A Fast and Simple Method for Whole-Genome Bisulfite Sequencing Library Preparation from Ultra-low DNA input: Pico-MethylSeq Library Preparation Kit

Karolyn Giang, TzuHung Chung, Xueguang Sun, & Xi-Yu Jia
Zymo Research Corporation, Irvine, CA

ZYMO RESEARCH
The Beauty of Science is to Make Things Simple

Abstract

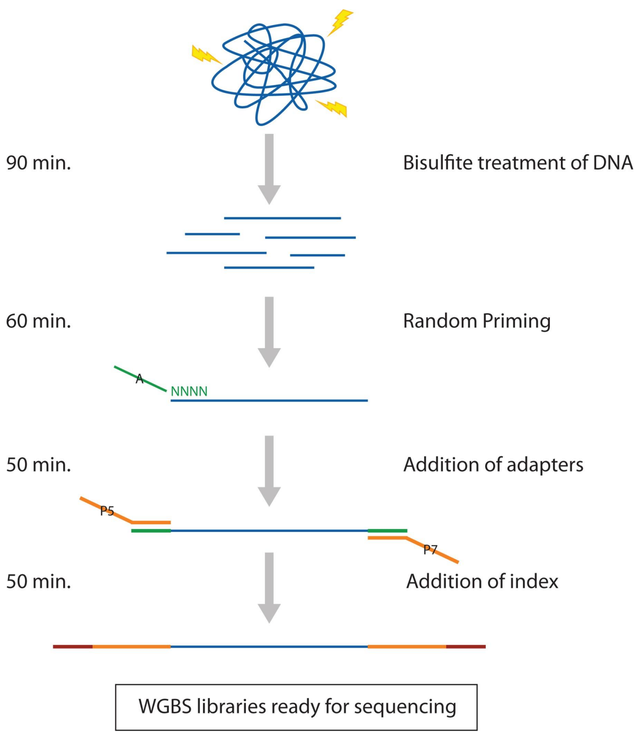
The distribution of 5-methylcytosine (5-mC) in DNA within the eukaryotic genome is known to greatly affect gene regulation and is currently a major topic of research. Studies on DNA methylation have been aided by advancements in bisulfite conversion and next-gen sequencing technologies which, when coupled, provide single-base resolution of 5-mC in the whole genome. Many whole-genome bisulfite sequencing (WGBS) library preparation protocols designed to analyze 5-mC distribution in the whole genome employ bisulfite to convert unmethylated cytosine bases to uracil after the library preparation and while these protocols produce reliable results, degradation of DNA is inherent to bisulfite conversion. As such, a large proportion of the adapterized library is fragmented and can no longer be amplified, which requires these protocols to call for large amounts of starting input DNA that is often times impossible to obtain. By rearranging the order of library preparation and bisulfite conversion, we developed a streamlined protocol that reveals whole-genome methylation patterns at single-base resolution. The work-flow leads with the degradation inherent to bisulfite conversion to randomly fragment the DNA prior to the library preparation, which allows the protocol to accommodate for pico-gram quantities of starting input, making it ideal for analysis in precious and limited samples. Comparisons of sequencing data from this WGBS library preparation method with the established Reduced Representation Bisulfite Sequencing method using human DNA showed a correlation coefficient of 0.95 for CpG sites with more than 10X coverage. With slight modifications, this protocol is versatile in its ability to prepare libraries for ChIP-seq and RNA-seq.

Introduction

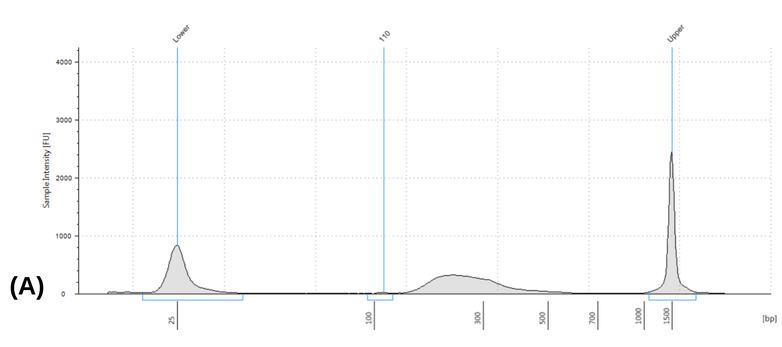
Conventional methods to make the whole genome bisulfite sequencing library typically involves fragmentation of the genomic DNA, end repairing, adapterization, bisulfite conversion and limited amplification. However, this requires large amounts of DNA input and multiple step purifications, which lead to loss of precious DNA sample throughout the library preparation and an under-representation of the methylated sites.

