DNA Methylation Guide

Overview, Methods, and Resources



Enabling Epigenetics Research

DNA Methylation Guide

Overview of DNA Methylation

Since first discovered in bacteria nearly 100 years ago, DNA methylation has been shown to be a crucial epigenetic regulator of gene expression and genomic organization in nearly all organisms. In eukaryotes, this modification of DNA is found in cytosines (C), where DNA Methyl Transferases (DNMTs) mediate the transfer of a methyl group to cytosines, converting them to 5-methylcytosine (5-mC). This minor DNA modification can change the activity of DNA without changing the primary sequence itself.

DNA methylation appears almost exclusively in the context of CpG dinucleotides. Overall, these dinucleotides are relatively rare in the mammalian genome and are mainly clustered in what are called CpG islands, ranging between 500-2,000 base pairs in length. In humans, >60% of these CpG islands are in gene promoters. While most CpG dinucleotides are methylated, those within promoter regions are almost always unmethylated and are associated with transcriptionally active genes. DNA methylation of promoters is linked with several fundamental processes including genomic imprinting, X-chromosome inactivation, and repression of transposable elements. Aberrant changes in DNA methylation patterns are associated with many diseases as well as aging.

DNA methylation is reversible. TET enzymes, a family of ten-eleven translocation (TET) methylcytosine dioxygenases are central for DNA demethylation. These enzymes convert 5-mC to 5-hmC. Their further action can catalyze the conversion of 5-hmC to 5-formylctyosine (5-fC) and then to 5-carboxyctyosine (5-caC). Both 5-fC and 5-caC can then be removed from the DNA sequence by base excision repair and replaced by an unmethylated cytosine. While 5-hmC was originally believed to be an incidental mark on the way to demethylation, evidence has since shown that the Tet family of proteins and 5-hmC are in fact involved in normal development as well as many disease states. In mammals, 5-hmC is found in high levels in the brain and in other tissues of the central nervous system.

Many scientists have engaged in studies to understand the distribution of DNA methylation, leading to the emergence of numerous methods for mapping 5-mC and 5-hmC. The sheer number of options can be overwhelming and make it difficult to know where to start. However, while there are many approaches for studying DNA methylation, the best tactic can be determined based on the goals and resources for a particular experiment, with each method having advantages and drawbacks.



What's Inside

Global Quantification of DNA Methylation 1

- Overview
- Methods and Comparison
- Comparison of Methods
- Sample Data and Tools to Study Global Quantification
- Products Available to Study Global DNA Methylation Levels

DNA Methylation Enrichment 3

- Overview
- Methods and Comparison
- Comparison of Methods
- Sample Data and Tools for DNA Enrichment
- Products & Services for Genome-Wide & Gene-Specific Analysis

Single-Nucleotide Resolution Analysis of DNA Methylation 7

- Overview
- Methods and Comparison
- Comparison of Methods
- Sample Data and Tools for Single-Nucleotide Analysis
- Products & Services for Single-Nucleotide Resolution 5-mc Analysis

Additional Resources 11

- Webinars
- Podcasts
- Blogs

Global Quantification of DNA Methylation

Overview

Global measurements of DNA methylation refer to the total level of 5-mC or 5-hmC content relative to the total cytosine content in sample DNA. Variations in overall genomic DNA methylation levels are observed across different developmental states and tissue types, as well as in various disease states. Research indicates a global decrease in 5-hmC levels associated with numerous cancers, while changes to the abundance of 5-hmC has been associated with neurodegenerative disorders making it a proposed molecular biomarker.

Methods and Comparison

Common methods used to measure global DNA methylation levels include chromatographic methods such as HPLC and mass spectrometry. However, while these methods are very sensitive and accurate, they require a high level of expertise, are low throughput and time-consuming. Global ELISA-based DNA methylation assays are a popular alternative method because they are easy to conduct and only require standard lab equipment. For these types of assays, DNA is first immobilized onto the wells of a 96-well plate. Methylated or hydroxymethylated cytosines in the DNA are then identified using either a 5-mC or a 5-hmC antibody, HRP-conjugated secondary antibody, and colorimetric detection reagents. By generating a standard curve using DNA with known methylation, the percentage of 5-mC or 5-hmC can be determined for each DNA sample. Clinical researchers can use these high-throughput assays when screening large numbers of patient samples from their study cohorts.

