DNA METHYLATION

DNA METHYLATION CONTROLS GENE EXPRESSION

FACTS:

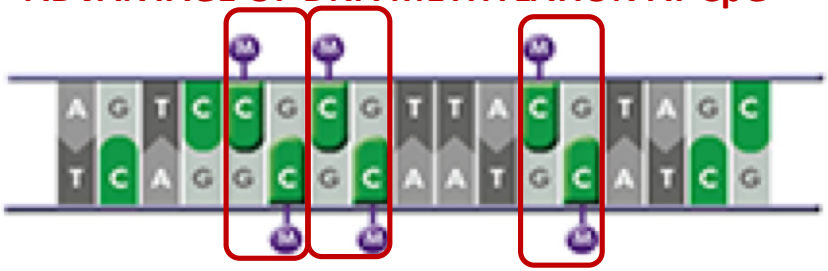
1. DNA methylation is created at CpG di-nucleotide motifs

2. An accumulation of CpG is called "CpG island" (CGI)

3. CpG islands are enriched at promotes and sequence elements that are important for gene expression control. In some cases, CpG islands can be also located in distant locations.

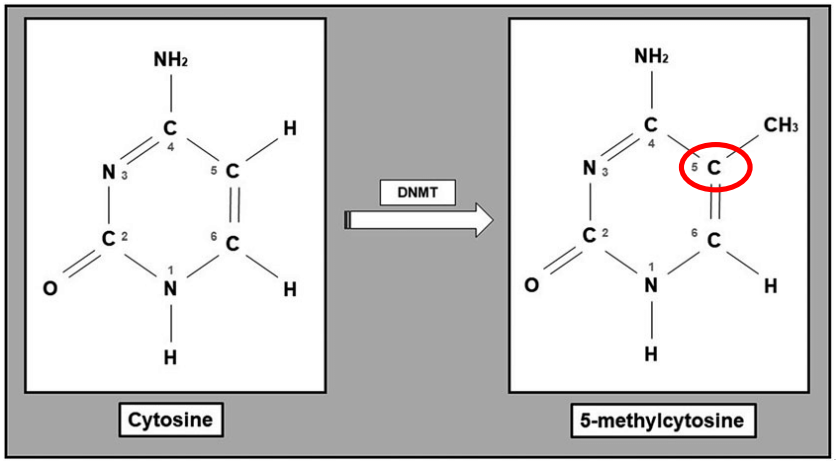
4. CpG methylation (="DNA methylation") is directly linked with stable, inheritable gene silencing

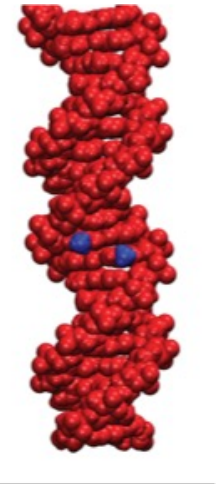
ADVANTAGE OF DNA METHYLATION AT CpG



CpGs are self-complementary
Di-nucleotide in paired stand also contains methylation
Methylation patterns can be maintained during DNA replication

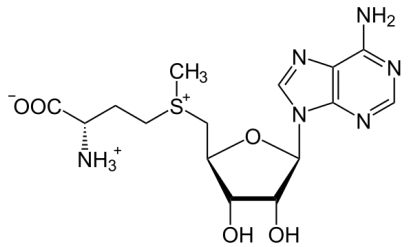
Epigenetic modification that is imposed on genomic DNA





DNA METHYLTRANSFERASES CATALYZE DNA METHYLATION

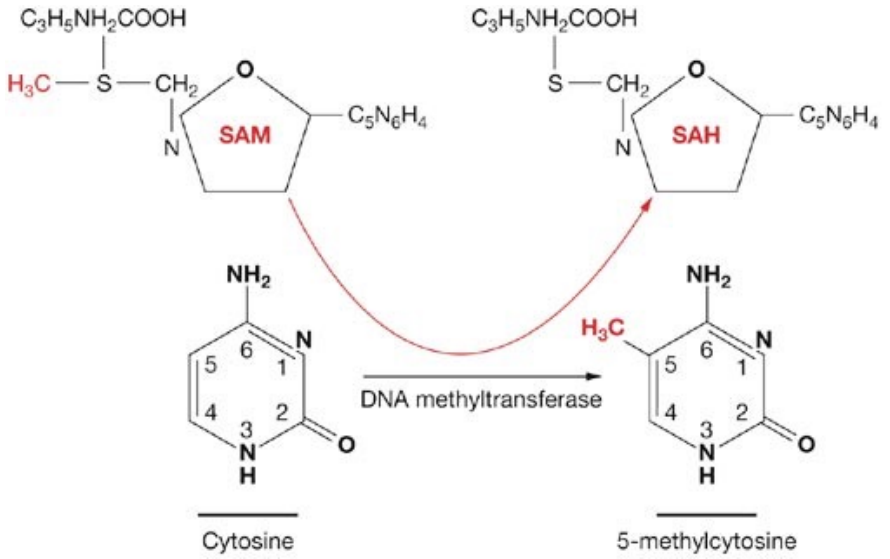
DNA methyltransferases (DNMTs) transfer a methyl-group from AdoMet (SAM) to Cytosine located in a CpG dinucleotide

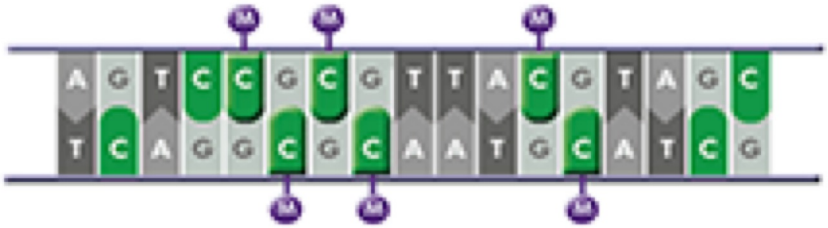


S-adenosyl-l-methionine (AdoMet) or (SAM),

The source of the methyl group is S-adenosyl-l-methionine (AdoMet) or (SAM), which is converted to S-adenosyl-l-homocysteine (AdoHcy) in the reaction.

S-Adenosyl methionine is a common cosubstrate involved in methyl group transfers, transsulfuration, and aminopropylation.
SAM = enzymatic cofactor
<u>SAM is after ATP the most commonly used cofactor used by the cell</u>
Although these anabolic reactions occur throughout the body, most **SAM-e is produced and consumed in the liver**. More than 40 methyl transfers from SAM-e are known, to various substrates such as nucleic acids, proteins, lipids and secondary metabolites. It is made from adenosine triphosphate (ATP) and methionine by methionine adenosyltransferase. SAM was first discovered in Italy by Giulio Cantoni in 1952.

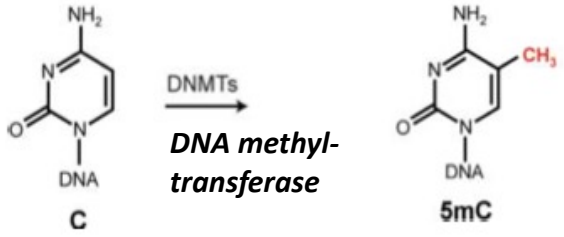




Mapping DNA methylation at CpG islands
BISULFITE SEQUENCING

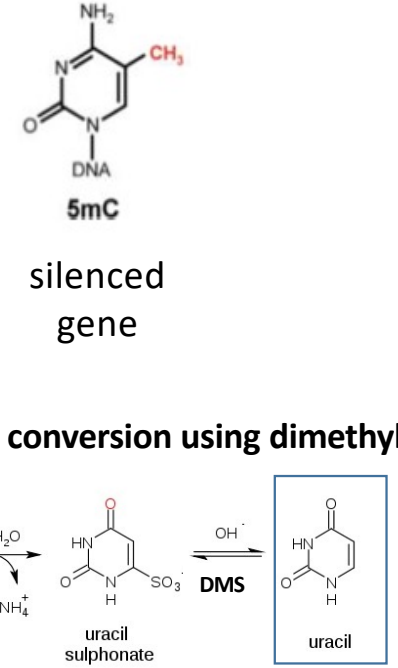
Methylation of cytosine at CpG dinucleotides is an important epigenetic regulatory modification in many eukaryotic genomes.

DNA methylation was found to be located genome-wide with a pattern of low methylation in proximity to promoters and high gene body-methylation in highly-expressed genes → methylation pattern can identify transcribed DNA (gene)

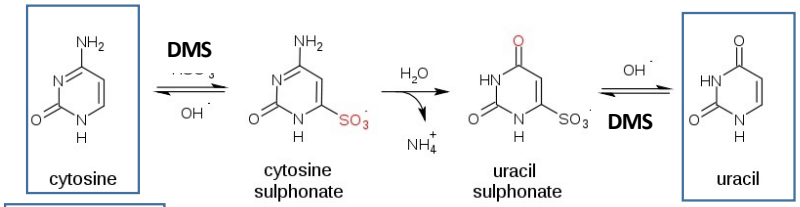


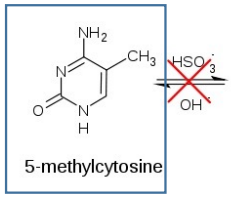
active gene

silenced gene



Bisulfite conversion: C→U conversion using dimethyl sulfate





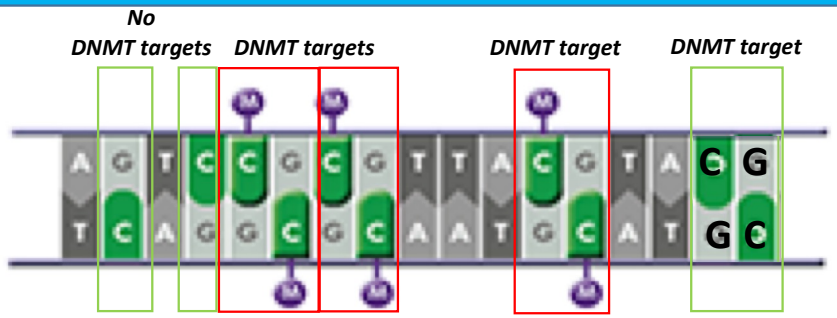
methylated C cannot be converted!!

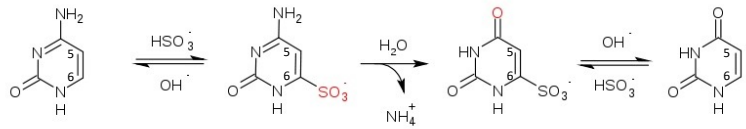
DMS = Dimethyl sulfate

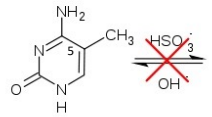
Treatment of DNA with bisulfite converts cytosine residues to uracil, but leaves 5-methylcytosine residues unaffected. Thus, bisulfite treatment introduces specific changes in the DNA sequence that depend on the methylation status of individual cytosine residues, yielding single-nucleotide resolution information about the methylation status of a segment of DNA.

Mapping DNA methylation at CpG islands
BISULFITE SEQUENCING

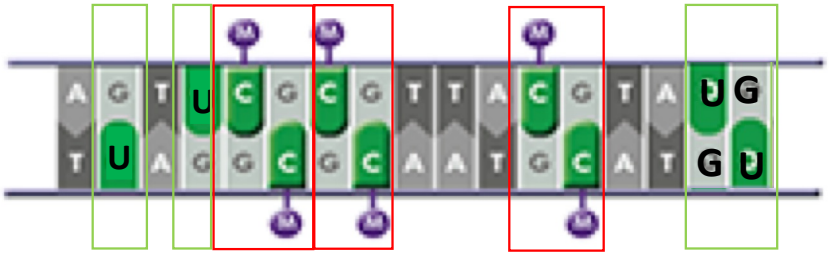
Genomic DNA



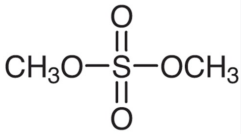




DNA for Sequencing



Bisulfite Conversion Using DMS (Dimethyl Sulfate; C2H6O4S)



Seqeuncing of both strands reveals C → U (T) transition

Compare with genomic sequence
C→U sequence change = no DNA methylation
C→C no sequence change = DNA methylation

DNA methyl transferases methylate DNA

Maintenance of DNA methylation patterns by DNMTs

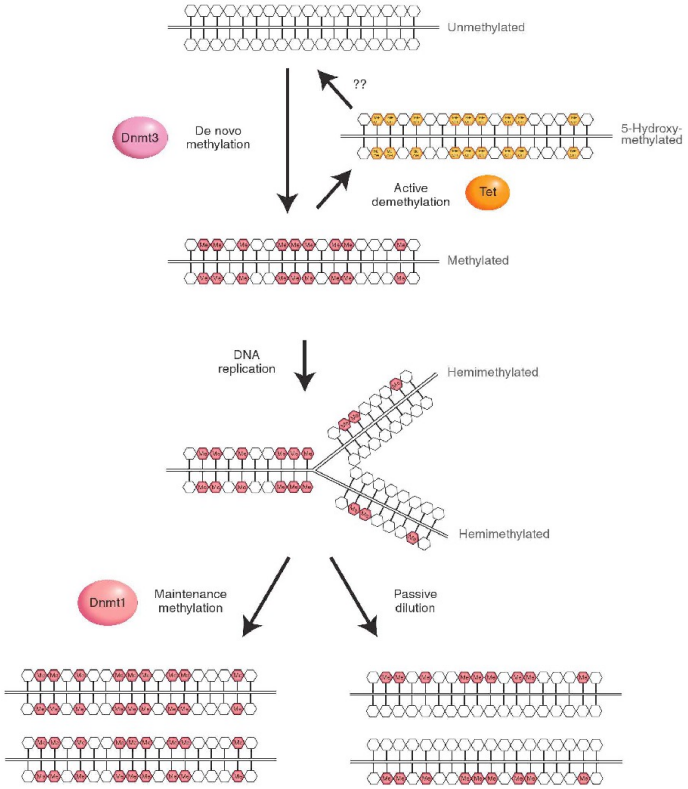


Figure 2. De novo methylation and maintenance methylation of DNA. A stretch of genomic DNA is shown as a line with self-complementary CpG pairs marked as vertical strokes. Unmethylated DNA (*top*) becomes methylated "de novo" by Dnmt3a and Dnmt3b to give symmetrical methylation at certain CpG pairs. On semiconservative DNA replication, a progeny DNA strand is base-paired with one of the methylated parental strands (the other replication product is not shown). Symmetry is restored by the maintenance DNA methyltransferase, Dnmt1, which completes half-methylated sites, but does not methylate unmodified CpGs.