The Pico Methyl-Seq Library Preparation kit provides a simple and fast workflow to make the next generation sequencing libraries for whole genome methylation analysis. The work flow leads with the degradation inherent to bisulfite conversion to randomly fragment the DNA and is coupled with three rounds of amplification using specially designed primers. The protocol can accommodate as little as 10 pg of genomic input in addition to DNA from FFPE samples, and is therefore ideal for methylation analysis in precious and limited samples or target-enriched samples.

Workflow



Validation and Analysis of Library



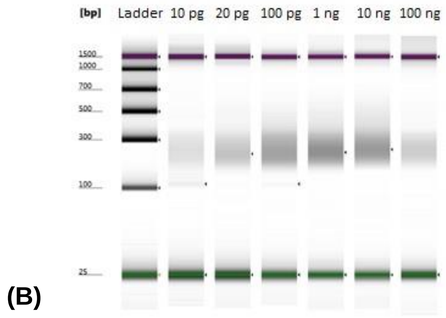


Figure 1. Library validation using the Agilent 2200 TapeStation D1K. (A) Electropherogram of a Pico Methyl-Seq (WGBS) library prepared using 100 pg of human gDNA. (B) Gel of libraries prepared from varied amounts of human gDNA. Fragments ranged between 150 and 350 bp.

	Sample 1		Sample 2	
Library Prep	**WGBS**	**RRBS**	**WGBS**	**RRBS**
Input	100 ng	500 ng	100 ng	500 ng
Total reads	939,332,602	23,840,624	784,068,249	24,720,198
Mapped reads	469,160,132	12,010,064	406,780,493	13,214,097
Unique CpG sites	23,854,084	4,137,561	24,696,217	4,262,581
Genome Coverage	81%	4%	86%	5%
Methylome Coverage	83%	14%	86%	15%
CpG Coverage	23x	7x	17x	7x

Table 1. Overview of sequencing results and quality of libraries. WGBS and RRBS libraries were prepared in parallel from the same starting input. WGBS libraries were sequenced over 4 lanes each and RRBS libraries were 6-plexed in one lane of 50 bp PE on an Illumina HiSeq 2500.

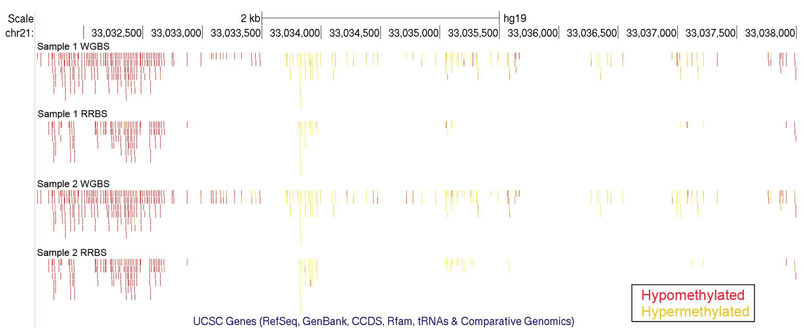
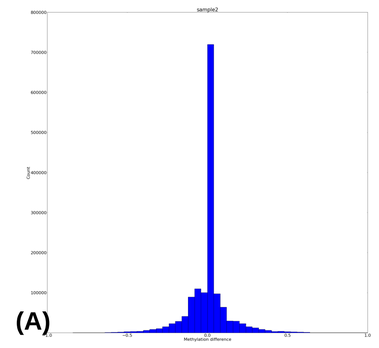


Figure 2. CpG methylation patterns from WGBS and RRBS libraries. Region above spans 5.5 kb human chromosome 21**.** WGBS and RRBS libraries were prepared in parallel from the same starting input.



