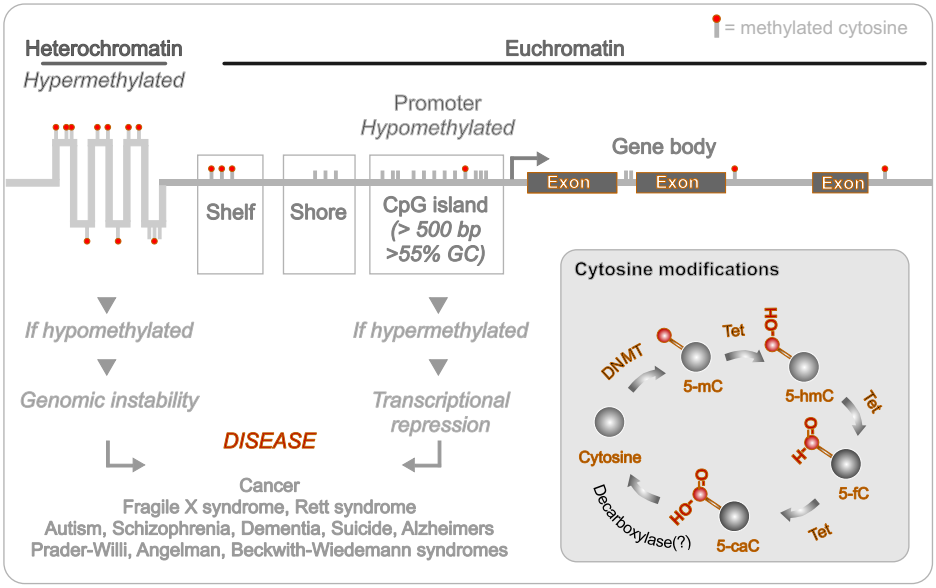
Field Guide to DNA methylation methods

Cytosine methylation (5-mC) has a significant effect on gene expression and chromatin re-modeling. Cytosine methylation (5-mC) and hydroxymethylation (5-hmC) regulate spatial and temporal gene expression, critical for embryonic development and cellular differentiation.

- *Modulation of gene expression.* Cytosine methylation in promoter regions often reduces gene expression. 88% of active promoters are associated with unmethylated CpGs in mammals. In comparison, hydroxymethylation often correlates with increased gene expression.
- *Cellular differentiation.* Methylation change usually drives one-way differentiation; that is, differentiated cells do not typically revert to stem cells stages.
- *Change in chromatin structure.* Heterochromatin formation is initiated by *hyper*methylation. Methyl-binding-domain proteins (MBD) specific for 5-mC recruit histone-modifying and chromatin-remodeling proteins. Abnormal *hypo*methylation of heterochromatin (usually hypermethylated) is linked to chromosomal instability and loss of imprinting. Controlled changes in chromatin structure are responsible for selective X chromosome inactivation and suppression of transposable elements.
- *Onset of diseases.* Perturbation of any of the above processes result in disease.



Methylation enzymes

Mammalian

DNMT1 (*D*NA *m*ethyl*t*ransferase) isoforms methylate hemimethylated CpGs, maintaining methylation patterns during DNA replication.
DNMT3a and *DNMT3b,* de novo methyltransferases that set up DNA methylation patterns early in development & can methylate unmethylated & methylated DNA.
DNMT3L currently unknown catalytic activity but facilitates de novo methyltransferases activity.
TET (*T*en *e*leven *t*ranslocation) enzymes are involved in oxidation of the methyl group with the production of 5-hydroxymethylation as an intermediate.

Plant

DRM2 (RNA directed DNA methylation) is homologous to *DNMT3. MET1 (Methyltransferase 1)* is homologous to *DNMT1. CMT3 is* unique to plants. Function for other methyl transferases in has not been assigned.

Bacteria

Bacterial DNA methylation is linked to virulence and antibiotic resistance.
Dam (*D*NA adenine *m*ethyltransferase), methylates 'A' in G**A**TC, key role in mismatch repair, DNA replication timing & gene regulation. *Dam* is independent of restriction modification systems.
Dcm (*D*NA cytosine *m*ethylase) produces 5-mC in CCAG and CCTGG sites.
EcoKI, methylates adenine in AAC(N6)GTGC and GCAC(N6)GTT.