Discovery of function and DNMT family members:

Discovery of the maintenance DNMT – DNMT1:

DNMT1: discovered first

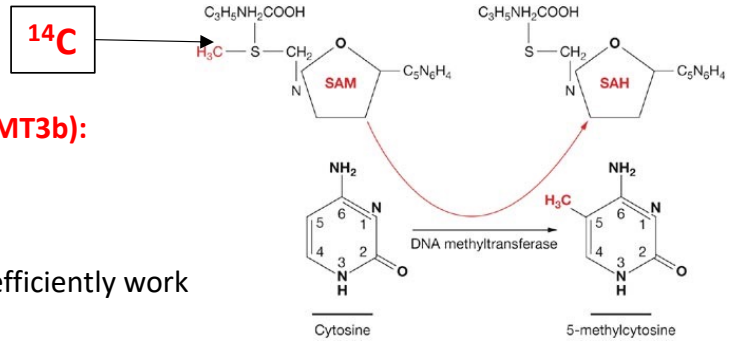
Cell extract + DNA containing CpG repeats + ^{14}C labelled -CH3 in AdoMet (SAM) → radioactive -CH3 transferred to DNA

Next step: Purification of enzymatic activity from cell extract → 200kDa complex containing a protein with specific DNA methyl transferase activity: **DNMT1**

Obtain sequence, deletion analysis, identification of the catalytic domaon

Biochemical characterization of substrate specificity:

This enzyme is 7– to 100-fold more active on hemi-methylated DNA as compared with un-methylated substrate *in vitro*



Discovery of *de novo* DNMTs (DNMT3a, DNMT3b):

Sequence of DNMT1 was used to identify
for genes with similar sequence
(sequence homology)

→ Discovery of de-novo DNMTs that work efficiently work on un-methylated DNA (DNMT3a, 3b)

→ De-novo DNMTs cannot efficiently methylate hemi-methylated DNA

LOSS OF DNA METHYLTRANSFERASES IS LETHAL DURING EMBRYONIC MOUSE DEVELOPMENT

Table 1. Function of mammalian DNA methyltransferases

DNA methyltransferase	Species	Major activity	Major phenotypes of loss of function
Dnmt1	Mouse	Maintenance methylation of CpG	Genome-wide loss of DNA methylation, embryonic lethality at embryonic day 9.5 (E9.5), abnormal expression of imprinted genes, ectopic X-chromosome inactivation, activation of silent retrotransposon. In cancer cell lines, it leads to cell cycle arrest and mitotic defects.
Dnmt3a	Mouse	De novo methylation of CpG	Postnatal lethality at 4–8 wk, male sterility, and failure to establish methylation imprints in both male and female germ cells
Dnmt3b	Mouse	De novo methylation of CpG	Demethylation of minor satellite DNA, embryonic lethality around E14.5 days with vascular and liver defects. (Embryos lacking both Dnmt3a and Dnmt3b fail to initiate de novo methylation after implantation and die at E9.5.)
DNMT3B	Human	De novo methylation of CpG	ICF syndrome: immunodeficiency, centromeric instability, and facial anomalies. Loss of methylation in repetitive elements and pericentromeric heterochromatin.

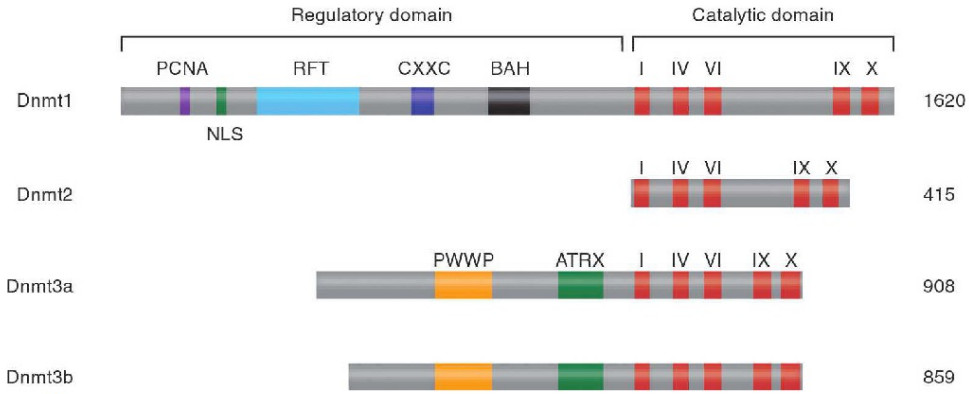


Figure 4. Mammalian DNA methyltransferases. The catalytic domains of Dnmt1, Dnmt2, and the Dnmt3 family members are conserved (the signature motifs, I, IV, VI, IX, and X, are most conserved in all cytosine methyltransferases), but there is little similarity among their amino-terminal regulatory domains. Domain abbreviations: PCNA, PCNA-interacting domain; NLS, nuclear localization signal; RFT, replication foci-targeting domain; CXXC, a cysteine-rich domain implicated in binding DNA sequences containing CpG dinucleotides; BAH, bromo-adjacent homology domain implicated in protein–protein interactions; PWWP, a domain containing a highly conserved "proline-tryptophan-tryptophan-proline" motif involved in heterochromatin association; ATRX, an ATRX-related cysteine-rich region containing a C2-C2 zinc finger and an atypical PHD domain implicated in protein–protein interactions.

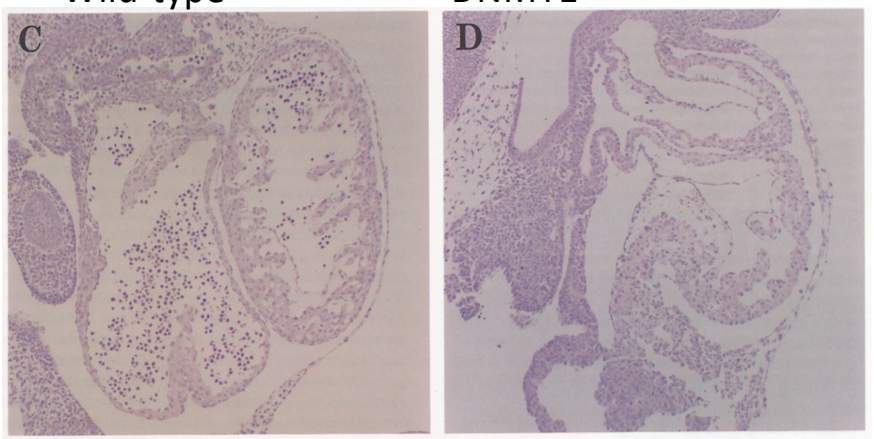
LOSS OF DNA METHYLTRANSFERASES IS LETHAL DURING EMBRYONIC MOUSE DEVELOPMENT

Embryo:
10.5 days after fertilization

Wild-type DNMT1$^{-/-}$ DNMT1$^{-/-}$

Heart

Wild-type DNMT1$^{-/-}$



ON THE SINGLE GENE LEVEL:
CpG islands (CGIs) are short sequences stretches with variable DNA methylation that regulate promoter activity

NOTE: across the entire genome, single CpGs are generally hyper-methylated (60-**90%)**

CpG islands are <u>differentially</u> methylated, but are generally hypo-methylated

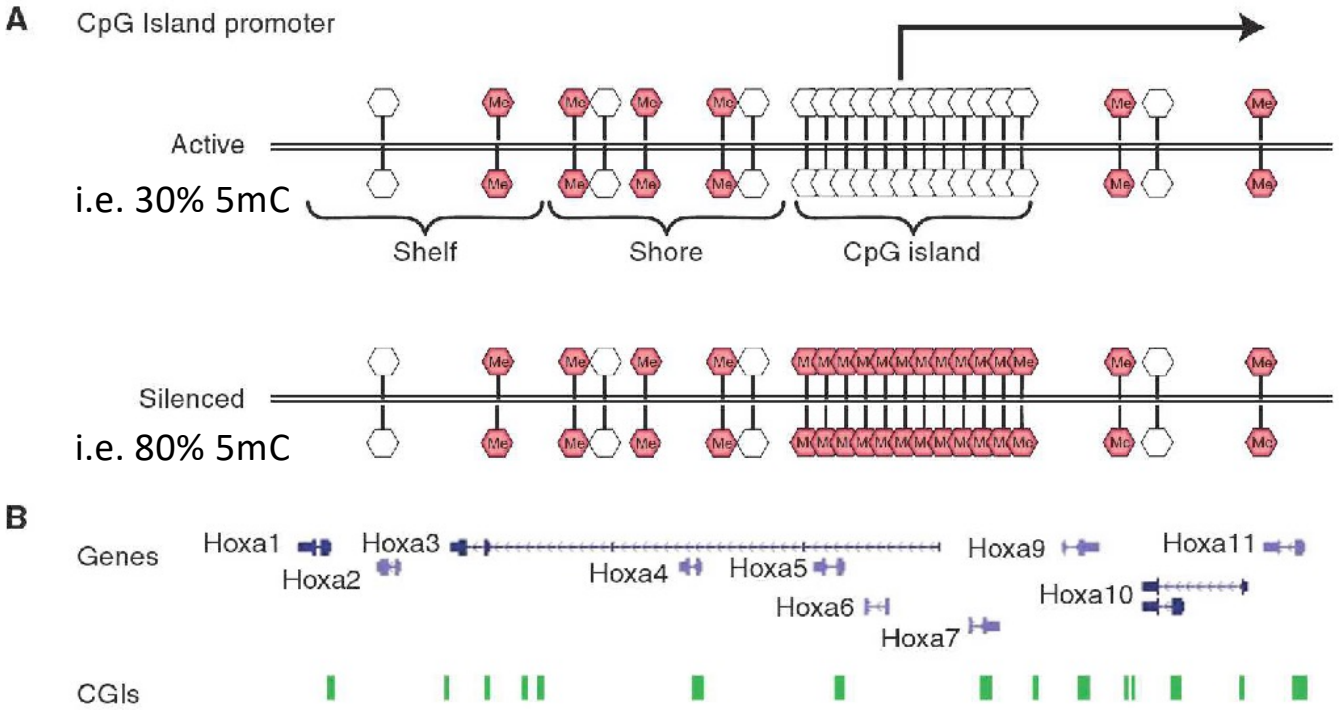
CpG islands (CGIs) have a length of ca. 1kb

60% of human genes are controlled by CGIs containing promoters that allow tissue/cell specific gene expression

CpG islands can overlap with the first exon (methylation level in 1st exon is good predictor of gene expression

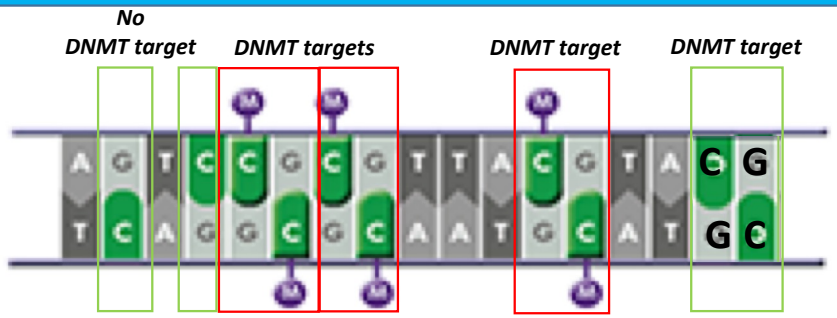
CpG islands located <2kb from promoter: shores

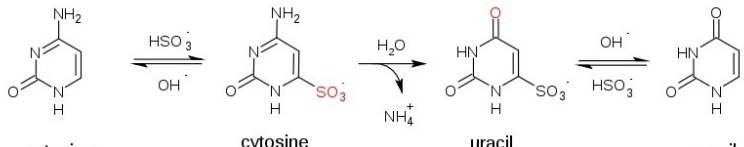
CpG islands located <2-4kb from promoter: shelf

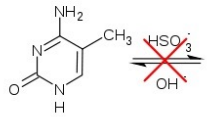


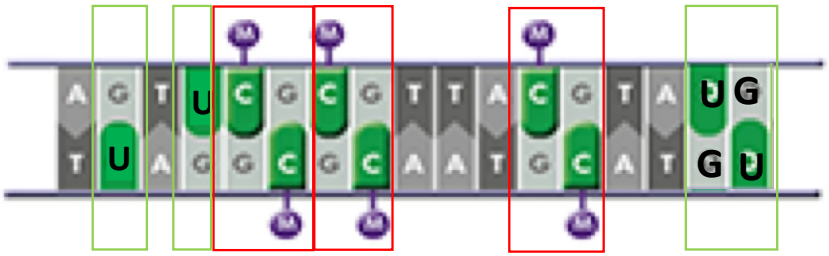
Mapping DNA methylation at CpG islands
BISULFITE SEQUENCING

Genomic DNA



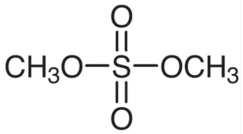






DNA for Sequencing

Bisulfite Conversion Using DMS (Dimethyl Sulfate; C2H6O4S)



Seqeuncing of both strands reveals C → U (T) transition

Compare with genomic sequence
C→U sequence change = no DNA methylation
C→C no sequence change = DNA methylation

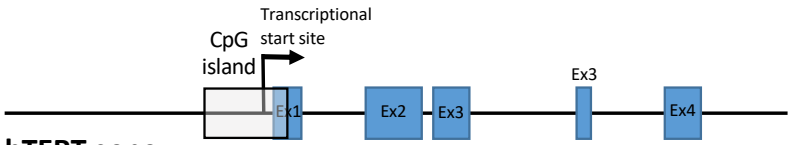
Mapping DNA methylation at CpG islands
BISULFITE SEQUENCING

An example:
hTERT encodes the telomerase gene
hTERT elongates telomeres thereby protecting cancer cells from replicative senescence

→ Cancer cells do not age and can proliferate forever!!