Method	Global 5mC/5hmC ELISA	LC-MS/MS
Method Overview	Total DNA or a subsampled surrogate is bound to wells of an ELISA plate. A capture antibody for 5mC/5hmC binds to methylated DNA. Secondary antibody and developer are used for colorimetric detection.	Liquid chromatography coupled with mass spectrometry is used on digested DNA.
Input	> 5-10 ng	50-100 ng
Resolution	Genome	Genome
Coverage	Global quantitation of methylation in DNA.	Global quantitation of methylation in DNA.
Advantages	Easy and inexpensive to use. HTP friendly.	Highly sensitive. Works with poor quality DNA (FFPE).
Disadvantages	Does not tell you which regions of the genome have methylation changes.	Requires specialized equipment. Low through-put.
Cost	\$	\$
Best Application	Useful for screening/monitoring samples for changes in DNA methylation related to disease, aging, or environmental factors like smoking or obesity.	Considered "gold-standard" for global measurements of DNA methylation quantification.

Comparison of Methods



LINE-1 Methylation in cfDNA from Human Serum

Global DNA Methylation – LINE-1 Assay shows changes in 5-mC levels in cfDNA with age.

cfDNA was isolated from human serum samples from healthy "normal" adults <25 yo (N=6, mean: 21.83) and >60 yo (N=9, mean:65.89) Following the Global DNA Methylation Line-1 Kit protocol, without Msel digestion, 25 ng of cfDNA was used per well. The % 5-mC for each sample was calculated as a percentage of detectable CpG residues in the LINE-1 probe region, using the methylated DNA standards included in the kit. All samples were done in triplicate. Boxplot shows 25%-75% (\Box), min-max (I), average (x), and median (-) % 5-mC.

Products Available to Study Global DNA Methylation Levels

Product	Format	Catalog Number
Global DNA Methylation Assay-LINE-1	1 x 96 rxns	55017
Global 5-hmC DNA ELISA Kit	1 x 96 rxns	55025

DNA Methylation Enrichment

Overview

Enrichment methods for mapping DNA methylation rely on the selective capture of methylated fragments from the rest of the genomic DNA. This is typically accomplished by antibody immunoprecipitation methods or with enzymes such as methyl-CpG binding domain (MBD) proteins or beta-glucosyltransferase which are used to selectively tag methylated fragments.

While the resolution for methylated (5-mC) and hydroxymethylated (5-hmC) detection is limited due to the size of the DNA fragments captured, typically between 100-300 nucleotides, the methods greatly reduce costs as only a small fraction of the DNA requires sequencing. This makes them great screening methods for identifying differently methylated regions of DNA between samples, especially when there is no prior knowledge of where it might occur. Once differently methylated regions are determined, methods with higher resolution can be used to focus on just those regions.

Methods and Comparison

MeDIP/hMeDIP

Methylated DNA Immunoprecipitation (MeDIP) and hydroxymethylated DNA Immunoprecipitation (hMeDIP) are antibody immunoprecipitation methods that utilize methylation specific antibodies to recognize and capture 5-mC and 5-hmC respectively. Antibodies specific for either 5-mC or 5-hmC are used to selectively tag methylated genomic DNA fragments. Antibody tagged fragments are then captured by magnetic beads that bind the FC region of the antibody. The enriched methylated DNA fragments can then be sequenced.

MeDIP and hMeDIP can capture any methylated cytosines regardless of the sequence context as the antibodies do not require CpGs for binding. However, the antibody will be biased towards capturing highly methylated regions of DNA as the antibodies generally require more than one methylated cytosine per fragment to bind.

MBD-Seq

MBD-capture or MBD-seq is an enzymatic based method for the enrichment of methylated DNA fragments. Instead of an antibody it uses recombinant methyl-binding protein MBD2b to selectively bind 5-methylcytosine DNA. Sheared DNA is mixed with His-tagged MBD2b protein so that methylated DNA can then be captured by nickel-coated magnetic beads.

