Protocol – BisPCR2-Seq

**Background**

The goal of this protocol is to prepare DNA sequencing libraries for targeted DNA methylation analysis. This strategy is useful for assessing DNA methylation across multiple target regions in multiple samples. The method uses high-throughput next-generation sequencing, meaning all CpGs within a target region, approximately 250 base pairs, can be measured with over 1000X coverage. The general procedure includes bisulfite conversion of genomic DNA followed by two PCR steps to generate DNA sequencing libraries. The goal of the first PCR (PCR#1) is "target enrichment" to amplify regions of interest from each sample. These primers have overhangs that are then used to amplify barcoded libraries in the second round of PCR. Target enrichment PCR products for each sample are pooled prior to PCR#2 to simultaneously add the same multiplexing index. Here we describe 48 primer pairs for multiplexing, although more can be designed as needed. Libraries are then pooled for sequencing on the Illumina Miseq. See Appendix A for a detailed scheme of library construction.

**Reagents**

* Target enrichment primers (custom primers for target regions modified with adaptor overhangs – see below)
* Library indexing primers (48 barcodes designed – Appendix B)
* Qiagen EpiTect Bisulfite Kit (48) (Cat. # 59104) - $212 from the cell center
* PyroMark PCR Kit (800) (Cat. # 978705) - $947.00 ea. from Qiagen (http://www.qiagen.com/products/catalog/assay-technologies/pyrosequencing/pyromark-pcr-kit)
* QIAquick PCR Purification Kit (50) (Cat. # 28104) - $96.14 ea. from the cell center (at least two kits required, in general)

**Target enrichment primer design**

* 1. Identify target regions for DNA methylation analysis. Amplicons should be 200-250 base pairs, so for regions greater than that, multiple amplicons will be required.
  2. Retrieve DNA sequences (i.e. from UCSC Genome Browswer).
     1. If interested in a CpG island, you can retrieve the DNA sequence by clicking on the CpG island.
     2. It is recommended to add 100bp of genomic sequence to the target region to allow for more efficient primer design.
     3. Copy and paste the sequence into a word document
  3. Design primers to amplify bisulfite converted gDNA. It is recommended to use Qiagen’s PyroMark Assay Design software, which allows for future pyrosequencing if necessary. Suggested settings for primer design:
     1. Minimum primer length = 18 bps
     2. Maximum primer length = 25 bps (total primer length < 60 for cost purposes)
     3. Optimal amplicon length = 250 to 300 bps
     4. Maximum amplicon length = 300 bps
     5. Allow primer over variable position 🡪 UNCHECK this box (CpGs in primer may interfere with NGS reaction)
  4. Modify primers by adding adaptor overhangs. (**NOTE**: All target enrichment primers will be modified with the same adaptor overhangs)

**PCR#1 Left Primer Overhang:**

5’ – ACACTCTTTCCCTACACGACGCTCTTCCGATCT – 3’

**PCR#1 Right Primer Overhang:**

5’ – GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT – 3’

* 1. Test primer amplification of bisulfite converted gDNA (PyroMark PCR kit)
     1. Note: If desired, PCR purify and check concentration to determine approximate amount of amplified DNA to be expected in future reactions.
     2. Note: If desired, test PCR#2 reaction to ensure the amplicon can be barcoded.

**Protocol**

1. **Target enrichment (PCR #1)**
   1. Bisulfite convert gDNA (Qiagen EpiTect Bisulfite Kit). Follow manufacturer’s instructions.
      1. It is recommended to start with 200-500ng of gDNA. However, any amount of starting material is suitable, as long as it is within the parameters of the Qiagen kit.
   2. Set up target enrichment PCR (PyroMark PCR)

|  |  |
| --- | --- |
|  | 1X |
| Bisulfite converted gDNA | 1ul |
| PyroMark (2X) | 12.5ul |
| Primer mix (10uM) | 1ul |
| Coral Load (10X) | 2.5ul |
| Water to 25ul | 8ul |
| Total | 25ul |

\*NOTE: bisulfite converted gDNA can be diluted to conserve material (i.e. 1:10). Adjust volume with water accordingly.

* + 1. IMPORTANT: Always include a water control sample for each primer
    2. It may be useful to use a 96-well format
  1. Run 5ul of each PCR reaction on a 1% agarose gel.
     1. NOTE: It is not necessary to run all PCR products for ALL samples. One sample should be sufficient to compare relative amounts across the different PCR products
     2. NOTE: be sure to keep exposure time consistent for each gel picture)
  2. Pool PCR products for each biological sample
     1. Determine amount of each PCR product to add based on the relative brightness of the band on the agarose gel.
        1. The goal is to have approximately the same amount of each PCR product in a given sample.
        2. Pool samples by purifying each PCR reaction, measure concentration, and pool the same quantity for each.