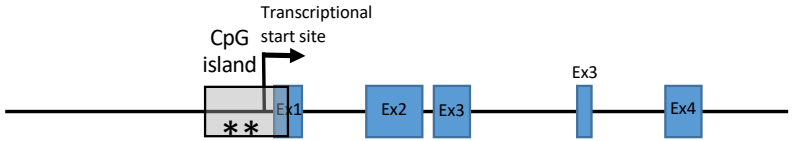
→ Therefore 90% of human cancers express telomerase. NOTE: telomerase gene is silenced by DNA methylation in differentiated cells.

→ Bisulfite sequencing showed that cancer cells have a de-methyalted CpG island located at the hTERT promoter. CpG islands can overlap with the 1st exon of the gene!!!!



hTERT gene
Cancer: ON

CANCER CELL
CpG
island
not methylated



hTERT gene
Normal: OFF

NORMAL CELL
CpG
island
is methylated (**)

Mapping DNA methylation at CpG islands of individual genes
BISULFITE SEQUENCING

Design PCR primers for target site

- avoid CpG dinucleotide in primer binding sites (differential methylation possible; C→U)
- consider that single C will be converted to U (place corresponding T in primer)
- target site should contain C without neighboring G → should be converted at 100% to U)

Prepare genomic DNA from normal cell and cancer cell

Purify DNA and perform bi-sulfite conversion (using DMS)
(Unmethylated C → U; Methylated C → C)

PCR Amplify your region of interest = CpG islands
in the TERT promoter using region specific primers;

Purify DNA fragment obtained by PCR

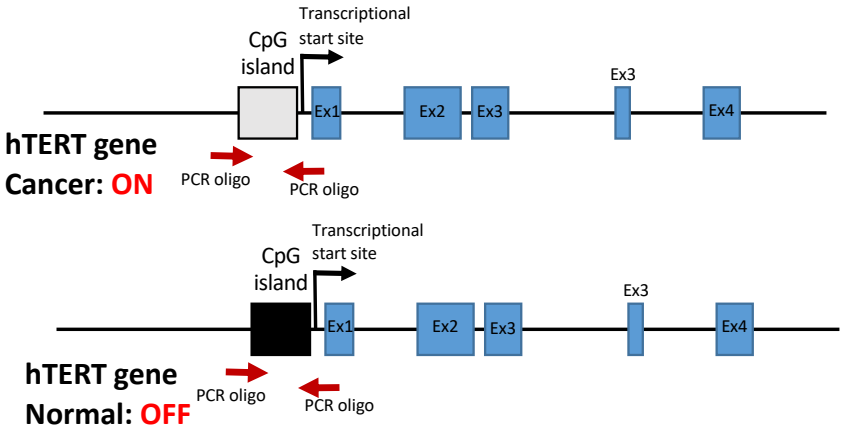
Clone fragment into Plasmid

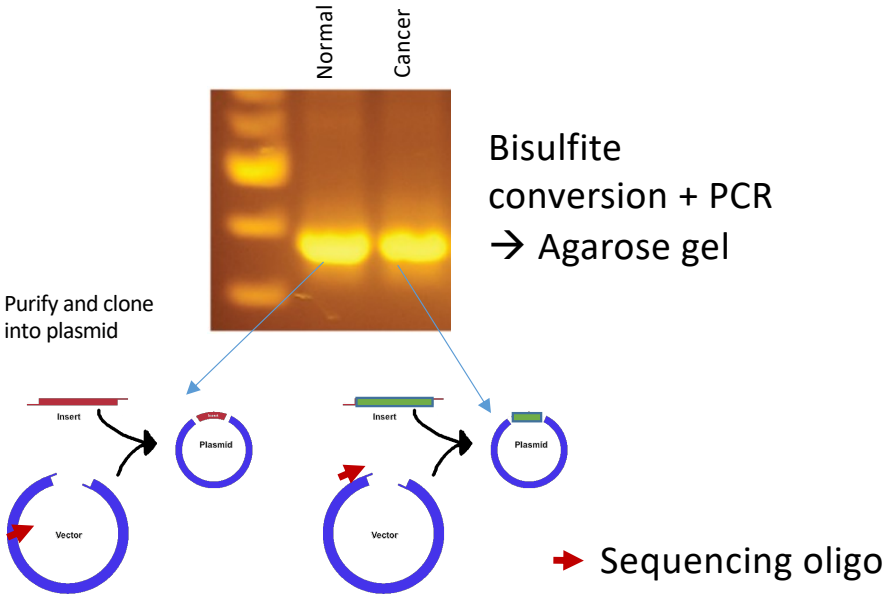
Transform bacteria with plasmid (one bacteria receives only one plasmid!!!)

Purify amplified plasmids from 10-15 individual bacterial cultures

Sequence inserts using a primer that anneals to the vector DNA, adjacent to the insertion site of the PCR product.

REMEMBER: only a single type of plasmid, generated by PCR cloning and subsequently amplified in a single transformed bacteria is analyzed. Single bacteria forms colonoy on agar plate → amplification in bacterial culture → DNA mini prep
To have a good representation, you need to sequence at least 10-15 clones





Mapping DNA methylation at CpG islands
BISULFITE SEQUENCING

Design PCR primers for target site

Prepare DNA from normal cell and cancer cell

Purify DNA and perform bi-sulfite conversion

PCR Amplify your region of interest = CpG islands
in the TERT promoter.
PCR primers should not contain CpG dinucleotides

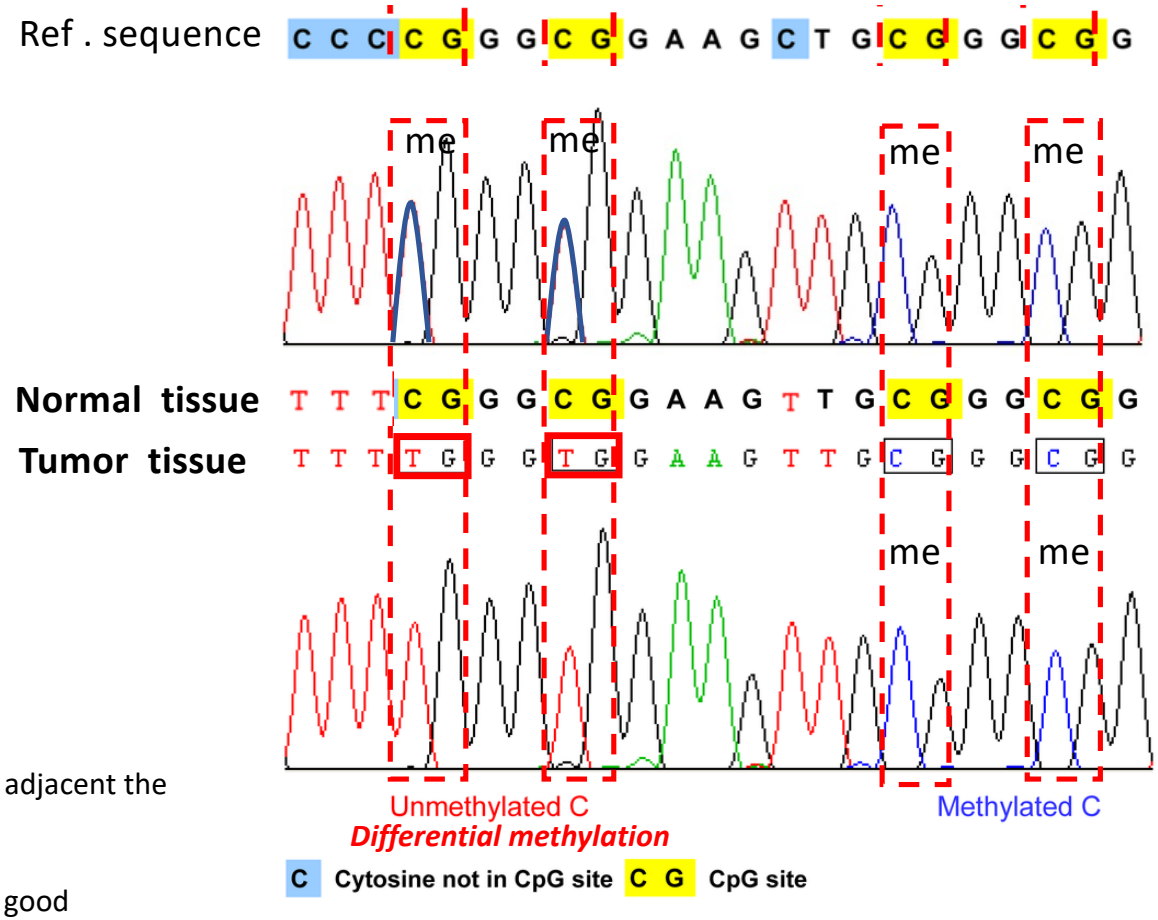
Purify DNA fragment obtained by PCR

Clone fragment into Plasmid

Transform bacteria with plasmid

Purify amplified plasmids

Sequence inserts using a primer that anneals to the vector DNA, adjacent the insertion site of the PCR product
REMEMBER: only a single type of plasmid is sequenced
This refers to the sequence of interest of a single cell!! To have a good representation, you need to sequence at least 10 clones
IMPORTANT: Quality control of your bisulfite converstion: ALL C that are not followed by G MUST have been converted to U!!!!



Compare with genomic sequence
Bisulfite conversion: C→U sequence change = DNA methylation
C→C no seqeunce change = no DNA methylation

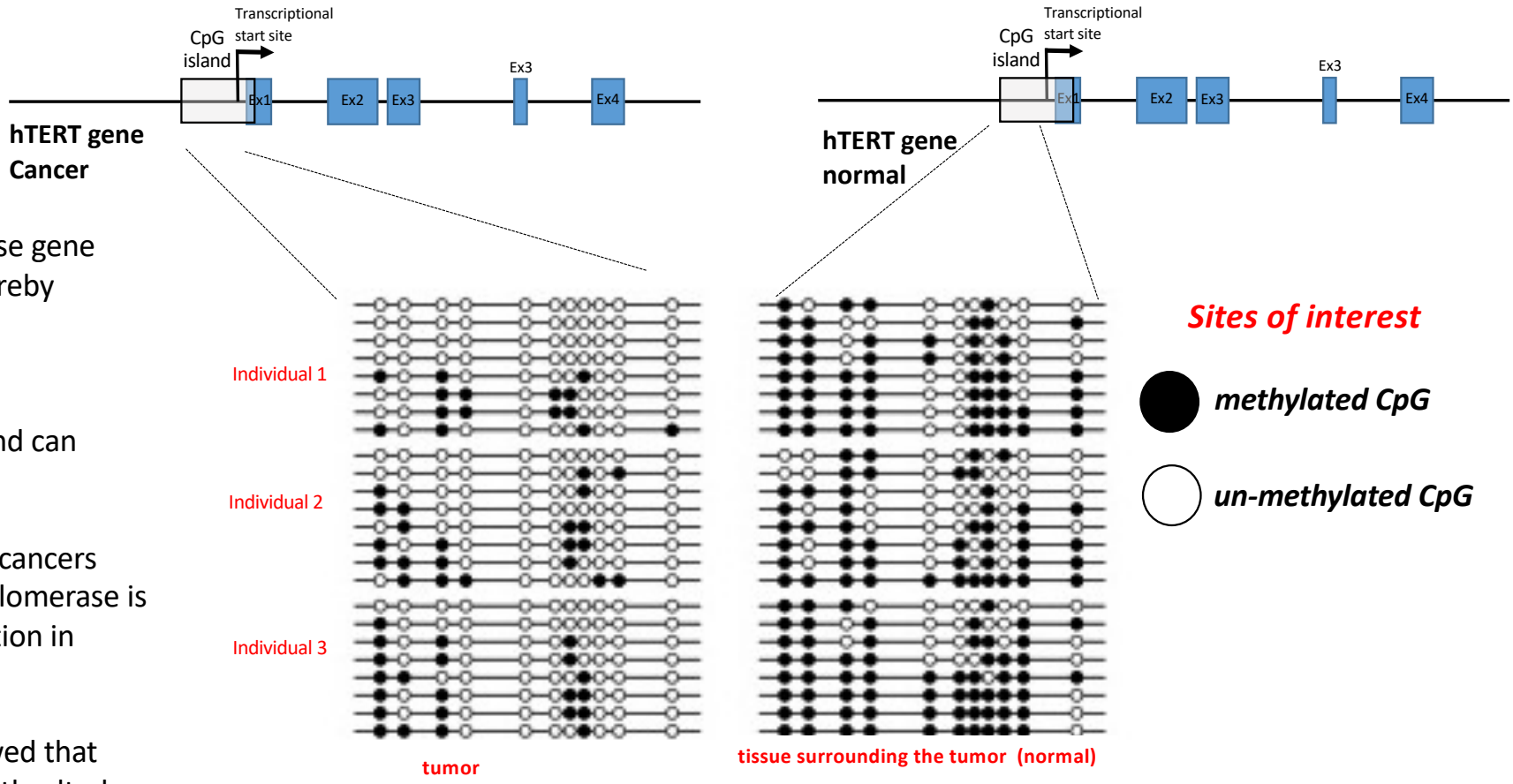
Mapping DNA methylation at CpG islands
BISULFITE SEQUENCING

An example:
hTERT encodes the telomerase gene Tert elongates telomeres thereby protecting cancer cells from replicative senescence

→ Cancer cells do not age and can proliferate forever!!

