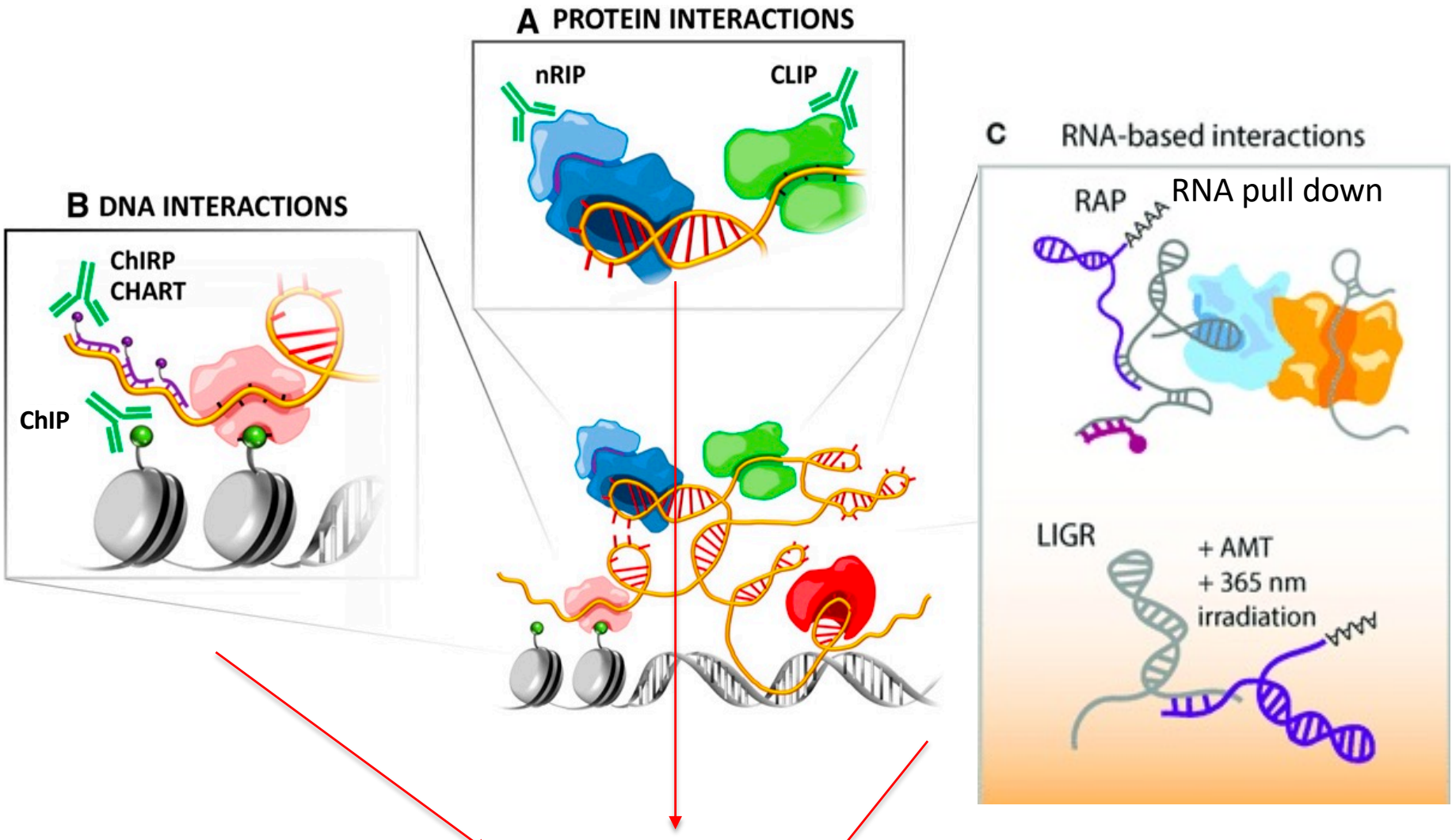
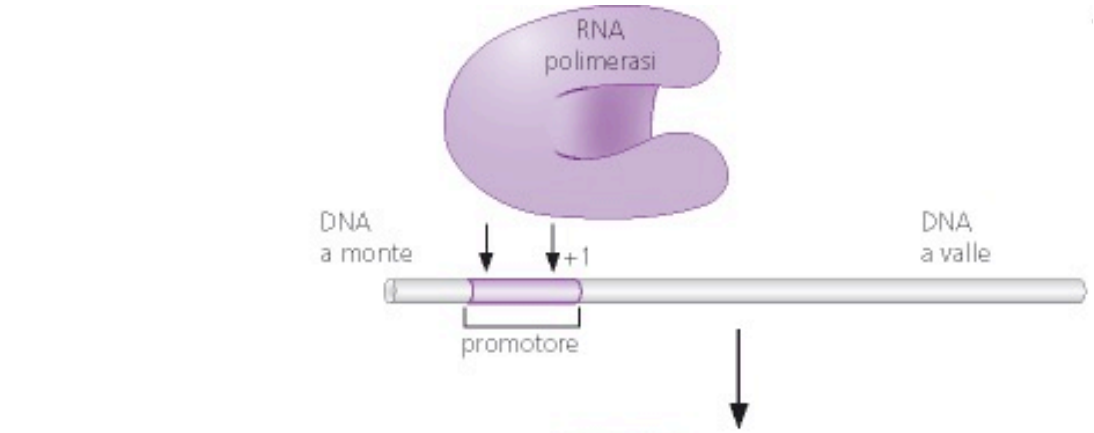
Methodologies for functional studies

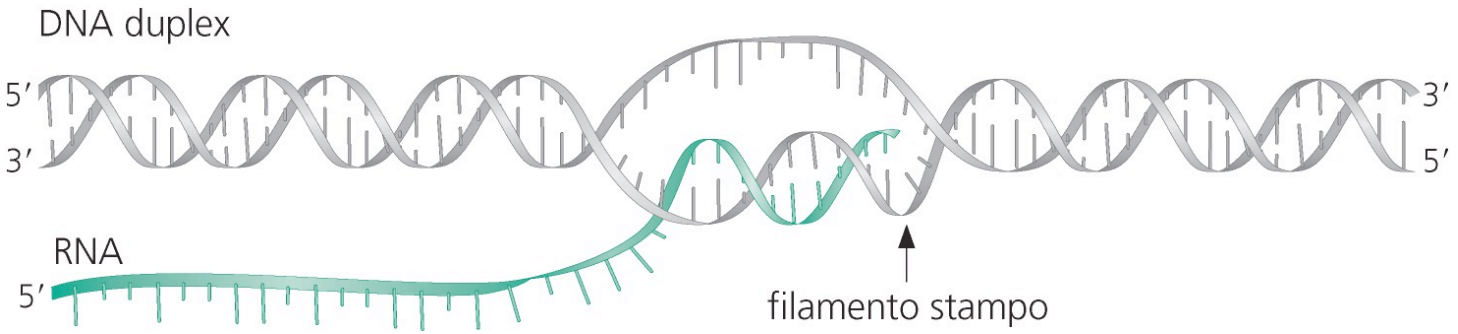


PCR and RT-PCR

DNA or RNA Sequencing

RNA transcription





Eukaryotic Transcriptional Regulation

1. Level of Chromatin (DNA accessibility)

- DNA methylation
- Histone modifications
- Histone remodeling
- Nucleosome composition

2. Level of DNA (Interaction with basal transcription machinery)

- Regulatory sequences (enhancers, silencers)
- Transcription factors (activators, repressors)

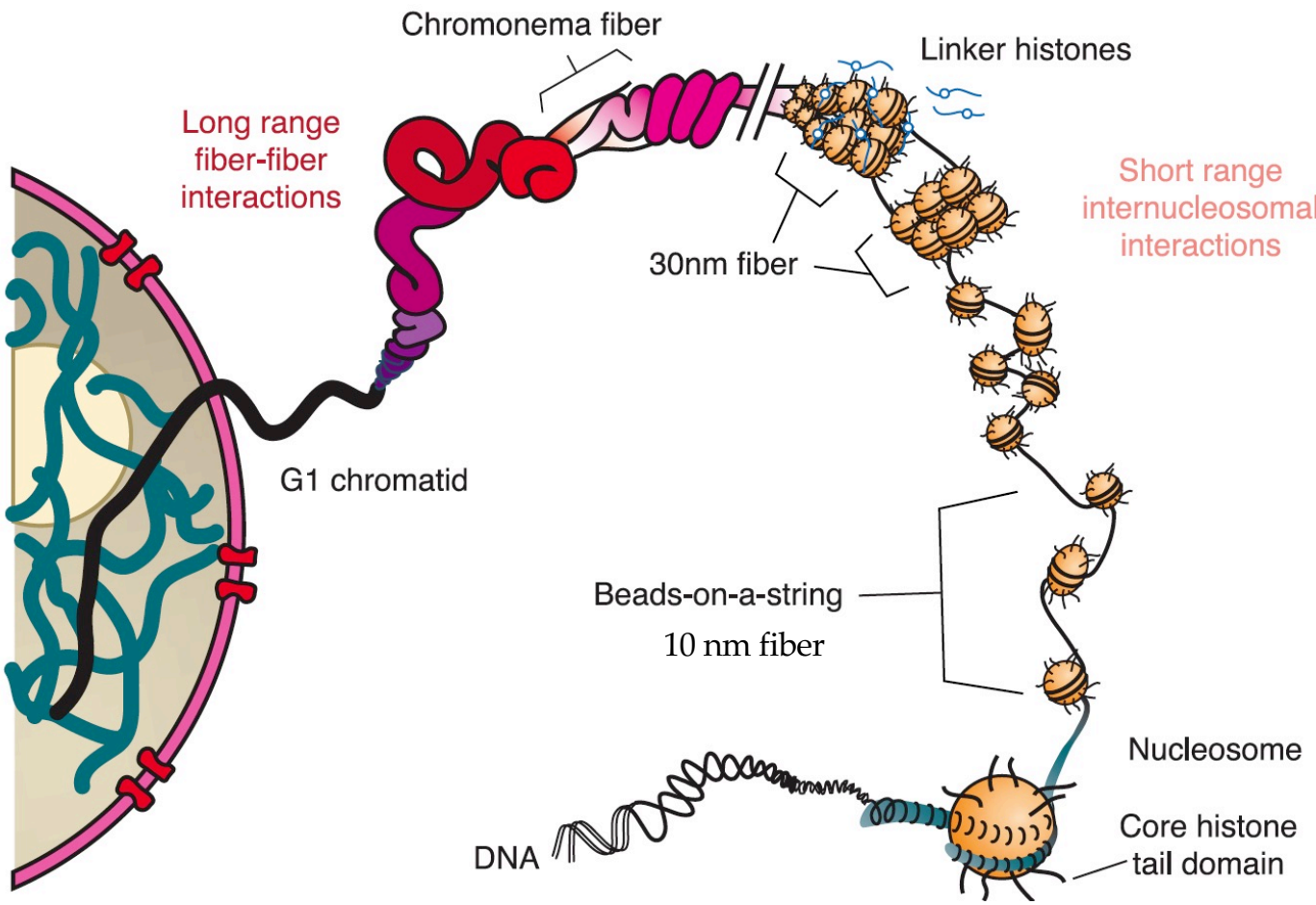
3. Level of Regulatory RNA (Interaction with DNA, RNA or protein)

- Small and long non-coding RNAs

Eukaryotic Transcriptional Regulation

-Level of Chromatin-

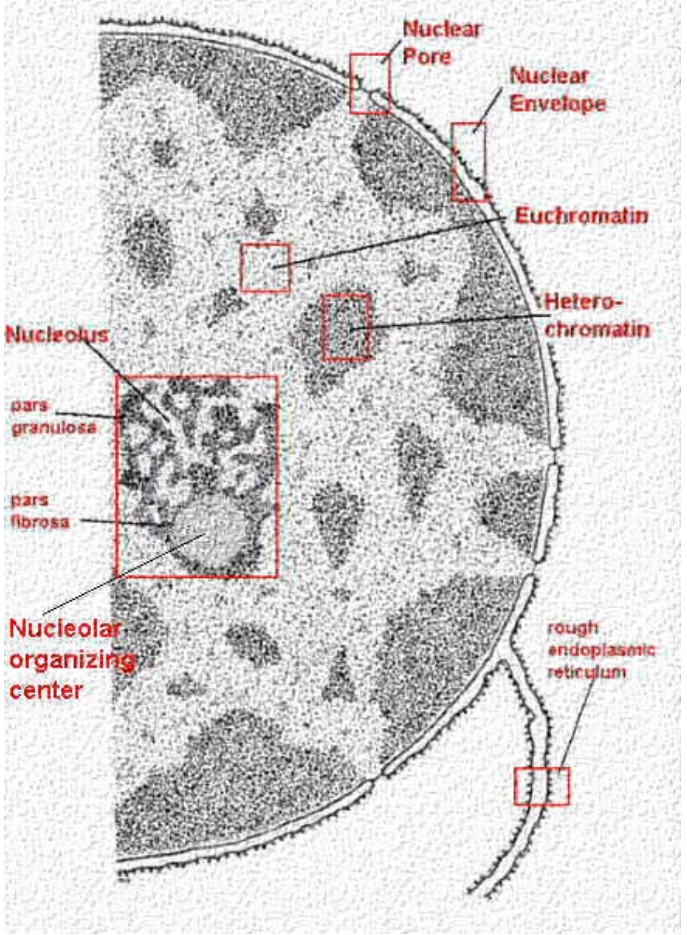
Chromatin-structure - reminder



- **Several levels of condensation**
  - 10 nm extended fiber with nucleosomes like beads-on-a-string (7x compaction)
  - 30 nm condensed fiber (40-50x compaction)
  - And higher

Chromatin compaction influences activity of DNA in transcription

- **Heterochromatin** -transcriptionally silent
- **Euchromatin** - transcriptionally active



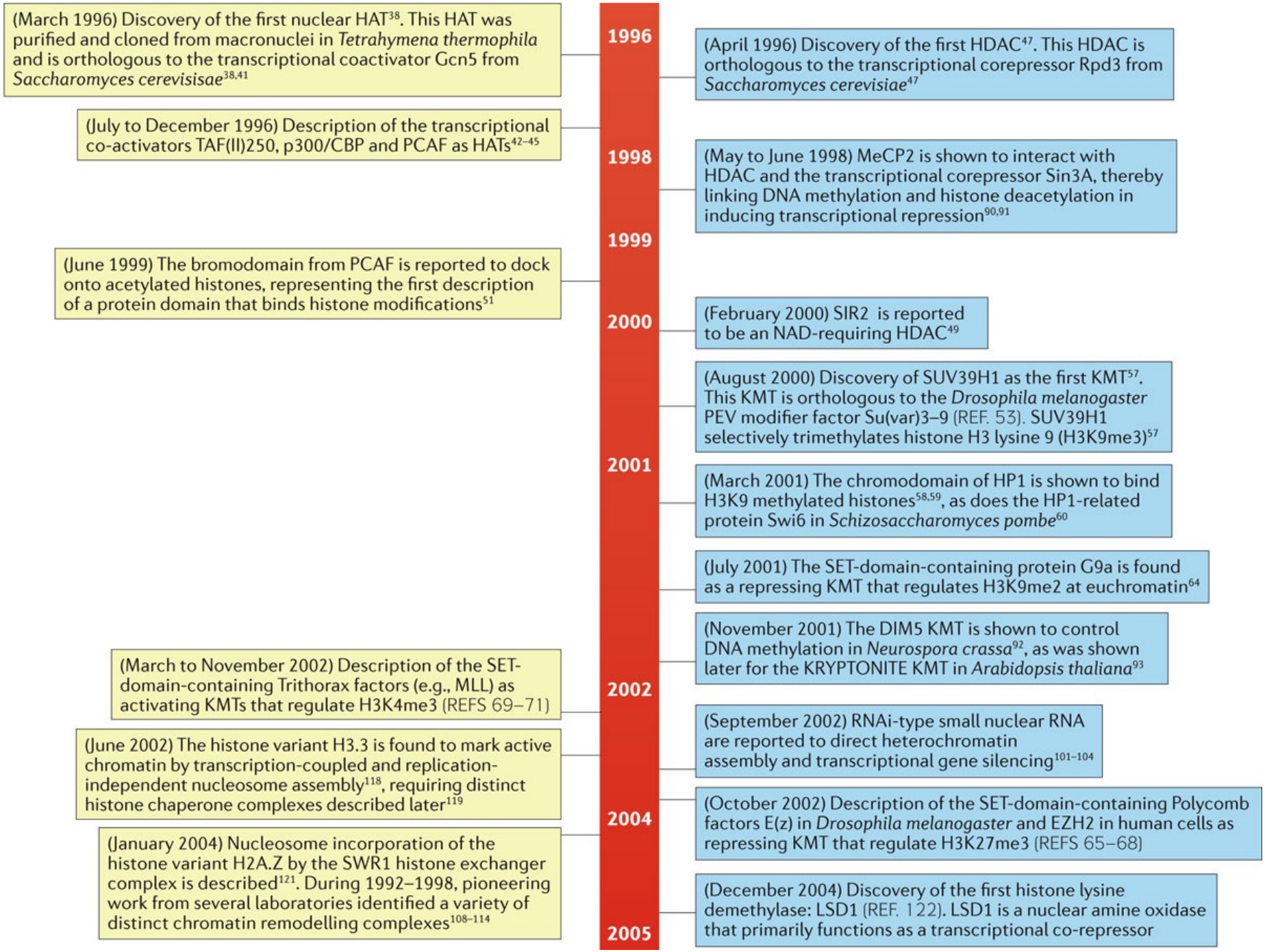
Miescher, Flemming, Kossel and Heitz defined nucleic acids, chromatin and histone proteins, which led to the cytological distinction between euchromatin and heterochromatin (from 1869 to 1928).

Epigenetic control

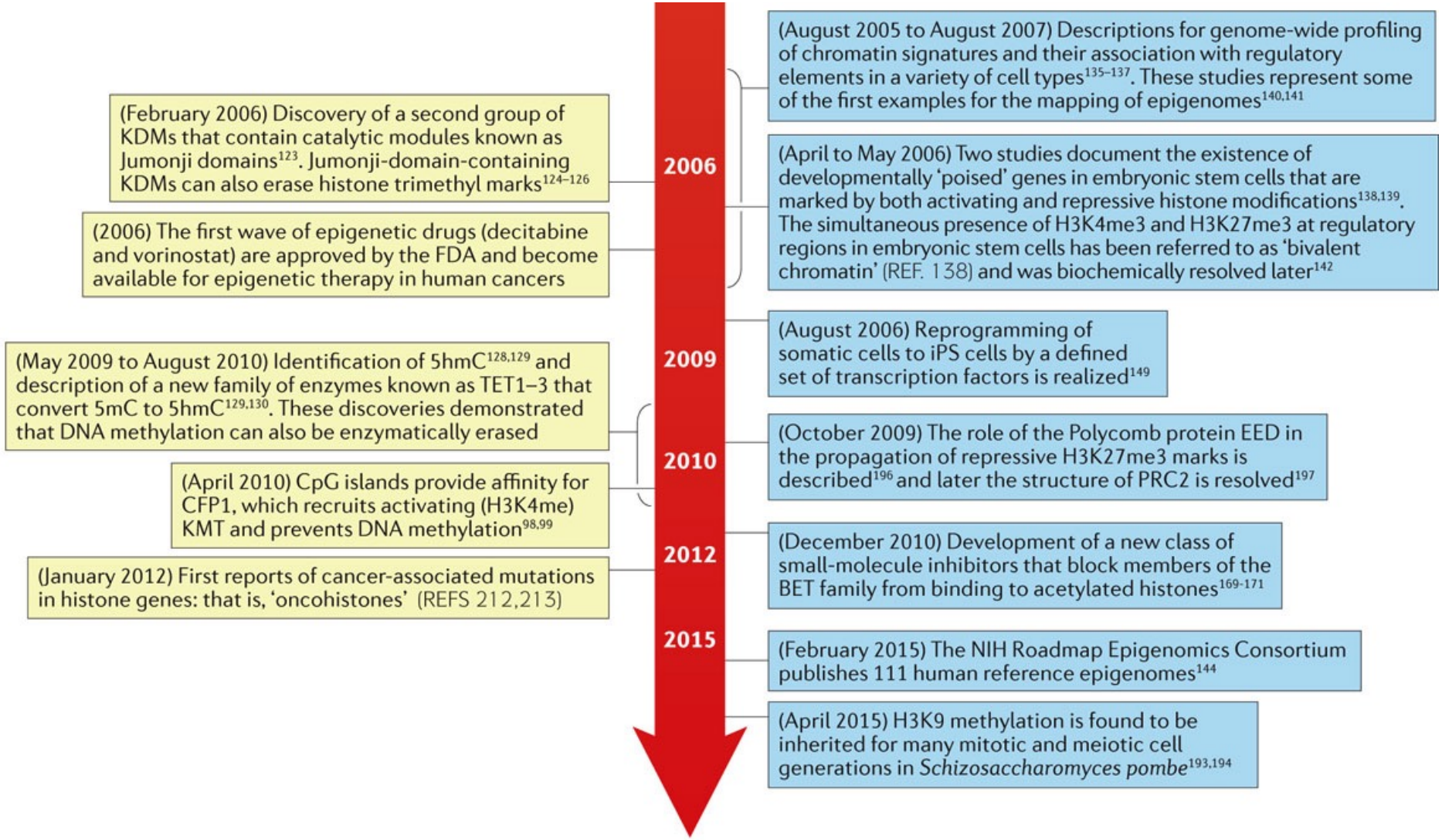
1942, Waddington coined the term 'epigenetics', which he defined as changes in phenotype without changes in genotype, to explain aspects of development for which there was little mechanistic understanding.

2016, epigenetic mechanisms transduce the inheritance of gene expression patterns without altering the underlying DNA sequence but by adapting chromatin, which is the physiological form of our genetic information.