Glossary

5-Hydroxymethylcytosine (5-hmC). First seen in bacteriophages in 1952. In 2009, found to be abundant in human and mouse brains and in embryonic stem cells. Formed by oxidation of 5-methyl cytosine by *TET* oxygenases.

5-Methylcytosine (5mC). Formed by addition of -CH3 to N5 position of cytosine. 70-80% of CpG dinucleotides are methylated (high proportion in repetitive elements). In plants, cytosine methylation is observed at CpG, CpHpG, and CpHpH (H = nucleotide other than guanine).

5-Formylcytosine. Oxidation product of 5-hmC. Detected in mouse embryonic stem cells.

Base modification. Addition of methyl- or other groups to nucleic acid bases. In addition to 5-methylcytosine, other common modifications are 7-methyl guanosine (5'-cap for RNA), and 6-methyl adenosine (common RNA modification).

CpG island. Defined as regions > 500 bp, > 55% GC and expected / observed CpG ratio of > 0.65. 40% of gene promoters contain islands.

CpG shelves. ≈ 4Kb from islands.

CpG shores. ≈ 2Kb from islands, > 75% of tissue-specific differentially methylated regions found in *shores.* Methylation in *shores* shows higher correlation with gene expression than CpG islands.

Differentially methylated regions (DMR). Cell-, tissue- and condition-specific differences in methylation.

Expected/observed CpG ratio. The human genome contains 1/4th the expected number of C-G pairs due to spontaneous deamination of meC to T over evolutionary time scales.

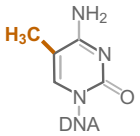
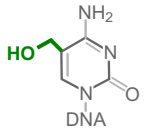
Genomic imprinting. An epigenetic process causing genes to be expressed only from one of the parental chromosomes.

Hypermethylation. Majority of cytsoines are methylated.
Hypomethylation. Majority of cytosines do not have 5-mC. Euchromatin and active gene promoters are hypomethylated.

RNA-directed DNA methylation (RdDM). dsRNAs directed de-novo DNA methylation at the original genomic location that produced the RNA.

Illumina sequencing: Field Guide to DNA methylation methods

Primary methods for detecting cystosine modifications

	Method	DNA prep	Notes
 5mC	**WGBS-Seq** Whole Genome Bisulfite-Seq	Na-bisulfite converts non-methylated cytosines to **U** (U read as **T**). 5mC and 5hmC are protected from conversion and read as **C**. Sequencing libraries are prepared by random-primer extension & sequenced with 2x75-100 bp reads lengths.	> 38 million CpG's queried in human. Requires > 90Gb sequence data/sample. Cost ≈ US$5,000/sample*.
	RRBS-Seq Reduced representation Bisulfite seq.	DNA digested with *MspI* (C|CGG), ≈100-150 bp fragments isolated, treated with bisulfite, libaries prepared and sequenced with 1x75 bp.	Queries ≈ 85% CpG islands (≈ 2M CpG's) & 60% of RefSeq promoters. Requires ≈ 40-50M reads, 3-5 Gb/sample. Cost ≈ US$300/sample*.
	MeDIP Methylated DNA immuno-precipitation	DNA sonicated to 100-300 bp, end-repaired, ligated to adapters, and denatured. 5mC containing fragments are captured with anti-5mC magnetic beads, and sequenced (1x75-100 bp). MeDIP may be combined with bisulfite conversion.	5**m**C detected at ≈150 bp resolution. Bias towards hypermethylated regions. Requires ≈ 60M reads, ≈ 5Gb/sample. Cost ≈ US$300/sample*.
	MIRA Methylated CpG island recovery assay	DNA sonicated to 100-300 bp, end-repaired, ligated to adapters and fragments captured with GST-labeled MBD2b and His-tagged MBD3L1 proteins. Purified by glutathione-magnetic beads. Sequenced with 1x75 bp reads.	Detects **m**CpGs at ≈ 150 bp resolution. Bias towards hypermethylated CpGs. Non-CpG methylation not detected. Requires ≈ 60M reads, ≈ 5Gb/sample. Cost ≈ US$300/sample*.
	HM450 array Human methylation 450k array	Bisulfite conversion of DNA. C to T changes at defined genomic positions detected by the Infinium assay.	Samples 96% CpG islands, 485,000 CpGs, 99% RefSeq promoters, 3' & 5' UTR, 1st exon, gene body, 3' UTR, shores and shelves. Cost ≈ US$240/sample*.
 5hmC	**Tab-Seq** Tet-assisted bisulfite seq	β-glucosyltransferase adds glucose to 5**h**mC but not 5mC. *Tet* enzymes oxidize 5mC to 5caC which converts to U on bisulfite treatment but glucosylated 5**h**mC cannot be oxidized. **h**mC will be read as **C** and mC will be read as **T**.	Detects genome-wide 5hmC at single base resolution. Requires > 90Gb/sample. Cost ≈ US$5,000/sample*.
	hMe-DIP Hydroxy Methylated DNA Immuno Precipitation	DNA sonicated to 100-300 bp, end repaired & ligated to adapters. hmC fragments captured with anti-5hmC labeled magnetic beads. Bound fragments sequenced with 100 bp reads.	Detects hydroxymethylated cytosines at about 150 bp resolution. Requires about 5Gb sequence data/sample. Cost ≈ US$300/sample*.
	RRHP-Seq Reduced Representation Hydroxymethyl-cytosine Profiling Sequencing	DNA digested with *MspI* (C|CGG) & ligated to adapters. hmC glucosylated with β-glucosyltransferase making it resistant to re-digestion with *MspI*. Treating the library with *MspI* selectively digests all non-hmC fragments.	≈ 85% of CpG islands (2 M CpG's) & 60% of RefSeq promoter regions queried. Requires ≈ 30-50M reads, 3-5 Gb/sample. Cost ≈ $300/sample*.