One advantage of the methyl-CpG binding protein enrichment strategy over MeDIP, is the input DNA sample does not require denaturing, as the protein can recognize methylated DNA in its native double-strand form. Also, the input required for MBD-seq is lower than for MeDIP, making this method ideal for cfDNA and looking for biomarkers related to diseases like cancer. Note that the MBD2b protein only binds to CpG methylated DNA, making this technique ideal for researchers studying CpG Islands but it will miss non CpG methylation.

hME-Seal

hME-Seal uses a beta-glucosyltransferase (βGT) enzyme to transfer an engineered glucose moiety, UDP-Azide-Glucose, onto the hydroxyl group of 5-hmC. A chemical labeling reaction then attaches a biotin to the glucose so that fragments can be captured by streptavidin beads.

Because β GT modifies 5-hmC regardless of sequence context, both CpG and non-CpG methylation can be identified. Like MBD-seq, the process works with dsDNA and works for low input samples. This makes the technique ideal for samples like cfDNA, which are often used when looking for biomarkers related to diseases like cancer.

Comparison of Methods

Method	MeDIP-Seq or hMeDIP-Seq	MBD-Capture/MBD-Seq	5hmC-Seal
Method Overview	Uses a methylation (5mC/5hmC) specific antibody to perform immunoprecipitation on fragmented DNA.	A tagged methyl-CpG binding domain (MBD) protein is used to capture methylated CpG from fragmented DNA.	Uses beta-Glucosyltransferase to selectively transfer glu- cose from a UDP-Azide-Glucose donor to 5hmC. A biotin conjugate is then attached to the modified glucosyl-5-hmC so that tagged fragments can be captured by streptavidin beads.
Input	>100 ng	>5-10 ng	>10 ng
Resolution	100-500 bp	100-500 bp	100-500 bp
Coverage	~88% CpG sites	Up to 27 million CpGs	>300k 5-hmCs
Advantages	Covers both CpG and non CpG sites.	Requires less sample than MeDIP. More sensitive and specific to CpG than MeDIP.	Requires less sample than hMeDIP. More sensitive and specific than antibody-based detection of 5hmC.
Disadvantages	MeDIP requires ssDNA. Antibody based techniques produce higher background than enzymatic based.	Misses non-CpG methylation.	
Cost	\$\$	\$\$	\$\$
Best Application	Cost effective method for determining differently methylated regions of the genome related to disease or environmental exposure.	Cost effective method for determining differently methylated regions of the genome related to disease or environmental exposure.	Commonly used to detect changes in hydroxymethylation in cfDNA samples that could be used for biomarkers for disease.

4

Sample Data and Tools for DNA Enrichment



MeDIP-Seq uncovers aging-associated differential methylation patterns in cfDNA.

Serum samples were collected in serum separator tubes (SST) from healthy "normal" adults, two <25 yo and two >60. Cell-Free DNA (cfDNA) was purified and 200 ng cfDNA was for MeDIPseq. MeDIP and Input libraries were sequenced (38bp PE chemistry on Next-Seq500) and processed as per Active Motif's standard pipeline for MeDIP-Seq. Peaks called using MACS2 relative to input (p-value cutoff = 1.00x10-⁷).

5



hMeDIP-Seq performed on human brain.

hMeDIP was performed with a 5-Hydroxymethylcytosine (5-hmC) antibody (Active Motif Catalog No. 39999) and 2 µg of human brain DNA. Enriched DNA was sequenced on the Illumina GA II and 19 million sequence tags were mapped to identify regions enriched for 5-hmC. The image above shows an 8 Mb region on chromosome 5 with 5-hmC enrichment across several gene bodies.

Products & Services for Genome-Wide & Gene-Specific Analysis

Product / Service	Format	Catalog number
MeDIP Kit	10 rxns	55009
hMeDIP Kit	10 rxns	55010
MethylPath [™] MeDIP-Seq Service	Custom-Contact Us	25025
MethylPath [™] hMeDIP-Seq Service	Custom-Contact Us	25052
MethylCollector™ MBD Capture Kit	30 rxns	55026
UDP-Azide-Glucose	1 vial	55020