→ Therefore 90% of human cancers express telomeres; but telomerase is silenced by DNA methylation in differentiated cells.

→ Bisulfite sequencing showed that cancer cells have a de-methyalted CpG island located at the hTERT promoter



Mapping DNA methylation at multiple CpG islands
METHYLATED DNA IMMUNOPRECIPITATION: METHYL-DIP

Methyl-DIP works similar to ChIP:

2 experimental samples:

- Control (normal cells)
- Experimental sample (cancer cells)

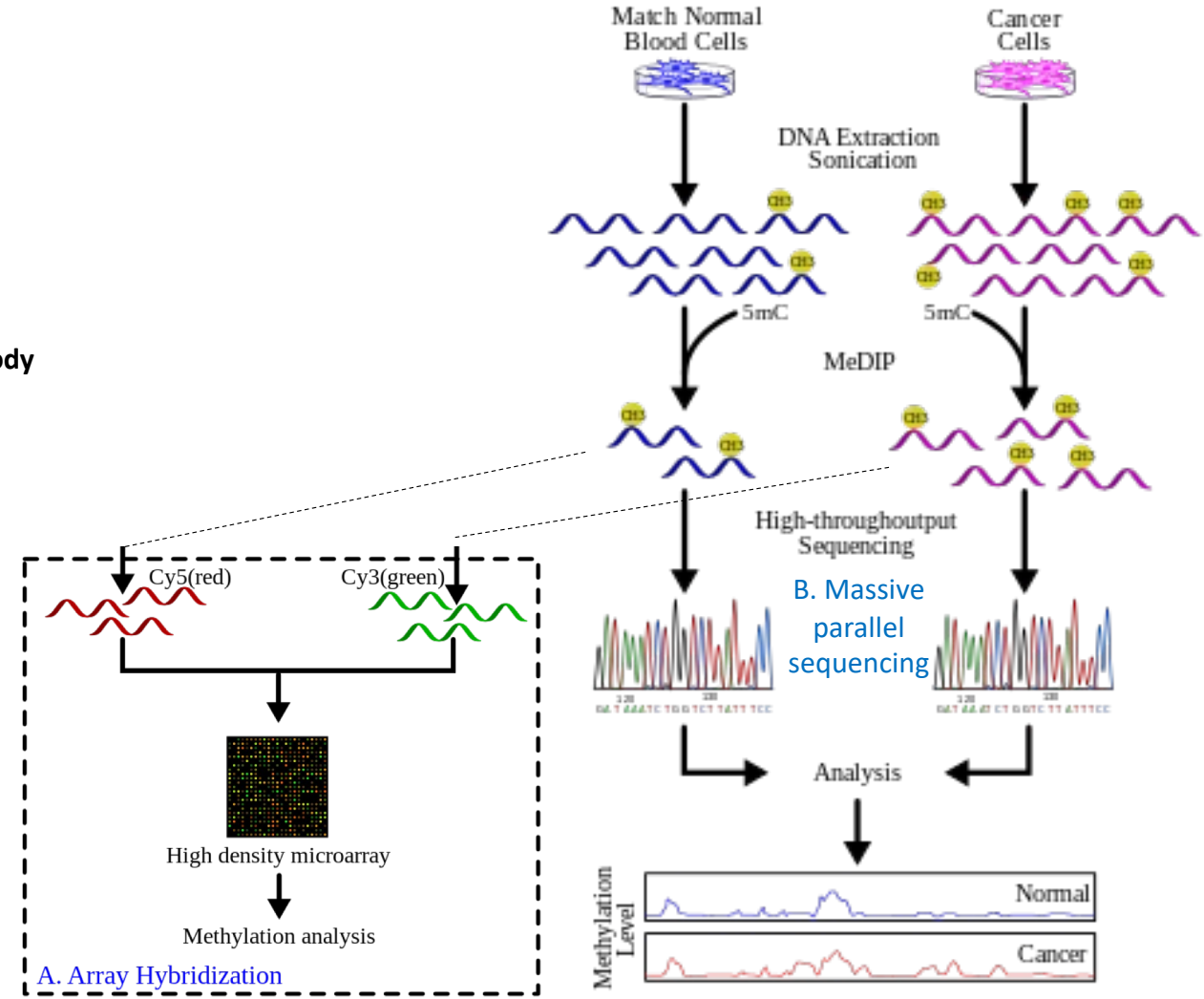
Prepare DNA – carries DNA methylation marks of CpG

Sonicate DNA

Immunoprecipitation using an anti-methyl-CpG specific antibody (monoclonal, discriminates between CpG and met-CpG)

Washing of precipitate

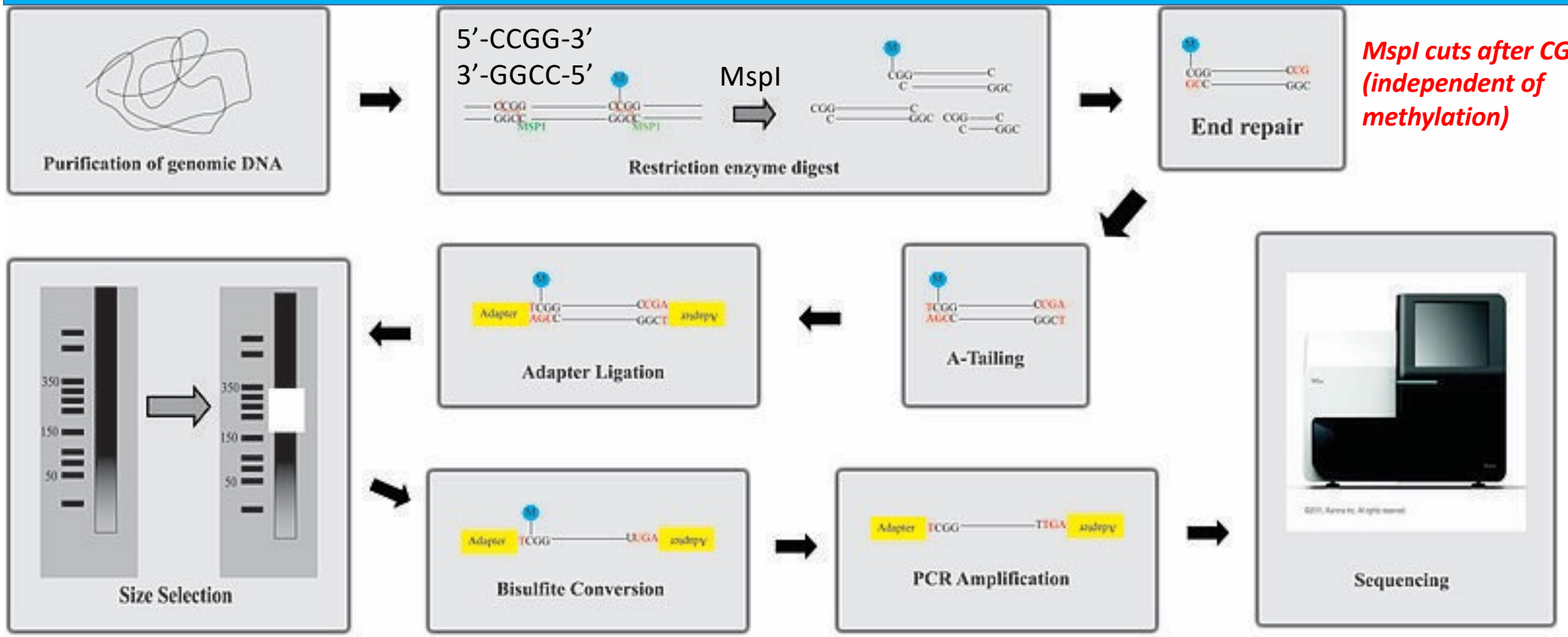
DATA ACQUISITION



Main methods to analyse DNA methylation

1. **PCR on specific CpG islands of interest**
2. **Differential labelling (Cy3-control; Cy5-cancer methyl-DNA) followed by hybridization to genome array**
3. **Massive parallel sequencing (different approaches)**

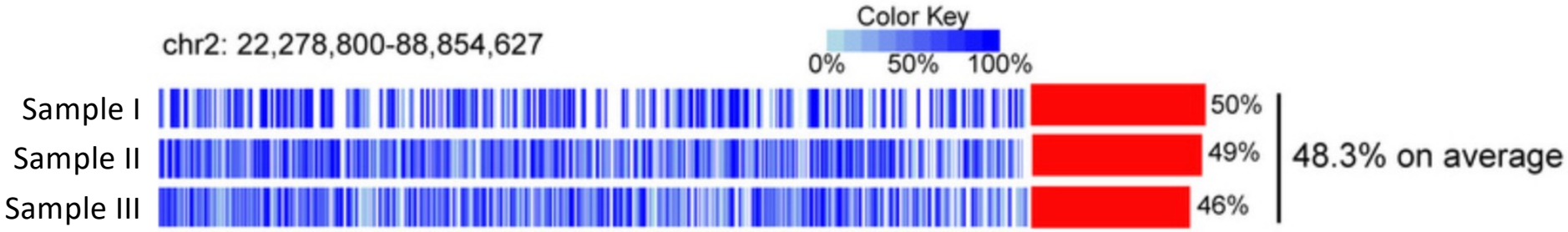
Mapping DNA methylation genome wide:
REDUCED REPRESENTATION BISULFITE SEQEUNCING (RRBS)



Reduced representation bisulfite sequencing (RRBS) is an efficient and high-throughput technique used to analyze the genome-wide methylation profiles on a single nucleotide level. This technique combines restriction enzymes and bisulfite sequencing in order to enrich for the areas of the genome that have a high CpG content. Due to the high cost and depth of sequencing needed to analyze methylation status in the entire genome. The fragments that comprise the reduced genome but still includes the majority of promoters, as well as regions such as repeated sequences that are difficult to profile using conventional bisulfite sequencing approaches.

Mapping DNA methylation genome wide:
REDUCED REPRESENTATION BISULFITE SEQEUNCING (RRBS)