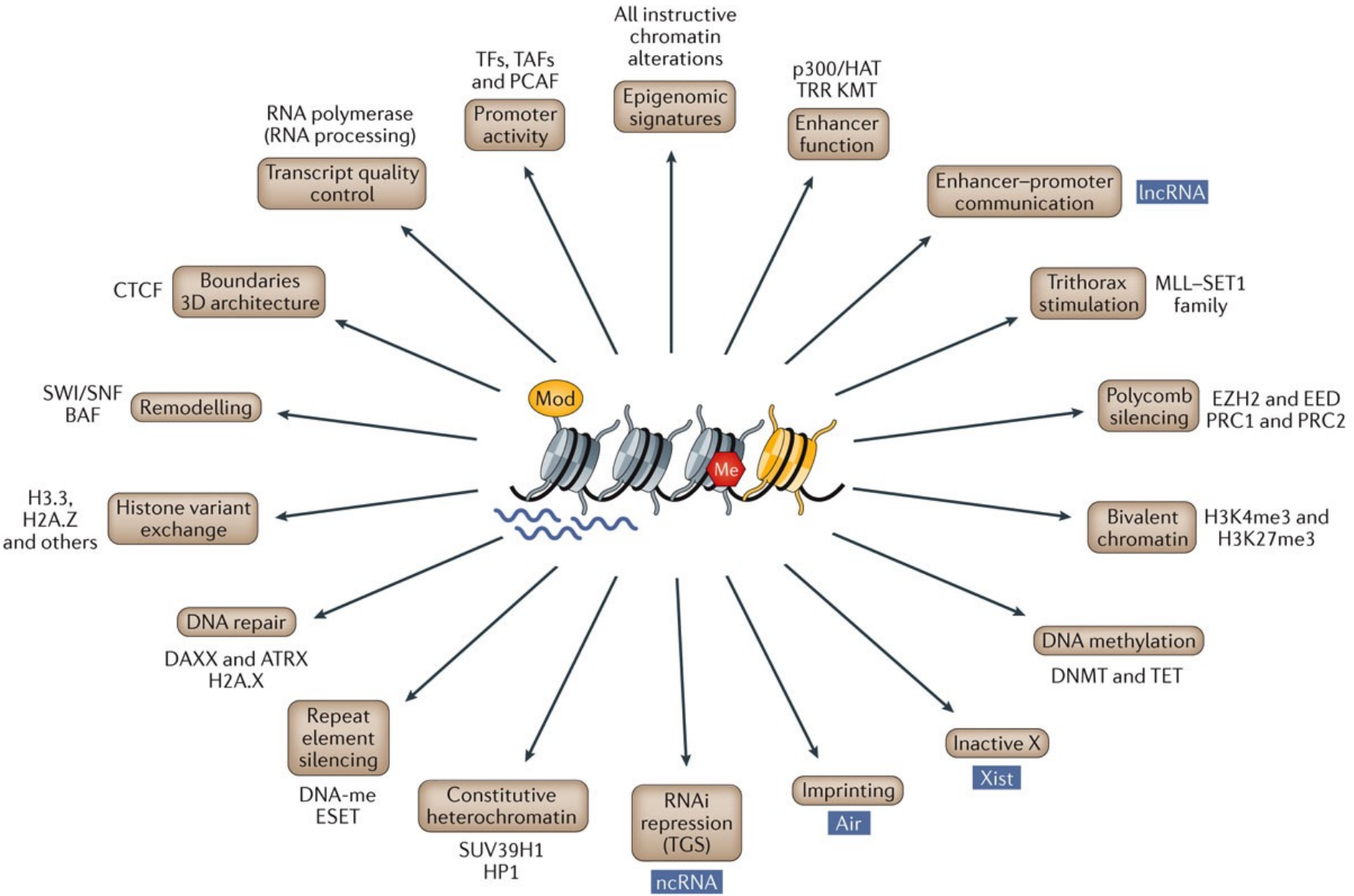
Timeline of major discoveries and advances in epigenetic research



Timeline of major discoveries and advances in epigenetic research



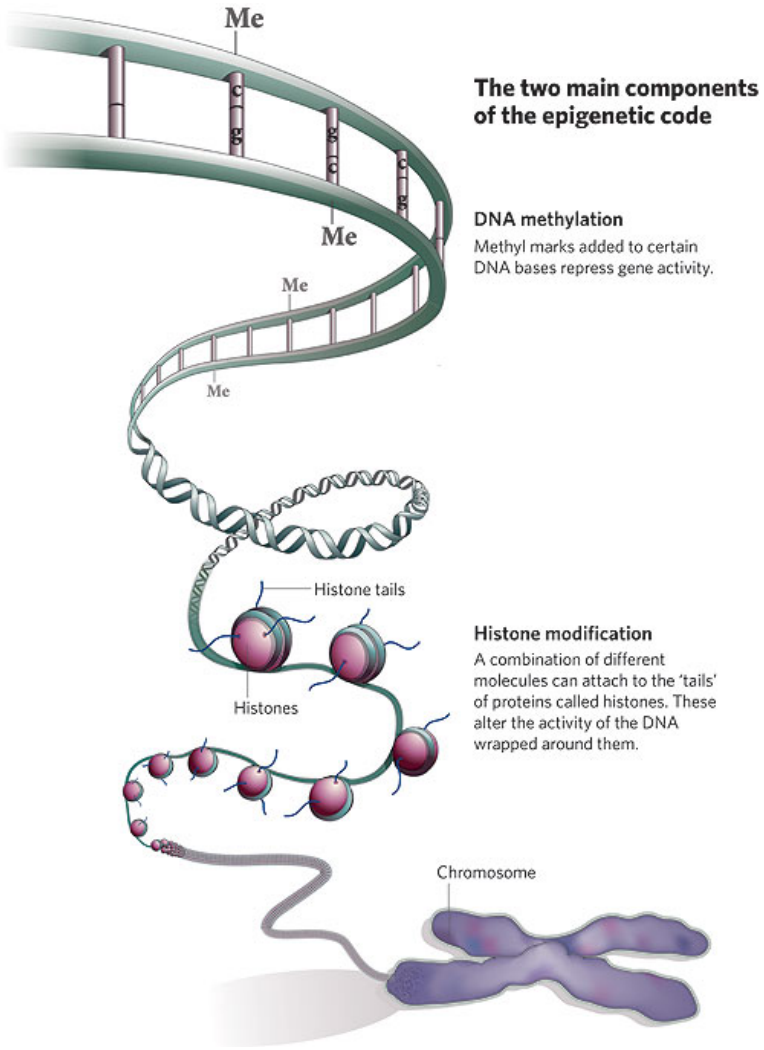
Key examples of chromatin contribution to epigenome function



Nature Reviews | Genetics

Molecular Mechanisms that regulate transcription

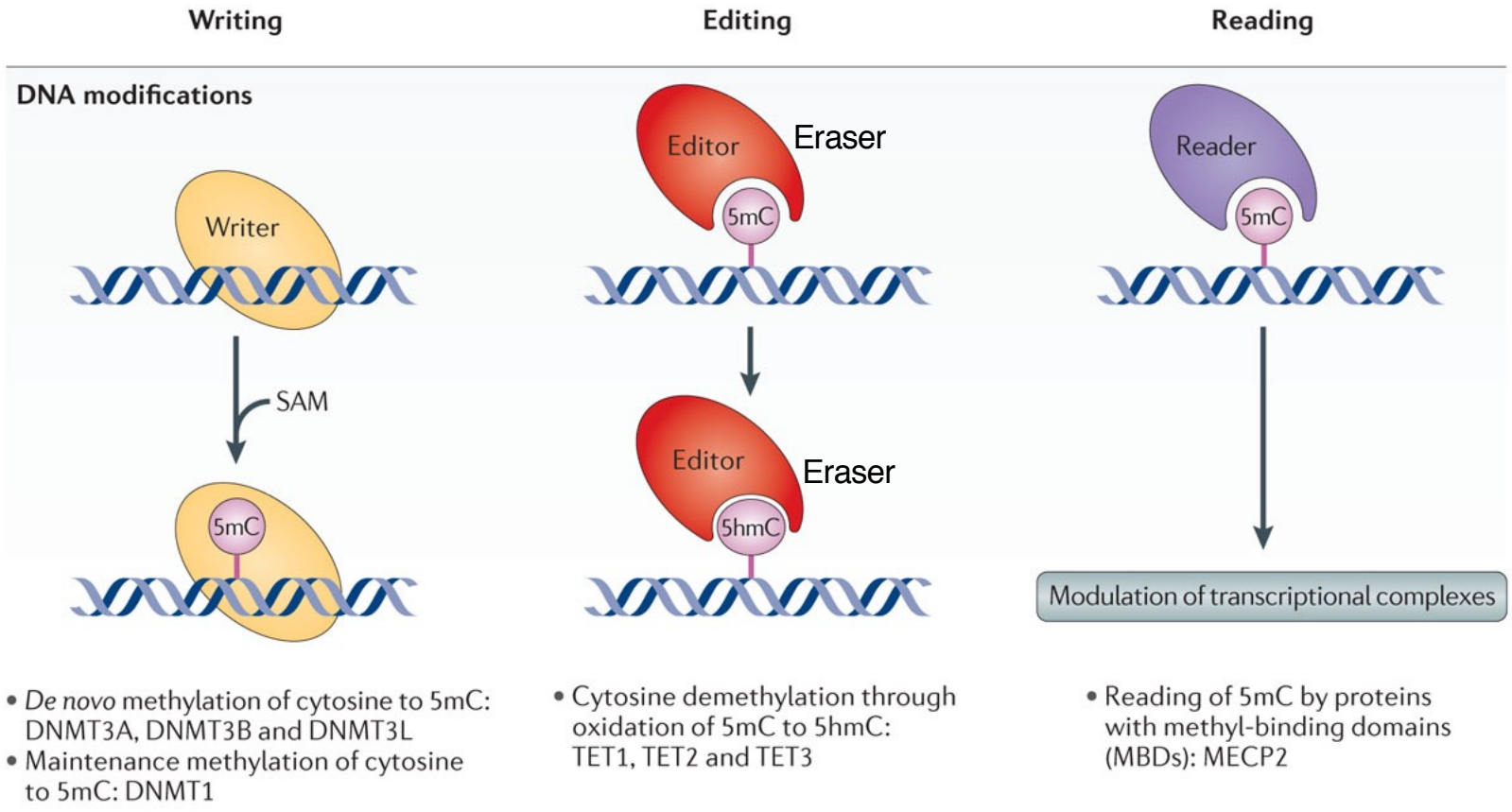
- DNA methylation
- Histone modifications



DNA methylation

DNA modifycation

DNA methylation recruited specific proteins that modulates transcription. Moreover, specific proteins (Erasers or Editor) can recognize this modification

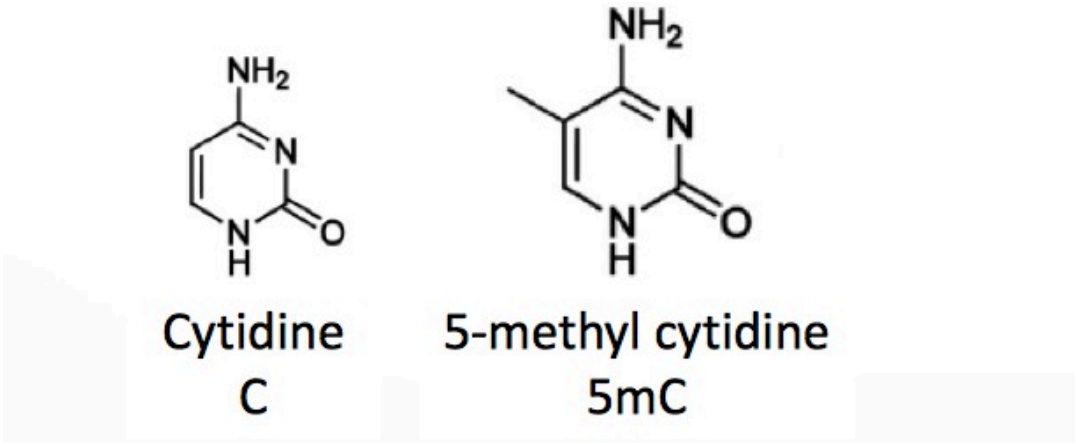


It is a frequent modification in plants (~14%) and mammals (~8%) , very rare in flies (~0.03%), not present in yeast and nematodes.

DNA methylation

- The predominant epigenetic modification of DNA in mammalian genomes is methylation of **cytosine nucleotides (5-MeC)**.

- The primary target sequence for DNA methylation in mammals is **5'- CpG-3' dinucleotides**.



Despite its ancient origins, DNA cytosine methylation has been lost in several eukaryotic lineages, including in many animals; common model organisms such as *Drosophila melanogaster*, *Caenorhabditis elegans*, fission yeasts and bakers' yeasts exhibit virtually no 5mC.

In fact, cytosine methylation comes at a cost: 5mC is inherently mutagenic because it can spontaneously undergo deamination, leading to C → T transitions. Thus, organisms with CpG methylation also have reduced CpG content. For example, **mammals have roughly 5-fold fewer CpG dinucleotides than expected from the nucleotide composition of their genome**.

Nevertheless, **mammalian genomes exhibit particularly high CpG methylation** levels; although there are some tissue-specific differences, **70–80% of CpGs are methylated**

The mammalian genome is generally CpG poor, with the exception of CpG islands (CGIs), which are relatively small genomic regions of roughly 1 kb on average.

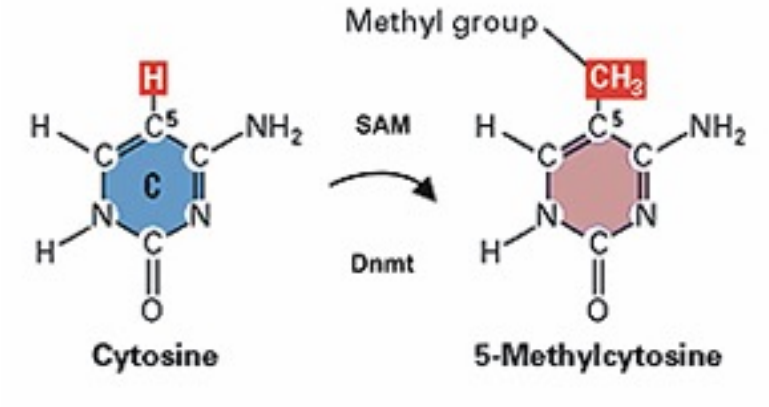
Over two-thirds of mammalian promoters are CGIs: virtually all housekeeping genes and tumor suppressor genes have CGI promoters, and so do several developmentally regulated genes.

CGIs are very rarely methylated; most inactive CGI promoters are silenced by Polycomb repressive complex 2-mediated H3K27 methylation, which is a more plastic mode of silencing than DNA methylation and therefore more amenable to gene (re)activation in response to developmental or environmental cues.

Nevertheless, there are three major classes of genes, in which stable, lifelong DNA-methylation-based silencing in somatic tissues is very important: **genes on the inactive X chromosome, imprinted genes, transposable elements and germline-specific genes.**

Roles of DNA methylation

Cytosine DNA methylation is a stable epigenetic mark that is crucial for diverse biological processes



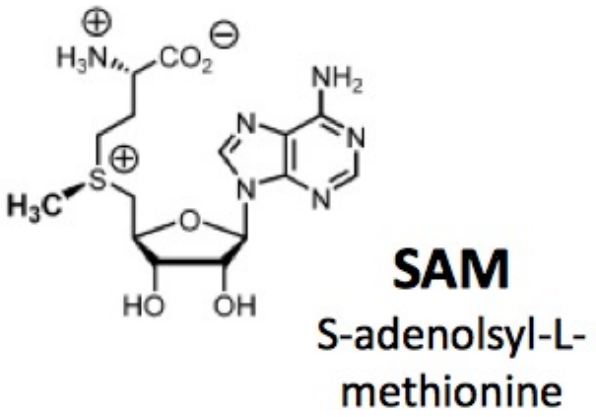
- Transcriptional silencing
- Protecting the genome from transposition
- Genomic imprinting
- X inactivation
- Tissue specific gene expression

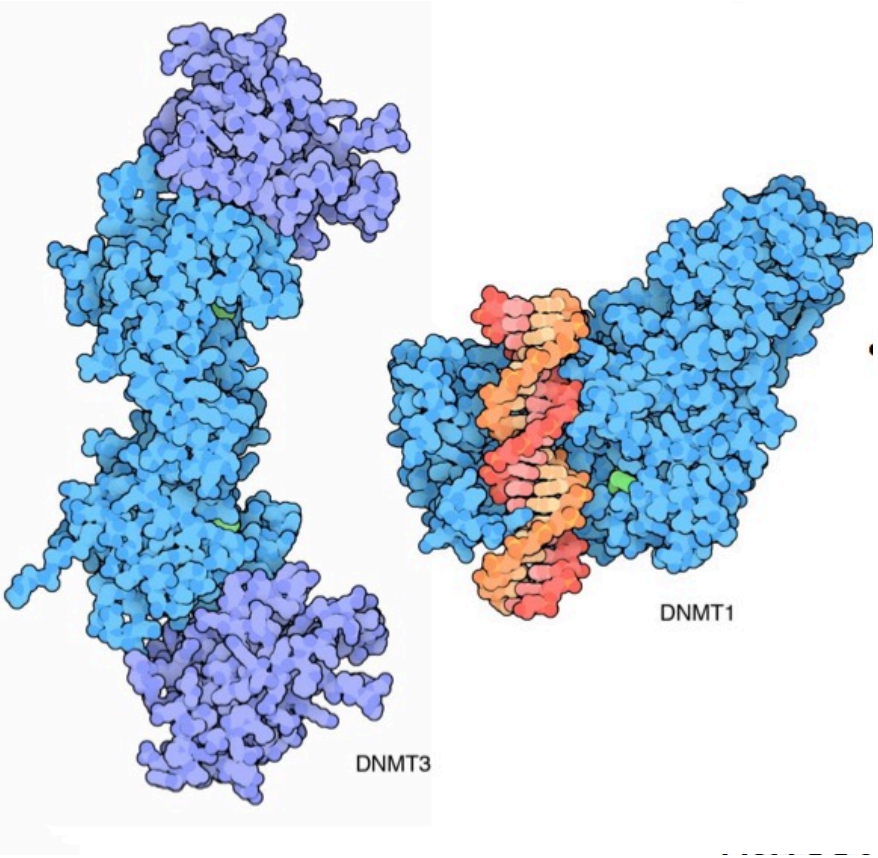
CpG Methyltransferases

DNMT1 -Maintenance methylation

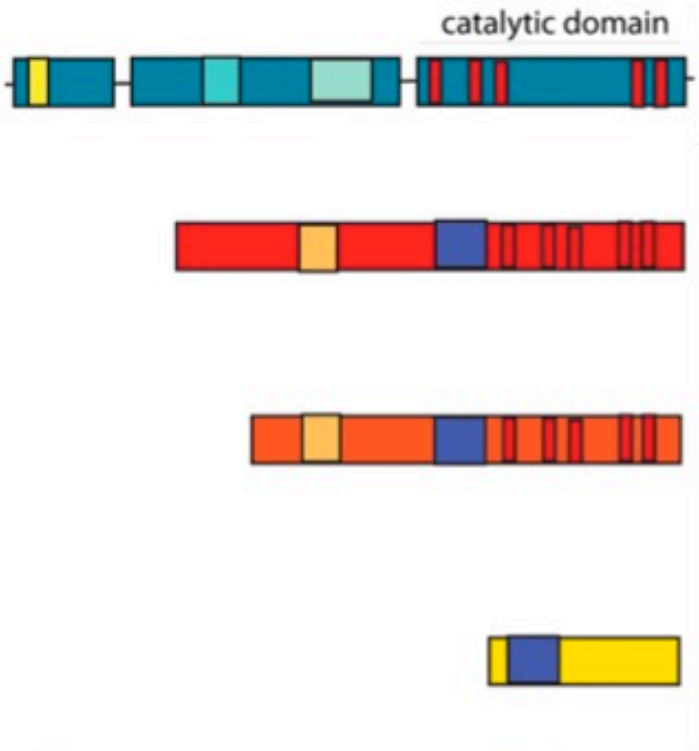
- Looks for hemimethylated CpG and maintains methylation pattern following replication

DNMT3 - De novo methylation





The DNMT family of enzymes



Family member	Function and effect on hematopoiesis with deletion
DNMT1	Maintenance DNA methyltransferase. Germline deletion embryonic lethal. Conditional deletion with *Mx*-Cre results in decreased HSC self-renewal and increased cycling of HSC's.
DNMT3A	*De novo* DNA methyltransferase. Germline deletion results in death by 4 weeks. Conditional deletion with Mx-Cre has minimal effects in primary mice or primarily transplanted mice. With serial transplantation however there is a huge increase in HSC's yet defect of differentiation .
DNMT3B	*De novo* DNA methyltransferase. Germline deletion embryonic lethal. Conditional deletion with *Mx*-Cre has minimal effects but Dnmt3a/Dnmt3b double knockout mice appear to have more pronounced effect on HSC's with some conflict in current reports.
DNMT3L	Catalytically inactive but stimulates enzymatic activity of DNTM3A. Germline knockout mice are viable but sterile. Hematopoietic phenotype of Dnmt3L knockout mice not reported.

Leucine zipper
Zinc finger
Polybromo domain
PWWP domains
PHD-like domains

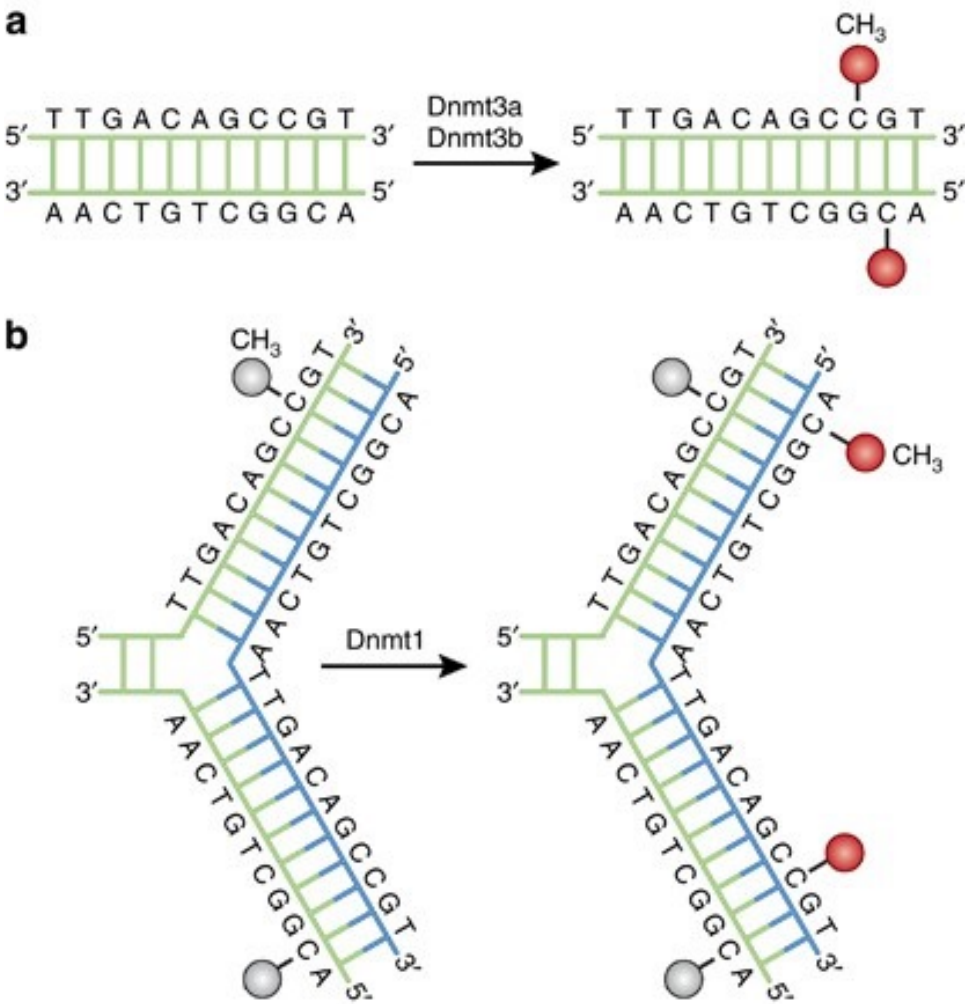
Although Dnmt3a is expressed relatively ubiquitously, Dnmt3b is poorly expressed by the majority of differentiated tissues with the exception of the thyroid, testes, and bone marrow.
Dnmt3L is mainly expressed in early development and is restricted to the germ cells and thymus in adulthood

Maintenance of DNA Methylation in Mammals

Dnmt1, which is highly expressed in mammalian tissues including the brain. Dnmt1 preferentially methylates hemimethylated DNA.

During DNA replication, Dnmt1 localizes to the replication fork where newly synthesized hemimethylated DNA is formed. Dnmt1 binds to the newly synthesized DNA and methylates it to precisely mimic the original methylation pattern present before DNA replication.
Additionally, Dnmt1 also has the ability to repair DNA methylation.

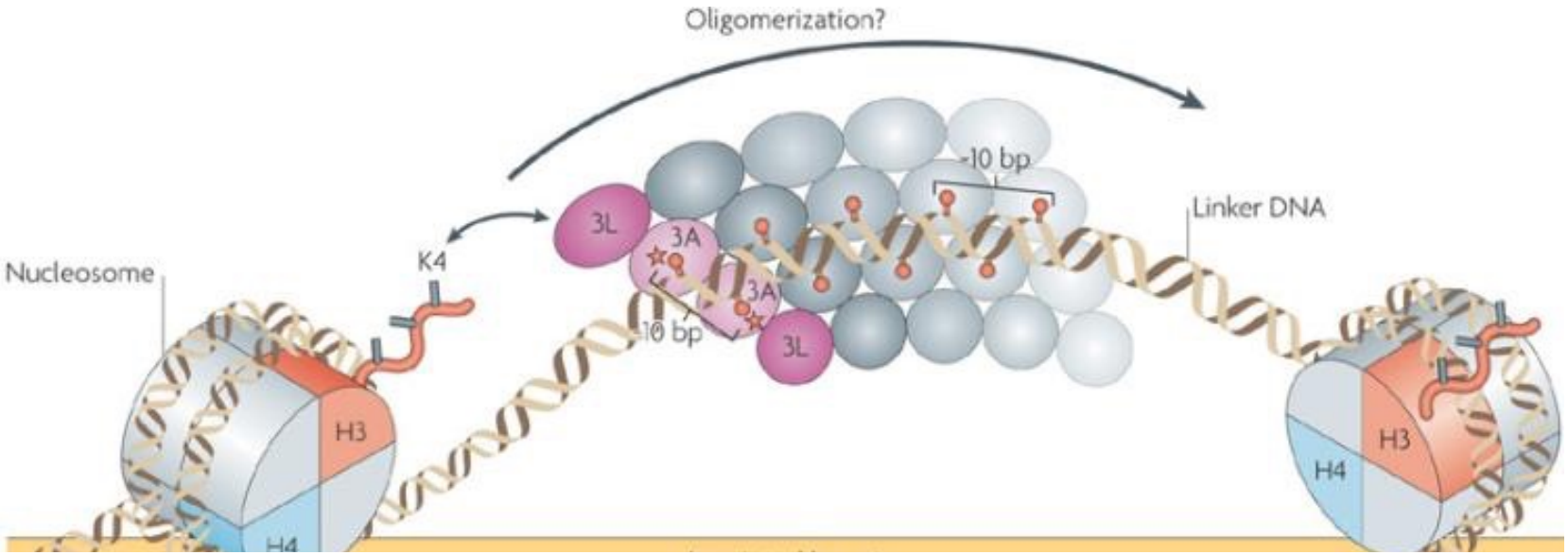
For this reason, **Dnmt1 is called the maintenance Dnmt** because it maintains the original pattern of DNA methylation in a cell lineage.



De Novo DNA Methylation in Mammals

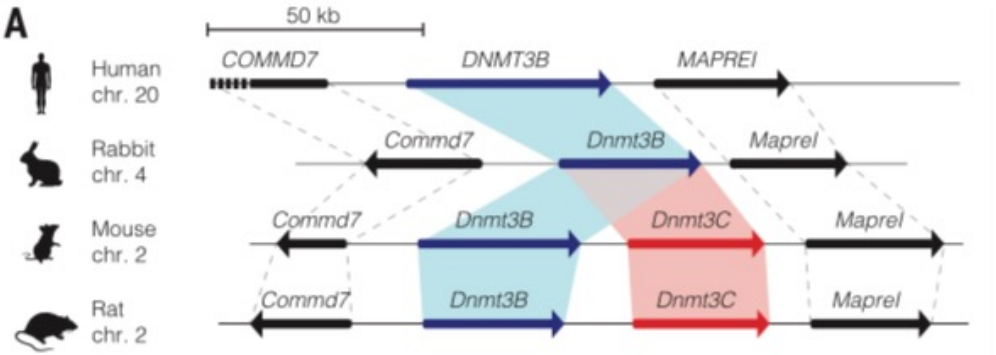
In mammals, DNA methylation patterns are established by the **DNA methyltransferase 3 (DNMT3)** family of *de novo* methyltransferases and maintained by the **maintenance methyltransferase DNMT1**. DNMT3L interacts with unmethylated H3K4 and recruits DNMT3A resulting in the formation of a tetrameric complex.

