Utilization of MBD-seq to Elucidate Differentially Methylated Regions in cfDNA Across Healthy and Cancer Patients

Camille Michelon[1], **Shuxiong Wang**[1], **Vidyodhaya Sundaram**[2], **Brian Egan**[1], **Benjamin Delatte**[1]
[1]**Active Motif, Inc.**, Carlsbad, CA 92008 [2]**BioChain, Inc.**, Newark, CA 94560

Poster # 4415

ACTIVE MOTIF®
Enabling Epigenetics Research

Introduction

5-Methylcytosine (5-mC) is the most abundant epigenetic DNA modification in eukaryotes. The presence of 5-mC alters gene expression and genome metabolism, and its deregulation is associated with many human diseases. Cell-Free DNA (cfDNA) released from apoptotic cells has recently gained attention, as mutations or epimutations (*e.g.* 5-mC variations) detected in cfDNA were shown to have high diagnostic potential to assess the presence, stage and outcome of several cancers. However, 5-mC detection in cfDNA remains a challenge as input material is often very scarce. MBD-Seq is a method that leverages the ability of Methyl-CpG-binding domain protein 2 (MBD2) to capture and detect highly methylated regions of DNA genome-wide. We evaluated MBD-Seq for the detection of aberrant methylation patterns in cfDNA from healthy and diseased patients. We found MBD-seq on cfDNA to be specific (> 100-fold mC:C detection ratio) and sensitive (5 ng of input cfDNA). Applying MBD-seq to samples from healthy vs colorectal (CRC-IV) and breast (B-IIIC, B-IV) cancer patients showed striking differences in methylation of key tumor-associated genes. Hence, MBD-Seq can elucidate aberrant methylation patterns from cfDNA, highlighting its potential as a diagnostics tool for the detection of cancer.

Schematic Workflow of Active Motif's cfMethylCollector Technology

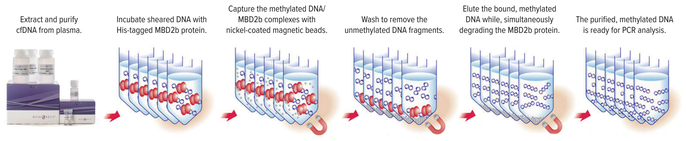


Figure 1. cfDNA is extracted and purified from plasma using the Active Motif Cell-Free DNA (cfDNA) Purification Kit™, spiked with 1% (ng) spike-in DNA, then introduced into a binding reaction alongside His-MBD2b, nickel beads, and high salt binding buffer for 1 hour. MBD2b binds to methylated spike-in DNA and methylated CpG islands (CGI) within cfDNA and forms DNA-Protein-Bead complexes. The complexes are pelleted on a magnet, unbound (unmethylated) material is washed and discarded. The enriched pulldown is eluted in a Proteinase K (PK) digestion buffer for 30 minutes. Silica beads are used to recover and purify DNA. The final pulldown undergoes qPCR QC on methylated and unmethylated spike-in DNA targets. If % methylated input ÷ % unmethylated input ratio is ≥ 50x, there is a quality enrichment suitable to proceed with library prep and sequencing. Assay time before library prep is approximately 4 hours (including qPCR QC).

Correlation with Published Low Input MBD-Seq Data

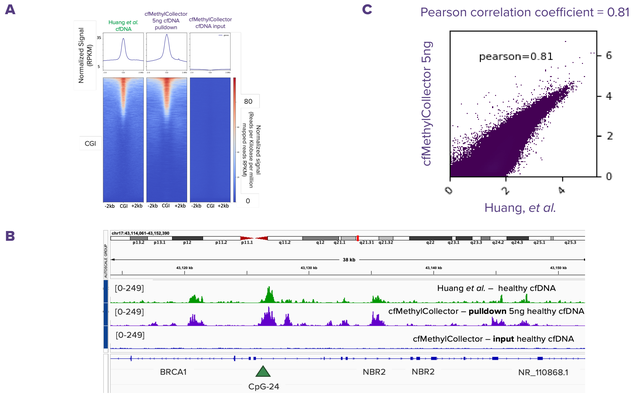


Figure 2 A. CGI Aggregation plot. Heatmap intensity indicates signal overlap between cfDNA sample and regions within reference CpG islands (CGI) file. Each row (y-axis) is a CGI in the reference file spanned over +/- 2kb (x-axis). Green is reference data [1], Purple is cfMethylCollector data.