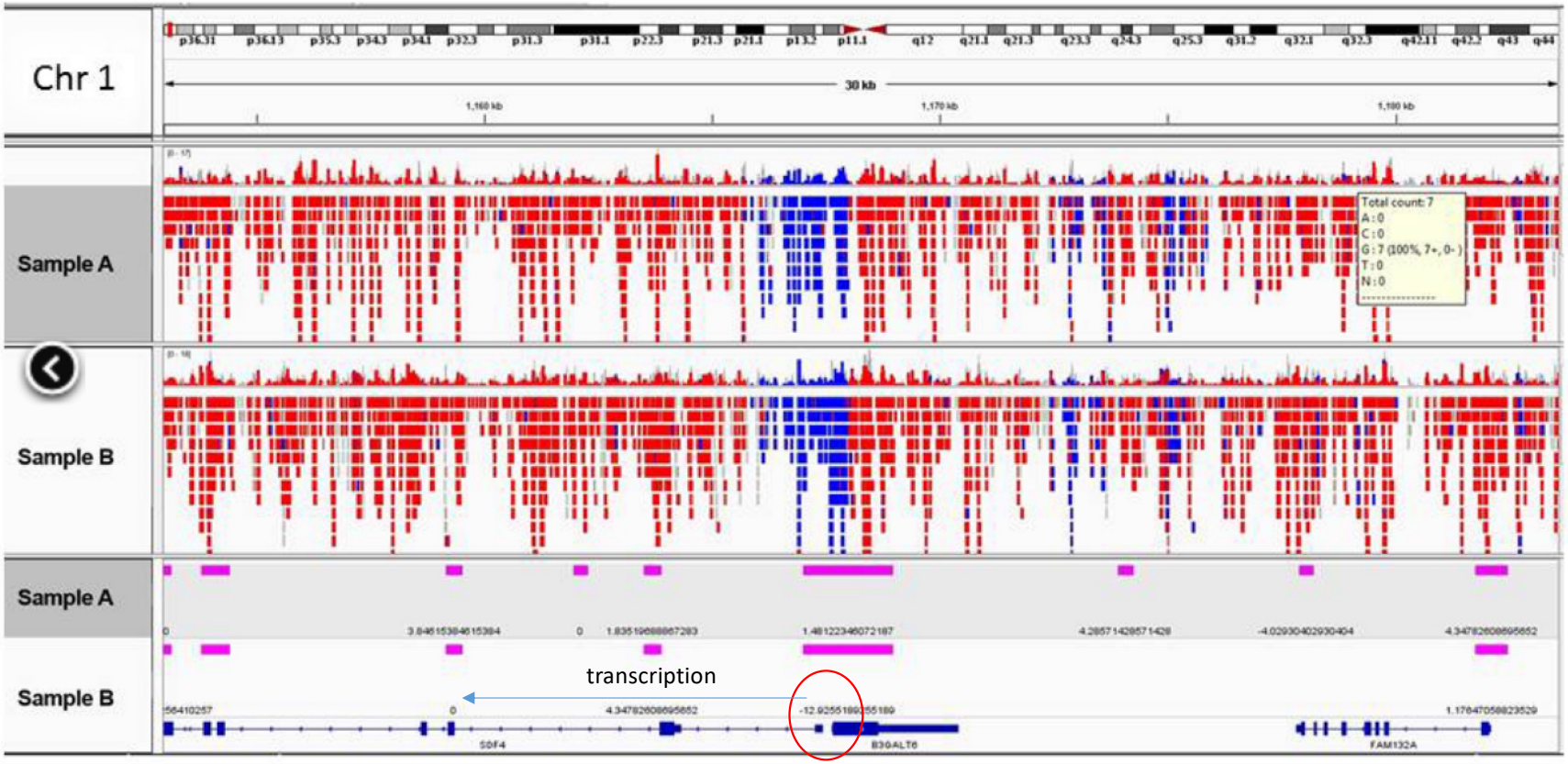
…precise information on global DNA methylation levels



The color key from light blue to dark blue indicates the DNA methylation level from low to high, respectively. The white regions in the left panels indicate a lack of DNA methylation information. The red bars in the right panel represent the average DNA methylation level of the corresponding genomic region. The DNA methylation levels were calculated and presented based on 30 kb windows, only if these windows have more than 5 CpG sites covered.

Mapping DNA methylation genome wide:
REDUCED REPRESENTATION BISULFITE SEQEUNCING (RRBS)

…precise information on locus specific DNA methylation levels



Hypo-methylated CGI

TRANSCRIPTIONAL REGULATION BY METHYL-DNA BINDING PROTEINS

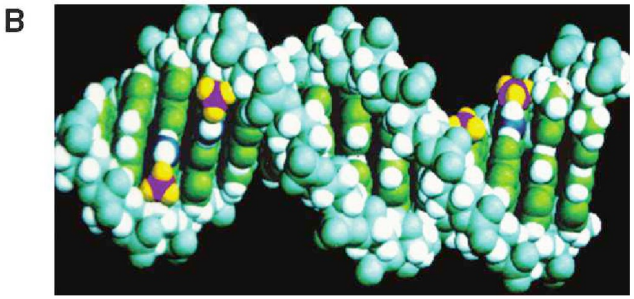
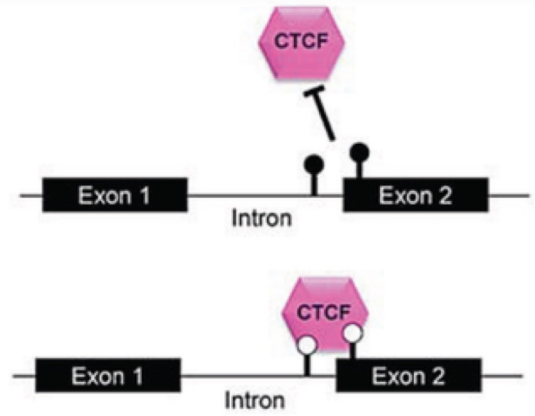


Figure 1. Cytosine methylation in DNA. (*A*) Addition of a methyl group, CH_3 (red), at the five position of the cytosine pyrimidine ring (black arrow) does not sterically interfere with GC base pairing (blue lines). DNA methyltransferases associate covalently with the carbon 6 position (straight green arrow) during methyl group transfer. (*B*) A model of B-form DNA methylated at cytosines in two self-complementary CpG sequences. The paired methyl moieties (magenta and yellow) lie in the major groove of the double helix.

Methylated DNA:
- obtains different structure
- Additional chemical group available for factor interaction

Consequence:
- Transcription factors cannot bind anymore
- Dedicated proteins bind methylated CpG dinuleotides

1. Interference with transcription factor binding



Example: CTCF
Unmethylated DNA CTCF binds → activation of expression
Methylated DNA: CTCF does not bind → no activation

Note: CTCF is a major epigenetic regulator that is involved in controlling genomic imprinting, enhance activation,…

EPIGENTIC READERS OF DNA METHYLATION

2. Transcriptional regulation by methyl-DNA binding proteins

Table 2. Functions of methyl-CpG binding proteins

MBP	Major activity	Species	Major phenotypes of loss-of-function mutations
MeCP2	Binds mCpG with adjacent run AT-rich run Transcriptional repressor	Mouse	Delayed onset neurological defects including inertia, hind-limb clasping, nonrhythmic breathing, and abnormal gait. Postnatal survival ~10 wk.
MECP2	Binds mCpG with adjacent AT run Transcriptional repressor	Human	Heterozygotes suffer from Rett syndrome, a profound neurological disorder characterized by apraxia, loss of purposeful hand use, breathing irregularities, and microcephaly
Mbd1	Binds mCpG via MBD; a major splice form is also able to bind CpG via a CxxC domain	Mouse	No overt phenotype, but subtle defects in neurogenesis detected
Mbd2	Binds mCpG Transcriptional repressor	Mouse	Viable and fertile, but show reduced maternal nurturing behavior. Defective gene regulation in T-helper cell differentiation leading to altered response to infection. Highly resistant to intestinal tumorigenesis.
Mbd3	Core component of NuRD corepressor complex Does not show strong binding to mCpG	Mouse	Early embryonic lethal
Mbd4	DNA repair protein that binds mCpG and T:G mismatches at mCpG sitesThymine DNA glycosylase that excises T from T:G mismatches	Mouse	Viable and fertile. three- to fourfold increase in mutations at CpG sites. Increased susceptibility to intestinal cancer correlates with C to T transitions within the *Apc* gene. Mbd4 functions to minimize the mutability of 5-methylcytosine.
Kaiso	Binds mCGmCG and CTGCNA Transcriptional repressor	Mouse	No overt phenotype. Small but significant delay in tumorigenesis on Min background.

DNA Glycosylase activity

Several proteins were identified to have affinity to methylated CpG
but do not have affinity to un-methylated CpG → mediate transcriptional silencing

→ CpG METHYL BINDING DOMAIN PROTEIN (MBD) FAMILY : MeCP1, MeCP2, MBD1, MBD2, MBD3, MBD4

→ Kaiso (unrelated protein)

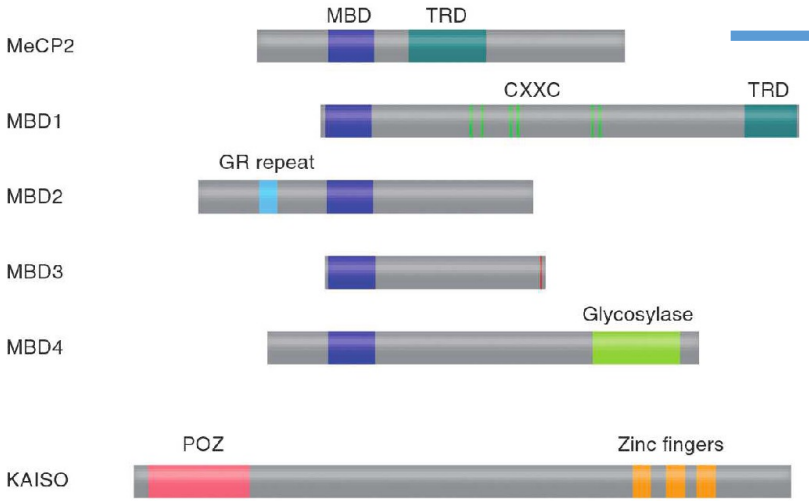
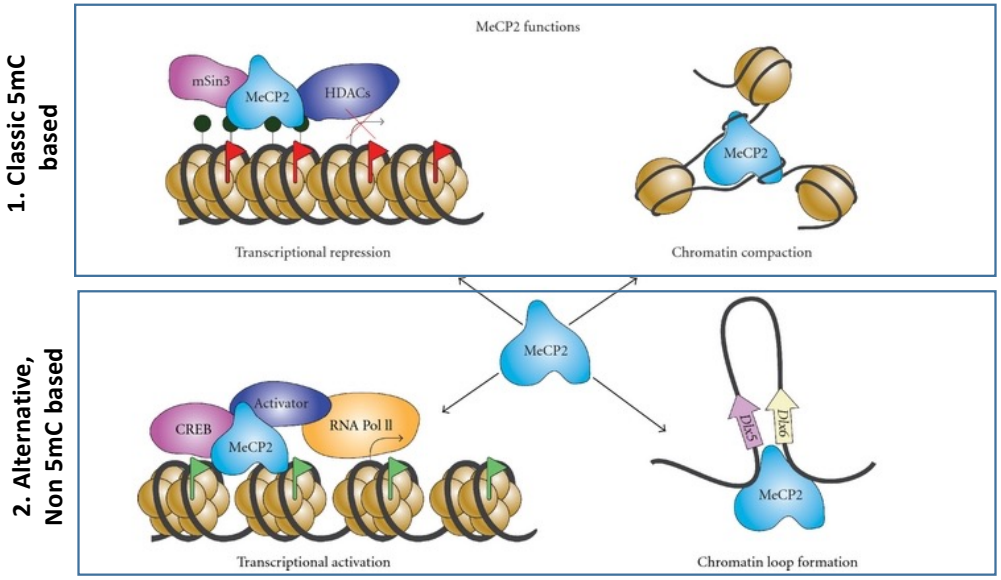


Figure 8. Proteins that bind methyl-CpG. Five members of the MBD protein family are aligned at their MBD domains (purple). Other domains are labeled and include TRD; CXXC domains, which are zinc fingers, some of which are implicated in binding to nonmethylated CpG; GR repeats that may bind; a T:G mismatch glycosylase domain that is involved in repair of 5mC deamination. Kaiso lacks the MBD domain, but binds methylated DNA via zinc fingers (orange) and possesses a POB/BTB domain that is shared with other transcriptional repressors. Domain abbreviations: MBD, methyl-CpG binding domain; TRD, transcriptional repression domain; POZ, poxvirus and zinc finger, a protein–protein interacting domain.

MeCP2 mutations cause Rett syndrome

Methylated CpG is bound by MeCP2



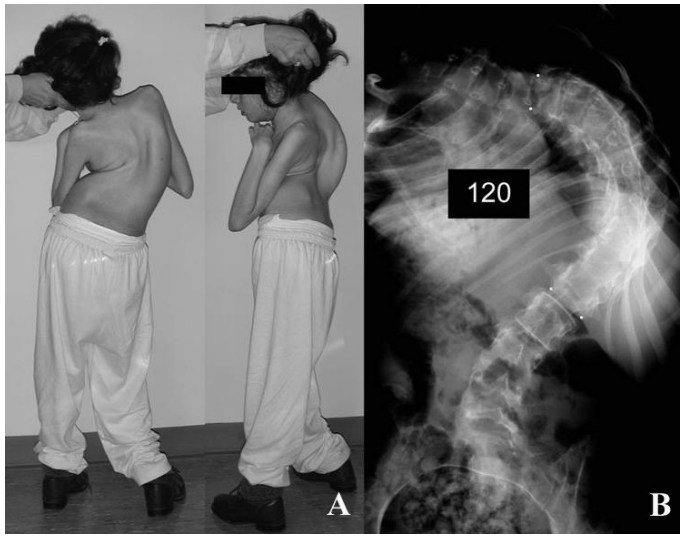
However MeCP2 was also found to be located on active genes

1. MeCP2 function acts as transcriptional repressor by binding methylated DNA and associattes **with corepressor complexes such as mSin3A and HDACs** .
2. Genomewide search for MeCP2 genomic distribution also showed DNA methylation independent role of MeCP2 in the assembly of HAT complexes

(i) MeCP2 was found to be associated with transcriptionally active genes

(ii) **Only 2.2%** of the most methylated promoters were bound by MeCP2.