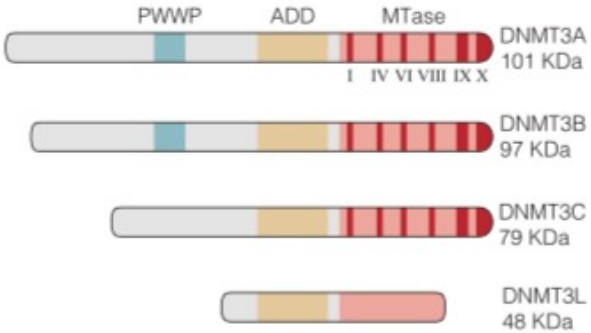
In addition to interactions with unmethylated H3K4 tails, other mechanisms for targeting DNA methylation to specific loci are interactions of DNMT3 with the **histone methyltransferases** G9a (H3K9me), Ezh2 (PRC2 component, H3K27me), suppressor of variegation 3-9 homologue 1 (SUV39H1) and SET domain bifurcated 1 (SETDB1) and **recruitment by lncRNAs**.



The DNMT3C methyltransferase protects male germ cells from transposon activity

Dnmt3C is a *de novo* DNA methyltransferase gene that evolved via a duplication of Dnmt3B in rodent genomes and was previously annotated as a pseudogene. DNMT3C is the enzyme responsible for methylating the promoters of evolutionarily young retrotransposons in the male germ line and this specialized activity is required for mouse fertility.





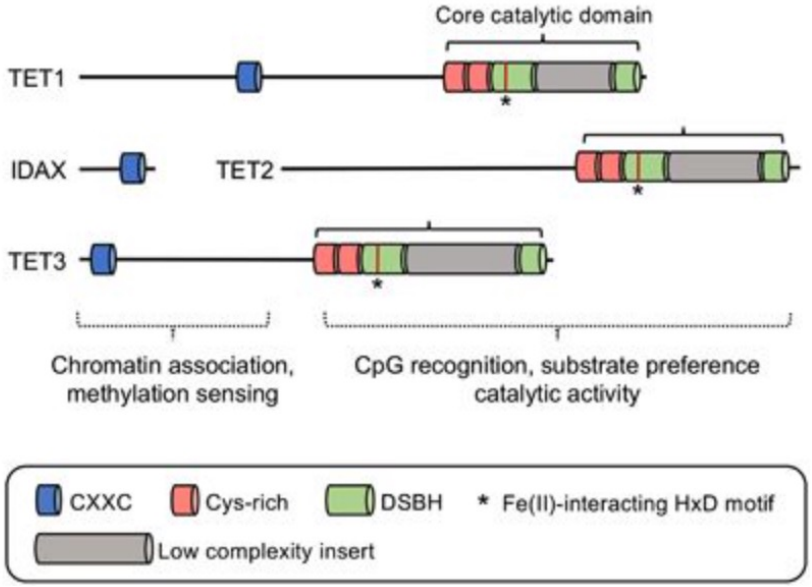
Barau et al., 2016 Nature

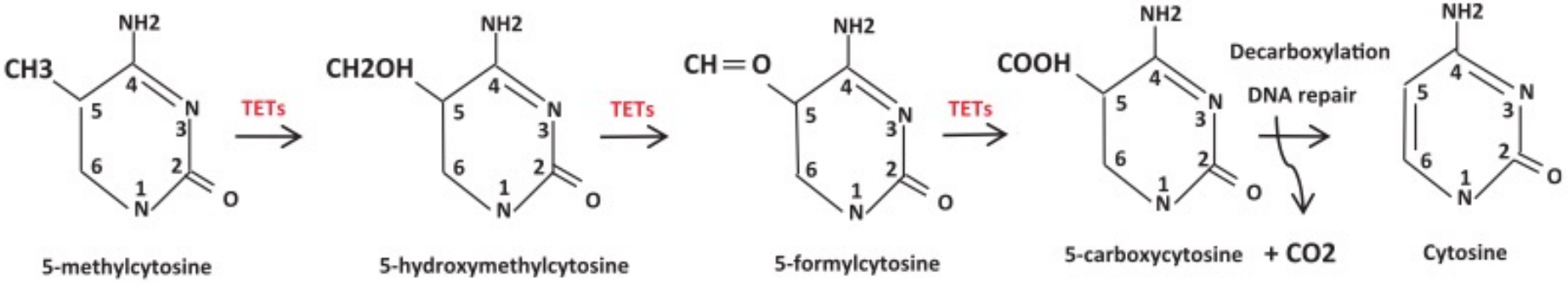
Removing the 5-methylcytosine (5mC)

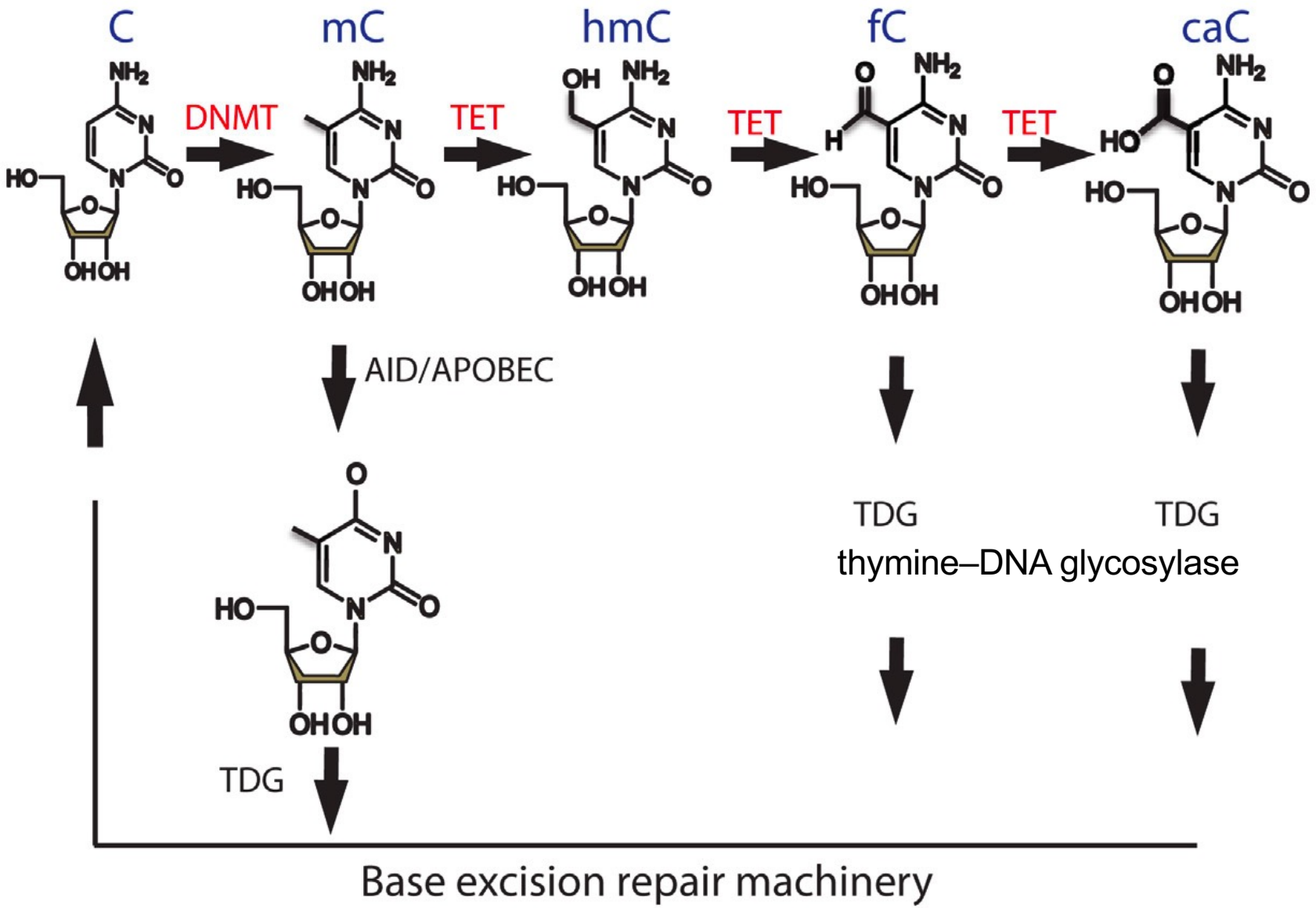
There is no known mechanism in mammalian cells that can cleave the strong covalent carbon-to-carbon bond that connects cytosine to a methyl group

Active demethylation of 5mC to cytosine is mediated through a series of oxidative steps mediated by the **ten-eleven-translocation (TET)** enzymes, which require **oxygen, iron and α-ketoglutarate**.

5mC is converted to **5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC)** and **5-carboxylcytosine (5-caC)**, which is then either decarboxylated to cytosine or recognized by thymine DNA glycosylase as part of the mismatch repair system (**BER**).



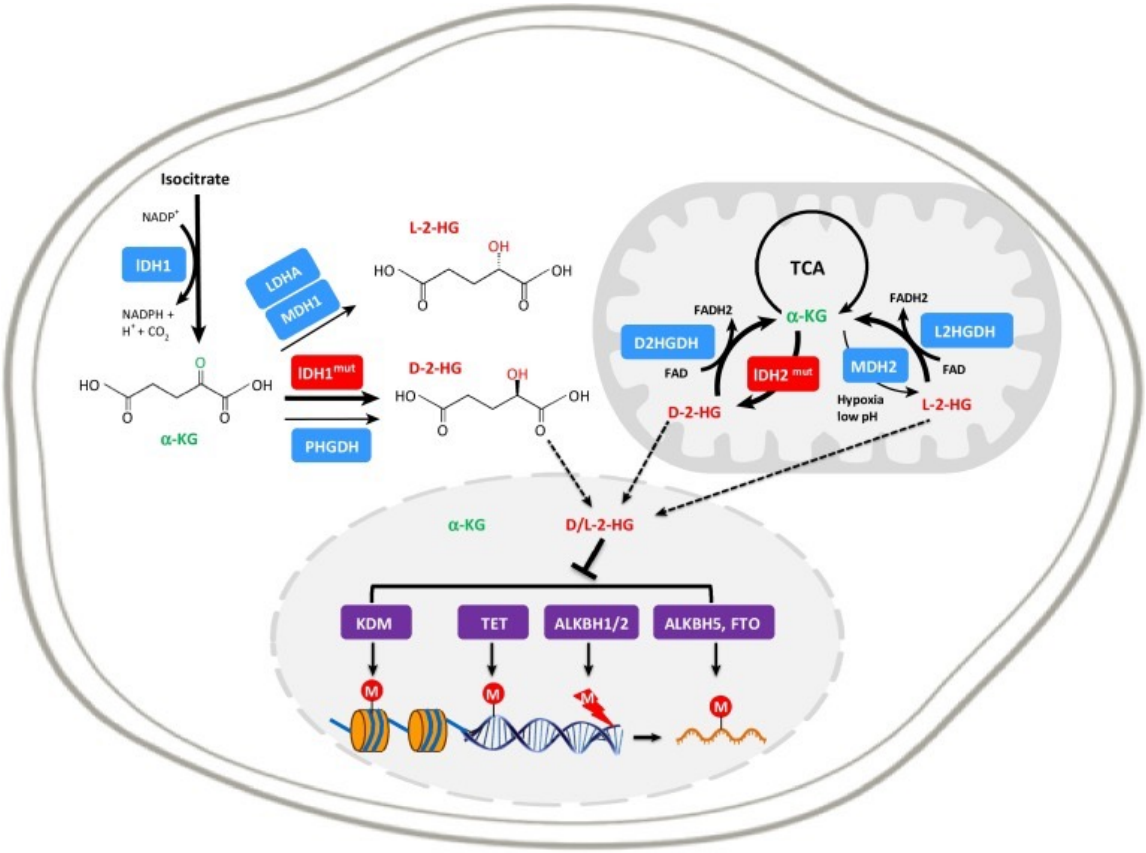


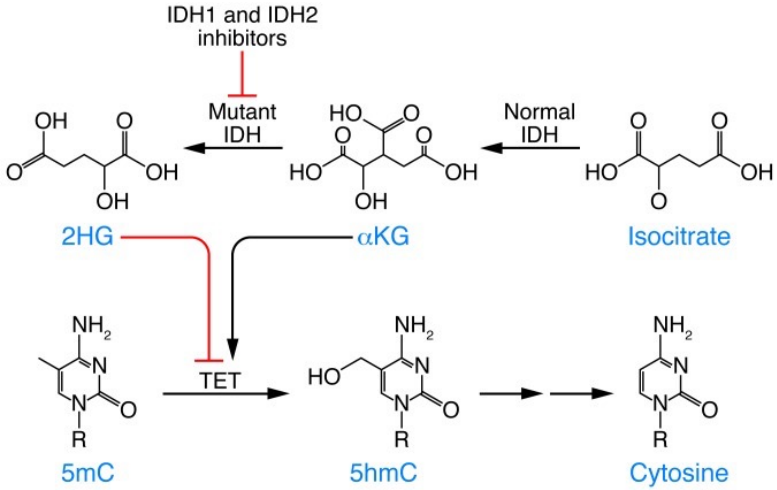


DNA is demethylated by stepwise oxidation (TET) or deamination (mediated by AID/APOBEC) or other pathways.

IDH1 and IDH2 mutations lead to epigenetic alteration in cancer

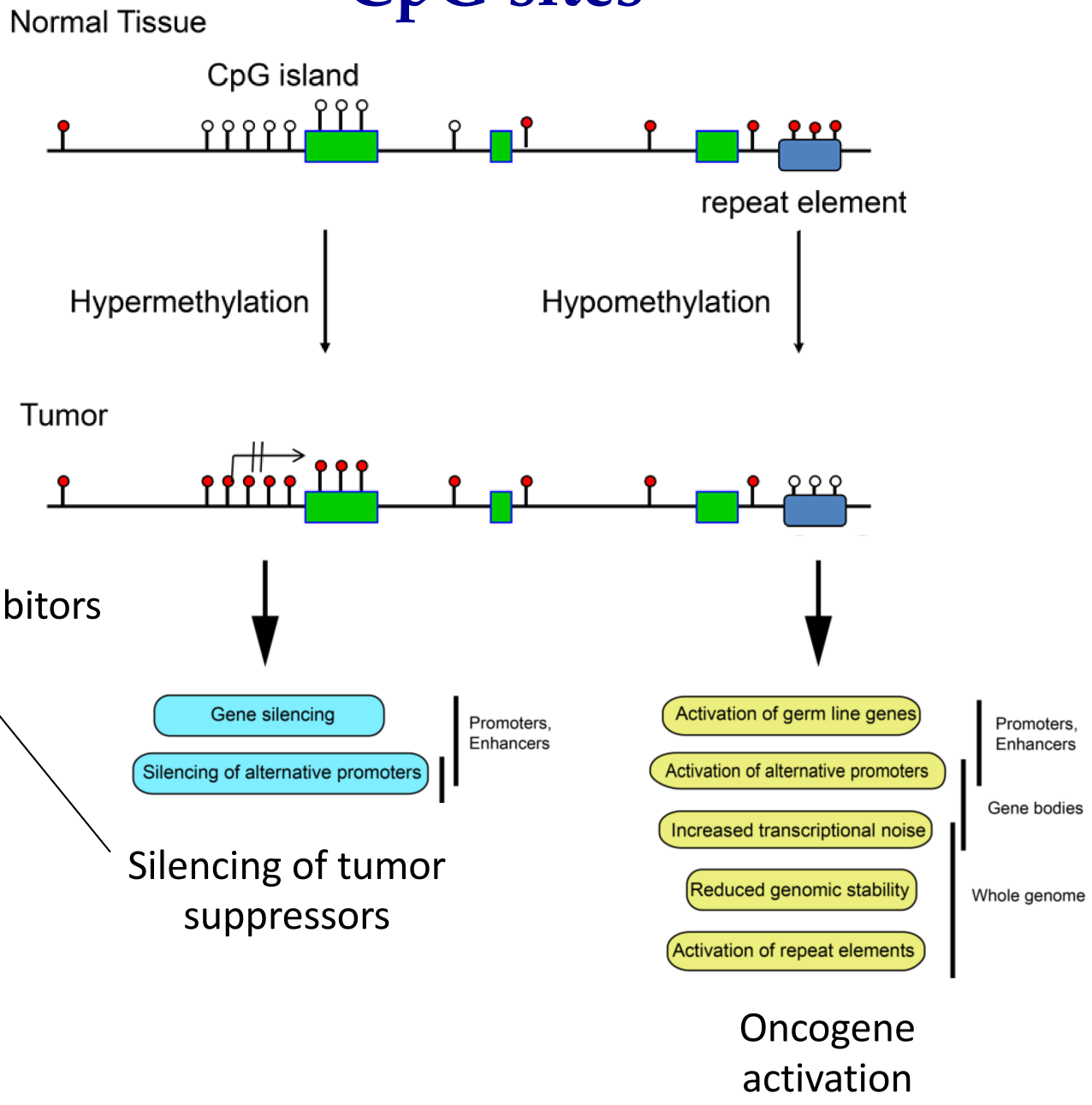
IDH1 and *IDH2* are the most frequently mutated metabolic genes in human cancer, resulting in neomorphic enzymes that convert α-ketoglutarate (α-KG) to 2-hydroxyglutarate (2-HG). 2-HG acts as an antagonist of α-KG to competitively inhibit the activity of α-KG-dependent dioxygenases, including those involved in histone and DNA demethylation.





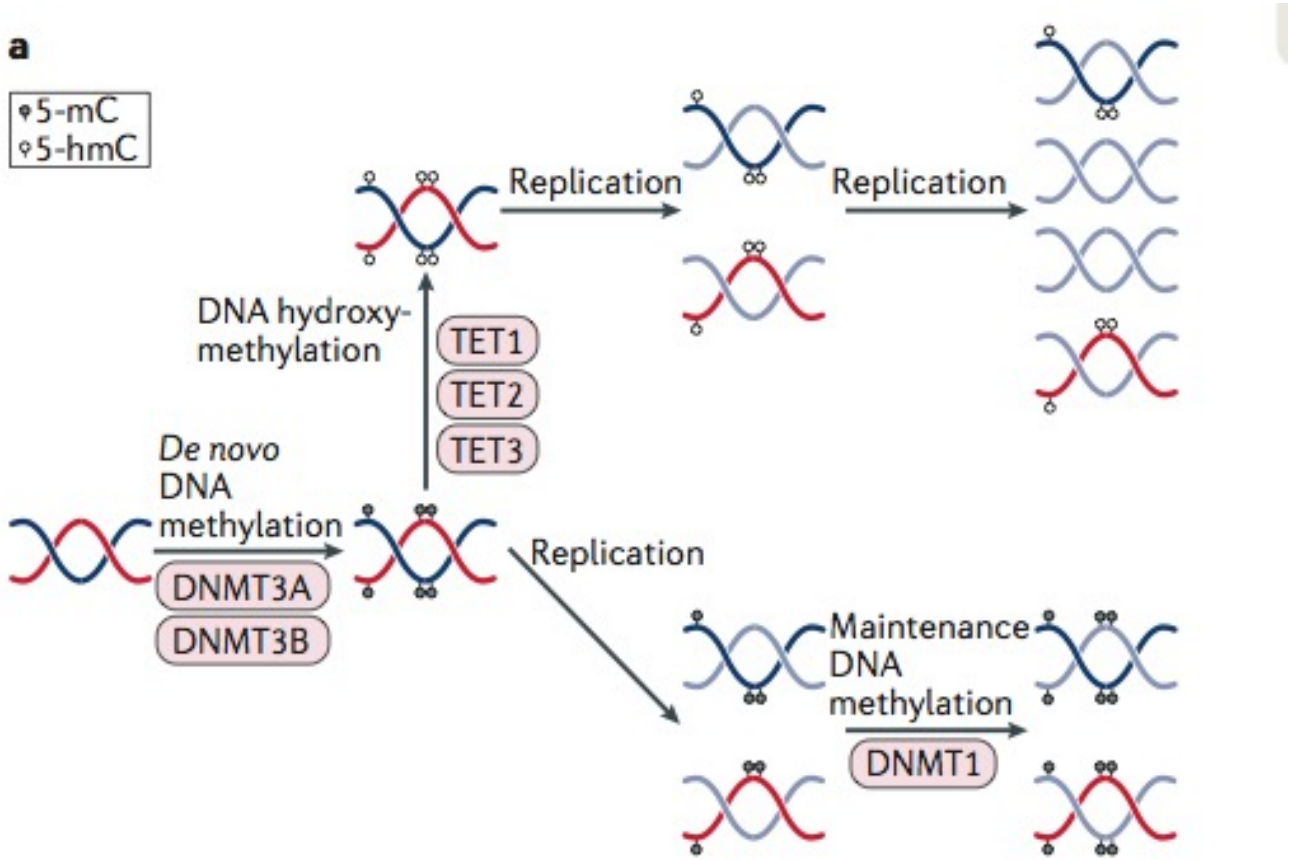
Acute myeloid leukemia

Carcinogenesis alters the methylation profile of CpG sites



Methylation patterns are reproduced at each round of cell division

DNMT3A and DNMT3B are *de novo* methyltransferases that mark unmethylated DNA, whereas DNMT1 maintains DNA methylation after DNA replication, 5-methylcytosine (5-mC) is oxidized to 5-hydroxymethylcytosine (5-hmC) by enzymes of the TET family (TET1, TET2 and TET3); 5-hmC is not recognized by DNMT1, which results in loss of DNA methylation during replication.



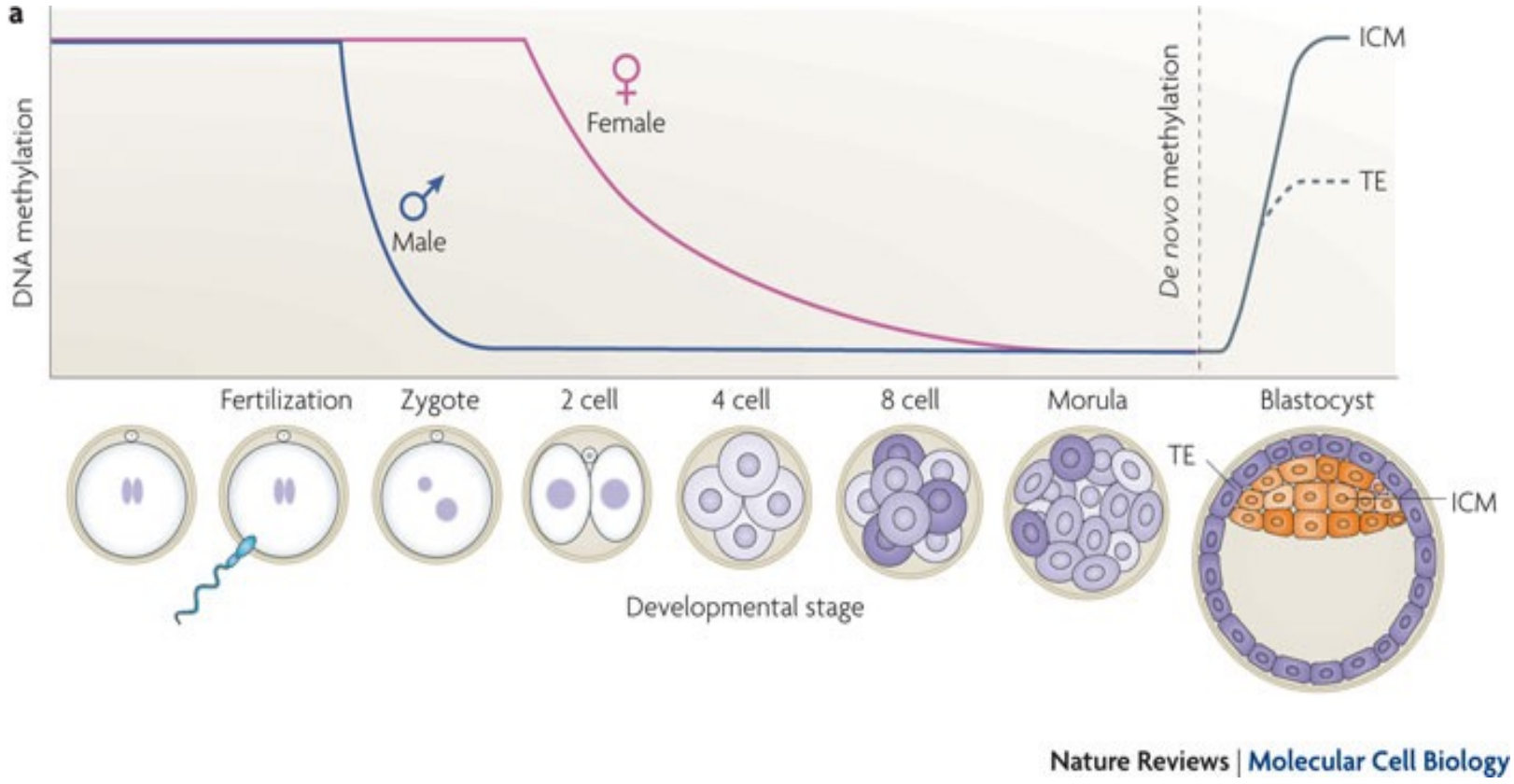
Readers of DNA Methylation

Whereas DNA methylation may itself reduce gene expression by impairing the binding of transcriptional activators, a second class of proteins with a high affinity for 5mC inhibits transcription factor binding.

