NEBNext® enzymatic solutions for DNA methylation profiling of highly damaged DNA inputs

Robin L. Armstrong, Vaishnavi Panchapakesa, Margaret R. Heider, Laura N. Blum, Brittany S. Sexton, Daniel J. Evanich, V. K. Chaithanya Ponnaluri, Bradley W. Langhorst, Louise Williams

New England Biolabs, Inc. 240 County Road, Ipswich, MA, 01938. USA

INTRODUCTION

Advances in Next Generation Sequencing (NGS) technologies have enabled large-scale quantification of DNA methylation. DNA isolated from archival material like formalin fixed paraffin embedded (FFPE) tissues is an invaluable resource for genome-wide methylation studies and has advanced the field of cancer genetics. However, FFPE DNA poses many notable challenges for preparing NGS libraries, including low input amounts and highly variable damage from fixation, storage, and extraction methods leading to insufficient coverage and sequencing artifacts.

To investigate the potential to improve DNA methylation profiling from FFPE samples, Illumina compatible libraries were constructed using FFPE DNA with DNA integrity numbers (DIN) ranging from 1.8 to 5.8. DNA was fragmented using mechanical shearing (Covaris) or enzymatic fragmentation (NEBNext UltraShear®). An enzyme-based methylation detection technology, NEBNext® Enzymatic Methyl-seq v2 (EM-seq v2), was used to detect DNA methylation. This system overcomes the limitations of chemical-based bisulfite conversion methods and minimizes damage to DNA, enabling longer insert sizes, lower duplication rates, and minimal GC bias resulting in more accurate quantification of DNA methylation. Additionally, DNA methylation was profiled in clinically relevant FFPE DNA from multiple tissue types using EM-seq v2 coupled with UltraShear fragmentation. Results for these challenging DNA types showed that the UltraShear fragmented EM-seq v2 libraries had lower duplication rates, lower chimeras, as well as higher percentages of mapped reads and increased library complexity compared to Covaris sheared libraries. These libraries exhibited higher unique coverage depth and uniformity, leading to more CpGs confidently detected.

EM-seq v2 coupled with UltraShear fragmentation produces reliable methylation profiles and improves library preparation success rates from challenging FFPE samples. The workflow is robust, flexible, and automation friendly, advancing the use of methylation profiling as a cancer biomarker.

METHODS

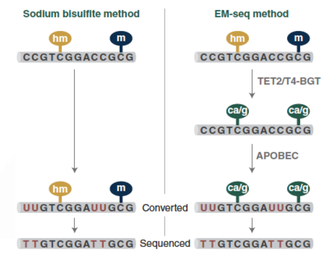
Experimental Overview

Illumina libraries were made to investigate the effects of different library prep methods on methylation in FFPE samples. Target capture was performed using the Twist Methylome panel.

DNA used:
- 10 ng normal liver FFPE DNA (DIN 3.5)
- 100 ng Tumor-normal matched FFPE DNA (DIN Range: 1.8 to 5.8)

DNA fragmentation and conversion method:
- Covaris fragmentation followed by bisulfite conversion
- Covaris fragmentation followed by EM-seq v2
- UltraShear fragmentation followed by EM-seq v2

Comparison of Enzymatic & Bisulfite Conversion methods

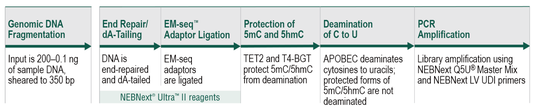


Bisulfite Conversion: Detection of 5mC & 5hmC

Library Construction Workflow:
- DNA was Covaris sheared, end-repaired and ligated to EM-seq adaptors
- Conversion using Zymo EZ DNA Methylation-Gold Kit (D5006)
- Libraries were amplified using NEBNext® Q5U® Master Mix

EM-seq v2: Enzymatic Detection of 5mC & 5hmC

Library Construction Workflow:
- DNA was fragmented using either NEBNext UltraShear or Covaris, followed by end-repair and ligation to EM-seq adaptors
- TET2 and T4-BGT were used to protect 5mC and 5hmC from deamination by APOBEC
- Libraries were amplified using NEBNext® Q5U® Master Mix



Sequencing and Data Analysis

- Reads were adaptor trimmed (fastp) then aligned to a composite human hs1 + methylation controls using bwa-meth
- 5mC and 5hmC information was extracted from the alignments using MethylDackel and levels were evaluated independently for each chromosome
- Picard was used to mark duplicates as well as calculate library insert size distributions and GC bias
- Differentially methylated CpGs and regions were called using defiant (P<0.05, 20% methylation)