Rett syndrome is caused by mutations in MeCP2



MeCP2 is X-linked
Mutations that mediate Rett syndrome are dominant (lethal in males)

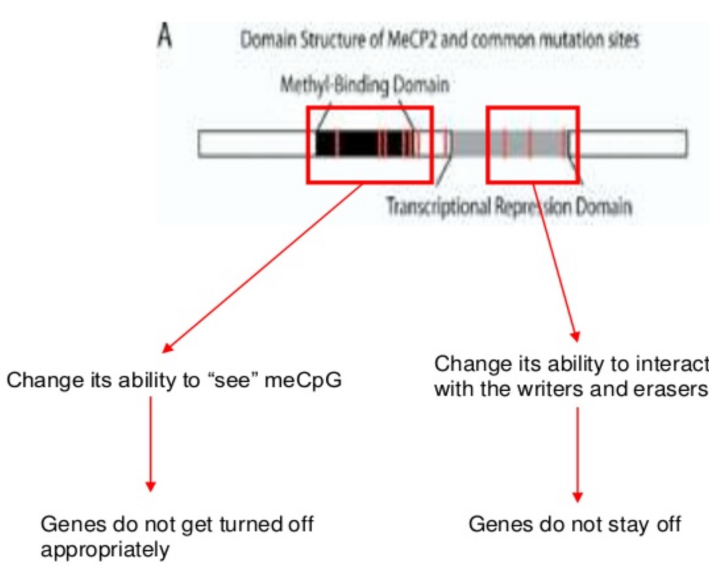
Affects predominantly neurons

Rett syndrome (RTT), originally termed cerebroatrophic hyperammonemia is a rare genetic postnatal neurological disorder of the grey matter of the brain that almost exclusively affects females but has also been found in male patients. The clinical features include small hands and feet and a deceleration of the rate of head growth (including microcephaly in some). Repetitive stereotyped hand movements, such as wringing and/or repeatedly putting hands into the mouth, are also noted. People with Rett syndrome are prone to gastrointestinal disorders and up to 80% have seizures. They typically have no verbal skills, and about 50% of affected individuals do not walk. Scoliosis, growth failure, and constipation are very common and can be problematic.

Loss of MeCP2 leads to alteration in gene expression –> Phenoptypic manifestation

How does MeCP2 effect the brain function?

- Through it's job as a reader of epigenetic bookmarks
- The wide array of functions that MeCP2 performs ALL contribute to Rett syndrome.
- The different mutations have different effects on the presentation of the disease.
- In addition since each person is different based on their personal epigenetics, the disease will be individual as well.



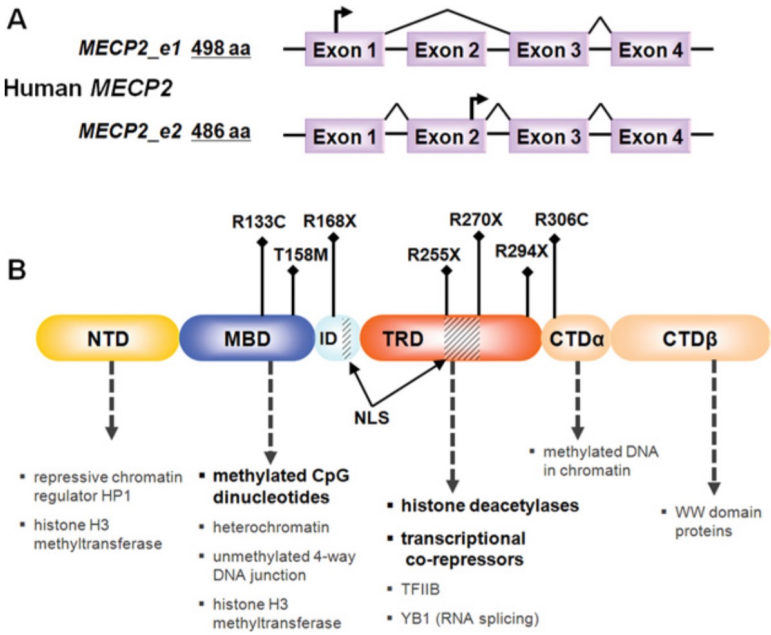


Figure 1 Composition of MeCP2: gene structure, splicing patterns and putative functional domains

(**A**) Splicing patterns generating the two mRNA isoforms of *MECP2*, _e1 and _e2. The two isoforms generate two protein isoforms of MeCP2 with differing N-termini due to the use of alternative translation start sites (bent arrows) and the absence or presence of exon 2 in the transcript. (**B**) Apart from the N-terminus, both MeCP2 isoforms are identical and contain several functionally distinct domains: NTD, N-terminal domain; MBD, methylated DNA-binding domain; ID, interdomain; TRD, transcription repression domain; CTD, C-terminal domain; NLS; nuclear localization signals. Locations of seven of the most common point mutations in RTT are indicated (◆). Below each domain are indicated major (bold) and other (grey) interactors and functions. HP1, heterochromatin protein 1; TFIIB, transcription factor IIB; YB1, Y-box-binding protein 1.

EPIGENTIC READERS OF DNA METHYLATION

Transcriptional regulation by methyl-dna binding proteins

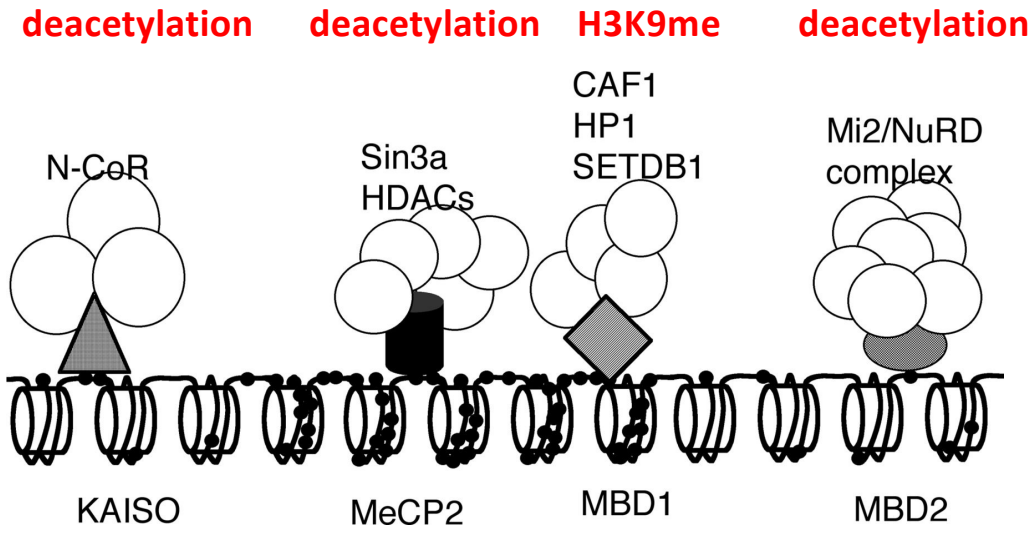
Table 2. Functions of methyl-CpG binding proteins

MBP	Major activity	Species	Major phenotypes of loss-of-function mutations
MeCP2	Binds mCpG with adjacent run AT-rich run Transcriptional repressor	Mouse	Delayed onset neurological defects including inertia, hind-limb clasping, nonrhythmic breathing, and abnormal gait. Postnatal survival ~10 wk.
MECP2	Binds mCpG with adjacent AT run Transcriptional repressor	Human	Heterozygotes suffer from Rett syndrome, a profound neurological disorder characterized by apraxia, loss of purposeful hand use, breathing irregularities, and microcephaly
Mbd1	Binds mCpG via MBD; a major splice form is also able to bind CpG via a CxxC domain	Mouse	No overt phenotype, but subtle defects in neurogenesis detected
Mbd2	Binds mCpG Transcriptional repressor	Mouse	Viable and fertile, but show reduced maternal nurturing behavior. Defective gene regulation in T-helper cell differentiation leading to altered response to infection. Highly resistant to intestinal tumorigenesis.
Mbd3	Core component of NuRD corepressor complex Does not show strong binding to mCpG	Mouse	Early embryonic lethal
Mbd4	DNA repair protein that binds mCpG and T:G mismatches at mCpG sitesThymine DNA glycosylase that excises T from T:G mismatches	Mouse	Viable and fertile. three- to fourfold increase in mutations at CpG sites. Increased susceptibility to intestinal cancer correlates with C to T transitions within the *Apc* gene. Mbd4 functions to minimize the mutability of 5-methylcytosine.
Kaiso	Binds mCGmCG and CTGCNA Transcriptional repressor	Mouse	No overt phenotype. Small but significant delay in tumorigenesis on Min background.

DNA Glycosylase activity

Methyl-CpG binding proteins are present in transcriptional co-repressor complexes

-MeCP2: component of the Sin3A HDAC complex
-MBD3: component of the NuRD HDAC complex
-MBD1: interacts with HDAC3. Mbd1 and HDAC3 are recruited by the PML-RARalpha hybrid protein to silence gene expression in Acute promyelocytic leukemia
-MBD1: interacts with the H3K9 HMTase SETDB1



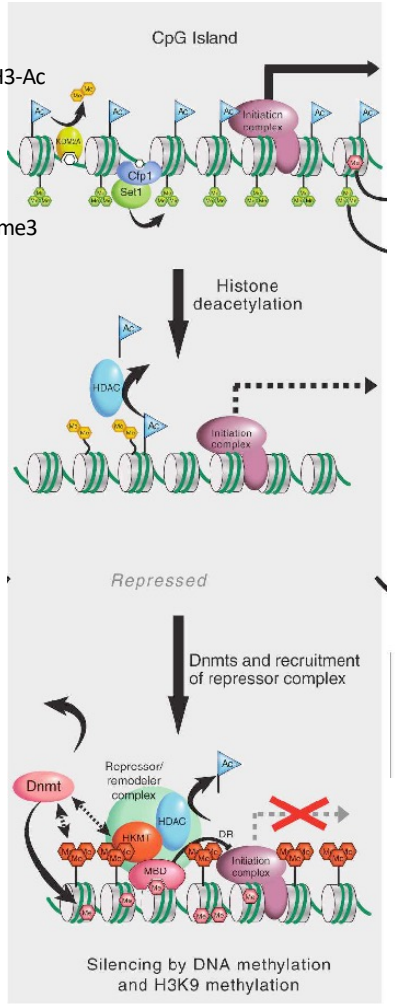
Collaboration to repress genes

TRANSCRIPTIONAL REGULATION BY METHYL-DNA BINDING PROTEINS
RECRUITMENT OF methyl-CpG binding proteins and co-repressor complexes

Active Status

Rapid Transition phase

Inactive Status

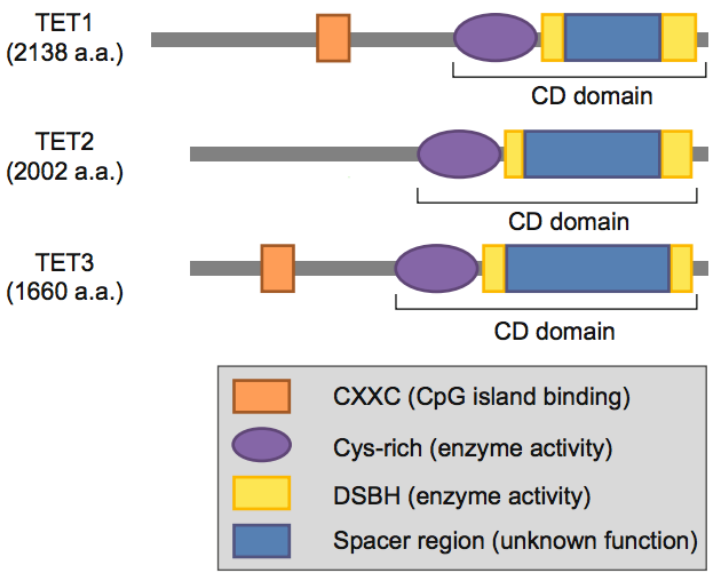


Absence of CGI methylation
Histone acetylation
Histone methylation (H3K4me3 by SET1)
collaborates to allow promoter activity and gene expression

Recruitment of transcriptional repressor that recruit:
DNMT3a/3b: de-novo DNA methylation (generation of local 5'methylcytisine
MBD1-HDAC3 complex binds methylated CpG → histone de-acetylation
MBD1 interacts with the HMTase SETDB1 → H3K9me3

CGI methylation
Histone hypoacetylation
Histone methylation (H3K9me3 or H3K27me3)

DNA METHYLATION IS REVERSIBLE – by OXIDATION:
DNA DEMETHYLATION BY Tet-family proteins = ACTIVE DE-METHYLATION



Critical domains and factors

- C-terminal CD domain (containing the Cys-rich and DSBH regions) with dioxygenase activity
- Cofactors: 2-oxoglutarate (2-OG) - and iron (II)-dependent dioxygenase activity
- Spacer region, the length of which varies between TET family proteins and its function remains unknown.
- CXXC domain, mediates their direct DNA-binding ability: the TET1 CXCC domain recognizes not only unmodified cytosine but also 5mC and 5hmC, and it favors binding to regions in the genome of high CpG content

DNA METHYLATION IS REVERSIBLE – by OXIDATION: DNA DEMETHYLATION BY Tet-family proteins

1. Demethylation by oxidation and absence of maintenance by DNMT1

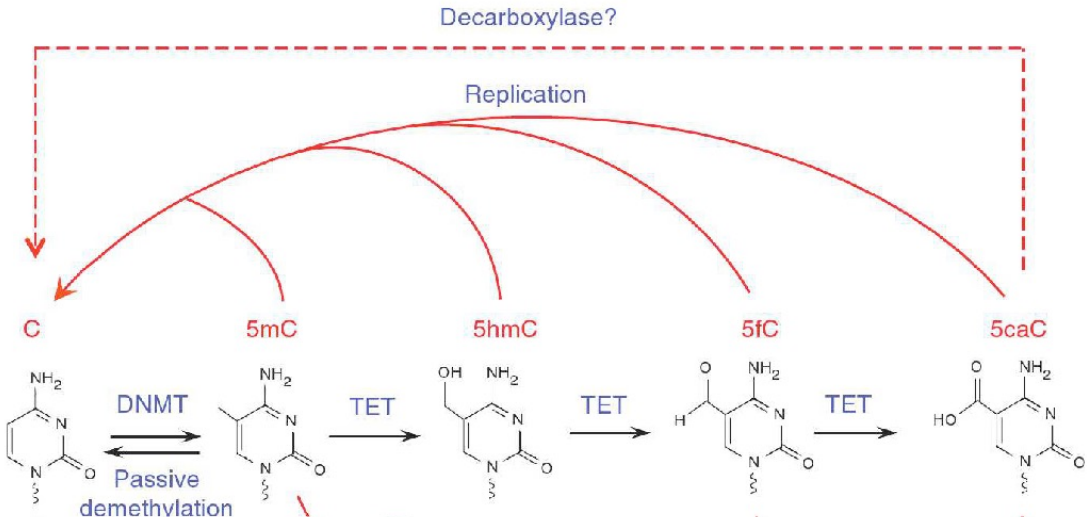


Figure 6. Model of Tet-initiated DNA demethylation pathways. DNA methylation (5mC) is established and maintained by DNMT. 5mC can by oxidized by Tet family of dioxygenases to generate 5hmC, 5fC, and 5caC. Because the oxidized 5mC derivatives cannot serve as substrates for DNMT1, they can be lost by replication-dependent passive demethylation. 5hmC can be deaminated by AID/APOBEC to become 5hmU, which together with 5fC and 5caC can be excised by glycosylases such as TDG, followed by DNA repair to generate C. Alternatively, a putative decarboxylase may convert 5caC to C.

