yield of the amplified DNA library using the Ovation RRBS Methyl-Seq System 1–16?
The expected yield is at least 200 ng, depending on the quality and quantity of the genomic DNA and the number of PCR cycles employed. This amount is in excess of the amount of DNA required for cluster generation.

Q13. Are the libraries directional?
Yes, the libraries are directional due to the way our library system is designed and the nature of bisulfite conversion. The forward sequencing reads will correspond to a bisulfite-converted version of either the original top or the original bottom strand (the C-to-T reads) and the reverse sequencing reads will correspond to the complement of the original top or the complement of the original bottom strand (the G-to-A reads). In contrast, a non-directional bisulfite converted library will have all four possible strands in the forward read (original top, original bottom, complement of original top and complement of original bottom).

Q14. What analysis software can be used for aligning, methylation calling, and visualization of your bisulfite sequencing data?
The number of analysis strategies and software tools for methylation-based sequencing studies is growing rapidly. The ideal analysis workflow for a given experiment depends on many variables, including the type of experiment and the goals of the study. Currently, NuGEN scientists use Bismark for aligning and determining methylation status. This program utilizes the Bowtie aligner (www.bioinformatics.bbsrc.ac.uk/projects/bismark/). The Broad IGV genome browser can be used to visualize the results of Bismark (www.broadinstitute.org/igv/).
Q15. How can I measure the efficiency of bisulfite conversion?
DNA material that is known to be unmethylated, such as lambda DNA, can be used to measure the efficiency of C-to-U conversion in the bisulfite conversion kit. This control DNA is not included with the Ovation RRBS Methyl-Seq System 1–16.

Q16. Can I use restriction enzymes other than, or in addition to, MspI?
Yes, however the restriction site must leave a similar 3’-GG overhang in order for the ligation to be effective. Please contact NuGEN technical support for further advice on integrating other enzymes into this protocol.

G. Update History
This document, the Ovation RRBS Methyl-Seq System 1–16 user guide (M01394 v3), is an update to address the following topics:

<table>
<thead>
<tr>
<th>Description</th>
<th>Section</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Added advisory and recommendations about sequencing with custom primers on Illumina platforms</td>
<td>III.C.</td>
<td>10</td>
</tr>
<tr>
<td>Added new section Data Analysis to Appendix</td>
<td>VII.C.</td>
<td>29</td>
</tr>
<tr>
<td>Added Next-Seq recommendations</td>
<td>VII.C.</td>
<td>12</td>
</tr>
<tr>
<td>Added new section Effects of Read Length on Mapping Rate</td>
<td>VII.F.</td>
<td>34</td>
</tr>
<tr>
<td>Reorganized content for better clarity</td>
<td>Throughout</td>
<td>Throughout</td>
</tr>
</tbody>
</table>