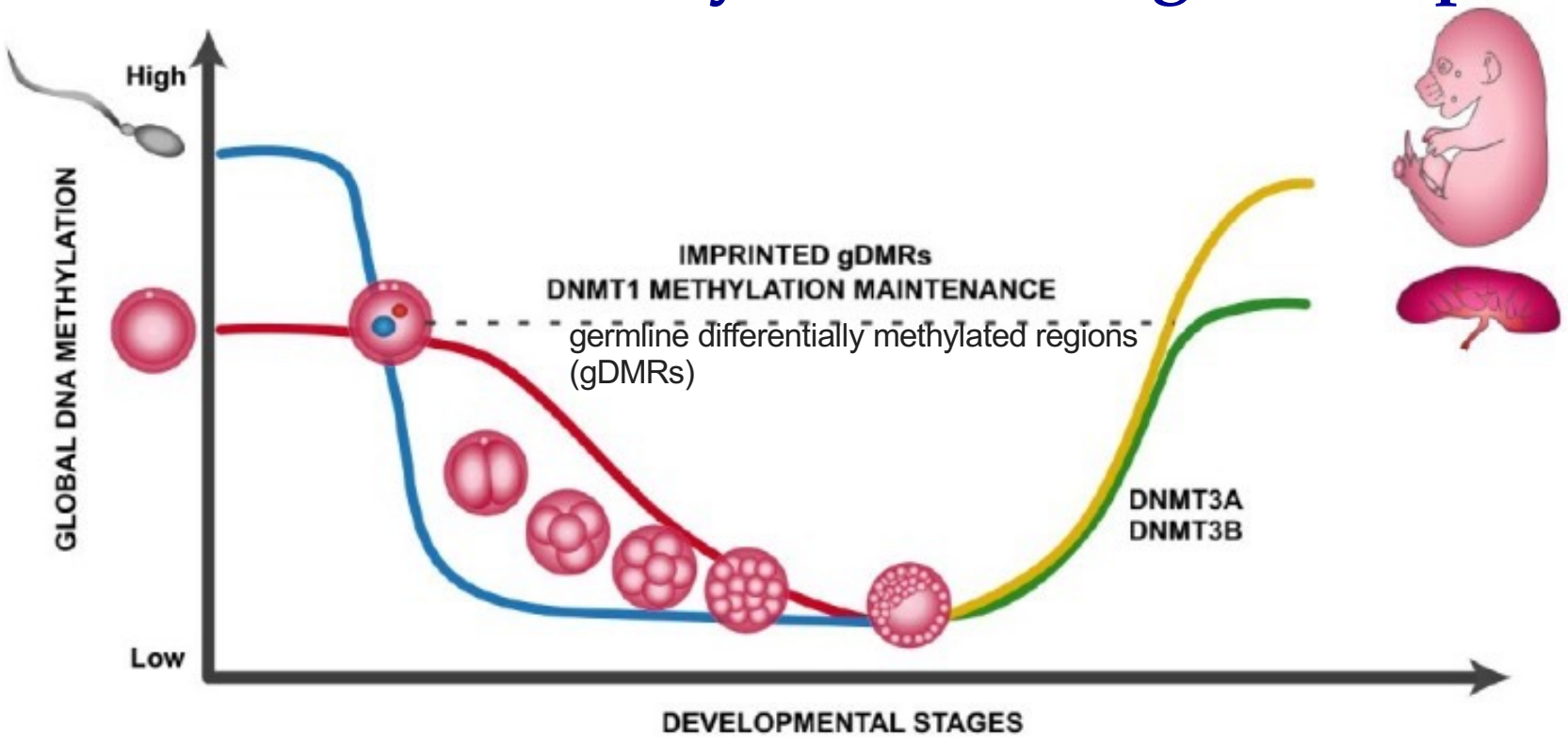
DNA methylation is recognized by three separate families of proteins:
1- **the MBD proteins** that contain a conserved methyl-CpG-binding domain (MBD) confering them a higher affinity for single methylated CpG sites. This family includes MeCP2, the first identified methyl-binding protein, along with MBD1-4. **MBDs are more highly expressed in the brain than in any other tissue**, and many MBDs are important for normal neuronal development and function
2- **The UHRF** (ubiquitin-like, containing PHD and RING finger domain) proteins, including UHRF1 and UHRF2, are multidomain proteins that bind to methylated cytosines. The UHRF protein family first binds to Dnmt1 and then targets it to hemimethylated DNA in order to maintain DNA methylation, especially during DNA replication
3- **the zinc-finger proteins:** Kaiso (binds to consecutive methylated CpG sites), ZBTB4 and ZBTB38 (bind to single methylated CpG site). Similar to the MBD family, repress transcription in a DNA methylation-dependent manner.

Dynamics of DNA methylation during development

Shortly after a sperm fertilizes an egg, the paternal genome rapidly undergoes genome-wide active DNA demethylation and remains demethylated following multiple rounds of cell division. During this time, the maternal genome experiences gradual, passive demethylation. *De novo* methylation patterns are established by the DNA methyltransferases DNMT3A and DNMT3B during the development of the blastocyst.

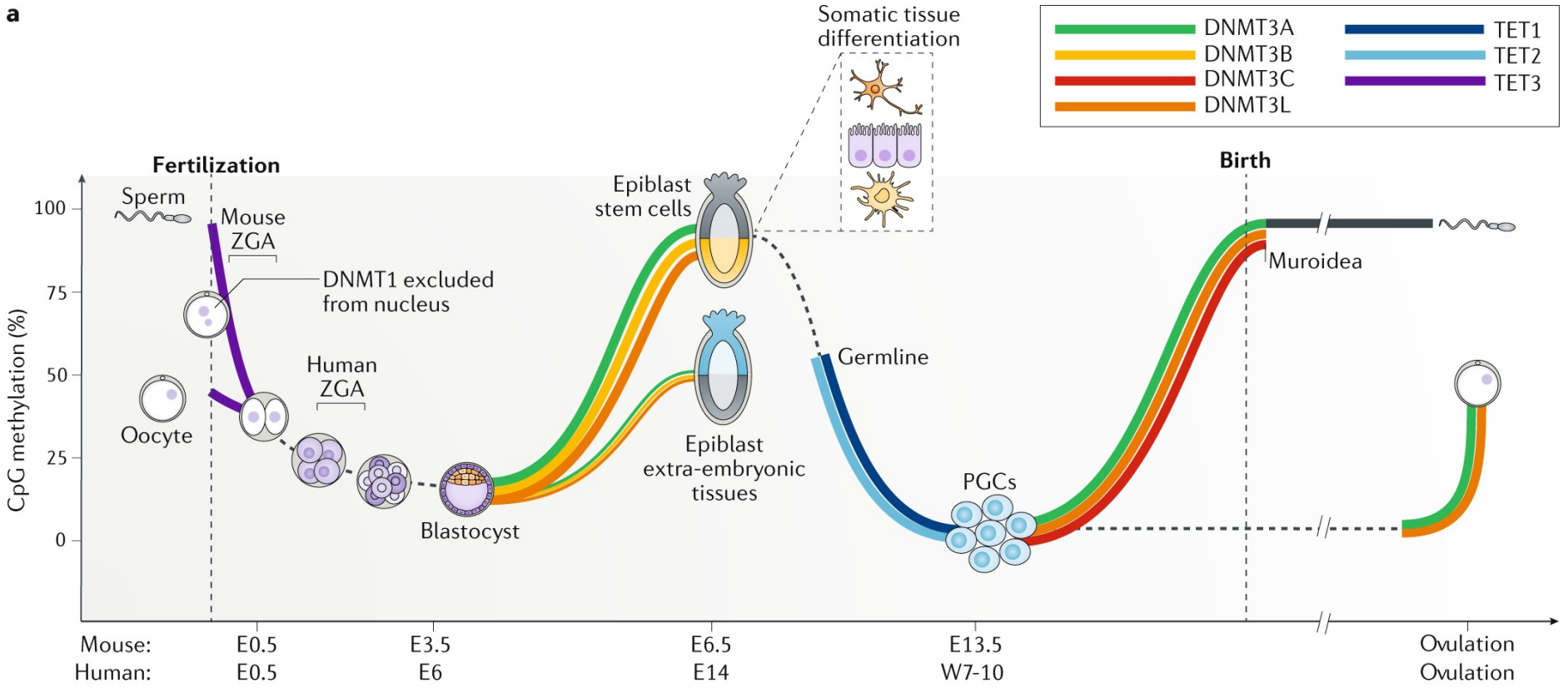


Dynamics of DNA methylation during development



Global DNA demethylation and remethylation during the epigenetic reprogramming of early embryogenesis in mice. Soon after fertilization, the zygotic paternal and maternal pronuclei undergo global demethylation during the pre-implantation stages, except for gDMRs which are maintained via DNMT1 activity. The paternal genome (blue) is initially actively demethylated by the TET3 enzyme followed by passive demethylation, whereas the maternal genome (red) demethylation is solely passive due to DNMT1 exclusion from the nucleus, hence the sharper demethylation slope for the paternal curve. After implantation, the blastocyst acquires *de novo* methylation patterns catalyzed by DNMT3A and DNMT3B to establish the embryonic and placental programs imperative for development initiation.

Dynamics of DNA methylation during development



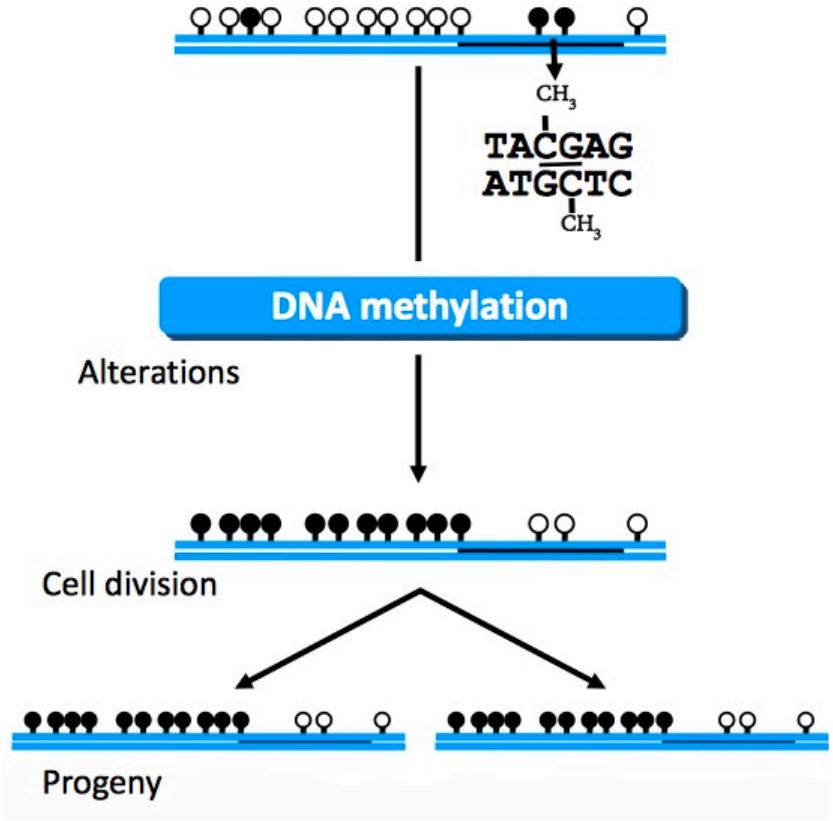
There are two waves of global DNA demethylation during embryogenesis. The first wave resets the epigenome for pluripotency at the blastocyst stage, with a steady decrease in DNA methylation from the zygote stage to the morula/blastocyst stage. The second wave of DNA demethylation occurs in the primordial germ cells *(PGCs)*. The DNA methylation is wide erased to restore germ line potency.

CpG Island

- **Definition**
  - ✓ Small stretches of about 300-3000bp
  - ✓ >50% GC content, regulatory regions

- **Genomic distribution**
  - ✓ ~70% of promoter regions contain CpG islands
  - ✓ Only 1% of remaining genome contains CpG islands
  - ✓ Frequently methylated in tumor cells

- **Methylation status**
  - ✓ Generally non-methylated in promoter of housekeeping and tumor suppessor genes
  - ✓ Typically methylated in transposable elements
  - ✓ Methylated in gene bodies influencing transcription and splicing

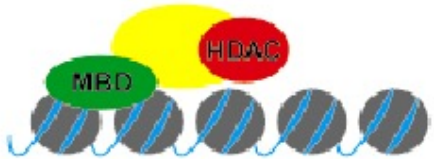


Methylated CpG Islands Inhibit Transcription

- Active promoters are usually unmethylated
- **Methylated DNA** recruits **methyl-CpG-binding domain proteins (MBD or MECP)** which may inhibit transcription by several mechanisms
- **Unmethylated CpG** islands recruit **Cfp1** which associates with a histone methyltransferase creating **H3K4me3 (transcriptional activation)**

Mechanisms for Repression Mediated by MBDs

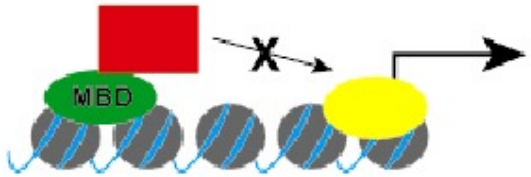
Model 1 - MBD proteins interact with HDAC to generate hypoacetylated, condensed chromatin



Model 2 - MBD proteins coat methylated loci occluding regulatory DNA

Model 3 - MBD proteins alter local DNA and or chromatin architecture

Model 4 - an MBD protein sequesters an essential transcription factor, preventing its function



nature genetics • volume 23 • october 1999

Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2

Ruthie E. Amir[1], Ignatia B. Van den Veyver[2,3], Mimi Wan[5], Charles Q. Tran[3], Uta Francke[5,6] & Huda Y. Zoghbi[1,2,4]

Rett syndrome1 (RTT) is a progressive neurodevelopmental disorder and one of the most common causes of mental retardation in females, with an incidence of 1 in 10,000–15,000. Patients with classic RTT appear to develop normally until 6–18 months of age, then gradually lose speech and purposeful hand use, and develop micro- cephaly, seizures, autism, ataxia, intermittent hyperventilation and stereotypic hand movements. After initial regression, the condition stabilizes and patients usually survive into adulthood. As RTT occurs almost exclusively in females, it has been proposed that RTT is caused by an X-linked dominant mutation with lethality in hemizygous males. Previous exclusion mapping studies using RTT families mapped the locus to Xq28. Using a systematic gene screening approach, we have identified mutations in the gene (*MECP2*) encoding X- linked methyl-CpG-binding protein 2 (MeCP2) as the cause of some cases of RTT. MeCP2 selectively binds CpG dinucleotides in the mammalian genome and mediates transcriptional repression through interaction with histone deacetylase and the corepressor SIN3A . In 5 of 21 sporadic patients, we found 3 *de novo* missense mutations in the region encoding the highly conserved methyl-binding domain (MBD) as well as a *de novo* frameshift and a *de novo* nonsense mutation, both of which disrupt the transcription repression domain (TRD). In two affected half-sisters of a RTT family, we found segregation of an additional missense mutation not detected in their obligate carrier mother. This suggests that the mother is a germline mosaic for this mutation. Our study reports the first disease-causing mutations in RTT and points to abnormal epigenetic regulation as the mechanism underlying the pathogenesis of RTT.

MeCP2 and Rett syndrome

Rett syndrome is a neurodevelopmental disorder linked to the X chromosome:

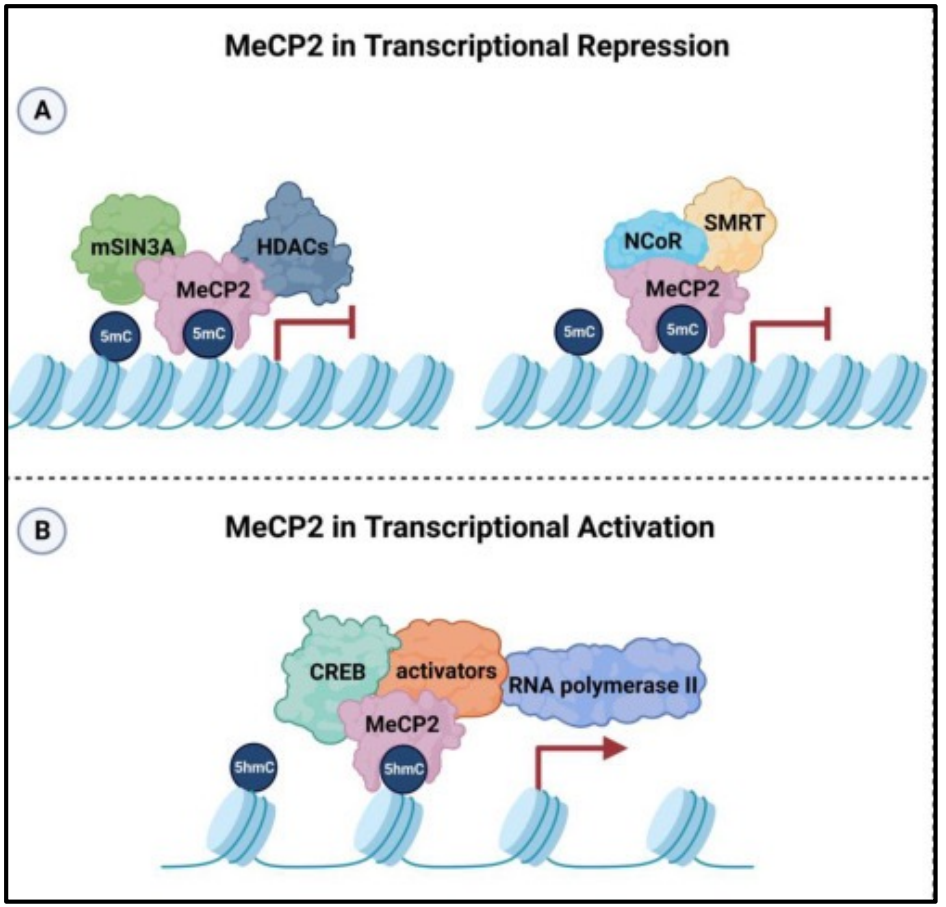
- Autistic traits
- Loss of acquired language
- Loss of motor skills

95% of Rett patients carries mutations on MECP2 gene

pre-genomic era:
MeCP2 represses specific genes in the brain

post-genomic era:
-60% MeCP2 binding sites fall outside the genes and only 6% in the CpG islands in the promoters.
-The modulation of gene expression mediated by MeCP2 is repressive for 377 genes while is activating for 2184!!

Mode of action of MeCP2



Studying DNA Methylation

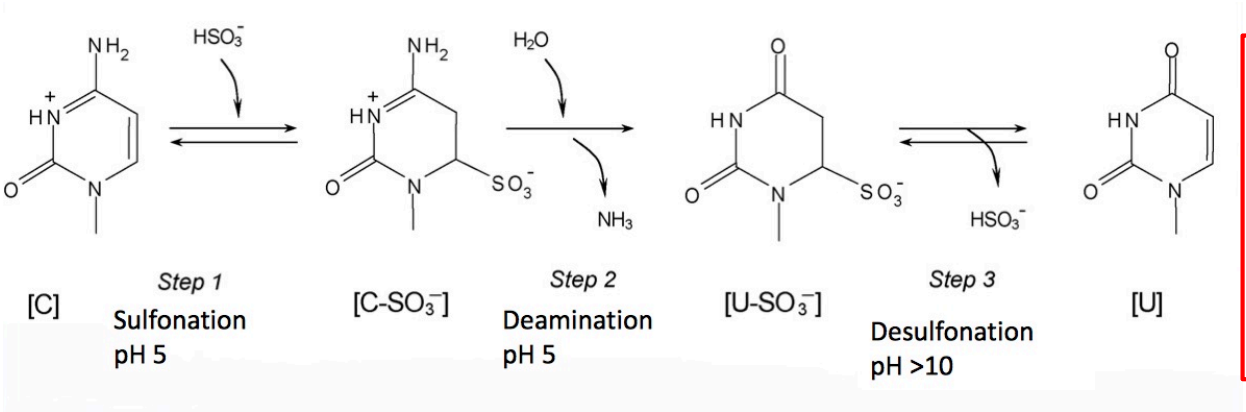
1. **Methylation-sensitive restriction enzymes combined with Southern blots or PCR**

   - Requires complete digestion of methylated DNA to avoid false positives and lacks sensitivity

2. **Methylation specific immunoprecipitation**

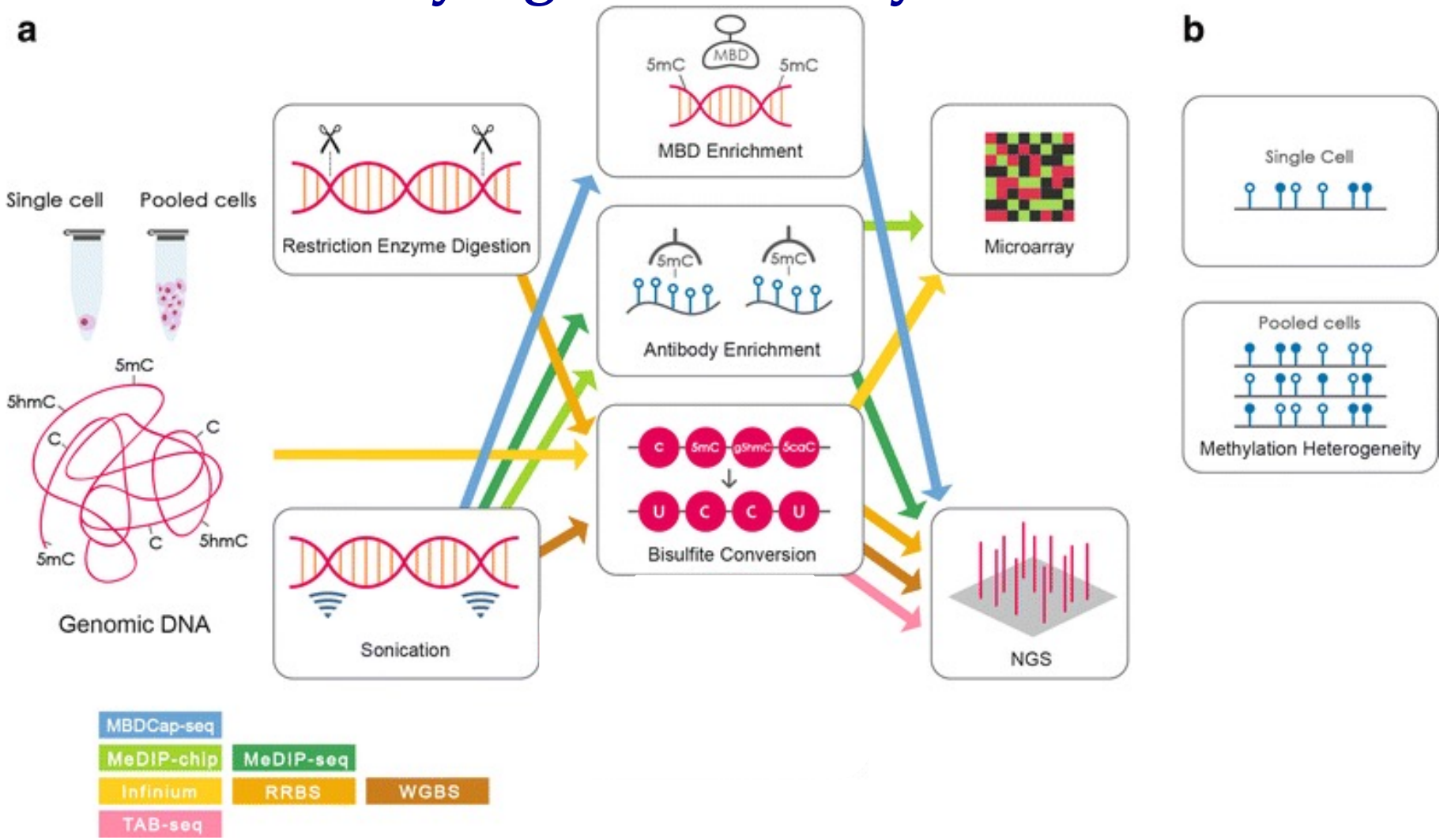
3. **<u>Bisulfite modification</u>** (<u>is the most widely used technique for studying DNA methylation</u>)

   - C's are converted to U's
   - <u>Methylated C's are NOT converted</u>



!!!!Standard molecular biology techniques such as PCR and cloning erase information about the position of methylated cytosines in DNA.