Tet-family proteins mediate DNA demethylation

CpG dinucleotides on template DNA strand → DNA replication → Parental DNA strand neo-synthesized DNA

5mC: 5-methylcytosine — GC-m3 / m3-CG

Tet family proteins → oxidation

5hmC: 5-hydroxymethylcytosine — GC-5hmC / 5hm-CG

Tet family proteins → oxidation

5fC: 5-formylcytisine — GC-5fcC / 5fc- CG

Tet family proteins → oxidation

5caC: 5-carboxylcytosine — GC-5caC / 5ca-CG

5mC, 5hmC and 5fC are abundant in the cell
5caC is present only at very low abundance

DNMT1 has exclusive specificity for 5mC
→ 5hMC, 5fC, 5caC: no substrate for DNMT1
→ Passive dillution of 5hMC, 5fC, 5caC during DNA replication

DNA METHYLATION IS REVERSIBLE:
DNA DEMETHYLATION BY Tet-family proteins

2. Demethylation by the action of Thymine DNA Glycosylases and Base Excision Repair

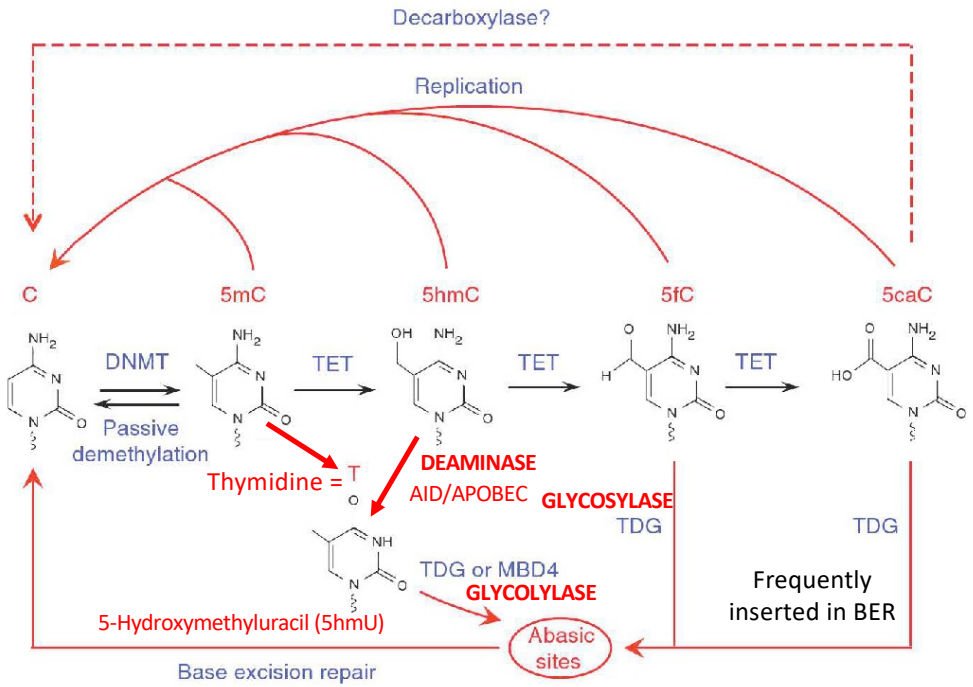


Figure 6. Model of Tet-initiated DNA demethylation pathways. DNA methylation (5mC) is established and maintained by DNMT. 5mC can by oxidized by Tet family of dioxygenases to generate 5hmC, 5fC, and 5caC. Because the oxidized 5mC derivatives cannot serve as substrates for DNMT1, they can be lost by replication-dependent passive demethylation. 5hmC can be deaminated by AID/APOBEC to become 5hmU, which together with 5fC and 5caC can be excised by glycosylases such as TDG, followed by DNA repair to generate C. Alternatively, a putative decarboxylase may convert 5caC to C.

DEAMINASE:

AID/APOBEC ("apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like")
is a conserved cytidine deaminase. Members of this family are /5mC→ T / 5hmC → 5hmU editing enzymes
Generation of mismatch!!

DNA GLYCOSYLASES:

Methyl-CpG Domain Protein 4 (MBD4):
G:T glycosylase - remove thymine generated by the deamination of 5-methylcytosine. Remove 5hmU produced by AID/AOPEC activity

Thymine-DNA glycosylase (TDG)
Clear involvment in gene expression control

- Thymine-DNA glycosylase (TDG) removes thymine moieties from G/T mismatches by hydrolyzing the carbon-nitrogen bond between the sugar-phosphate backbone of DNA and the mispaired thymidine.
- With lower activity, this enzyme also removes thymine from C/T and T/T misspairings.
- TDG can also remove uracil and 5-bromouracil from mispairings with guanine.
- **-Efficiently excise 5-Hydroxymethyluracil (5hmU, 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), the key oxidation products of 5-methylcytosine in genomic DNA**

INote, TDG knockout mouse models showed no increase in mispairing frequency suggesting that other enzymes, like the functional homologue MBD4, may provide functional redundancy.

.

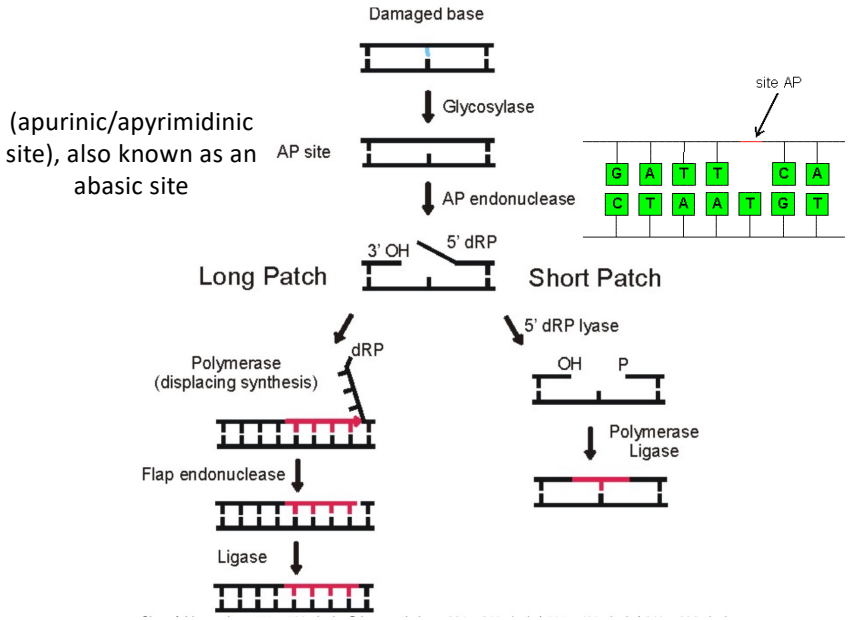
→ RESULT: Activation of BER pathway

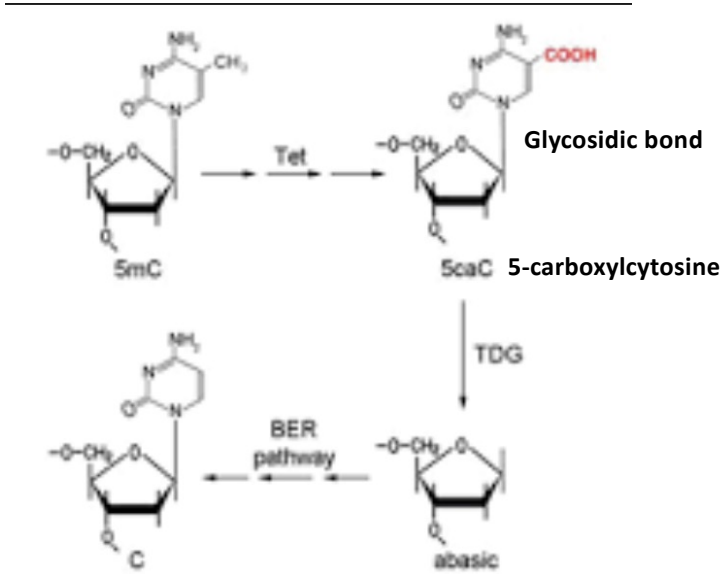
Check textbooks: glycosilases cleave off bases from sugar → apyrimidic/apurinic site → BER pthway

DNA METHYLATION IS REVERSIBLE:
DNA DEMETHYLATION BY Tet-family proteins

2. Demethylation by the action of Thymine DNA Glycosylases and Base Excision Repair

Base excision pathway and loss of DNA methylation



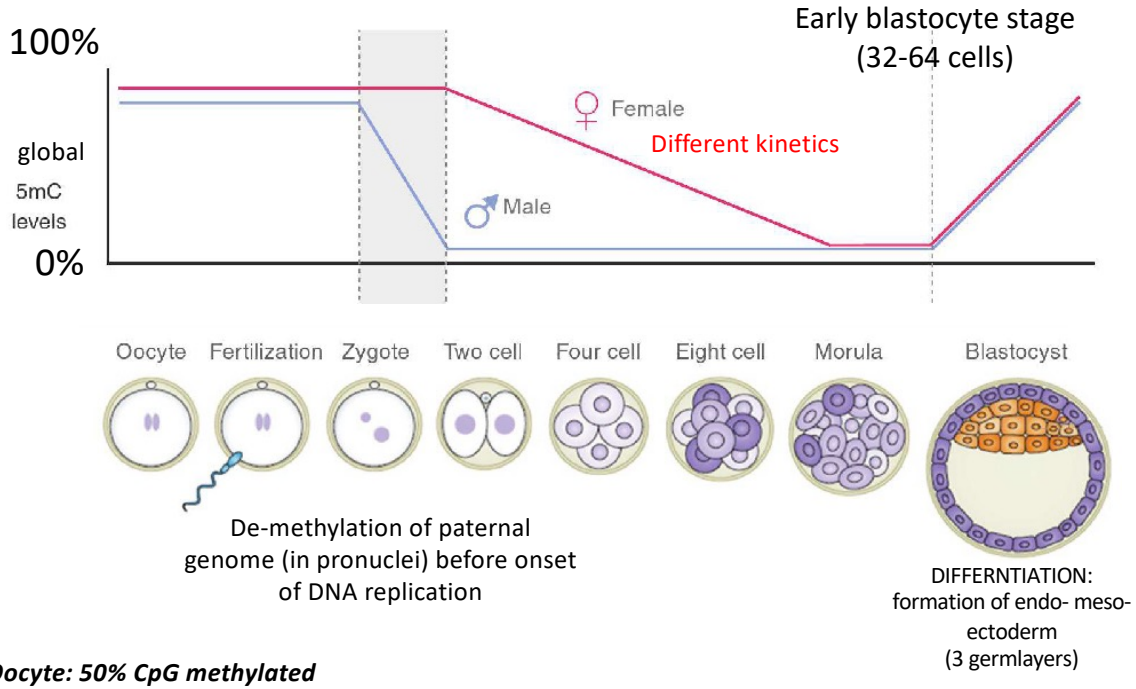


Unmethylated, on both strands

→ Activation of BER pathway
→ Insertion of unmethylated C
→ Methylated C in CpG dinucleotides lost
→ **ACTIVE DNA DE-METHYLTION**