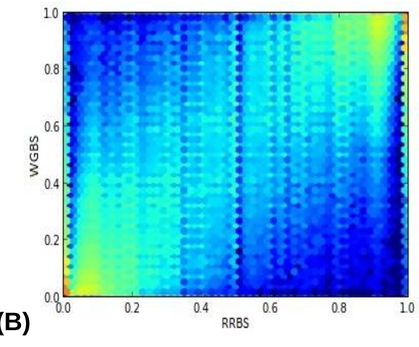


Figure 3. Comparison of methylation values between WGBS and RRBS library prep using the same sample . (A) Distribution of difference in methylation values at individual CpG sites compared between WGBS and RRBS. (B) r = 0.9554

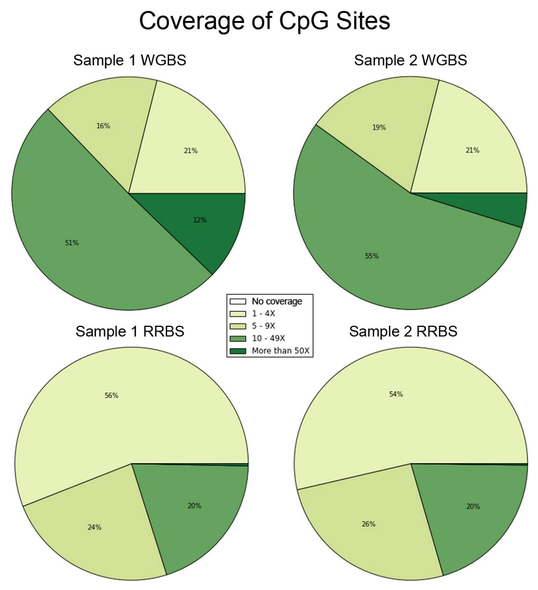


Figure 4. Pie chart representation of overall CpG sequencing reads coverage.

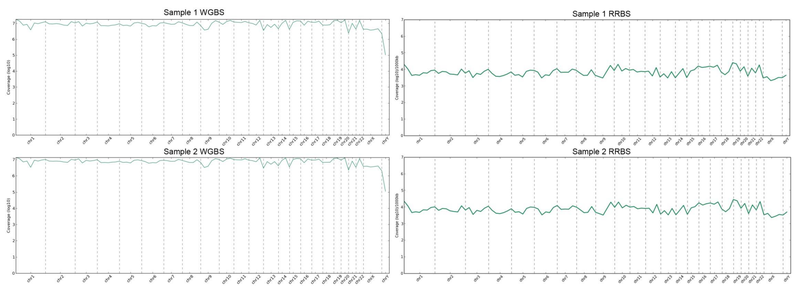


Figure 5. Representation of CpG site coverage across the whole human genome.

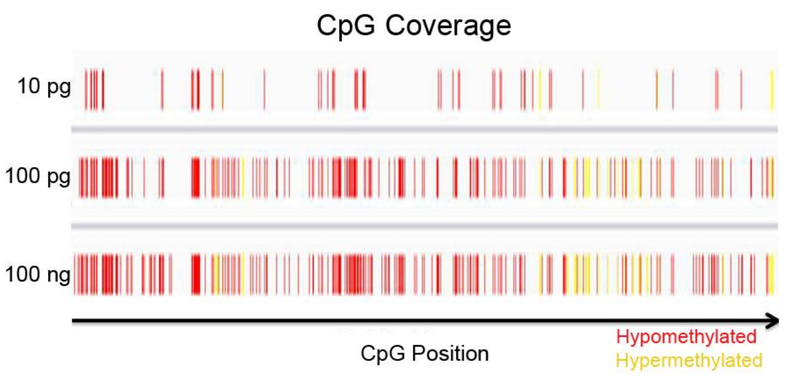


Figure 6. CpG methylation patterns from WGBS libraries generated from varied amounts of human gDNA inputs. Input amount was obtained through serial dilutions. Libraries were 6-plexed in one lane of 50 bp PE on an Illumina HiSeq 2500.

Key features

- ✓ Post-bisulfite library preparation for Whole Genome Bisulfite Sequencing (WGBS)
- ✓ Simple, ligation-free workflow that can be completed in six hours
- ✓ No size-selection of library necessary
- ✓ Accommodates ultra-low DNA input
- ✓ Compatible with fragmented, FFPE, or ssDNA
- ✓ Can be used for enriched DNA (i.e. ChIP DNA, Agilent SureSelect Methyl-Seq)