1. **Barcode library (PCR #2)**
   1. Set-up the following PCR reaction for each PCR pool:

|  |  |
| --- | --- |
|  | 1X |
| Pooled PCR products (20-50ng) | x |
| PyroMark (2X) | 12.5ul |
| Library primer (10uM) | 1ul |
| Indexing primer (10uM) | 1ul |
| Coral Load (10X) | 2.5ul |
| Water to 25ul | x |
| Total | 25ul |

**NOTE: Currently Indeces #31 and #41 should NOT be used in the same pool.** If 48 (or more) libraries are desired, additional indeces are required and should be checked for compatibility.

Follow manufacturers instructions, EXCEPT amplify with 11 cycles instead of 45. The Goal is to use minimal amplification cycles to add barcodes. Cycle number can be adjusted to fit individual experiment requirements.

* 1. Purify each library using QIAquick PCR purification kit.

1. **Library quality control (QC) and Sample Pooling**

NOTE: For samples outside of the Kaestner Lab, library QC and sample pooling steps should be done by the NGSC.

* 1. Qubit
     1. Measure quantity of DNA in each library
  2. BioAnalyzer
     1. Measure fragment length; average length should equal average PCR product length + 122bps (adapters)
  3. KAPA
     1. Measure the amount of library (versus total DNA content)
     2. Determine molarity for subsequent pooling
  4. Pool libraries at a concentration of 2nM

Appendix A – BisPCR2-Seq Primer Design



Appendix B – PCR#2 Primer Sequences

NOTE: All reactions require the same Forward Primer (Library\_Primer1)

**NOTE: Currently Indeces #31 and #41 should NOT be used in the same pool.** If 48 (or more) libraries are desired, additional indeces are required and should be checked for compatibility.