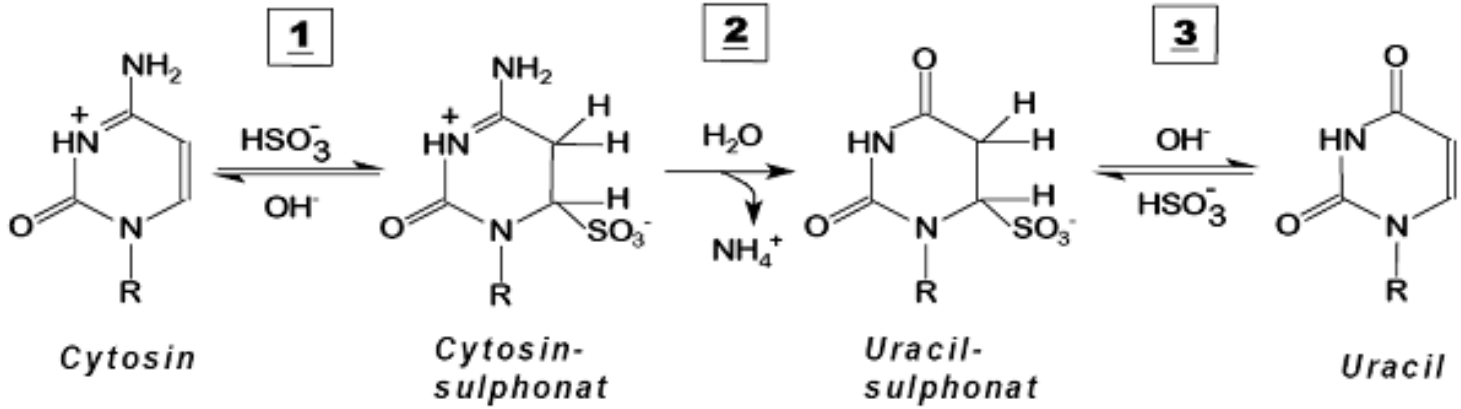
Studying DNA Methylation



1. fragmentation of genomic DNA by restriction enzyme digestion or sonication.

2. Affinity enrichment-based methods use either methyl-CpG-binding domain (MBD) proteins or antibodies specific for 5mC (as in MeDIP) to enrich methylated DNA region

3. sodium bisulfite deaminates unmethylated C to uracil (U), while methylated C residues remain unaffected. The U eventually converts to thymine (T) in a subsequent polymerase chain reaction (PCR).



Allele 1 (metilato)

---ACTCCACGG---TCCATCGCT---
---TGAGGTGCC---AGGTAGCGA---

Allele 2 (non metilato)

---ACTCCACGG---TCCATCGCT---
---TGAGGTGCC---AGGTAGCGA---

bisulfite treatment
Alkylation
denaturation

---AUTUUAUGG---TUUATCGCT---
---TGAGGTGUU---AGGTAGCGA---

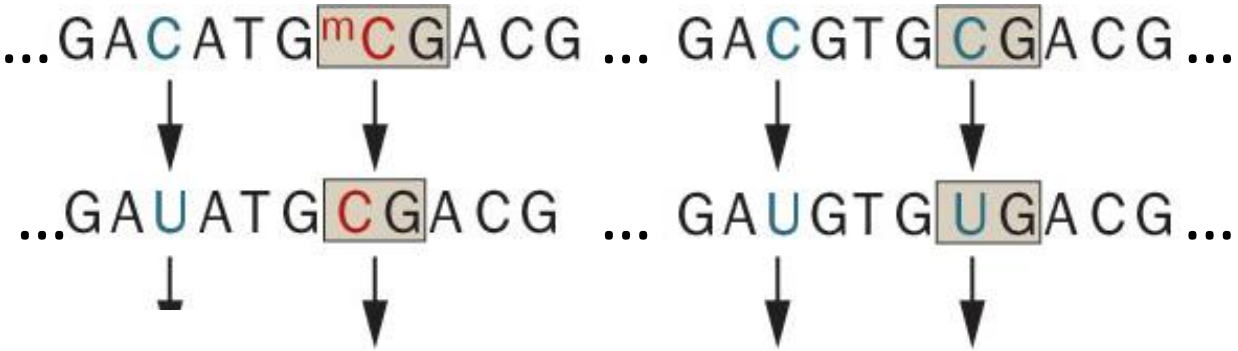
---AUTUUAUGG---TUUATUGCT---
---TGAGGTGUU---AGGTAGCGA---

PCR and fragment sequencing to identify
polymorphisms generated by bisulfite treatment.

Methyl-Seq

Genomic DNA

Converted DNA



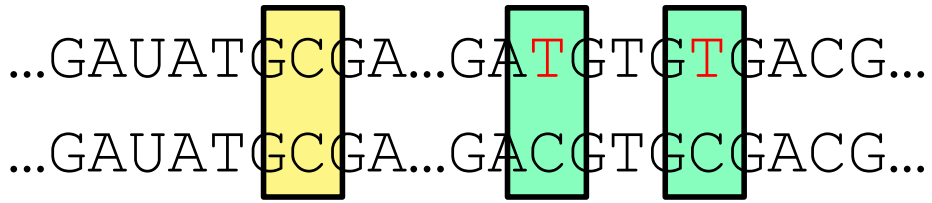
Methylated C remain C while the unmethylated C are converted in U (this will be read as T by the sequencing machine)

DNA sequencing

Sequences will be aligned on reference genome. C-T substitution means that the C was not methylated, conversely all the other kept Cs were originally methylated

Original converted DNA

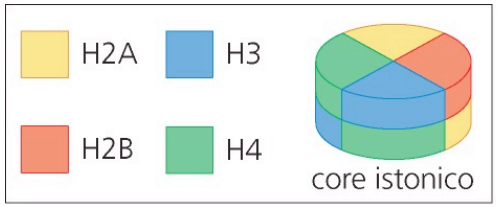
Reference Genome

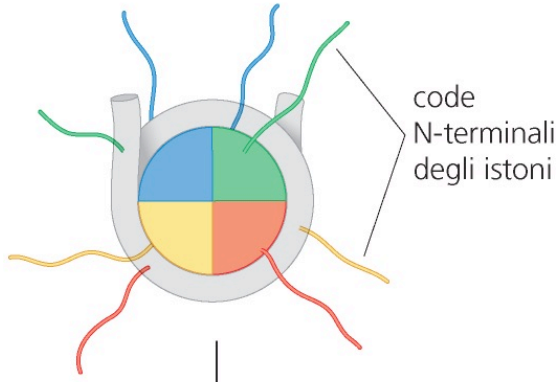


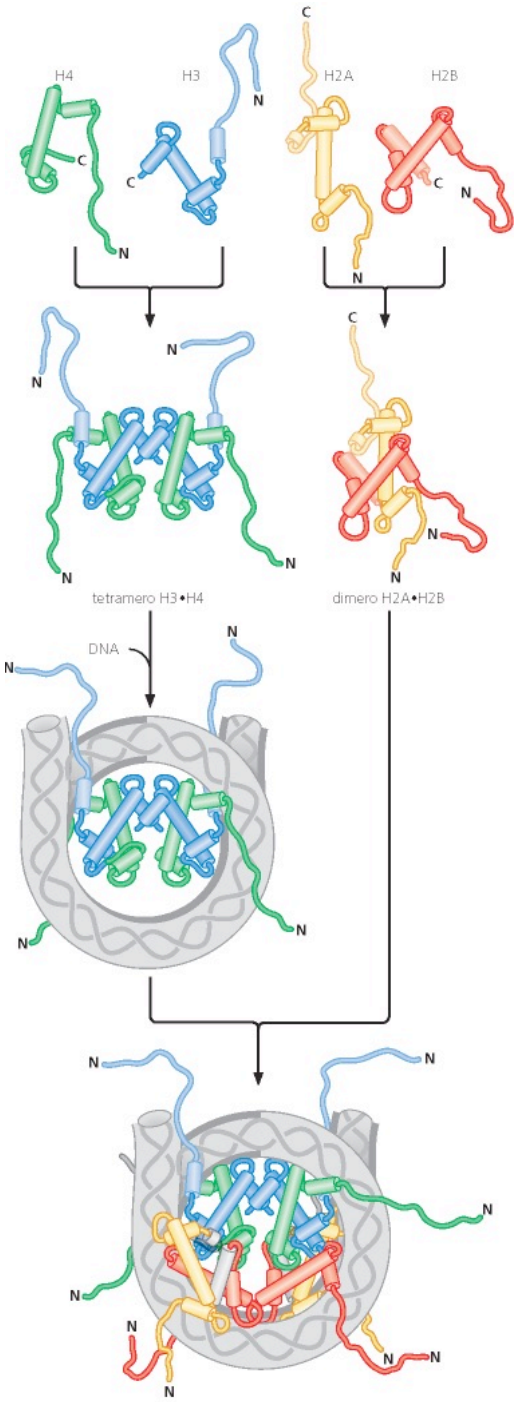
Histone modifications

The nucleosome
- repeat unit of chromatin

- 146 bp of DNA wrapped around a histone octamer core
- so that it forms 1.7 turns of a left-handed superhelix within the nucleosome core particle
- while the histon core is involved in histone-histone packing and DNA-contact, the N-terminal tail point outwards and is available for interaction

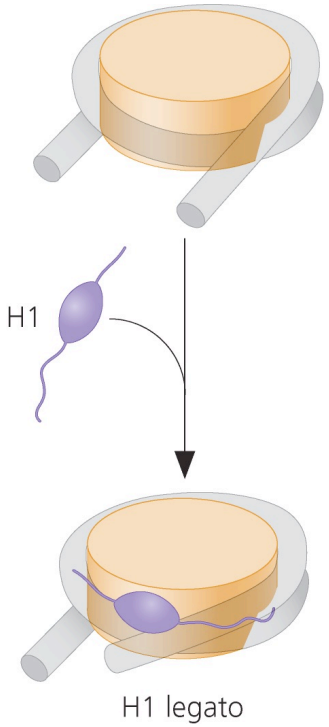


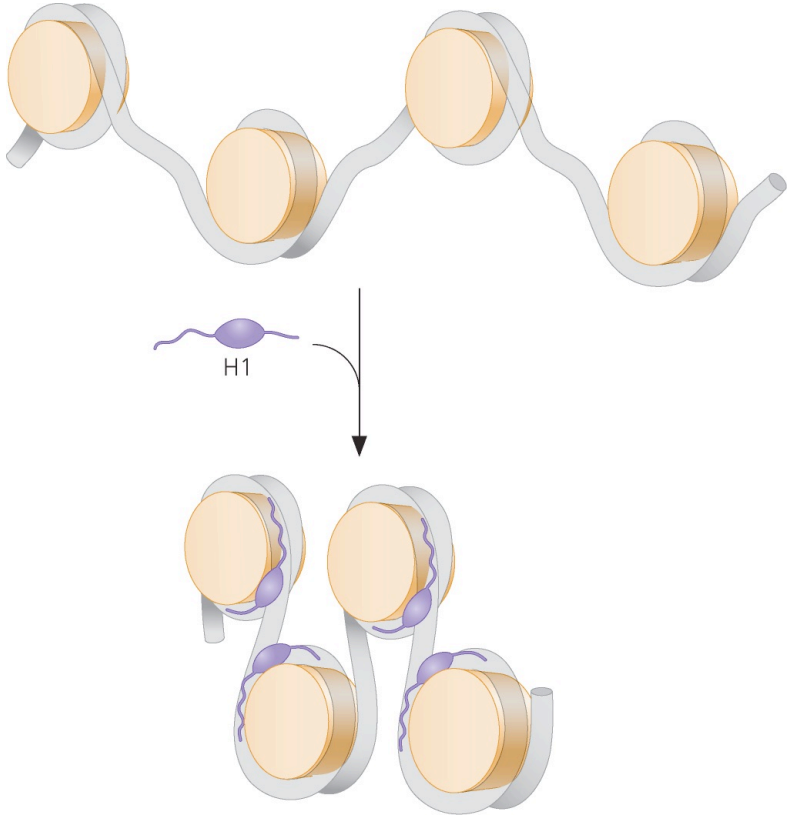


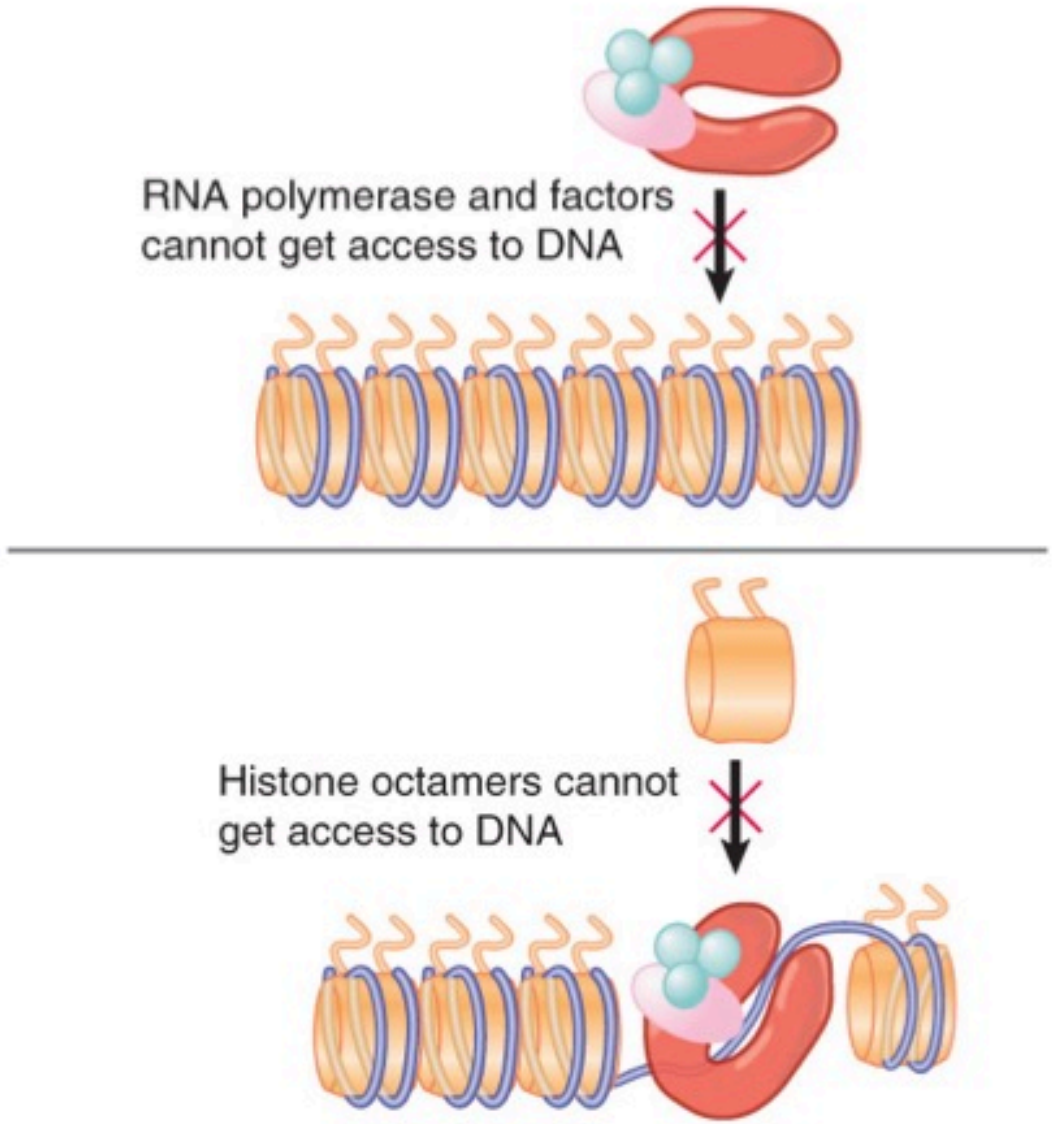


Histone H1 - linker histone

- **H1 linker histone**
  - associated with linker DNA between nucleosomes (about one H1 per nucleosome)
    - Binds DNA at entry/exit
    - stimulates folding 10 nm ➜ 30 nm fiber
    - repressive effect on transcription
  - H1 binds weaker to acetylated nucleosomes






RNA polymerase and factors
cannot get access to DNA
Histone octamers cannot
get access to DNA

Regulation of transcription by chromatin

1. Histone variants

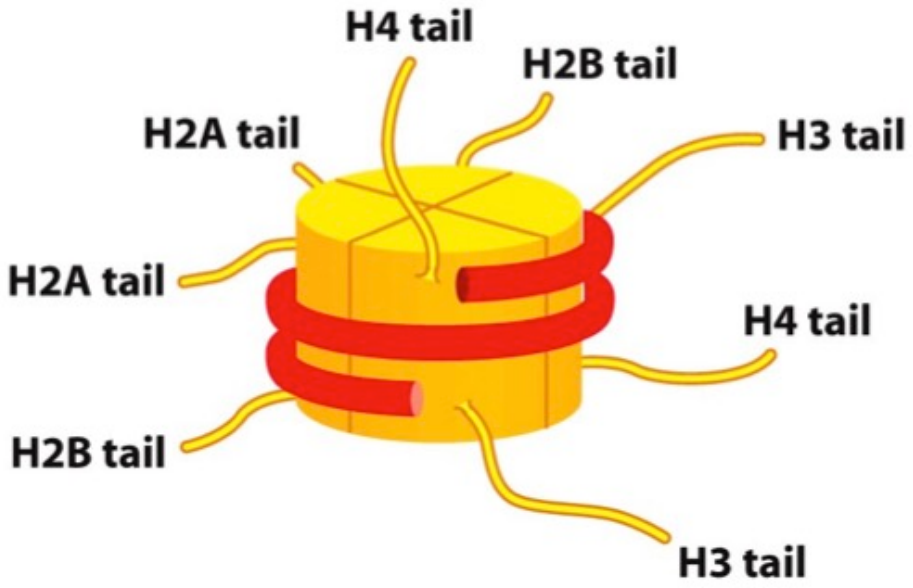
2. Opening of chromatin through directed **modication of histone tails** (e.g. acetylation and methylation)

3. Opening of chromatin through directed **nucleosome mobilization** by **remodeling complexes**

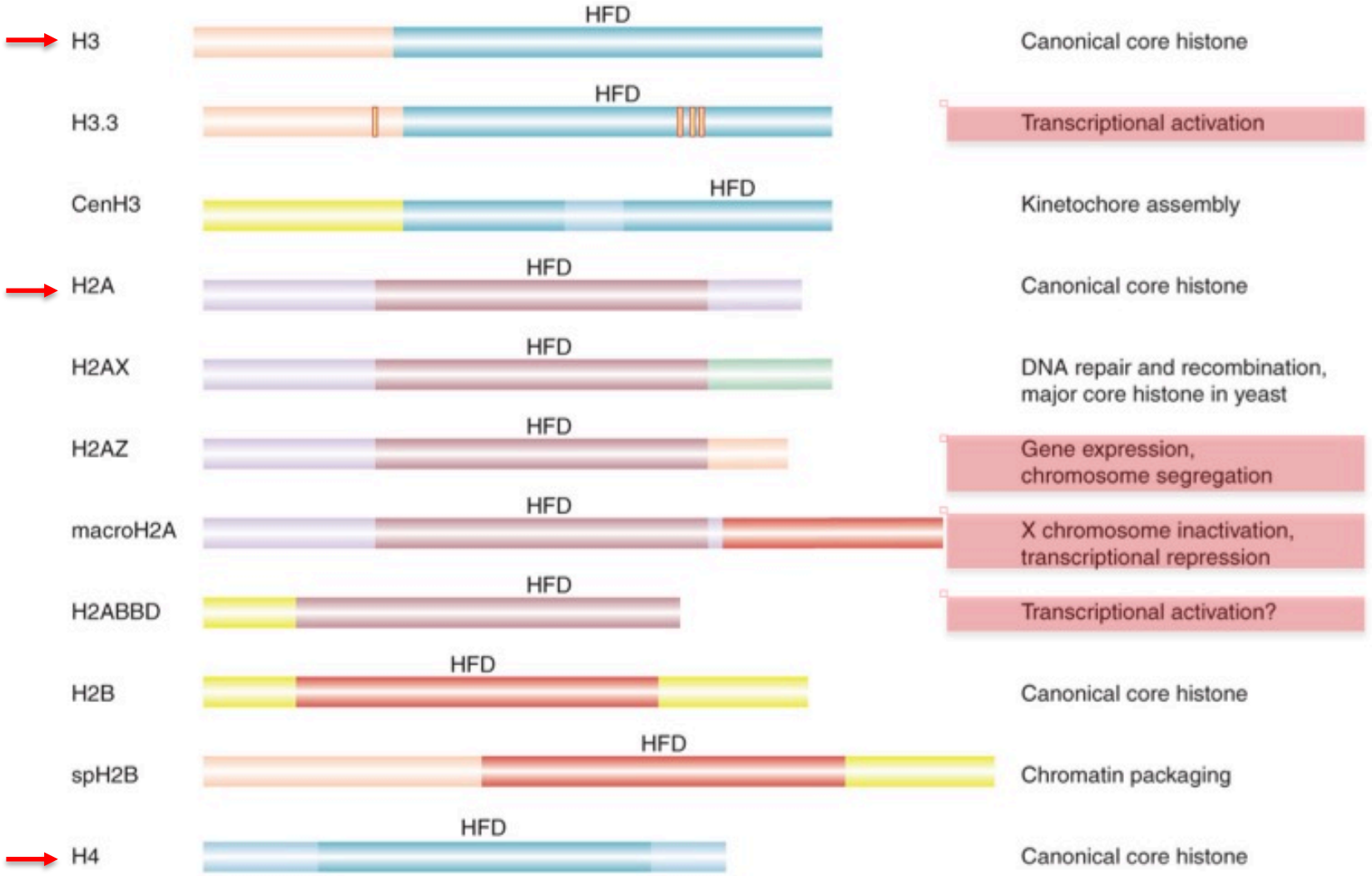
4. **Positioning of nucleosomes** creates promoters with different requirement for remodeling

Histone variants and modifications influence transcriptional activity

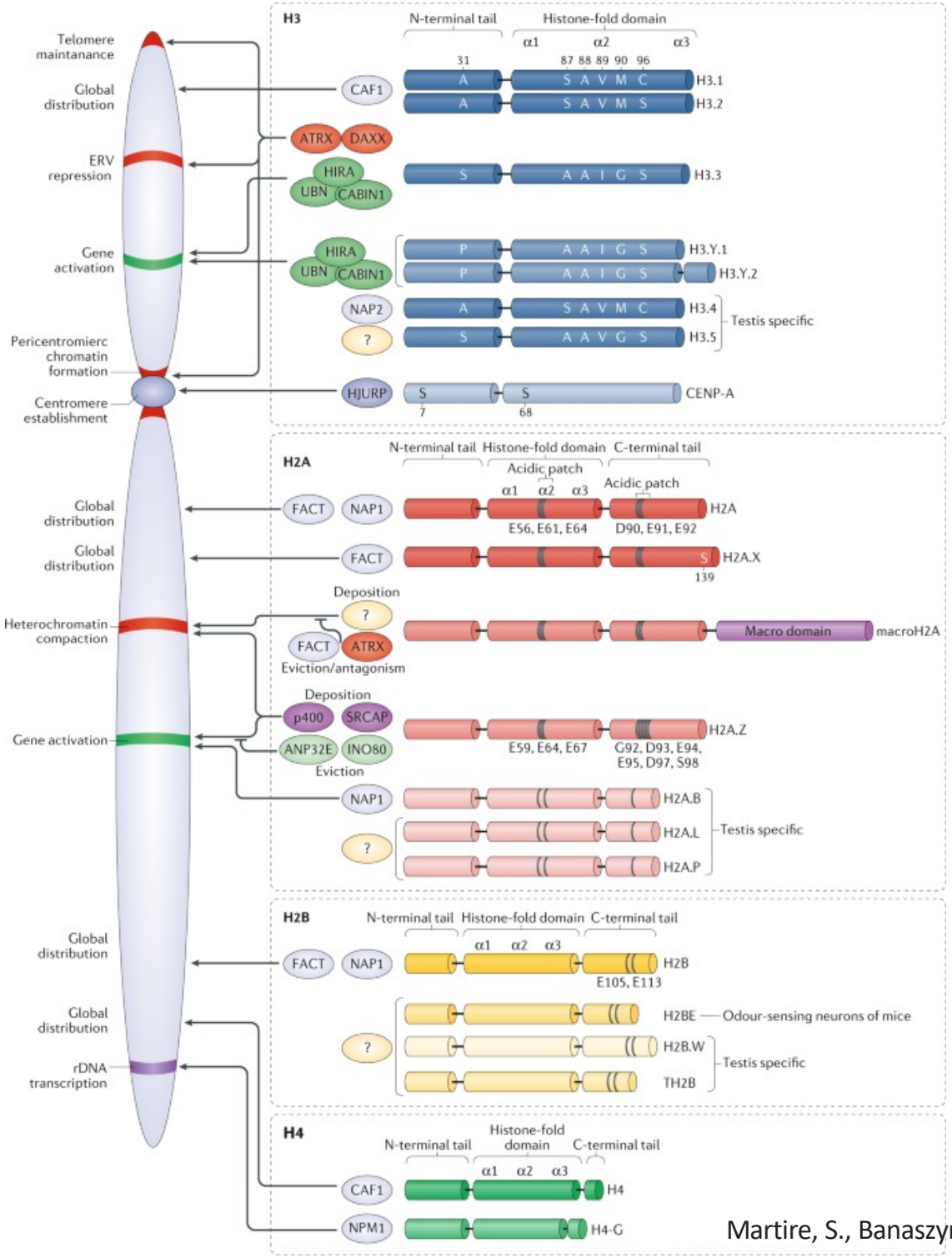
- Histone tails can be covalently modified to facilitate or prevent transcription
- Variants for Histone 2 and 3 mediate transcriptional control