Single-Nucleotide Resolution Analysis of DNA Methylation

Overview

When researchers need to know precisely which cytosine residues are methylated a DNA methylation assay with single-nucleotide resolution is required. Historically these methods involve performing bisulfite conversion. Bisulfite Conversion reveals the methylation state of each cytosine in DNA. DNA is first denatured (made single-stranded) and then treated with sodium bisulfite. Sodium bisulfite selectively changes unmethylated cytosines into uracils through deamination, while leaving methylated cytosines (both 5-mC and 5-hmC) unchanged. By then amplifying the treated DNA with PCR, uracils are further converted to thymines. At this point all nucleotides that were originally unmethylated cytosines become thymines while those that were methylated cytosines. The result is a clear nucleotide difference between methylated and unmethylated cytosines that can be easily identified by comparison to the original reference genome. A newer alternative to bisulfite conversion is to use enzymes such as TET2 and APOBEC3A to selectively convert cytosines to uracils. This approach avoids the damage that bisulfite conversion does to DNA but does not change resolution, cost, or coverage between the different techniques below.

Methods and Comparison

Whole Genome Bisulfite Sequencing (WGBS)

Often referred to as the "gold standard" for analysis of DNA methylation, Whole Genome Bisulfite Sequencing (WGBS) provides the greatest coverage and resolution of any DNA methylation analysis technique. Genomic DNA is first isolated and undergoes bisulfite conversion. The DNA is then adapter ligated and used to prepare a sequencing library. Following sequencing, DNA is bioinformatically analyzed to reveal methylated cytosines across the entire genome. The drawback to such comprehensive coverage is the high cost of sequencing and the complex data output. To achieve 15x coverage of the human methylome, a minimum of 800 million reads (100PE) is recommended. Despite the higher cost, WGBS may be preferred in situations where methylation in low CpG density regions is required or when looking at non-GpG methylation. The study of methylation in stem cells, neurons, and oocytes are some such examples.

Reduced Representation Bisulfite Sequencing (RRBS)

Reduced Representation Bisulfite Sequencing (RRBS) uses restriction enzymes together with DNA size selection to greatly reduce the percentage of the genome that must be sequenced. DNA is first digested using methylation-insensitive restriction enzymes (commonly Mspl) which generate fragments containing CpG dinucleotides at the ends. Fragments are then adapter ligated and size selected so that CpG rich regions of DNA small enough to be sequenced remain. Enriched fragments can then be bisulfite converted and amplified to reveal methylated cytosines and sequenced. RRBS captures 80-85% of CpG islands and 50-60% of human promoters while sequencing only 3% of the genome. This offers a huge cost savings over WBGS. RRBS is a good method to look for single base resolution changes in CpG rich regions like promotors which are commonly abnormally methylated in diseases such as cancer. However, it will miss 85-90% of CpGs in the genome.

Targeted Bisulfite Sequencing

Targeted bisulfite sequencing allows for very accurate analysis of methylation at single nucleotide resolution for specific regions of interest. It begins with bisulfite conversion of DNA and then uses primers to PCR amplify targeted loci. Primers must be designed to cover each region. Amplicons undergo library construction and are sequenced. By limiting sequencing to a small number of multiplexed amplicons, high sequencing depth of cytosines is achieved for those regions at low cost. Targeted bisulfite sequencing is a good method to use when looking at just a hand-full of loci. Much more than this, and the costs will stack up quickly and it's better to use an assay with more global coverage.

Methylation Arrays

Methylation arrays allow for quantification of methylation at selected cytosines in multiplexed samples on a chip. DNA is first bisulfite treated and PCR amplified to reveal the methylation pattern of the DNA. The DNA is then combined with complementary probes which target specific methylation sites. Two probes per target can be used, one complimentary to the methylated state and the other to the unmethylated state of the cytosine. Alternatively, a single probe can be used which distinguishes between methylation states based on a single base extension over the methylation site. Differently tagged nucleotides reveal levels of methylation from a fluorescent signal that is specific and quantifiable for the unmethylated or methylated state at each site. Methylation arrays are easy to use and reproducible. Coverage is very high (>96) for CpG islands, but they will miss most other (>98%) CpG sites in the genome. They also can only currently be used for human and mouse samples. These are great for clinical screening of samples for aberrant methylation differences in CpG islands.