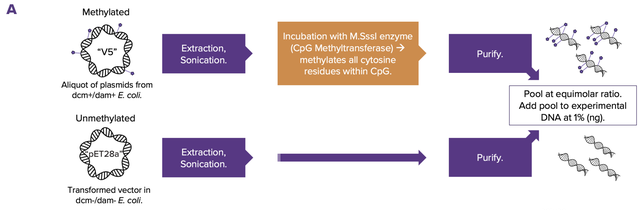
Figure 2 B. IGV Tracks. Integrative genome viewer at canonical methylated locus, *NBR2*, arrow indicates CpG island. Both samples have CpG-specific signal of similar amplitude.

Figure 2 C. Pearson Correlation. cfMethylCollector pulldown and reference cfDNA public data for low input MBD-seq have a good correlation coefficient. Bin size is 5kb.

References

1. Huang J et al. 2021 Mar 16. Cell-free DNA methylome profiling by MBD-seq with ultra-low input. Epigenetics.:1–14. doi:https://doi.org/10.1080/15592294.2021.1896984. SRR13040548
2. Zhang S et al. 2017. Distinct prognostic values of S100 mRNA expression in breast cancer. Scientific Reports. 7(1). doi:https://doi.org/10.1038/srep39786.
3. Torres-Martinez Z et al. 2021. Key genes and drug delivery systems to improve the efficiency of chemotherapy. Cancer drug resistance (Alhambra, Calif). 4:163–191. doi:https://doi.org/10.20517/cdr.2020.64.

QC on Spike-In DNA



B

Sample	Cancer	Color
1	Colorectal cancer stage 4 (CRC-IV)	Red
2	Colorectal cancer stage 4 (CRC-IV)	Green
3	Breast cancer stage 3C (B-IIIC)	Yellow
4	Breast cancer stage 4 (B-IV)	Purple

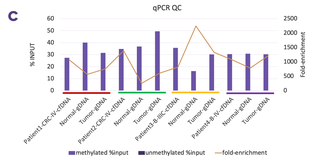


Figure 3 A. Spike-In DNA Production. Methylated spike-in DNA originates from v5 plasmid propagated in dcm+/dam+ (DNA methyltransferases) *E. coli*, extracted, then sonicated using PIXUL™. M.SssI incubation was performed to methylate all cytosine residues in CpG dinucleotide context. Unmethylated Spike-In DNA was produced from pET28 plasmid propagated in dcm-/dam- *E. coli*, extracted, sonicated, and purified using same parameters. After quantification, methylated and unmethylated DNA were combined equimolarly and spiked at 1%.

Figure 3 B. Sample Overview. Cancer type and stage are specified across 4 patients. Red (patient 1, CRC-IV), Green (patient 2, CRC-IV), Yellow (patient 3, B-IIIC), Purple (patient 4, B-IV).

Figure 3 C. qPCR QC. Percent input and fold-enrichment for cfDNA, normal gDNA, and tumor gDNA are displayed across 4 patients. Purple bars represent % input of methylated spike-in DNA, Black bars represent % input of unmethylated spike-in DNA.

$$\%\ \text{input} = \left(2^{-(Avg\ Pulldown\ Ct - Avg\ Input\ Ct)}\right) x\ 100$$

Orange line represents fold-enrichment of spike-in DNA.

$$\text{Fold-enrichment} = \frac{Methylated\ plasmid\ \%\ input}{Unmethylated\ plasmid\ \%input}$$

To meet QC criteria for a quality enrichment (depletion of unmethylated, enrichment of methylated spike-in DNA), fold-enrichment must be ≥ 50x. All samples meet QC criteria.

Sequencing Statistics

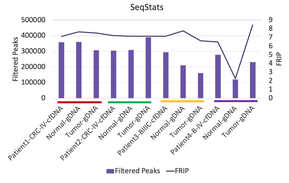


Figure 4. Sequencing Statistics. FRiP (Fraction of reads in peaks) and filtered peaks are plotted for each sample. Red (patient 1, CRC-IV), Green (patient 2, CRC-IV), Yellow (patient 3, B-IIIC), Purple (patient 4, B-IV).

CpG Island (CGI) Aggregation Plots

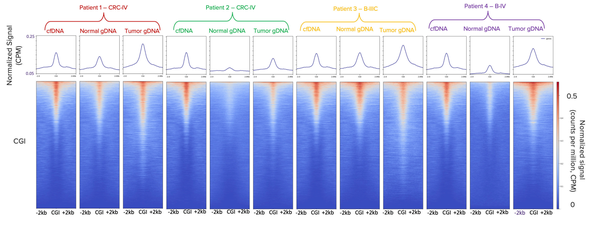


Figure 5. CpG Island (CGI) Aggregation Plots.

Top plot, Aggregate Signal. Mean signal overlap between sample and reference CGI file across the genome. Signal intensity (y-axis) is centralized and spanned over +/- 2kb of CGI regions (x-axis).