Special functions of histone variants



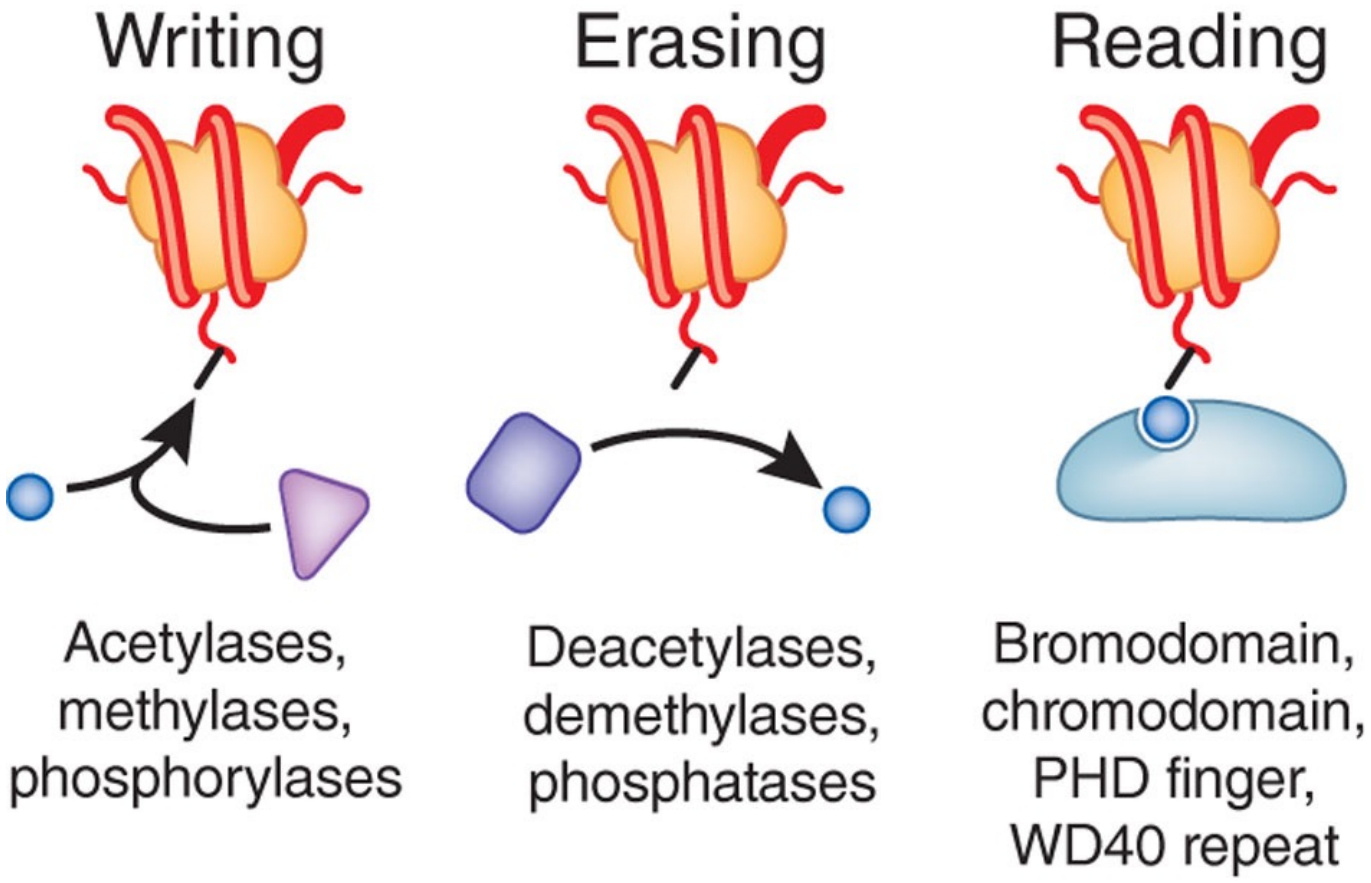
Special functions of histone variants



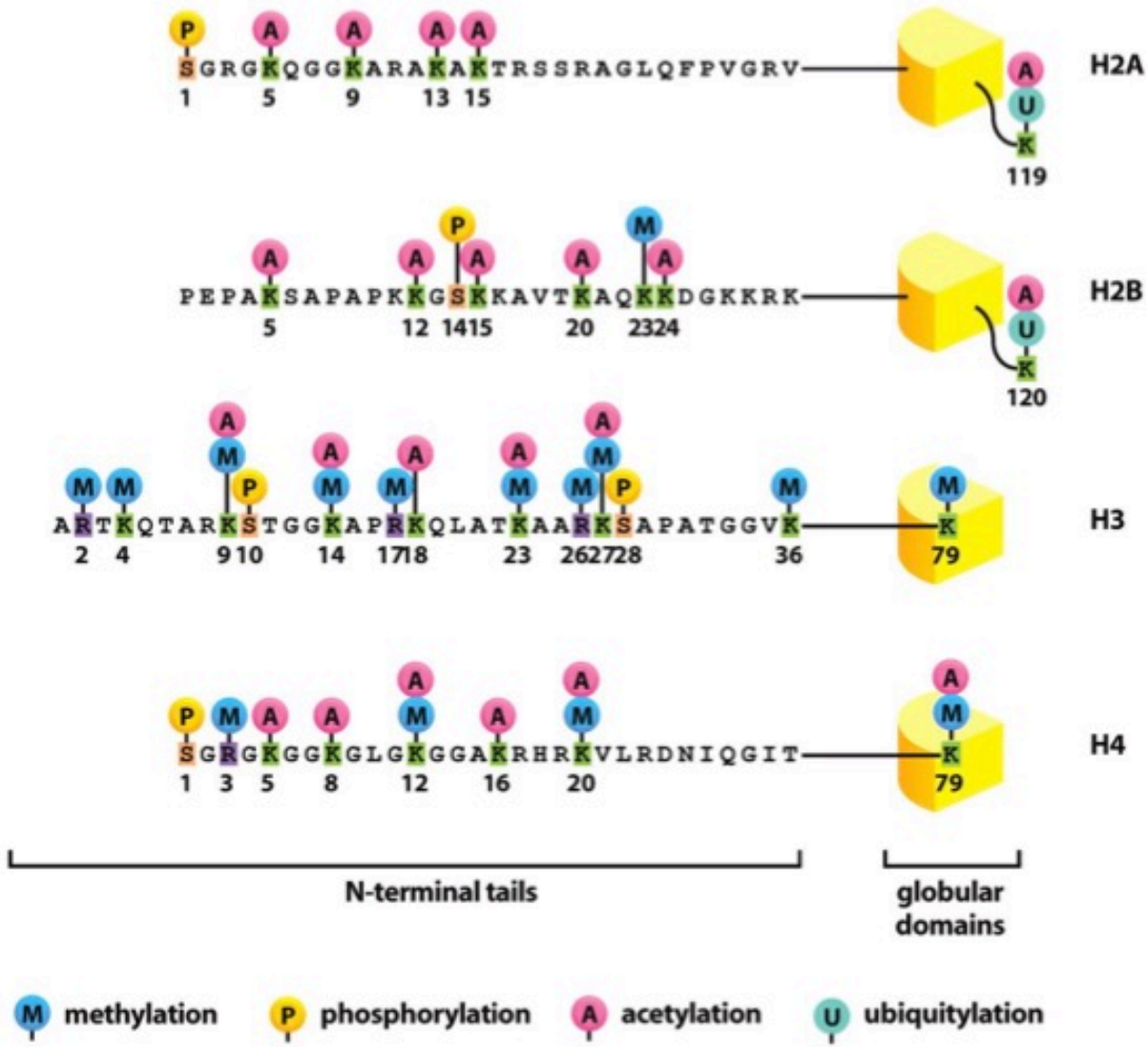
Martire, S., Banaszynski, L.A. *Nat Rev Mol Cell Biol* **21**, 522–541 (2020)

Histone post translational modifications (PTMs)

'**Writers**' introduce histone marks (circles), '**erasers**' (or **editors**) take them out and '**readers**' can recognize a particular form of histone modification.



Histone post translational modifications (PTMs)

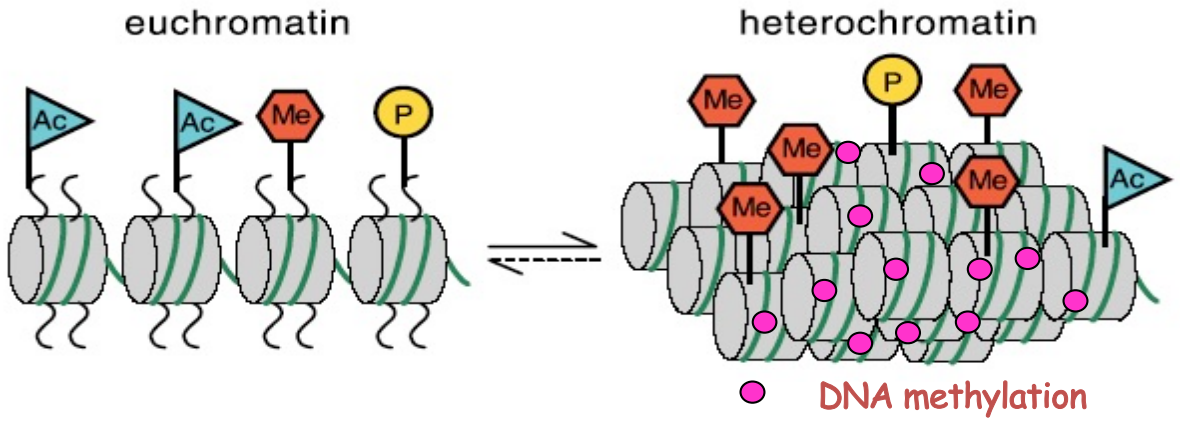


Histone PTMs regulate transcription

Table 1. Different Classes of Modifications Identified on Histones

Chromatin Modifications	Residues Modified	Functions Regulated
Acetylation	**K**-ac	Transcription, Repair, Replication, Condensation
Methylation (lysines)	**K**-me1 **K**-me2 **K**-me3	Transcription, Repair
Methylation (arginines)	**R**-me1 **R**-me2a **R**-me2s	Transcription
Phosphorylation	**S**-ph **T**-ph	Transcription, Repair, Condensation
Ubiquitylation	**K**-ub	Transcription, Repair
Sumoylation	**K**-su	Transcription
ADP ribosylation	**E**-ar	Transcription
Deimination	**R > Cit**	Transcription
Proline Isomerization	**P-cis > P-trans**	Transcription

Histone Modifications Associated with Heterochromatin and Euchromatin



Heterochromatin (inactive/condensed)

H3 ARTKQTARK^{Me_3}STGGKAPRKQLATKAARKSAPAT
9

H3K9Me3

H3 ARTKQTARKSTGGKAPRKQLATKAARK^{Me_3}SAPAT
27

H3K27Me3

Euchromatin (active/open)

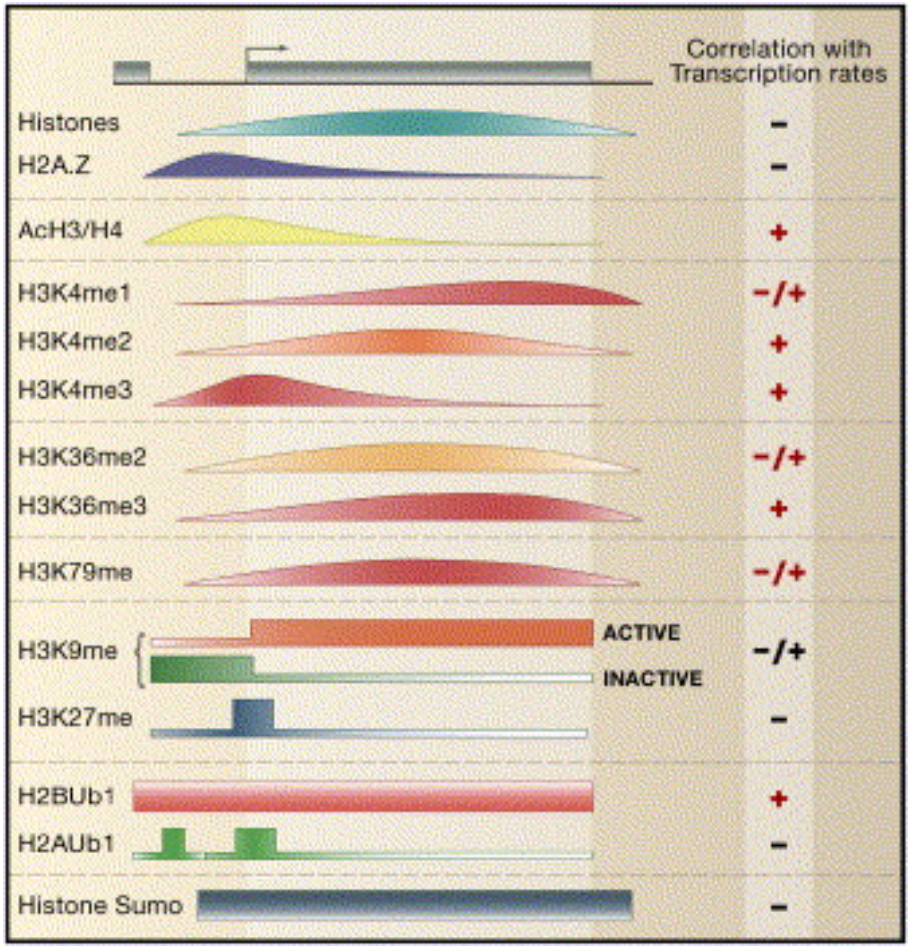
H3 ARTKQTARK^{Ac}S^{P}TGGK^{Ac}APRKQLATKAARKSAPAT
9 10 14

H3K9Ac H3S10P H3K14Ac

H3 AR^{Me_3}TKQTARKSTGGK^{Ac}APRKQLATKAARKSAPAT
4 14

H3K4Me3 H3K14Ac

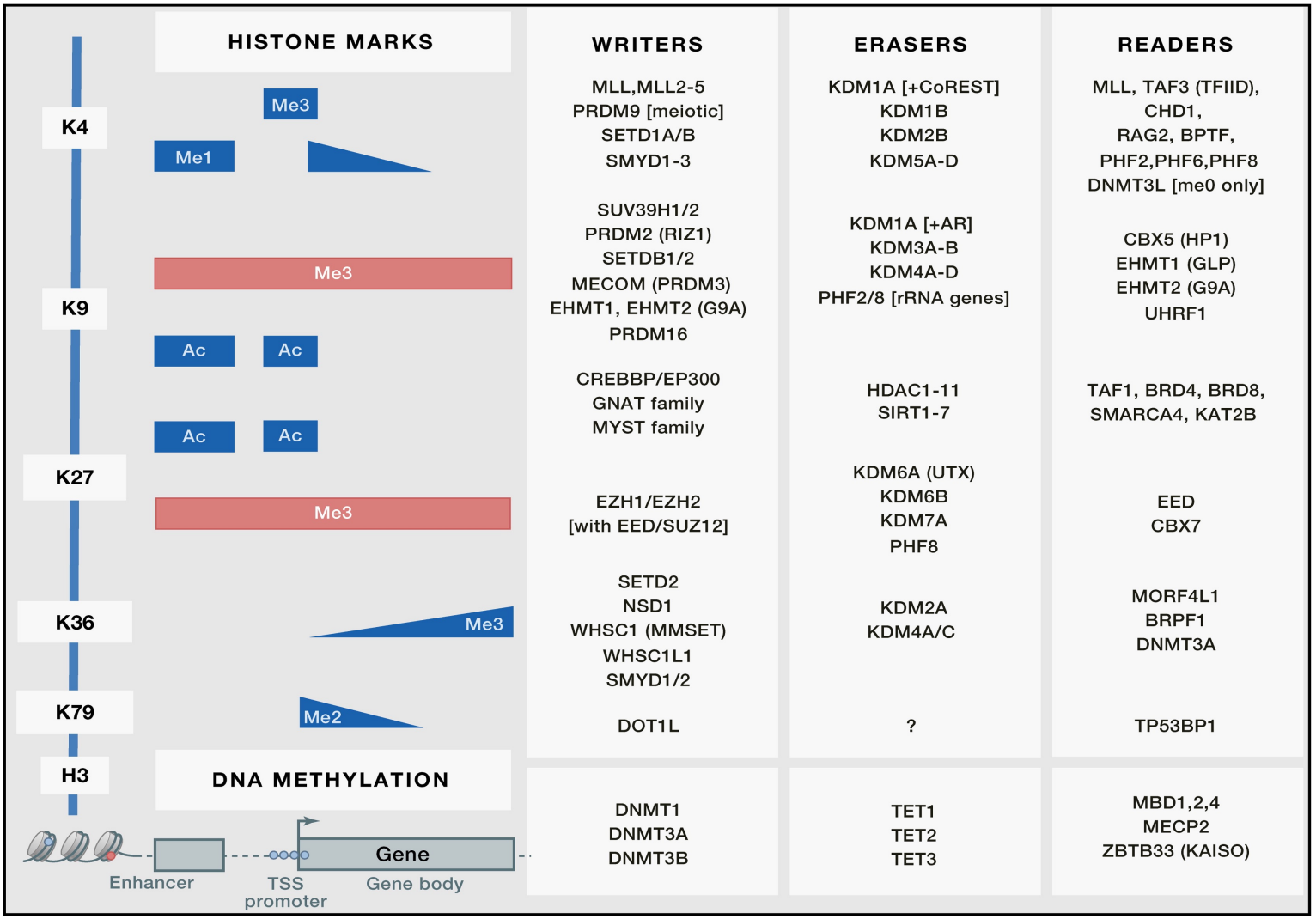
Histone PTMs regulate transcription



- Histone modifying enzymes (**writers**) generate "the histone code". They usally act in large complexes.
- PTMs dictate chromatin structure and serve as a scaffold for additional regulatory proteins (**readers**).
- Most modifications have been found to be dynamic, and enzymes that remove the modification have been identified (**erasers**).

Histone H3 Lysine Writers, Erasers, and Readers

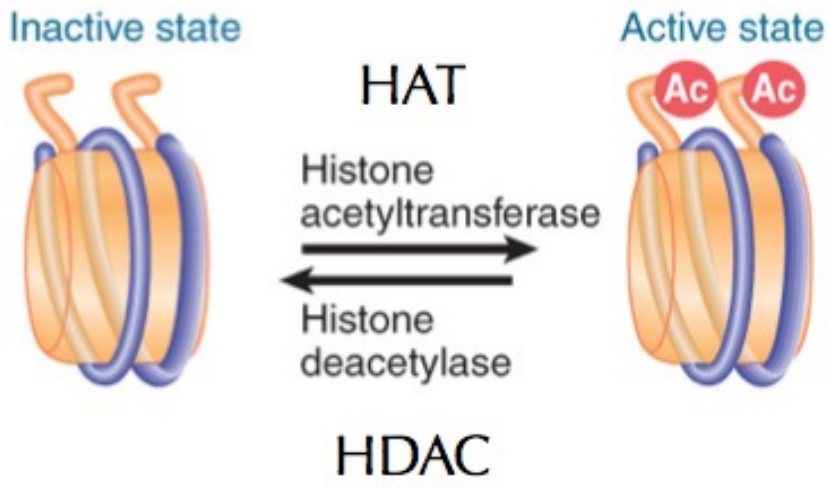
Post-translational modifications of histones are coordinated by counteracting **writers** and **erasers**.



Histone PTMs regulate transcription

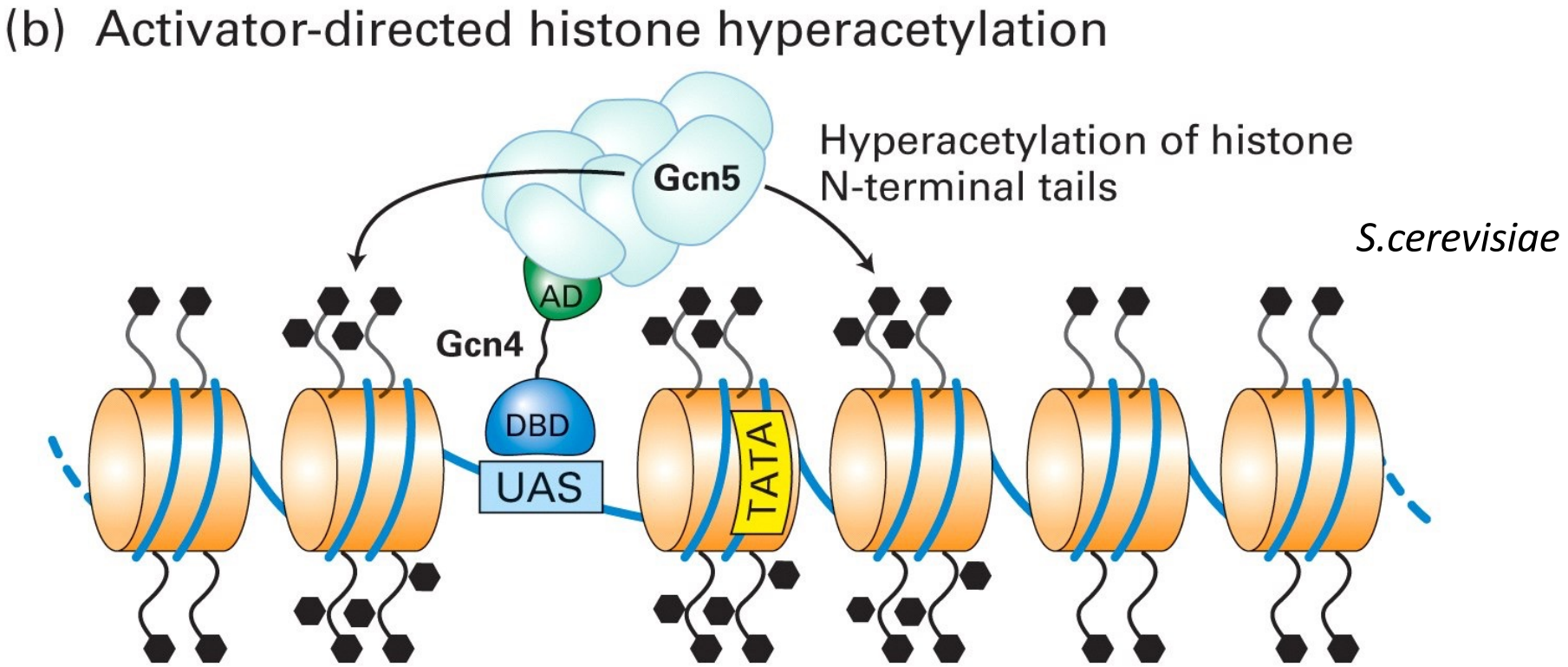
- **Acetylation** is associated with activation of transcription and absent in heterochromatin

- **Histone acetyltransferases (HATs)** and **histone deacetylases (HDACs)** bind to transcription factors or coactivators often forming large regulatory complexes



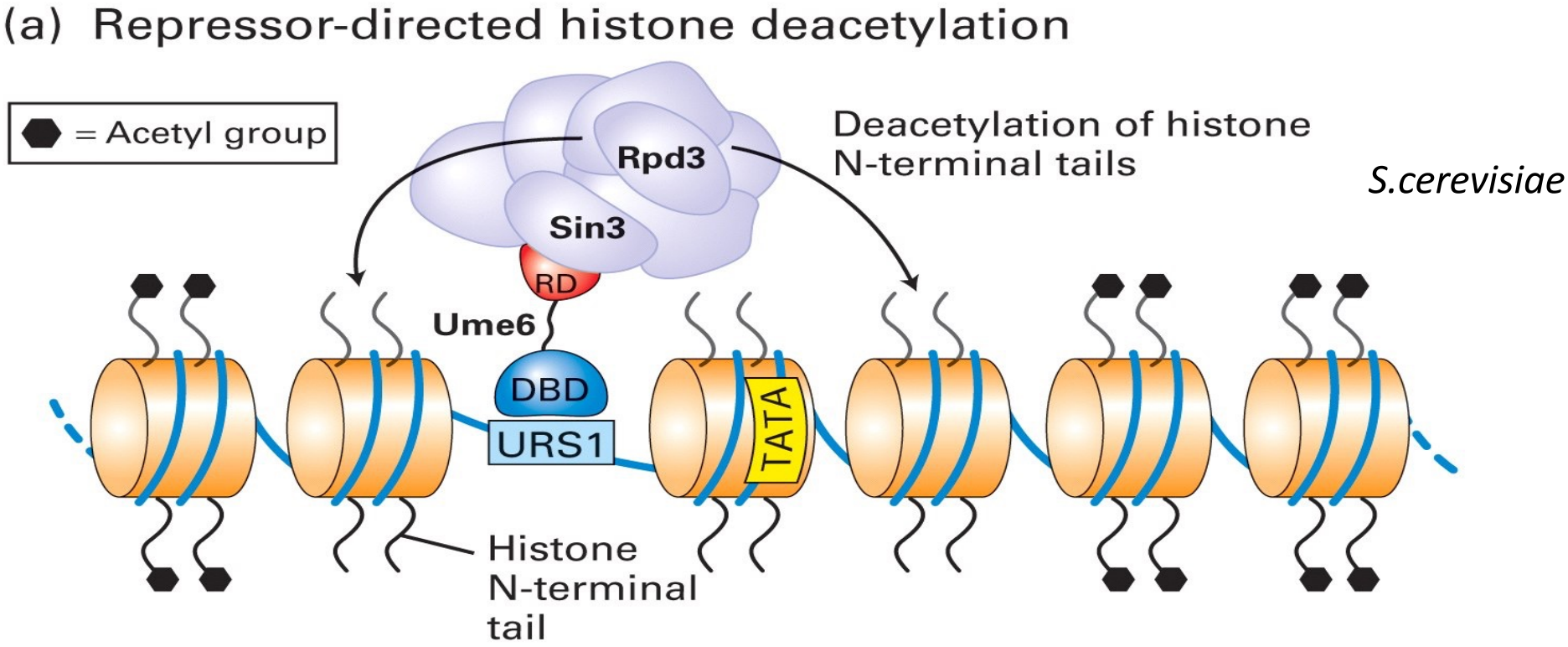
- chemical inhibitors (i.e. trichostatin) that inhibits HDACs resulted in increased gene expression

Activators Regulate Gene Expression by Modulating Chromatin Structure to be Open



- The DBD of Activators like Gcn4 bind their Upstream Activating Sequence (UAS).
- Activation Domain (AD) attracts protein complexes containing **histone acetylases** (Gcn5)
- Subsequent acetylation of histone tails serve to open up chromatin.