DYNAMICS OF DNA METHYLATION: Example epigenetic reprogramming post-fertilization

DNA METHYLATION IS ABUNANT IN THE GENOME AND IS SUBJECTED TO DRAMATIC ALTERATIONS DURING EMBRYOGENESIS



Oocyte: 50% CpG methylated
Sperm: 80% CpG methylated

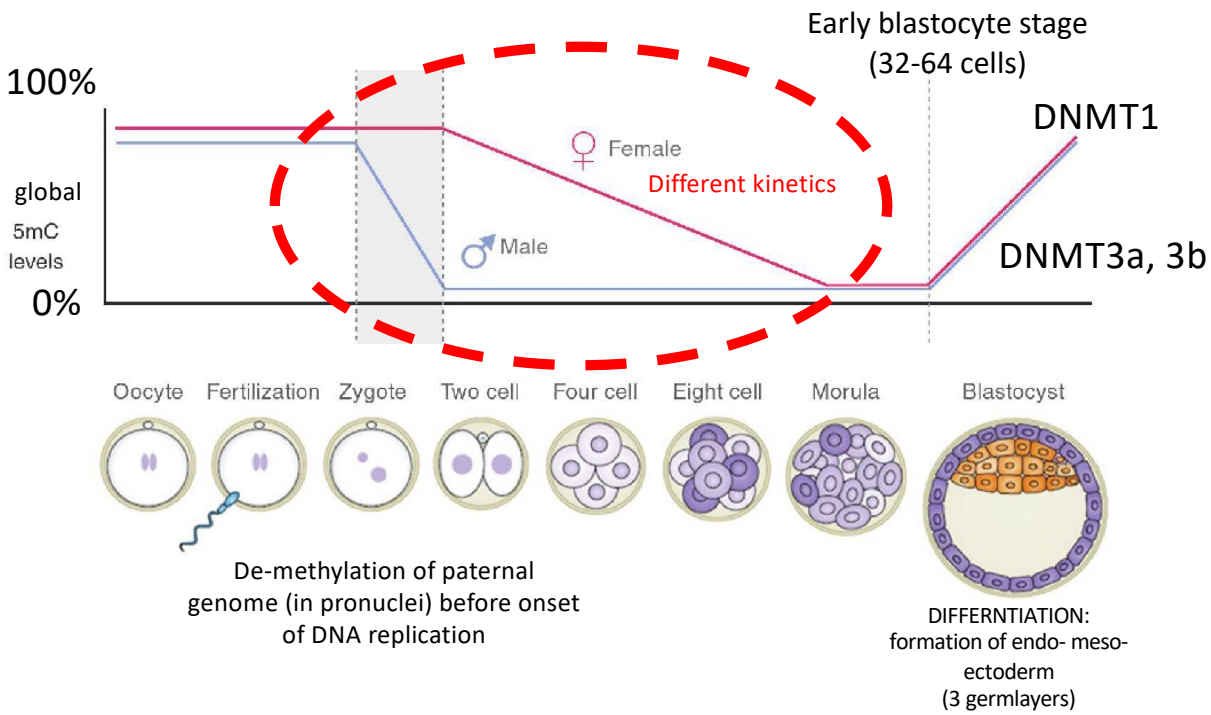
Experimental approach – perform DNA methylation mapping in embryonic development:

- Prepare DNA from sperm and oocyte
- Prepare DNA from male and female pronucleus
- Prepare DNA from 2, 4, 8 cell embryos
- Prepare DNA from Morula
- Prepare DNA from Blastocyst

Quantify 5'-Methylcytopsine content considering regions containing **male – female polymorphisms**

50%- 80% of CpG di-nucleotides are methylated in the human genome in sperm/oocytes
DNA methylation is erased during early embryonic development
Different kinetics
Re-establishment of DNA methylation patterns during differentiation

DNA METHYLATION IS ABUNDANT IN THE GENOME AND IS SUBJECTED TO DRAMATIC ALTERATIONS DURING EMBRYOGENESIS



DNA methylation levels are high in fertilized Oocytes that contain the paternal and maternal genome (carries characteristic methylation patterns)

Paternal and maternal methylation patterns are erased with different kinetics (exception: imprinted genes maintain paternal and maternal methylation information). → the paternal and maternal methylation epigenome is cancelled

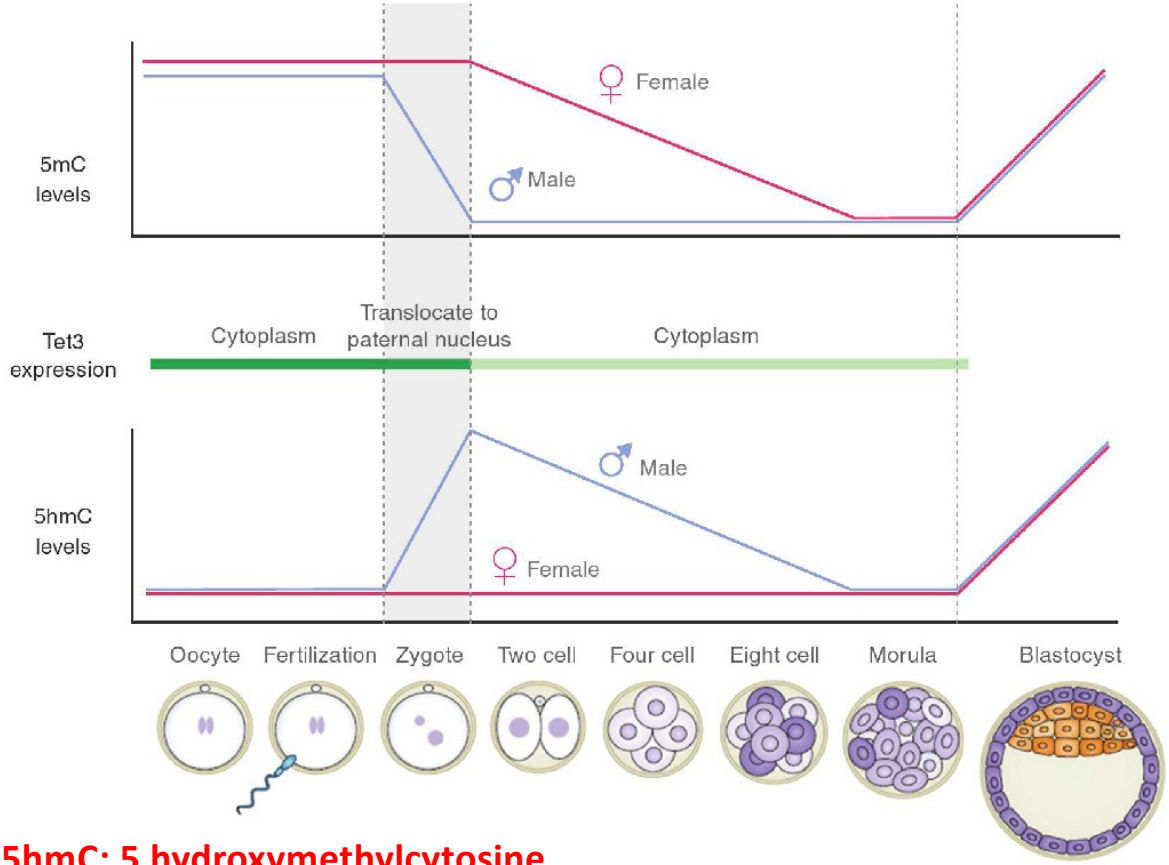
DNA methylation levels remain low during the first cell division events until the blastocyst stage

In the blastocyst stage cell differentiation programs are activated and genes need to be regulated on the epigenetic level → DNA methylation is increasing due to the action of DNMT3a, DNMT3b
(remember: loss of DNMT1, DNMT3a or DNMT3b is lethal → establishment and maintenance of DNA methylation is impaired)

70%- 80% of CpG di-nucleotides are methylated in the human genome!
Remember only 2% of the genome encode for mRNAs
98% is noncoding DNA that contains a large proportion of transposable elements, repeat sequences, etc…

DNA METHYLATION IS REVERSIBLE:
ACTIVE AND PASSIVE DNA DEMETHYLATION

DNA de-methylation of the paternal and maternal genome has different kinetics



5hmC: 5 hydroxymethylcytosine

ACTIVE DNA DEMETHYLATION – PATERNAL GENOME

Enzymatic activity rapidly de-methylates 5mC → increase of 5hmC

PATERNAL GENOME: fast de-methylation of DNA

- → ***In zygotes Tet3 is localized to the PATERNAL pronucleus***
- → *Paternal DNA is demethylated*
- → *High levels of* 5hmC: 5-hydroxymethylcytosine, 5fc: 5-formylcytisine and 5caC: 5-carboxylcytosine *were detected at high levels in the paternal nucleus*
- → *BER machinery concentrated in pronucleus*

PASSIVE DNA DEMETHYLATION – MATERNAL GENOME

Successive rounds of DNA methylation reduce the amount of 5mC.

DNMT1 and Tet3 are excluded from the nucleus in early phase of development (until blastocyst)

MATERNAL GENOME: slow de-methylation of DNA

(PATERNAL GENOME: already demethylated by Tet3)

DNA METHYLATION IS REVERSIBLE:
ACTIVE AND PASSIVE DNA DEMETHYLATION

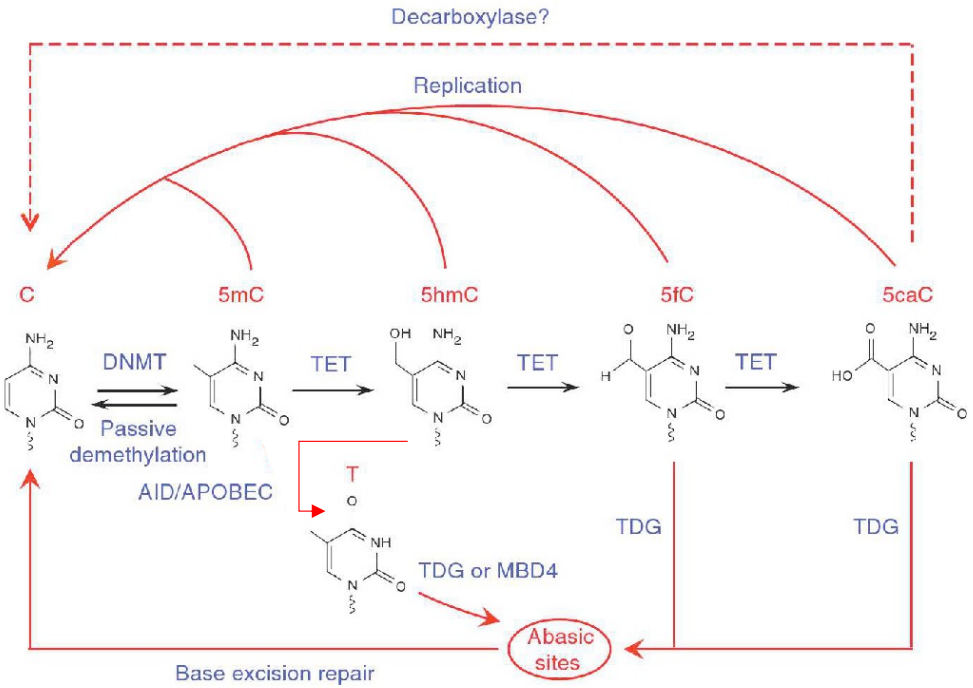
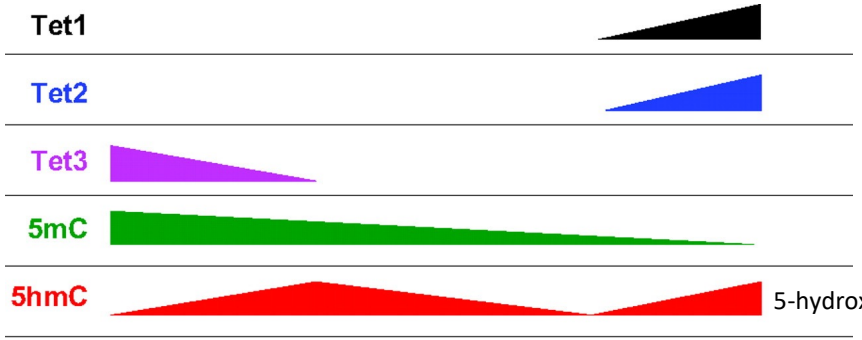


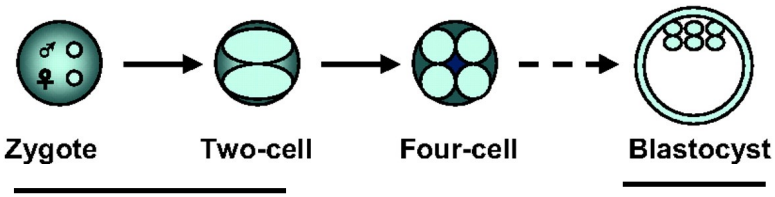
Figure 6. Model of Tet-initiated DNA demethylation pathways. DNA methylation (5mC) is established and maintained by DNMT. 5mC can by oxidized by Tet family of dioxygenases to generate 5hmC, 5fC, and 5caC. Because the oxidized 5mC derivatives cannot serve as substrates for DNMT1, they can be lost by replication-dependent passive demethylation. 5hmC can be deaminated by AID/APOBEC to become 5hmU, which together with 5fC and 5caC can be excised by glycosylases such as TDG, followed by DNA repair to generate C. Alternatively, a putative decarboxylase may convert 5caC to C.

ACTIVE DNA DEMETHYLATION – PATERNAL GENOME

Enzymatic activity rapidly de-methylates 5mC
PATERNAL GENOME: fast de-methylation of DNA

- → ***In zygotes Tet3 is localized to the PATERNAL nucleus***
- → *Paternal DNA is demethylated*
- → *High levels of* 5hmC: 5-hydroxymethylcytosine, 5fc: 5-formylcytisine and 5caC: 5-carboxylcytosine *were detected at high levels in the paternal nucleus*
- → *BER machinery concentrated in pronucleus*





wave of 5mC oxidation

Changes in gene regulation
Changes in DNA metylation
5hmC levels increase