Bottom plot, Heatmap. Heatmap intensity indicates signal overlap between cfDNA sample and regions within reference CpG islands (CGI) file. Each row (y-axis) is a CGI in the reference file spanned over +/- 2kb (x-axis).

Strong CGI-specific signal demonstrates assay specificity. Generally, tumor gDNA and cfDNA exhibits higher aggregate signal than normal gDNA. Red (patient 1, CRC-IV), Green (patient 2, CRC-IV), Yellow (patient 3, B-IIIC), Purple (patient 4, B-IV).

Enrichment of Methylated DNA by cfMethylCollector

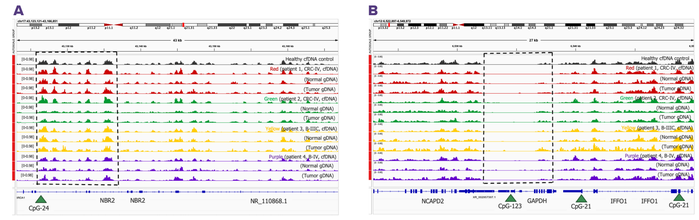


Figure 6 A. *NBR2*. IGV at canonical methylated locus, *NBR2*, arrow indicates CpG island. cfDNA, normal gDNA, and tumor gDNA across four cancer patients are displayed. Signal at methylated locus indicates specificity for enrichment of methylated DNA. Red (patient 1, CRC-IV), Green (patient 2, CRC-IV), Yellow (patient 3, B-IIIC), Purple (patient 4, B-IV).

Figure 6 B. *GAPDH*. Lack of signal at canonical unmethylated locus, *GAPDH*, indicates cfMethylCollector specificity for enrichment of methylated DNA.

Differential Methylation Region (DMR) Analysis

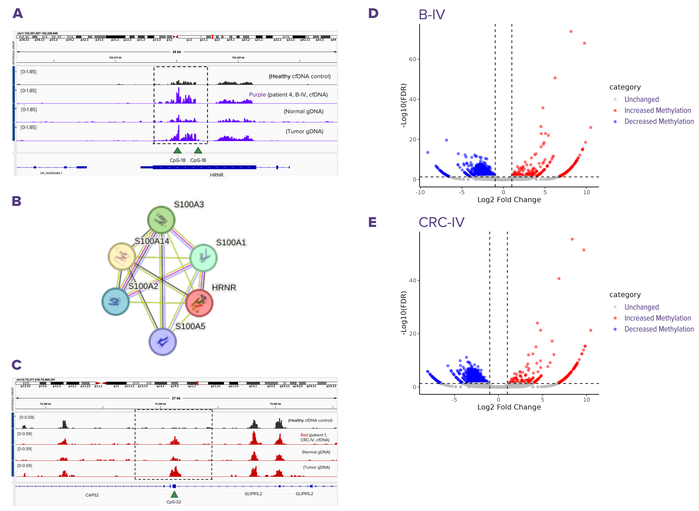


Figure 7 A. DMR Analysis, B-IV. *HRNR* gene identified as significant (p-value = 1.91E-05, FDR = 0.01) DMR between healthy cfDNA control and cfDNA from patient 4, B-IV (Purple).

Figure 7 B. STRING Analysis. *HRNR* belongs to S100 gene family, its suppression is associated with breast cancer progression in transformation from preinvasive to invasive carcinoma [2].

Figure 7 C. DMR Analysis, CRC-IV. *GLIPR1L2* gene identified as significant (p-value = 2.17E-05, FDR = 0.01) DMR between healthy cfDNA control and cfDNA from patient 1, CRC-IV (Red). *GLIPR1L2* is a paralog of *GLIPR1*, a tumor-suppressor gene with apoptosis-inducing activities [3].

Figure 7 D, E. Volcano Plots. Distribution of significant (FDR) methylation fold-changes for D. B-IV (patient 4, Purple) and E. CRC-IV (patient 1, Red) relative to healthy cfDNA control.

Summary

1. cfMethylCollector utilizes MBD2b to bind and pulldown methylated CpG-islands (CGI) and prepare high quality libraries with as little as 5ng cfDNA.
2. qPCR on spike-in DNA targets enables post-enrichment QC to ensure quality enrichment before library preparation and sequencing.
3. cfMethylCollector elucidates differentially methylated regions (DMRs) in cfDNA. Analogous methylation patterns and DMRs are found between cfDNA and tumor gDNA relative to normal gDNA and cfDNA controls.
4. cfDNA can be a surrogate molecule to assess aberrant methylation patterns in cancer patients without the need for invasive tumor biopsy.

BioChain®

Patient genomic DNA was extracted using BioChain AnaPrep 12 Dx instrument. cfDNA was collected by consenting donors courtesy of BioChain. Healthy cfDNA control obtained from BioChain Cat No. #Z2121001.