Comparison of Methods

Method	WGBS - Whole-Genome Bisulfite Sequencing	RRBS - Reduced Representation Bisulfite Sequencing	Targeted Bisulfite Sequencing	Methylation Bead Array
Method Overview	DNA is bisulfite treated and PCR amplified to reveal the methylation pattern of the DNA. The DNA is then used to prepare a sequencing library.	A restriction enzyme, typically Mspl, is first used to cut DNA fragments in regions of high CG, which are common in CpG islands and promotors. DNA is then size selected for 40-220 bp fragments, adapter ligated, and bisulfite converted.	DNA is bisulfite treated and PCR amplified to reveal the methylation pattern of the DNA. PCR primers are then used to amplify targeted loci. These amplicons undergo library construction and are sequenced.	DNA is bisulfite treated and PCR amplified to reveal the methylation pattern of the DNA. It is then combined with DNA probes which target specific methylation sites.
Input	>500 ng	>200 ng	>1-50 ng	250 ng
Resolution	Single base	Single base	Single base	Single base
Coverage	28 million CpGs	3-5 million CpGs, 80-85% CpG Islands	Regions of selected amplicons; typically <10.	850,000 CpGs, >96% CpG islands
Advantages	Comprehensive coverage of methylated and unmethylated regions.	Can sequence regions dense in CpG methylation genome wide at single-base resolution for lower cost than WGBS.	Once primers are designed, high sequenc- ing depth at low cost for regions of interest.	Easy, reproducible, works with FFPE samples
Disadvantages	Expensive & time consuming. Not practical for high throughput. Does not distinguish 5mC from 5hmC.	Biased towards regions enriched in CCGG. Genes which have no or sparse CCGG motif not covered. Does not distinguish 5mC from 5hmC.	Primers must be designed for every region. Does not distinguish 5mC from 5hmC.	For human and mouse only. Misses 95% of CpG sites. Does not distinguish 5mC from 5hmC.
Cost	\$\$\$\$	\$\$\$	\$\$	\$\$\$
Best Application	Ideal when interested in Iow-CpG-density regions, intergenic "gene deserts", or partially methylated domains and distal regulatory elements, that would be missed by RRBS	Profiling CpG islands and promoters	Great when interested in methylation in a small number (<8) of regions of interest, especially if not in CpG islands.	Screening of clinical samples for aberrant methylation in CpG islands.

Sample Data and Tools for Single-Nucleotide Analysis



RRBS data from human samples.

The displayed regions are representative regions from the genome-wide data set and shows differential DNA methylation at an exon of CD19. Each block is a separate data point with red representing a methylated cytosine and blue representing an unmethylated base.

9



Heat map of Targeted Next-Gen Bisulfite Sequencing data.

MBD-Seq was first performed on 3 samples of interest to identify differentially methylated regions. One identified region was used for Targeted Next-Gen Bisulfite Sequencing on a broader population of 9 samples. This heat map shows the single base pair resolution of the 72 CpGs in this region across the 9 samples.

10

Products & Services for Single-Nucleotide Resolution 5-mc Analysis

Product / Service	Format	Catalog Number
Reduced Representation Bisulfite Sequencing (RRBS)	Custom-Contact Us	25069
Targeted Bisulfite Sequencing Service	Custom-Contact Us	25035
FFPE Bisulfite Conversion Kit	40 rxns	55021
Bisulfite Conversion Kit	50 rxns	55016

Additional Resources

Webinars

• Anchor-Based Bisulfite Sequencing (ABBS) Determines Genome-wide DNA Methylation

Podcasts

- CpG Islands, DNA Methylation, and Disease (Sir Adrian Bird)
- The Role of DNA Methylation in Epilepsy (Katja Kobow)
- DNA Methylation and Mammalian Development (Déborah Bourc'his)
- Effects of DNA Methylation on Diabetes (Charlotte Ling)
- Effects of DNA Methylation on Chromatin Structure and Transcription (Dirk Schübeler)

Blogs

- ABBS A Novel Method to Map DNA Methylation Genome-Wide at Base Resolution
- DNA Methylation in Cell-free DNA (cfDNA): Benefits, Limitations & Future Potential for Precision Medicine
- Complete Guide to Using Reduced Representation Bisulfite Sequencing (RRBS) for Genome-Wide DNA Methylation Analysis