RESULTS

NEBNext UltraShear coupled with EM-seq v2 enables high quality libraries from FFPE DNA

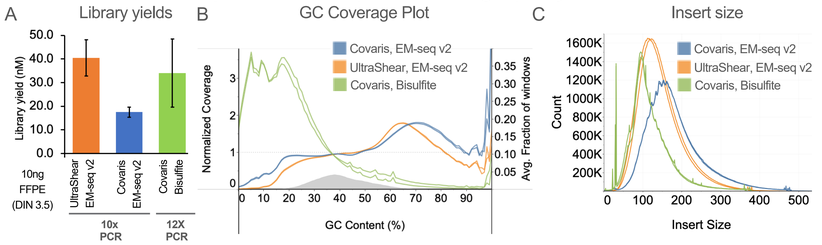


Figure 1. Libraries were prepared in duplicate from 10 ng of normal liver FFPE tissue DNA (DIN 3.5; BioChain Institute). FFPE DNA was fragmented by NEBNext UltraShear (enzymatic) or Covaris (mechanical) followed by EM-seq v2 conversion and 10 PCR cycles or Covaris shearing followed by sodium bisulfite conversion and 12 PCR cycles. Libraries were sequenced (2x100 bases) on an Illumina NovaSeq 6000. 387 million total reads were used for the analysis and aligned to the human hs1+controls genome. (A) UltraShear fragmented libraries produce higher yields. (B) GC coverage plot showing better GC representation for the EM-seq v2 libraries compared to bisulfite libraries. (C) Covaris and UltraShear EM-seq v2 libraries have longer inserts than bisulfite converted libraries.

RESULTS

Methylation and library quality metrics

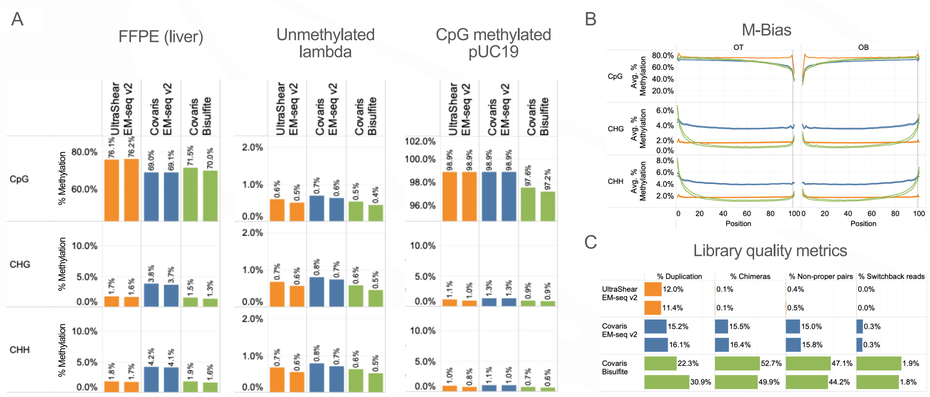


Figure 2. (A) Percent methylation detected in CpG, CHG and CHH contexts for EM-seq v2 and bisulfite converted liver FFPE libraries. Unmethylated Lambda control levels indicate efficient deamination. CpG methylated pUC19 controls show efficient 5mC/5hmC protection for EM-seq v2 converted libraries. (B) M-bias plot showing the level of methylation observed across the read in the CpG, CHG and CHH contexts for Original Top (OT) and Original Bottom (OB) strands. UltraShear fragmented libraries demonstrate improved end-repair effect on M-bias compared to Covaris sheared samples. (C) Library quality metrics were assessed using Picard Alignment summary metrics and fgbio FindSwitchbackreads. UltraShear fragmented FFPE DNA exhibits lower chimeras, duplication, non-proper pairs and fold-back reads resulting in higher quality libraries with higher fraction of useable reads.

