Amplicon Bisulfite Sequencing Library Creation Protocol and Notes

Note: the following protocol is optimized for creation of libraries from a single target region of interest that was 392 bp long. Comments on how to adapt this protocol to sequence multiple regions from the same sample simultaneously follow the protocol.

1. Extract gDNA with Zymo Research Qucik gDNA MiniPrep kit.
2. Measure DNA concentration with the dsDNA Broad Range Qubit assay.
3. Bisulfite convert 300 ng of gDNA with the Zymo Research EZ DNA methylation kit. Yield is generally about 3 to 4 ng/uL in 10 uL elution volume. We have found that concentration can be increased to about 10 ng/uL in 12 uL elution volume when converting 1 ug of gDNA. Conversion of greater masses of DNA did not lead to a significantly increased yield of converted gDNA.
4. Amplify the region of interest from the bisulfite converted gDNA with the Qiagen PyroMark PCR kit. Use 10 ng of bisulfite converted gDNA. Note: the specific reaction conditions for each region will need to be optimized (see kit protocol).
5. Use 1.8X reaction volume of SPRI (Ampure XP) beads to purify the PCR reactions. Note this assumes that a single PCR product is produced and has been verified by a diagnostic gel.
   a. Yield is generally ~5 to 10 ng/uL in 15 uL.
   b. Note: I usually limit myself to 4 SPRI cleanups at once so for higher throughput gel extraction may be preferable.
6. Ligate on Illumina adapters with NEB Quick ligase.
   a. See Illumina adapter preparation protocol.
      Adapter A: 5' – ACACTCTTTCCCTACACGACGCTCTTCCGATCT – 3'
      Adapter B: 5' -- /5phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCAC – 3'
   b. The final adapter concentration should be 10X the molar concentration of the amplified DNA.
   c. Example ligation reaction:

Reagent	Volume for 1 Reaction
Library PCR Product	Variable*
2X Quick Ligase Buffer	10 uL
15 uM Y-Adapter	0.13 uL**
Quick Ligase	1 uL
Water	Variable
Total Volume	20 uL

      *I aim for approximately 100 ng of PCR product per ligation. One ligation generally gives me sufficient DNA for subsequent steps.
      **This gives about 10-fold molar excess for about 50 ng of a 392 bp PCR product.
   d. I recommend that the ligase be heat inactivated. Without inactivation, the ligations often produce smears on the gel and are harder to purify for downstream use.
7. Run the ligations on a 1.5% agarose gel (1X TBE) to ensure that the ligation was successful.
   a. If the ligation is incomplete three bands should be observed: unligated insert, insert plus 1 adapter, and insert plus 2 adapters. The latter two bands should run higher than their predicted length due to the ssDNA in the Y-adapter. While not necessary, we find that purifying the topmost of the three bands improves library performance.
   b. Runs at 90V (about 7.5 V/cm) work well in our hands. We use a higher quality, low melting point agarose gel on this step to increase band resolution (specifically Agarose II from ISC BioExpress).
8. Extract the ligation products from the gel with the Qiagen QiaQuick Gel Extraction kit.
   a. Perform the gel dissolution step at room temperature for 1 hour to ensure that the gel completely dissolves.
   b. After the Buffer PE wash, discard the flow-through and dry the column by centfiguation for 5 minutes.
   c. Heat the Buffer EB to 65 °C before elution. Use 30 uL EB when eluting.
9. Amplify adapter ligated DNA while adding the index primers using NEB Phusion High Fidelity DNA Polymerase (M0530). The number of cycles and expected yield will vary with the amount of input library

DNA, but 12 cycles is a good estimate. Splitting the adapter ligated DNA from step 8 into two PCR reactions per library gives higher yields. An example reaction is below (I now typically set up two, 50 uL PCRs per sample):

Reagent	Volume for 1 Reaction
DNA	Variable
5X HF Buffer	10 uL
Illumina Primer 1.0 (25 uM)	1 uL
Illumina Primer 2.0 (0.5 uM)***	1 uL
Index Primer (25 uM)****	1 uL
dNTPs (10 mM)	1.5 uL
HF Phusion DNA Pol	0.5 uL
Water	Variable
Total Volume	50 uL

***Prepare the 0.5 uM dilution of Illumina Primer 2.0 fresh before each reaction to prevent any absorption on the side of the tube or freeze-thaw degradation from altering this very low concentration.
Primer 1.0: 5' – AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT – 3'
Primer 2.0: 5' – GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT – 3'
****An example index primer is included below with the index sequence highlighted in blue.
5' – CAAGCAGAAGACGGCATACGAGATTCGTGACAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGA – 3'

Thermocycler Protocol		
1	98 C	30 sec
2	98 C	10 sec
3	64 C	30 sec
4	72 C	30 sec
5	Go To Step 2	14 Times (15 cycles)*****
6	72 C	5 min
7	4 C	Hold

*****Number of cycles variable based on input DNA. 12 cycles is a good starting point if your yields are comparable to ours.

10. Gel purify the library on a 1.5% Agarose II gel (1X TBE). Extract the libraries from the gel with the Qiagen Minelute kit. Be sure the Minelute columns warm to room temperature before use.
11. Measure the library DNA concentration via dsDNA High Sensitivity Qubit assay. Pool libraries as appropriate and submit for paired-end 250 bp read next generation sequencing. We recommend spiking these libraries into other libraries that occupy a full lane on an Illumina sequencer. This will produce several thousand reads per sample.
12. Upon receiving the FASTQ data files, trim the reads with fqtrim (http://ccb.jhu.edu/software/fqtrim/) using a quality threshold of 20.
13. Map reads with Bismark (http://www.bioinformatics.babraham.ac.uk/projects/bismark/). The output will contain the cytosine coordinate, the % methylation, and the number of methylated/unmethylated reads.
14. Visualize the data according to your lab's needs and capabilities.

Because the analysis pipeline uses read mapping software, multiple PCR regions from the same sample can be analyzed with the same sequencing barcode. The ability to sequence many regions simultaneously drastically increases throughput but causes some changes to the protocol. These changes are the following:

1. In step 3, always convert 1 ug of gDNA. This is critical for preparing more libraries because the increased number of bisulfite PCR reactions requires more bisulfite converted gDNA.
2. In step 5, the SPRI bead purification can be replaced by the Qiagen QiaQuick PCR Purification protocol, eluting with 30 uL of Buffer EB. This change is recommended because SPRI bead purification is slow.

3. In step 6, when ligating the Y-adapters to multiple regions, we made an equimolar mix of each region that totaled 5.72×10^{-13} moles. This value is given as a guideline for expected yield and can be scaled up or down according to the needs of alternative protocols.
4. In step 7, when gel purifying the ligation products, the gel runs are shortened to 30 or 40 minutes at 80 to 90 V. This is done because the ligation triplet bands will overlap when multiple regions of similar length are included in the ligation. We recommend preventing the bands from separating to ensure purification of all final ligation products and lower agarose loads on the purification columns.
5. In step 10, when gel purifying the final library PCR product, the gel run is shortened to 30 or 40 minutes at 80 to 90 V in order to ensure library products are not separated.
6. In step 11, when mixing libraries containing multiple regions for analysis, we use the average length of all library products and the approximation of 650 g/mol bp to calculate the molecular weight for each library.