References

Deaton AM, Bird A (2011). "CpG islands and the regulation of transcription". Genes Dev. 25 (10): 1010–22. doi:10.1101/gad.2037511. PMC 3093116. PMID 21576262. (CGI)

Jeziorska DM, Murray RJS, De Gobbi M, Gaentzsch R, Garrick D, Ayyub H, Chen T, Li E, Telenius J, Lynch M, Graham B, Smith AJH, Lund JN, Hughes JR, Higgs DR, Tufarelli C. DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. Proc Natl Acad Sci U S A. 2017 Sep 5;114(36):E7526-E7535. doi: 10.1073/pnas.1703087114. Epub 2017 Aug 21. PMID: 28827334; PMCID: PMC5594649. (CGI)

Kurdyukov S, Bullock M. DNA Methylation Analysis: Choosing the Right Method. Biology (Basel). 2016 Jan 6;5(1):3. doi: 10.3390/biology5010003. PMID: 26751487; PMCID: PMC4810160. (Global Methylation)

Sun Z, Cunningham J, Slager S, Kocher JP. Base resolution methylome profiling: considerations in platform selection, data preprocessing and analysis. Epigenomics. 2015 Aug;7(5):813-28. doi: 10.2217/epi.15.21. Epub 2015 Sep 14. PMID: 26366945; PMCID: PMC4790440. (WBGS/RRBS)

Xiao, Y., Yu, F., Pang, L. et al. MeSiC: A Model-Based Method for Estimating 5 mC Levels at Single-CpG Resolution from MeDIP-seq. Sci Rep 5, 14699 (2015). https://doi. org/10.1038/srep14699 (MeDIP coverage)

Sun Z, Cunningham J, Slager S, Kocher JP. Base resolution methylome profiling: considerations in platform selection, data preprocessing and analysis. Epigenomics. 2015 Aug;7(5):813-28. doi: 10.2217/epi.15.21. Epub 2015 Sep 14. PMID: 26366945; PMCID: PMC4790440. (RRBS coverage)

Aberg KA, Chan RF, van den Oord EJCG. MBD-seq - realities of a misunderstood method for high-quality methylome-wide association studies. Epigenetics. 2020 Apr;15(4):431-438. doi: 10.1080/15592294.2019.1695339. Epub 2019 Nov 25. PMID: 31739727; PMCID: PMC7153543. (MBD-Seq)

Clark C, Palta P, Joyce CJ, Scott C, Grundberg E, Deloukas P, Palotie A, Coffey AJ. A comparison of the whole genome approach of MeDIP-seq to the targeted approach of the Infinium HumanMethylation450 BeadChip(®) for methylome profiling. PLoS One. 2012;7(11):e50233. doi: 10.1371/journal.pone.0050233. Epub 2012 Nov 29. PMID: 23209683; PMCID: PMC3510246. (Me-DIP)

Beck D, Ben Maamar M, Skinner MK. Genome-wide CpG density and DNA methylation analysis method (MeDIP, RRBS, and WGBS) comparisons. Epigenetics. 2022 May;17(5):518-530. doi: 10.1080/15592294.2021.1924970. Epub 2021 May 11. PMID: 33975521; PMCID: PMC9067529. (MeDIP, RRBS, WGBS)

Song CX, Yi C, He C. Mapping recently identified nucleotide variants in the genome and transcriptome. Nat Biotechnol. 2012 Nov;30(11):1107-16. doi: 10.1038/nbt.2398. PMID: 23138310; PMCID: PMC3537840. (hMeDIP/hmE-Seal)



NORTH AMERICA

Fax: 760.431.1351

JAPAN

Toll Free: 877.222.9543 Direct: 760.431.1263

japantech@activemotif.com

sales@activemotif.com tech_service@activemotif.com Direct: +81 (0)3 5225 3638 Fax: +81 (0)3 5261 8733

Germany: 0800/181 99 10 United Kingdom: 0800/169 31 47 France: 0800/90 99 79 Other Countries, Direct: +32 (0)2 653 0001 Fax: +32 (0)2 653 0050

eurotech@activemotif.com

CHINA

Hotline: 400 018 8123 Direct: +86 21 2092 6090

techchina@activemotif.com



EUROPE