HIGHER QUALITY SEQUENCING DATA WITH EM-seq LIBRARIES

DATA ANALYSIS

RESULTS

SAMPLE PREPARATION

METHODS

EM-seq was used to identify cDNA from a healthy human donor. EM-seq libraries demonstrated lower duplication rates, higher percentages of mapped reads and less GC bias compared to WGBS libraries. These libraries also identified a higher number of CpGs and demonstrated higher correlation between the EM-seq libraries compared to WGBS libraries. In conclusion, EM-seq libraries have superior sequencing metrics resulting in robust methylation profiling for these types of challenging DNA samples.

SIMILAR GLOBAL METHYLATION LEVELS BETWEEN EM-seq AND WGBS

EM-seq Enables Accurate and Precise Methylome Analysis of Challenging DNA Samples

V. K. Chaithanya Ponnaluri, Louise Williams, Brittany Sexton, Lana Saleh, Katherine Marks, Mala Samarayake, Bradley Langhorst, Zhiiy Sun, C. Thomas Evans, Jr., Romualdas Vaisvila, Eileen Dimalta, Theodore B. Davis

New England Biolabs, Ipswich, MA 01938, USA

The cytosine modifications 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC), are key epigenetic factors that play an important role in cellular processes. Misregulation of these cytosine forms results in disease states like cancer. Advances in sample preparation from biological matrices like blood (cell-free DNA (cDNA)) as well as improvements in sequencing library preparation methods have enabled cancer biomarker identification based on methylation profiling in a minimally invasive manner.

Bisulfite sequencing is the standard method to detect methylation and has been employed for both targeted and whole genome methylation analysis. However, the chemical based bisulfite conversion of cytosines to uracils also results in DNA damage which subsequently results in shorter DNA insert sizes as well as introducing bias into the data. Therefore, analysis of DNA methylation from cDNA is challenging as the DNA is typically of low quality and quantity. Robust biomarker detection relies primarily on the ability to profile methylation accurately. To overcome the drawbacks of bisulfite sequencing, we developed an enzyme based methylation detection technology, called NEBNext Enzymatic Methyl-Seq (EM-Seq-), that minimizes damage to DNA requiring less PCR cycles, lower duplication rates and minimal GC bias resulting in more accurate quantification of methylation in the sample DNA.

EM-seq and bisulfite sequencing were compared using Illumina NovaSeq 6000. (A) EM-seq libraries have higher yield using fewer PCR cycles compared to WGBS. (B) Library duplication percentages are lower for EM-seq. (C) Insert size distribution is similar between EM-seq and WGBS libraries and (D) the GC distribution of EM-seq and bisulfite libraries indicate that EM-seq libraries show more even coverage than bisulfite libraries. The bisulfite libraries are AT rich and have lower GC coverage.

EM-seq and bisulfite libraries at 1x minimum coverage were performed using methylKit 1.4.0. (E) EM-seq libraries identified more unique CpGs than bisulfite libraries at 1x minimum coverage. EM-seq libraries have higher coverage at all DNA inputs across these genomic features.

Identification of CpGs using the EM-seq method is robust compared to whole genome bisulfite sequencing. EM-seq compared to WGBS for cDNA:

• Higher library yields with less PCR cycles
• Lower percent duplication
• Detects more CpGs with fewer reads

Provides new method to evaluate low input cDNA data with high concordance between the replicates for accurate methylation based biomarker detection.

We thank Laurie Mazebold, Danielle Fuchs and Abe Quinby from the NEB Sequencing Core for their assistance.

WGBS WGBS EM 0.75 ± 0.07 0.75 ± 0.07 100 WGBS EM 70 30 90 60 21 EM 6 21 EM 30 EM 100 50 76.45 76.45 0.75 ± 0.07 0.75 ± 0.07 80 80