Enzymatic Methyl-seq enables accurate and robust methylation detection

Genomics

vaishnavi panchapakesa (panchapakesa@neb.com), New England Biolabs, **chaithanya ponnaluri**, NewEngland Biolabs, **Louise Williams**, New England Biolabs, **Matthew Campbell**, New England Biolabs, **Bradley Langhorst**, New England Biolabs, **Eileen Dimalanta**, New England Biolabs, **Theodore Davis**, New England Biolabs

New England Biolabs, Ipswich, MA 01938, USA

DNA methylation is one of the most important epigenetic regulatory mechanisms. The ability to accurately identify 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) gives us greater insight into potential gene regulatory mechanisms. Bisulfite sequencing (BS) is traditionally used to detect methylated cytosines, however, the chemical based conversion of cytosines to uracils leads to DNA damage which subsequently translates to shorter DNA insert sizes as well as biases in the data. To overcome these limitations, we developed NEBNext® Enzymatic Methyl-seq (EM-seq®), an enzymatic approach for detecting cytosine methylation.

EM-seq and BS Illumina libraries were prepared using 10 ng to 200 ng NA12878 DNA. EM-seq libraries have longer inserts and less GC bias compared to bisulfite converted libraries. Global methylation levels are similar between the two methods, indicating overall detection of methylated Cs is similar. However, CpG correlation plots demonstrated higher correlation coefficients indicating that EM-seq libraries are more consistent than BS across replicates and input amount. GC Bias and dinucleotide distribution showed that EM-seq has more even dinucleotide representation compared to the AT rich representation observed for BS. EM-seq libraries exhibit more even coverage allowing for a higher percentage of CpGs to be assessed and therefore leading to more consistent evaluation of methylation across key genomic features (TSS, CpG island, etc.).

There is increasing interest in the diagnostic applications of circulating cell-free DNA (cfDNA). Analysis of DNA methylation from cfDNA is challenging as the DNA is typically of low quantity and quality. EM-Seq and BS libraries were made using cfDNA. EM-seq libraries had longer inserts, lower duplication rates, higher percentages of mapped reads and less GC bias compared to BS libraries. These libraries also identified a higher number of CpGs resulting in enhanced coverage across genomic features, such as transcription start sites (TSS) and CpG islands. EM-seq is robust and reproducible, facilitating the generation of libraries with superior sequencing metrics for these challenging DNA samples.