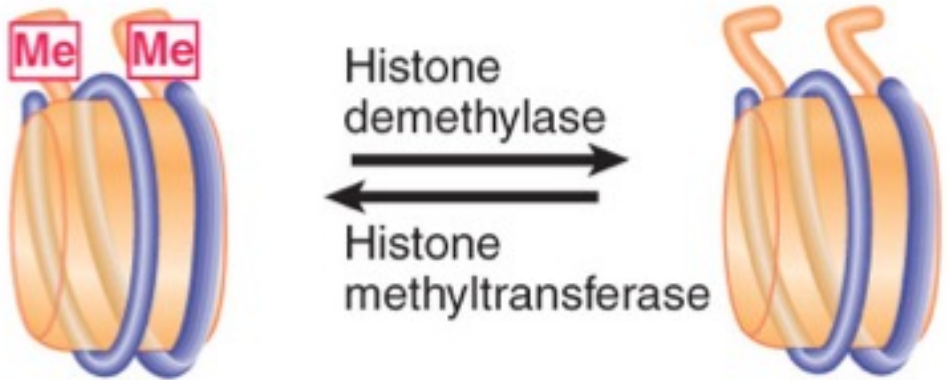
Repressors Regulate Gene Expression by Modulating Chromatin Structure to be Closed



- The DBD of repressors (like Ume6) bind a DNA element (URS1) and the Repression Domain (RD) recruits a protein complex containing a **histone deacetylase** like Rpd 3.
- The subsequent deacetylation of histone N-terminal tails results in chromatin condensation which promotes gene repression.

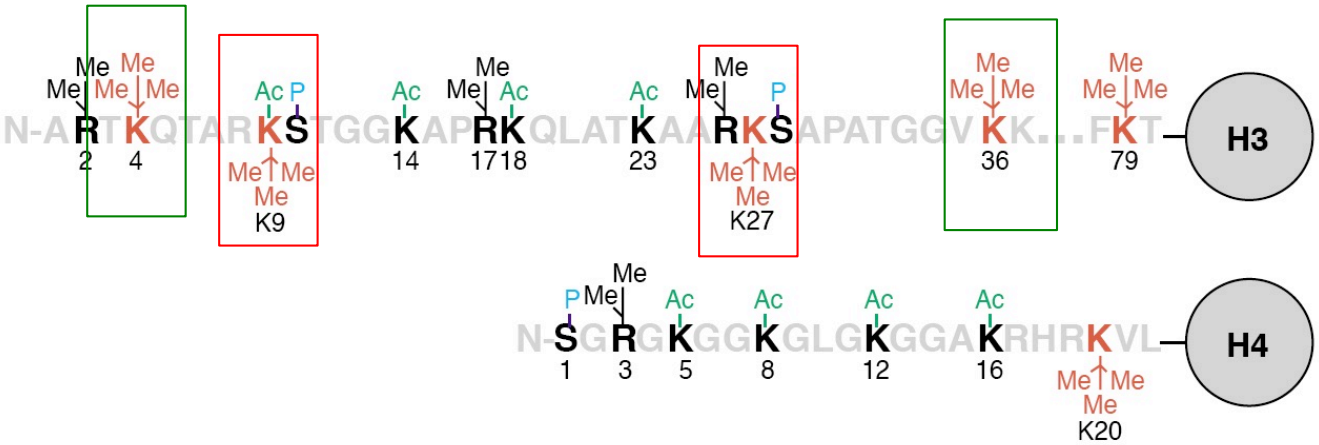
Histone methylation

- Both **Lys (K)** and **Arg (R)** can be methylated at more than one methyl-group
- Histone methylation is a relatively stable modification with a slow turnover rate.
- An ideal epigenetic mark for more long-term maintenance of chromatin states.
- Methylated residues are present both in eu- and heterochromatin.
- Acetylated H3K9 prevents methylation, and prevents HP1 binding for heterochromatin formation



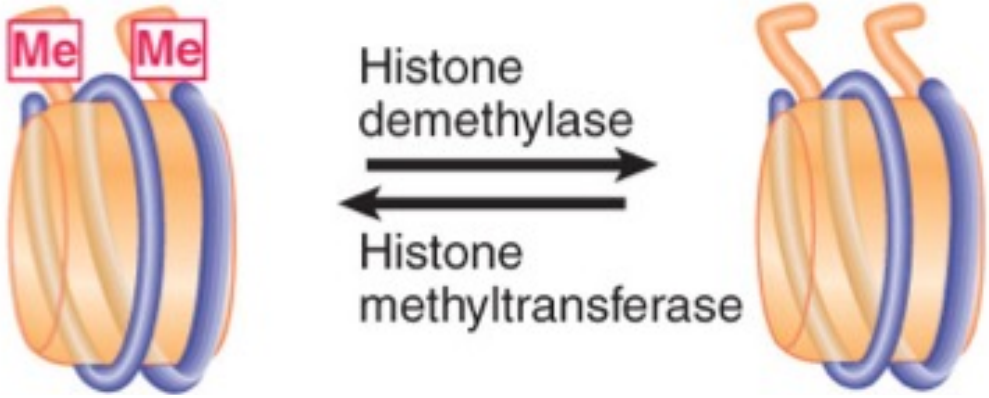
The subtrates:
Histone tails - multiple methylations

N-term tails reversible methylated in **Arginine** (**R**) and **Lysine** (K), particularly in H3 and H4.



- Above ≈ activation
- Below ≈ repression

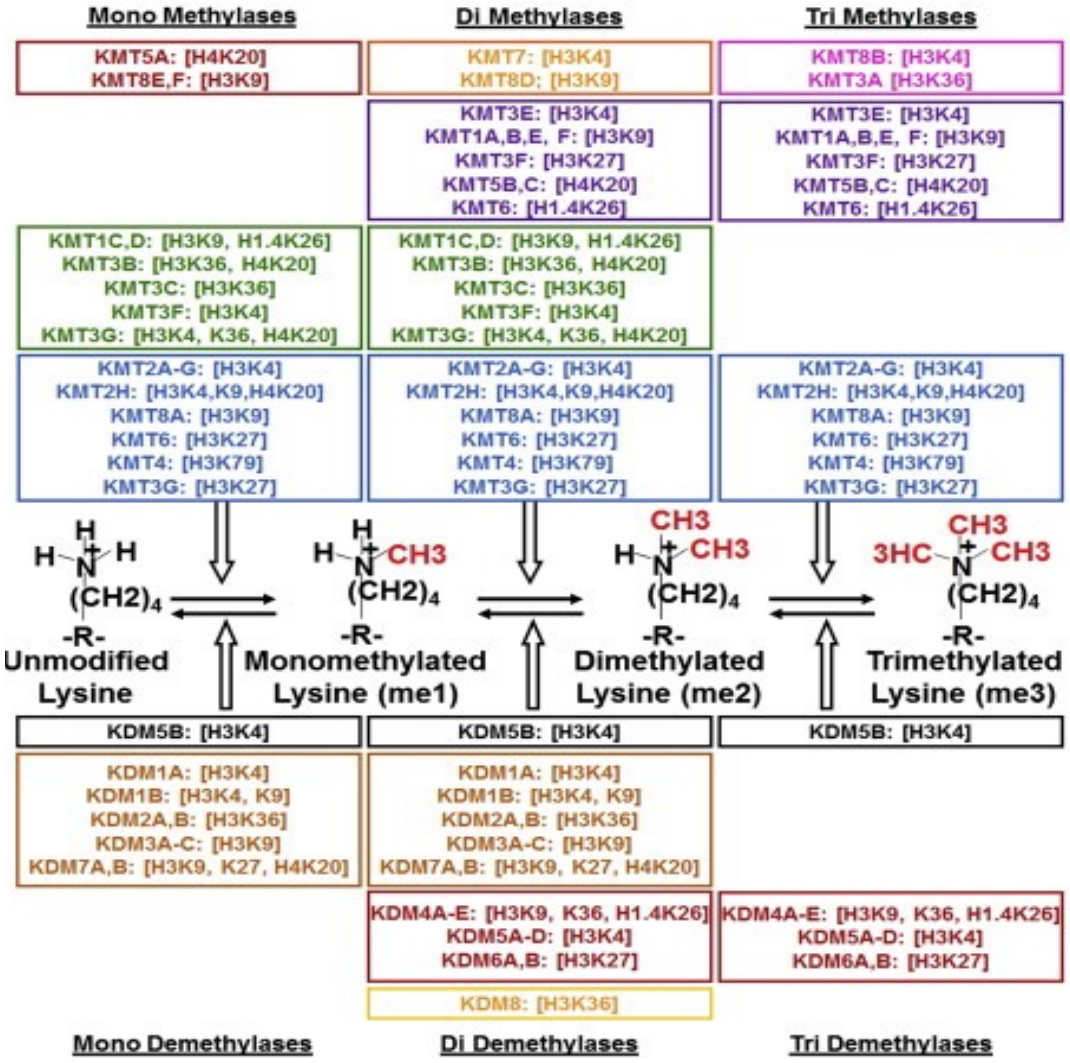
Histone PTMs regulate transcription



	Transcription
H3K4me1	–/+
H3K4me2	+
H3K4me3	+
H3K36me2	–/+
H3K36me3	+
H3K79me	–/+
H3K9me	–/+
H3K27me	–

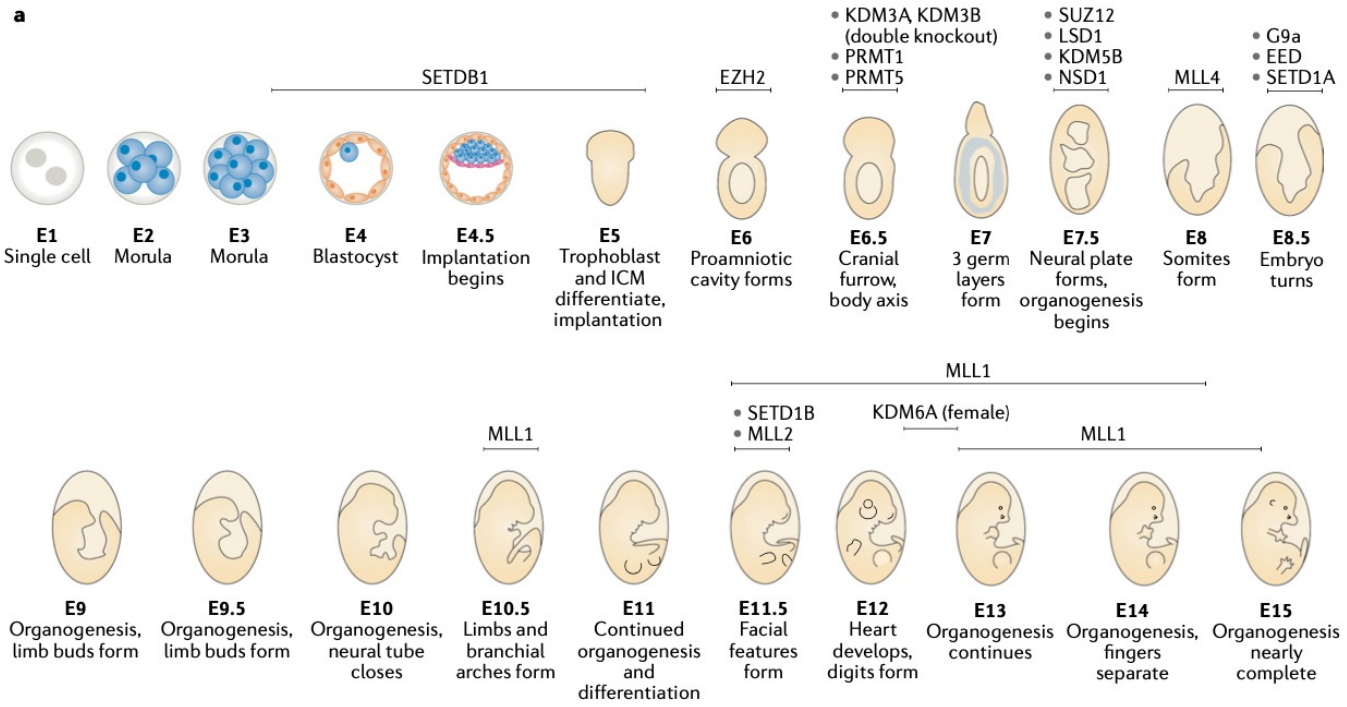
Activation/silencing determined by the counteracting action of KMTs and KDMs

- **KMT**s (histone lysine methylase) and **KDM**s (histone lysine demethylases) have a high degree of specificity for particular lysine residues and the degree of methylation.



The importance of histone methylation regulators in mammalian development

Loss of some histone methylation regulators causes very early lethality, before or during implantation (for example, SETDB1), whereas other regulators are required at later stages of organogenesis, with the majority exhibiting lethality between embryonic day 7 (E7) and E12. For some regulators (MLL1, SETDB1), lethality was observed at different stages, depending on the report.



Polycomb and MLL/Trithorax Complexes

Epigenetic regulation of gene expression is one of the key mechanisms regulating cell-fate choices and cell identity during development. One of the **most prominent** and enigmatic epigenetic **regulatory systems involves the evolutionarily conserved Polycomb group (PcG) and Trithorax group (TrxG) components,** acting antagonistically to orchestrate the expression of key genes in cell differentiation and developmental processes.

H3K27me3

Polycomb-group Proteins

- Maintains a silenced state
- Prevents chromatin remodelling

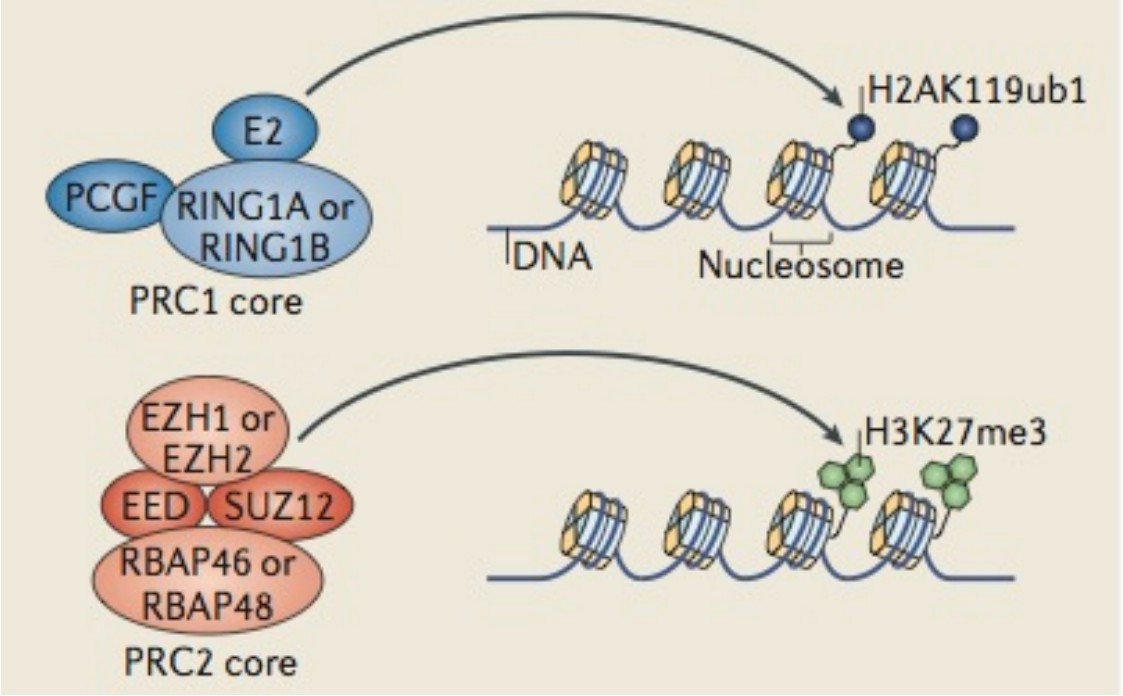
H3K4me3

Trithorax-group Proteins

- Maintains an active state
- Counteracts the action of PcG proteins

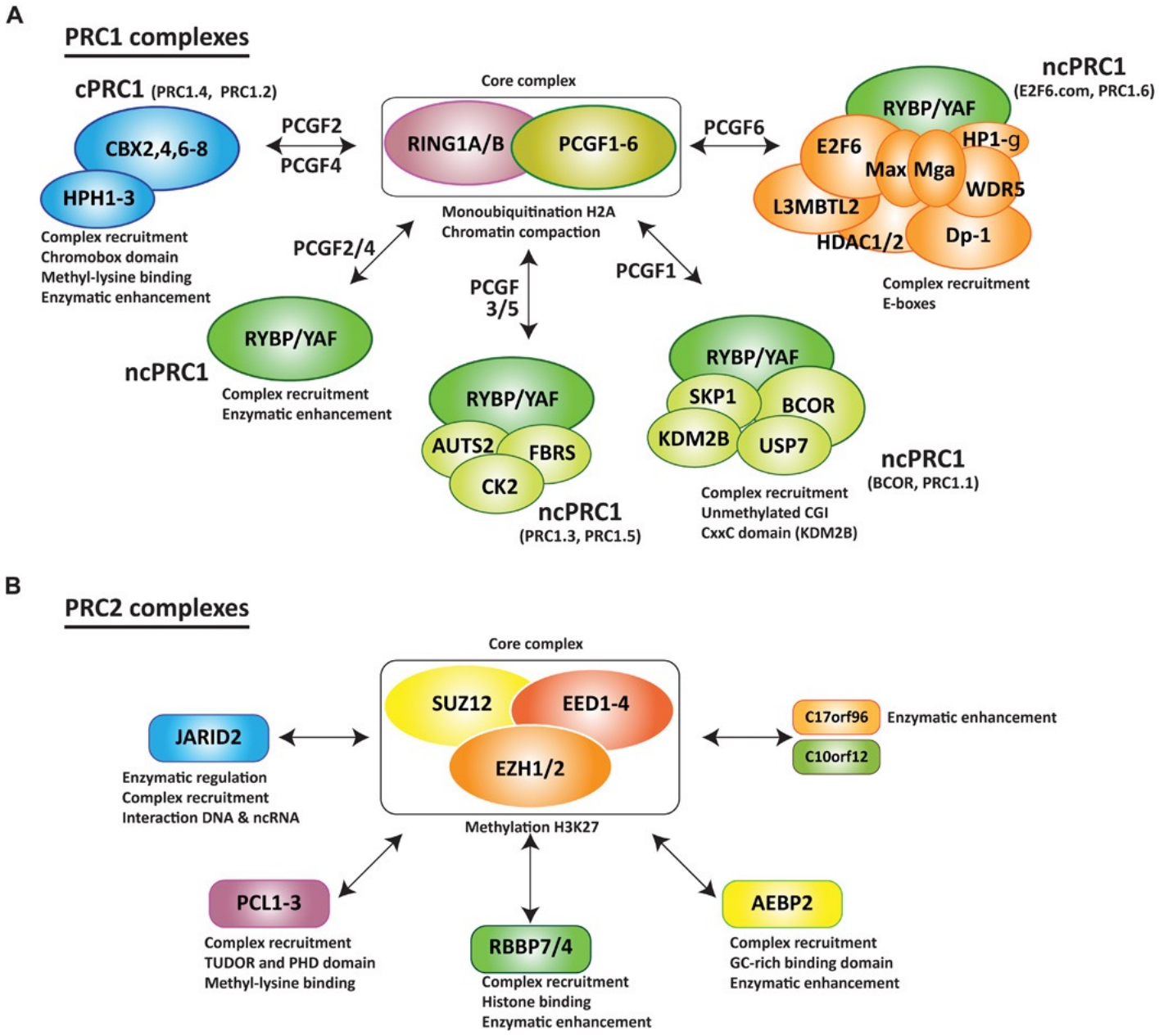
Core PRCs and their chromatin-modifying activities

•**Polycomb repressive complexes** (PRCs) repress transcription by a mechanism that involves the modification of chromatin.



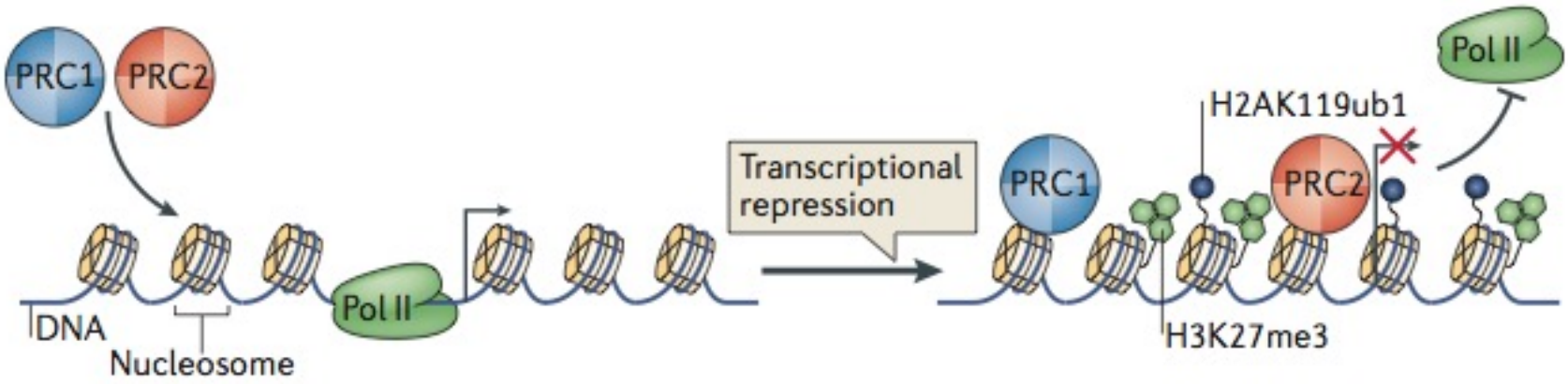
PcG complexes in mammals.

The interaction of the core complex with other accessory proteins defines the complete composition of each subcomplex. These accessory proteins have been found to regulate recruitment to specific chromatin domains and/or to modulate the catalytic activity of the core complex.



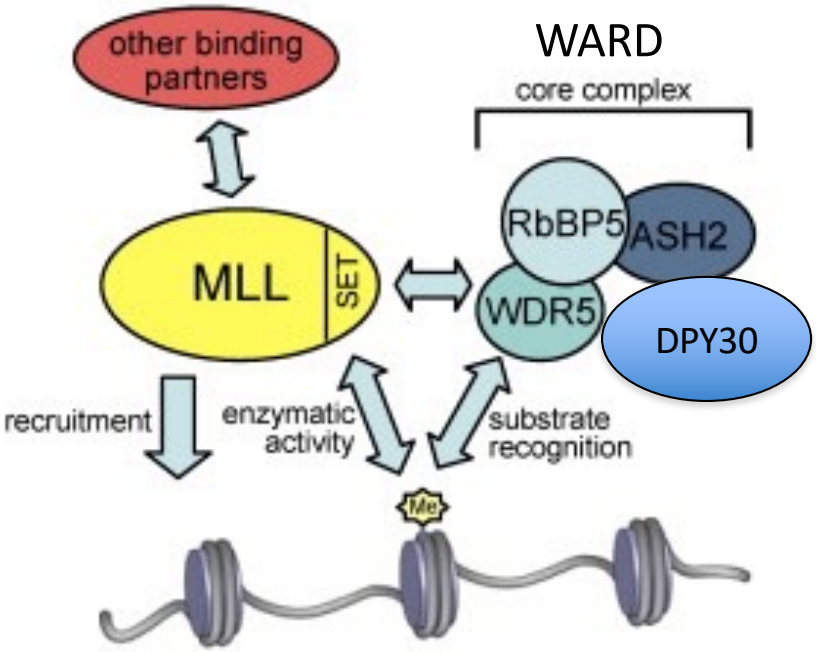
Polycomb systems and gene regulation

Polycomb complexes lead to Polycomb chromatin domain formation, which then directs repression of transcription by RNA polymerase II (Pol II) at the associated gene.