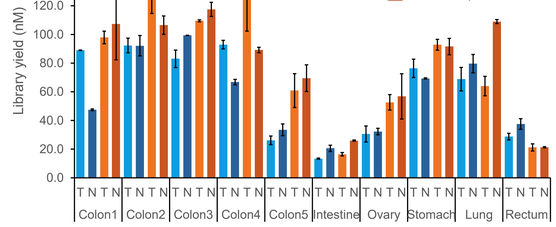
Robust library preparation from varying qualities of tumor-normal FFPE patient samples

A

Matched tumor-normal sample information

Sample name	Sample type	Sex	Sample type	Tumor/ Normal	DIN
Colon1	Colon adenocarcinoma	F	FFPE	Tumor	4.5
				Normal	4
Colon2	Colon adenocarcinoma	M	FFPE	Tumor	5.7
				Normal	5.8
Colon3	Colon adenocarcinoma	M	FFPE	Tumor	5
				Normal	5.6
Colon4	Colon adenocarcinoma	M	FFPE	Tumor	5.1
				Normal	5
Colon5	Colon adenocarcinoma	F	FFPE	Tumor	4.5
				Normal	4.5
Intestine	Intestine adenocarcinoma	M	FFPE	Tumor	1.9
				Normal	2.4
Ovary	Ovary serous surface papillary carcinoma	F	FFPE	Tumor	3.6
				Normal	3.9
Stomach	Stomach signet ring cell carcinoma	F	FFPE	Tumor	3.9
				Normal	4.6
Lung	Lung small cell carcinoma	M	FFPE	Tumor	2.6
				Normal	4.3
Rectum	Rectum adenocarcinoma	F	FFPE	Tumor	2.1
				Normal	1.8

B



C

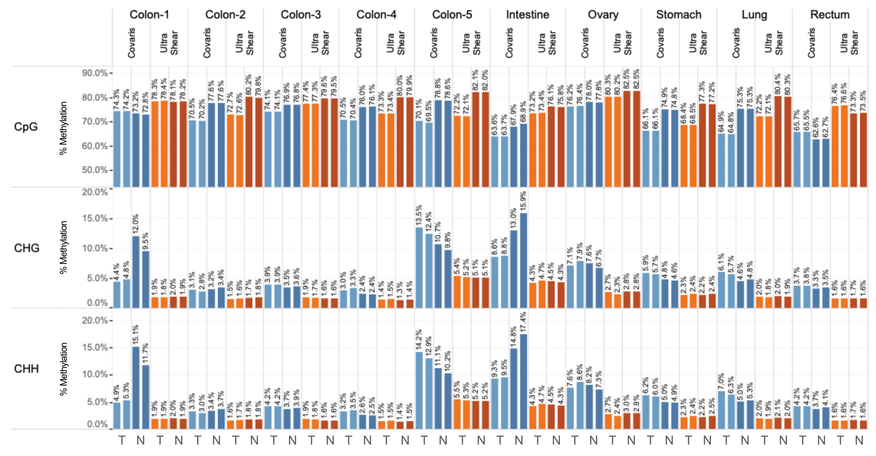


Figure 3. Libraries were prepared using 100 ng of tumor-normal matched FFPE patient sample DNA. FFPE DNA was fragmented by NEBNext UltraShear or Covaris followed by EM-seq v2 and 8 PCR cycles. Libraries were sequenced (2x100 bases) on an Illumina NovaSeq 6000. 18 million total reads were used for analysis and aligned to the human hs1+controls genome. (A) Ten different patient tumor-normal matched samples with DINs ranging from 1.8-5.8. (B) Average library yields for each matched tumor/normal sample. C) Percent methylation detected in the CpG, CHG and CHH contexts for EM-seq v2 libraries. UltraShear libraries exhibit higher overall CpG methylation and lower CHG and CHH levels compared to Covaris libraries.

RESULTS

Robust CpG correlations in target enriched tumor-normal matched EM-seq v2 libraries

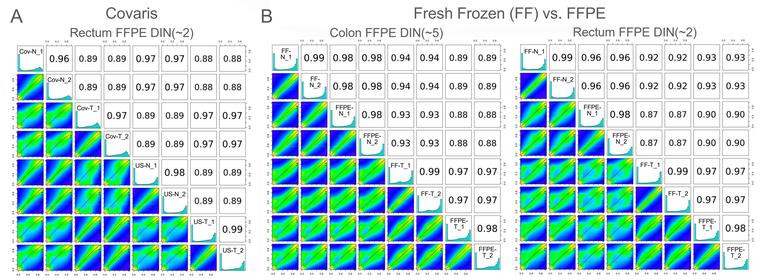


Figure 4. Libraries were prepared from 100 ng of tumor-normal matched Fresh Frozen (FF) and FFPE patient samples. DNA was fragmented by NEBNext UltraShear or Covaris followed by EM-seq v2 and target capture using the Twist Methylome panel. Reads were aligned to the human T2T genome. Pearson correlations of single base methylation levels at 5X minimum coverage between (A) Covaris and UltraShear fragmentation of tumor-normal matched Rectum FFPE DNA samples using 751M reads (9,636,392 CpGs). (B) UltraShear fragmentation of tumor-normal matched Colon (Colon4) and Rectum fresh frozen and FFPE DNA samples using 825M reads (11,935,735 CpGs Colon, 10,649,544 CpGs Rectum).