* Cost is provided for guidance only & may not reflect your actual cost.

Techniques in brief

COBRA. *C*ombined *B*isulfite followed by *R*estriction *A*nalysis. RE recognition affected by bisulfite treatment.

C-Subtraction. *MseI* (methylation independent) digestion, ligation to linkers, then digestion with methylation sensitive *BstUI* or *HpaII* to reduce unmethylated DNA.

HELP. HpaII enrichment by ligation PCR. Differential restriction of methylated and unmethylated CpG sites.

MCA. Methylated CpG-island amplification. DNA cut with *SmaI* (blunt ends, cannot cut meC) followed with *XmaI* (sticky ends, cuts meC and C). Sticky end adapters select for meC fragments.

MeDIP. Methylated DNA immunoprecipitation with anti-methylcytosine antibody.

Methyl light. Methylation specific qPCR (see MSP).

Methylation trapping of methyl transferases onto DNA with 5-Aza-2'-deoxycytidine (decitabine) followed by immunoprecipitation.

MIRA. Methylated CpG island recovery assay. Capture with MBD2b and MBD3L1 protein heterodimer.

MSDK. Methylation specific digital karyotyping. Cleavage with methylation sensitive *AscI*. Sequence tags sequenced and mapped like SAGE.

MSP. Methylation specific PCR on bisulfite converted DNA with two sets of primers to amplify either -C or -T base.

OxBS. K-perruthenate oxidizes only 5hmC to 5fC that is converted to U by bisulfite.

RLGS. Restriction Landmark Genome Scanning with 2D gel of methylation sensitive *NotI* and *AscI* restricted DNA.

RRHP. Reduced Representation Hydroxymethylation Profiling.

RRBS. Reduced Representation Bisulfite Sequencing.

Tab-Seq. Tet-assisted bisulfite sequencing of 5-hmC.

WGBS-Seq. Na-bisulfite conversion of C -> T. mC and hmC are not converted and are read as C by whole genome next generation sequencing.

References:

- Harris RA et al (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotech* 28, 1097-1105.
- Guibert S, Weber M (2013) Functions of DNA methylation and hydroxymethylation in mammalian development. *Curr Top Dev Biol* 104, 47-83.
- Laird PW (2010) Principles and challenges of genome-wide DNA methylation analysis. *Nature Rev Genetics* 11, 191-203.
- Meissner A (2012) Epigenomics, special issue. *Genome Biol* 13(10).
- Stevens M et al (2013) Estimating absolute methylation levels at single CpG resolution from methylation enrichment and restriction enzyme sequencing methods. *Genome Res.* 10.1101/gr.152231.112.
- Xie W et al (2013) Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* 153, 1134-48.
- Yu P (2013) Spatiotemporal clustering of the epigenome reveals rules of dynamic gene regulation. *Genome Res* 23, 352-64.
- Yu M et al (2012) Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nature Protocols* 7, 2159-70.

FOR RESEARCH USE ONLY

illumina®