|  |  |  |
| --- | --- | --- |
| **Primer Name** | **Primer Sequence** | Barcode |
| **Library\_Primer1** | **AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC** | **ALL libraries** |
| Primer2\_Index\_1 | CAAGCAGAAGACGGCATACGAGAT**CGTGAT**GTGACTGGAGTTCAGACGTGT | ATCACG |
| Primer2\_Index\_2 | CAAGCAGAAGACGGCATACGAGAT**ACATCG**GTGACTGGAGTTCAGACGTGT | CGATGT |
| Primer2\_Index\_3 | CAAGCAGAAGACGGCATACGAGAT**GCCTAA**GTGACTGGAGTTCAGACGTGT | TTAGGC |
| Primer2\_Index\_4 | CAAGCAGAAGACGGCATACGAGAT**TGGTCA**GTGACTGGAGTTCAGACGTGT | TGACCA |
| Primer2\_Index\_5 | CAAGCAGAAGACGGCATACGAGAT**CACTGT**GTGACTGGAGTTCAGACGTGT | ACAGTG |
| Primer2\_Index\_6 | CAAGCAGAAGACGGCATACGAGAT**ATTGGC**GTGACTGGAGTTCAGACGTGT | GCCAAT |
| Primer2\_Index\_7 | CAAGCAGAAGACGGCATACGAGAT**GATCTG**GTGACTGGAGTTCAGACGTGT | CAGATC |
| Primer2\_Index\_8 | CAAGCAGAAGACGGCATACGAGAT**TCAAGT**GTGACTGGAGTTCAGACGTGT | ACTTGA |
| Primer2\_Index\_9 | CAAGCAGAAGACGGCATACGAGAT**CTGATC**GTGACTGGAGTTCAGACGTGT | GATCAG |
| Primer2\_Index\_10 | CAAGCAGAAGACGGCATACGAGAT**AAGCTA**GTGACTGGAGTTCAGACGTGT | TAGCTT |
| Primer2\_Index\_11 | CAAGCAGAAGACGGCATACGAGAT**GTAGCC**GTGACTGGAGTTCAGACGTGT | GGCTAC |
| Primer2\_Index\_12 | CAAGCAGAAGACGGCATACGAGAT**TACAAG**GTGACTGGAGTTCAGACGTGT | CTTGTA |
| Primer2\_Index\_13 | CAAGCAGAAGACGGCATACGAGAT**TTGACT**GTGACTGGAGTTCAGACGTGT | AGTCAA |
| Primer2\_Index\_14 | CAAGCAGAAGACGGCATACGAGAT**GGAACT**GTGACTGGAGTTCAGACGTGT | AGTTCC |
| Primer2\_Index\_15 | CAAGCAGAAGACGGCATACGAGAT**TGACAT**GTGACTGGAGTTCAGACGTGT | ATGTCA |
| Primer2\_Index\_16 | CAAGCAGAAGACGGCATACGAGAT**GGACGG**GTGACTGGAGTTCAGACGTGT | CCGTCC |
| Primer2\_Index\_17 | CAAGCAGAAGACGGCATACGAGAT**CTCTAC**GTGACTGGAGTTCAGACGTGT | GTAGAG |
| Primer2\_Index\_18 | CAAGCAGAAGACGGCATACGAGAT**GCGGAC**GTGACTGGAGTTCAGACGTGT | GTCCGC |
| Primer2\_Index\_19 | CAAGCAGAAGACGGCATACGAGAT**TTTCAC**GTGACTGGAGTTCAGACGTGT | GTGAAA |
| Primer2\_Index\_20 | CAAGCAGAAGACGGCATACGAGAT**GGCCAC**GTGACTGGAGTTCAGACGTGT | GTGGCC |
| Primer2\_Index\_21 | CAAGCAGAAGACGGCATACGAGAT**CGAAAC**GTGACTGGAGTTCAGACGTGT | GTTTCG |
| Primer2\_Index\_22 | CAAGCAGAAGACGGCATACGAGAT**CGTACG**GTGACTGGAGTTCAGACGTGT | CGTACG |
| Primer2\_Index\_23 | CAAGCAGAAGACGGCATACGAGAT**CCACTC**GTGACTGGAGTTCAGACGTGT | GAGTGG |
| Primer2\_Index\_24 | CAAGCAGAAGACGGCATACGAGAT**GCTACC**GTGACTGGAGTTCAGACGTGT | GGTAGC |
| Primer2\_Index\_25 | CAAGCAGAAGACGGCATACGAGAT**ATCAGT**GTGACTGGAGTTCAGACGTGT | ACTGAT |
| Primer2\_Index\_26 | CAAGCAGAAGACGGCATACGAGAT**GCTCAT**GTGACTGGAGTTCAGACGTGT | ATGAGC |
| Primer2\_Index\_27 | CAAGCAGAAGACGGCATACGAGAT**AGGAAT**GTGACTGGAGTTCAGACGTGT | ATTCCT |
| Primer2\_Index\_28 | CAAGCAGAAGACGGCATACGAGAT**CTTTTG**GTGACTGGAGTTCAGACGTGT | CAAAAG |
| Primer2\_Index\_29 | CAAGCAGAAGACGGCATACGAGAT**TAGTTG**GTGACTGGAGTTCAGACGTGT | CAACTA |
| Primer2\_Index\_30 | CAAGCAGAAGACGGCATACGAGAT**CCGGTG**GTGACTGGAGTTCAGACGTGT | CACCGG |
| Primer2\_Index\_31 | CAAGCAGAAGACGGCATACGAGAT**ATCGTG**GTGACTGGAGTTCAGACGTGT | CACGAT |
| Primer2\_Index\_32 | CAAGCAGAAGACGGCATACGAGAT**TGAGTG**GTGACTGGAGTTCAGACGTGT | CACTCA |
| Primer2\_Index\_33 | CAAGCAGAAGACGGCATACGAGAT**CGCCTG**GTGACTGGAGTTCAGACGTGT | CAGGCG |
| Primer2\_Index\_34 | CAAGCAGAAGACGGCATACGAGAT**GCCATG**GTGACTGGAGTTCAGACGTGT | CATGGC |
| Primer2\_Index\_35 | CAAGCAGAAGACGGCATACGAGAT**AAAATG**GTGACTGGAGTTCAGACGTGT | CATTTT |
| Primer2\_Index\_36 | CAAGCAGAAGACGGCATACGAGAT**TGTTGG**GTGACTGGAGTTCAGACGTGT | CCAACA |
| Primer2\_Index\_37 | CAAGCAGAAGACGGCATACGAGAT**ATTCCG**GTGACTGGAGTTCAGACGTGT | CGGAAT |
| Primer2\_Index\_38 | CAAGCAGAAGACGGCATACGAGAT**AGCTAG**GTGACTGGAGTTCAGACGTGT | CTAGCT |
| Primer2\_Index\_39 | CAAGCAGAAGACGGCATACGAGAT**GTATAG**GTGACTGGAGTTCAGACGTGT | CTATAC |
| Primer2\_Index\_40 | CAAGCAGAAGACGGCATACGAGAT**TCTGAG**GTGACTGGAGTTCAGACGTGT | CTCAGA |
| Primer2\_Index\_41 | CAAGCAGAAGACGGCATACGAGAT**GTCGTC**GTGACTGGAGTTCAGACGTGT | GACGAC |
| Primer2\_Index\_42 | CAAGCAGAAGACGGCATACGAGAT**CGATTA**GTGACTGGAGTTCAGACGTGT | TAATCG |
| Primer2\_Index\_43 | CAAGCAGAAGACGGCATACGAGAT**GCTGTA**GTGACTGGAGTTCAGACGTGT | TACAGC |
| Primer2\_Index\_44 | CAAGCAGAAGACGGCATACGAGAT**ATTATA**GTGACTGGAGTTCAGACGTGT | TATAAT |
| Primer2\_Index\_45 | CAAGCAGAAGACGGCATACGAGAT**GAATGA**GTGACTGGAGTTCAGACGTGT | TCATTC |
| Primer2\_Index\_46 | CAAGCAGAAGACGGCATACGAGAT**TCGGGA**GTGACTGGAGTTCAGACGTGT | TCCCGA |
| Primer2\_Index\_47 | CAAGCAGAAGACGGCATACGAGAT**CTTCGA**GTGACTGGAGTTCAGACGTGT | TCGAAG |
| Primer2\_Index\_48 | CAAGCAGAAGACGGCATACGAGAT**TGCCGA**GTGACTGGAGTTCAGACGTGT | TCGGCA |