Hypomethylation and hypermethylation detected in colon and rectum tumor samples

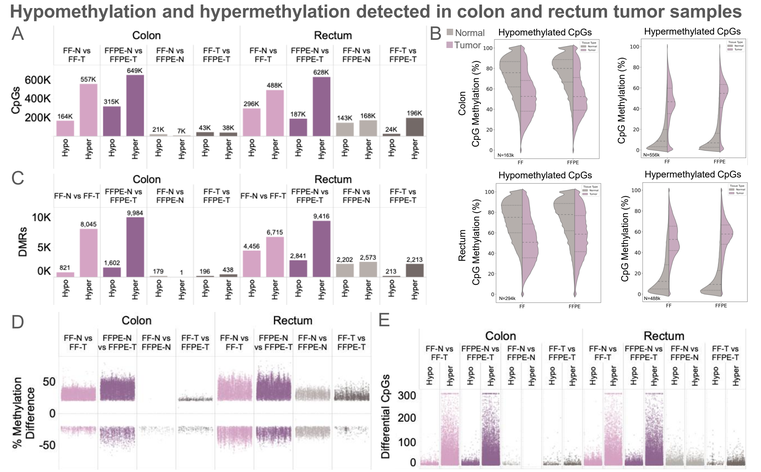


Figure 5. CpG methylation of tumor-normal matched Colon and Rectum fresh frozen and FFPE DNA samples using 825M reads each with median target coverage ~200X except for rectum FFPE (~100X). (A) Hypo- and hyper-differentially methylated CpGs with at least 20% difference in methylation (P<0.05, 5X coverage). (B) Methylation distribution of hypo- and hyper-methylated CpGs identified in the fresh frozen tumor/normal samples, shown for both fresh frozen and FFPE. (C) Hypo- and hyper-differentially methylated regions (DMRs) with at least 20% mean methylation difference and 10 differential CpGs (P<0.05, 5X coverage). (D) Mean methylation difference of DMRs. (E) Number of differential CpGs per DMR. DMRs with more than 300 differential CpGs are shown at 300.

Hypermethylated DMRs are associated with clinically-relevant genes

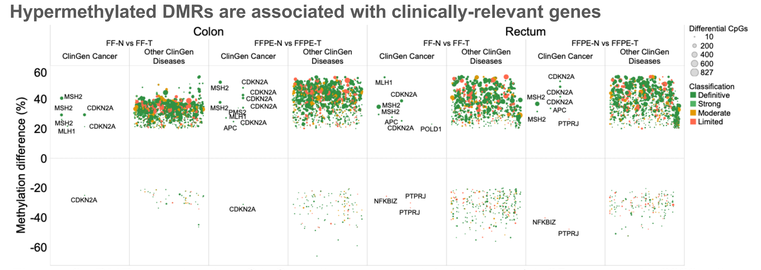


Figure 6. DMRs overlapping ClinGen disease-associated genes (N=35 cancer, N=2660 other diseases), sized by the number of differential CpGs per region and colored by classification of association evaluated by expert panels. DMRs were intersected with genes including 1kb upstream of the gene start. 235 genes with DMR(s) where at least 20 differential CpGs were found in the FF normal vs. FFPE normal or FF tumor vs. FFPE tumor comparisons were interpreted as possible preservation artifacts and excluded for clarity.

CONCLUSIONS

- EM-seq v2 libraries produce higher library yield and data quality compared to bisulfite libraries.
- UltraShear fragmented libraries exhibit lower duplication, chimeras and non-proper pairs resulting in higher quality libraries with higher fraction of useable reads.
- Robust, high-quality EM-seq v2 libraries can be prepared with varying quality of tumor-normal FFPE matched samples and used to detect differentially methylated sites in genes of interest.

ACKNOWLEDGEMENTS

We thank Kristen Augulewicz, Harry Bell, Rebecca Gawron, and Dora Posfai for their technical support and members of NEBNext for productive discussions.

Poster # 403