Writing the H3K4 Methylation Mark

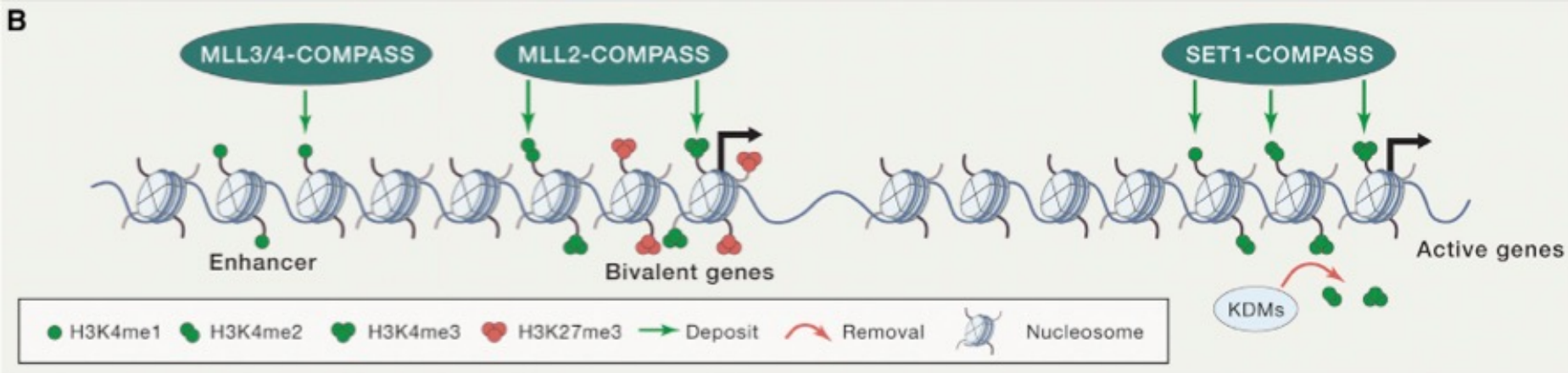
- **Trithorax** (MLL or COMPASS in mammals) methylates H3K4 and recruits HAT and remodeling complexes



MLL-family HMTs associate with the core complex containing, **WDR5**, and **ASH2, RbBP5** and **DPY30 (**abbreviated in **WARD)**. The core complex cooperates with the catalytic SET domain to methylate H3K4, whereas other regions of the MLL protein are involved in association with other protein partners and in recruitment of the MLL complex to the target genes. **WDR5** plays a role in substrate recognition and presentation, with preferential, but not exclusive, binding to the H3K4me2 substrate.

Writing the H3K4 Methylation Mark

Set1A/B COMPASS complexes catalyze mono-, di-, and trimethylation on H3K4 at active promoters. The activity of the partially redundant COMPASS complexes containing MLL3/KMT2C and MLL4/KMT2D leads to the deposition of H3K4me1 at enhancers, facilitating the recruitment of other activators such as CBP/p300. The deposition of methyl marks on H3K4 at bivalent regions is performed by MLL2/COMPASS. Multiple histone demethylases, including members of the KDM1/LSD, KDM2A/FBXL11, and KDM5/JARID families, are implicated in the removal of methyl groups on H3K4.

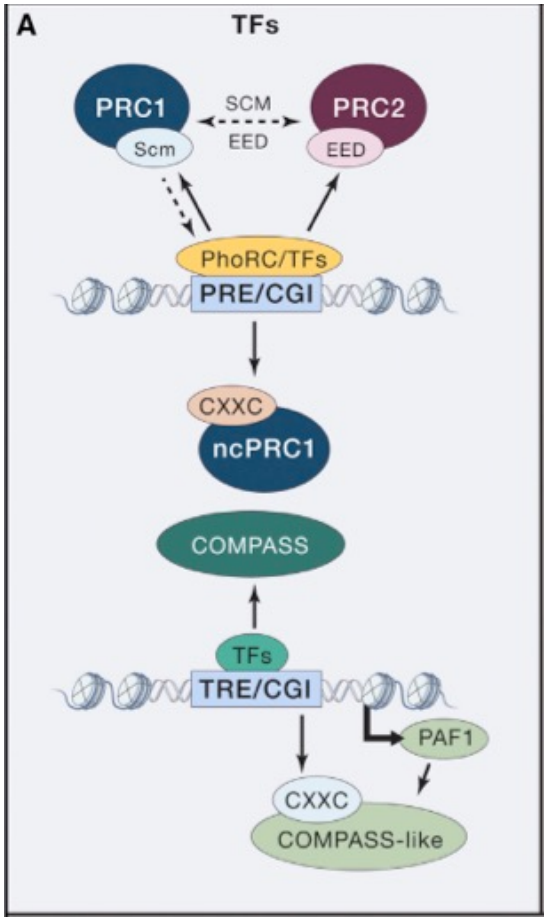


Recruitment of histone modifying complexes

Transcription factor (TF)-mediated recruitment

- Cooperative interactions between TFs and PRC1 and PRC2 complexes contribute to stable PcG binding. The PRC1 accessory protein Scm is a key protein in mediating PRC1 interaction and PhoRC. Similarly, SCMH1/2 and EED might mediate interactions between PRC1 and PRC2 complexes, thereby facilitating stable binding to chromatin. KDM2B, which is part of ncPRC1.1, binds to CGIs (hypomethylated CpG islands) through its zinc-finger CxxC domain.

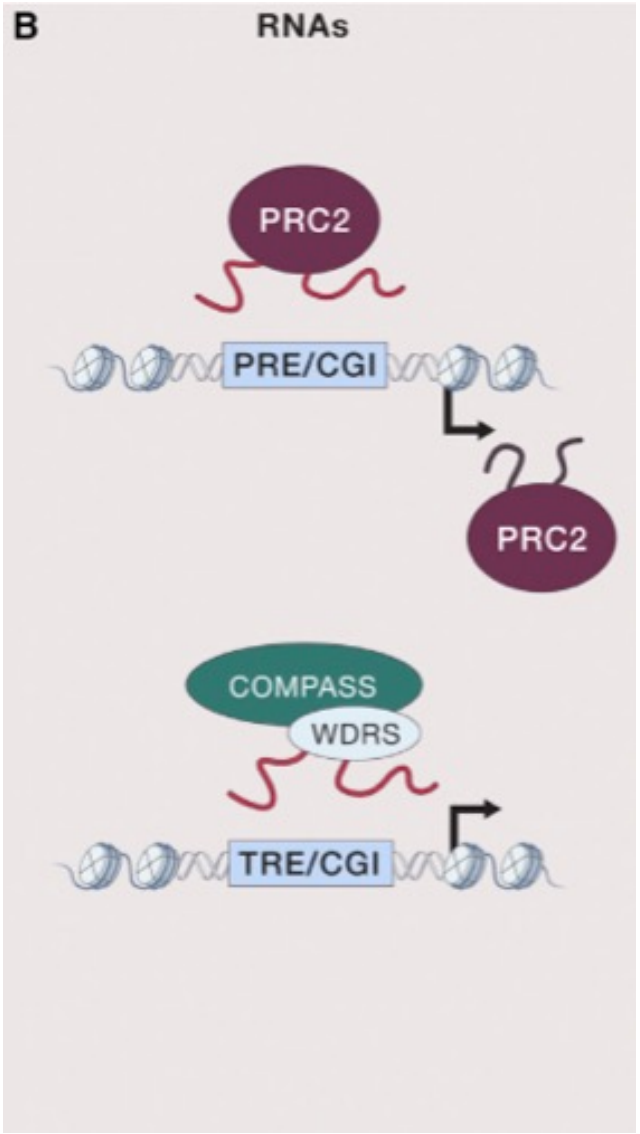
- Similarly, COMPASS(-like) complexes can bind sequence specifically to their target sites via the CxxC domain of MLL1/2 or can be recruited by TFs such as CFP1 or FOXA1 or via their interaction with PAF1.



RNA-mediated recruitment

- RNA-mediated recruitment. ncRNAs (like XIST, HOTAIR, or KCNQ1, shown in red), or short RNAs transcribed from repressed genes that form stem- loop structures (in black), can induce recruitment of PcG complexes.

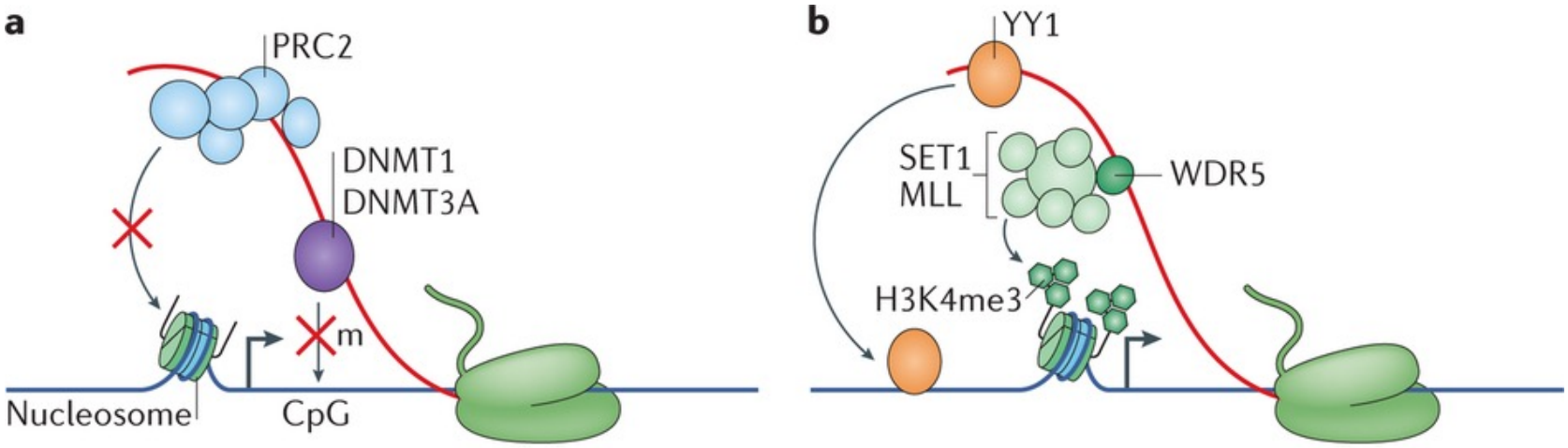
- In a similar manner, HOTTIP ncRNA interacts with the WDR5 subunit of COMPASS to recruit it to the HOXA locus.



Nascent RNA modulates the association of regulatory factors with chromatin to maintain gene activity

a) **Nascent RNA can compete** with chromatin for binding of repressive chromatin modifiers, such as PRC2, and DNMT1 and DNMT3A, which primarily methylate the DNA at CpG dinucleotides.

b) Interaction of the transcription factor YY1 with **nascent RNA facilitates** its transfer to chromatin. Similarly, the interaction of WDR5, which is a component of MLL, with nascent RNA facilitates their transfer to chromatin and H3K4me3, thereby forming a positive-feedback loop that promotes gene expression.

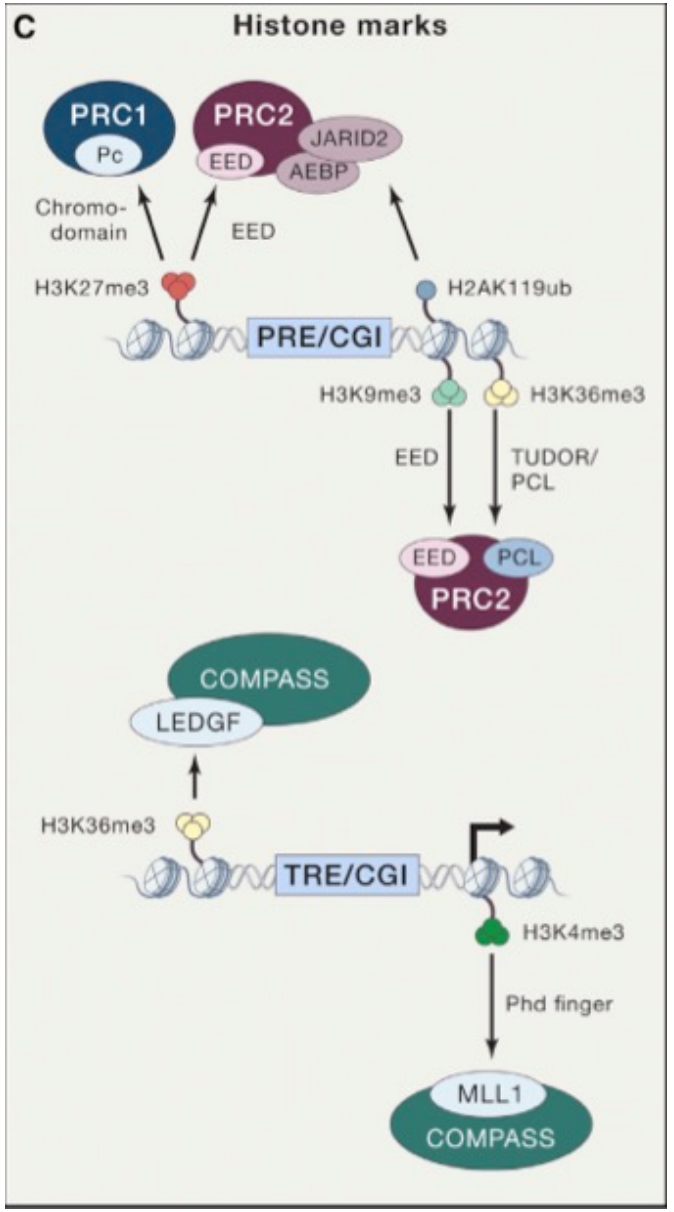


Nature Reviews | Molecular Cell Biology

Chromatin-mediated recruitment

- H3K27me3 can increase the affinity of PRC1 to chromatin via interaction with the chromodomain of Pc/CBX. The EED subunit of PRC2 can bind to methylated H3K9 and H3K27, thereby stabilizing its binding to chromatin. H2AK119ub can induce PRC2 recruitment mediated by AEBP2 and JARID2. H3K36me3 can stabilize binding of PRC2 via the TUDOR domain of PCL proteins.

- TrxG complexes containing MLL1 can be recruited to their target sites via their interaction of the PHD finger with H3K4me3. H3K36me2, deposited by ASH1L, can promote binding of COMPASS complexes via the epigenetic reader protein LEDGF.



Model for Role of H3K9 Methylation in Heterochromatin Formation

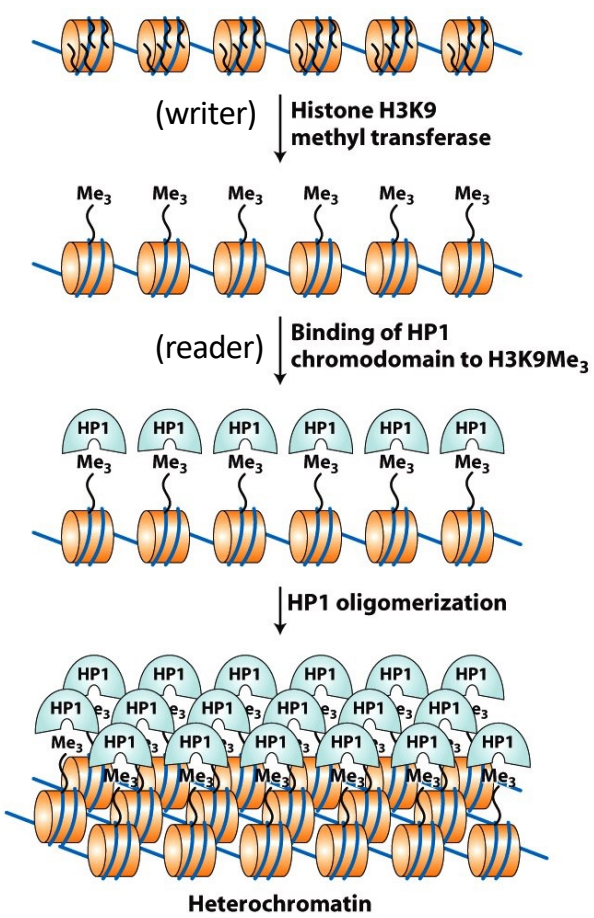


Figure 6-34a
Molecular Cell Biology, Sixth Edition

- HP1 (heterochromatin protein 1) binds to H3K9-Me_3
- HP1 oligomerization
- Spreads along chromatin
- Condenses chromatin into heterochromatin

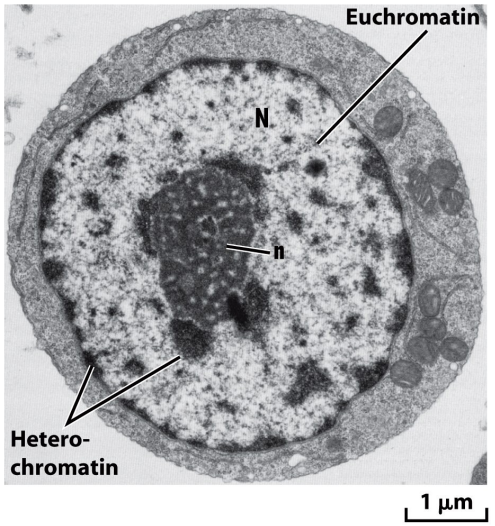
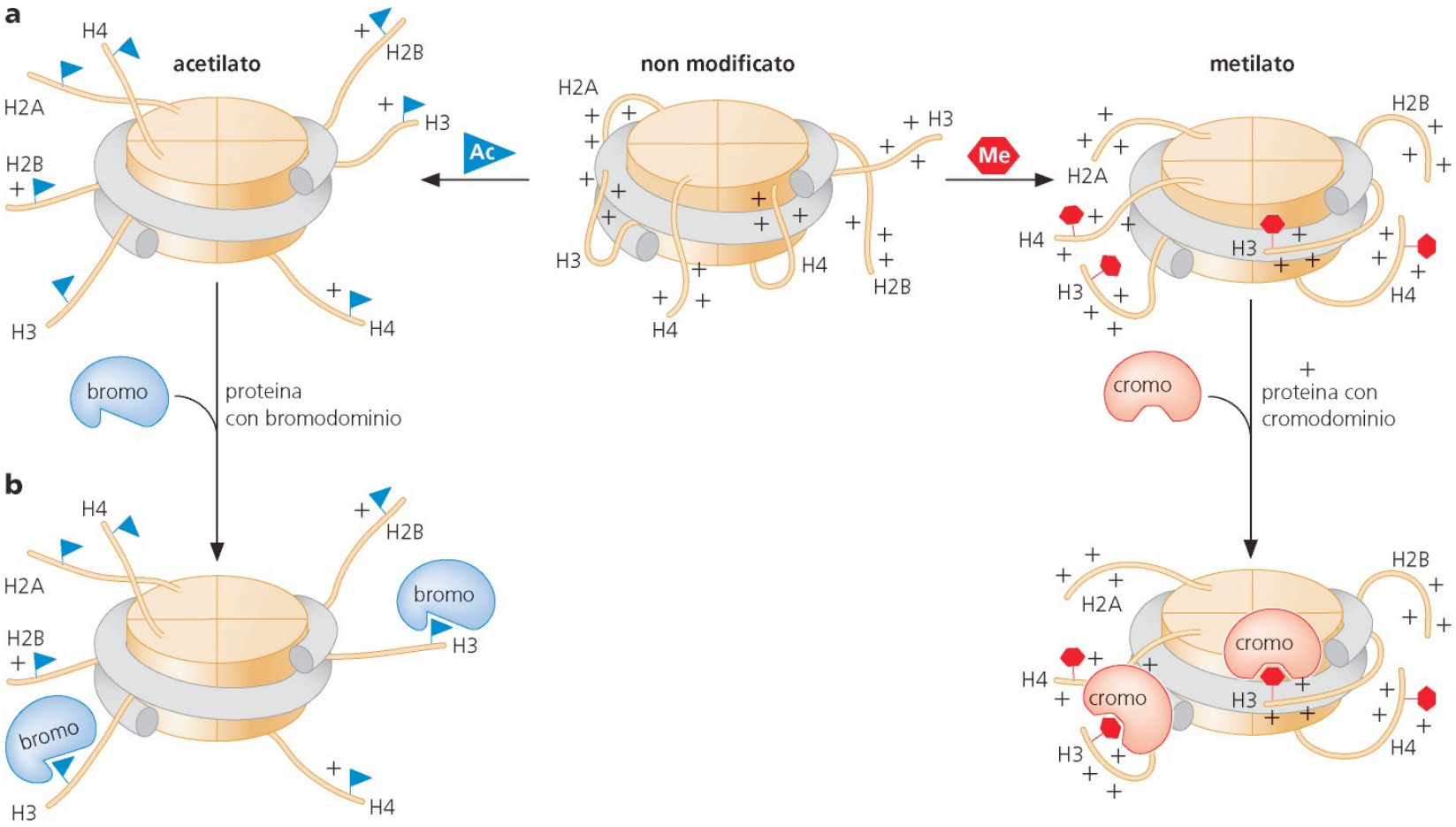
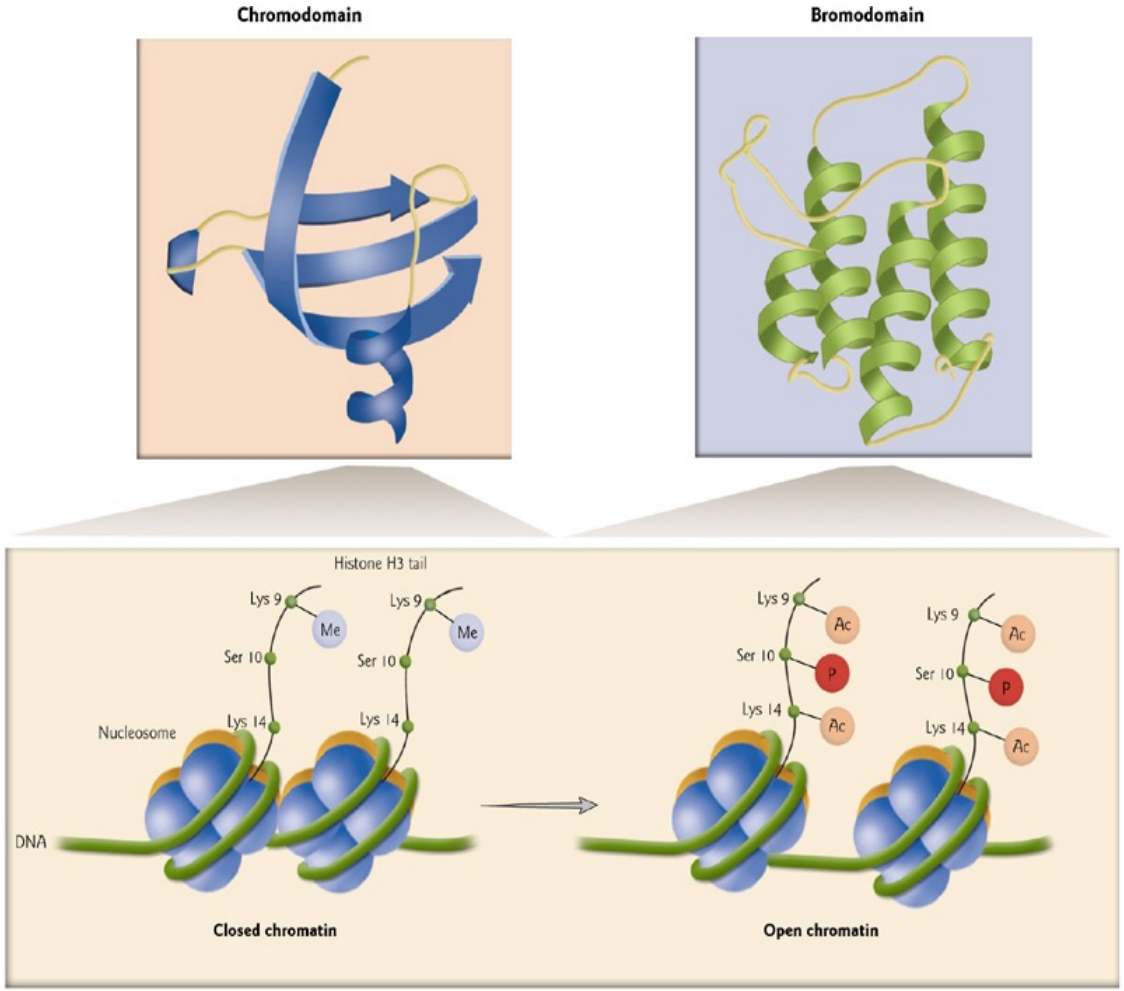


Figure 6-33a
Molecular Cell Biology, Sixth Edition

Histone Modifications are recognized by specific regulative proteins (readers)



Landmarks for Chromatin-binding Proteins



<u>Chromodomain</u>

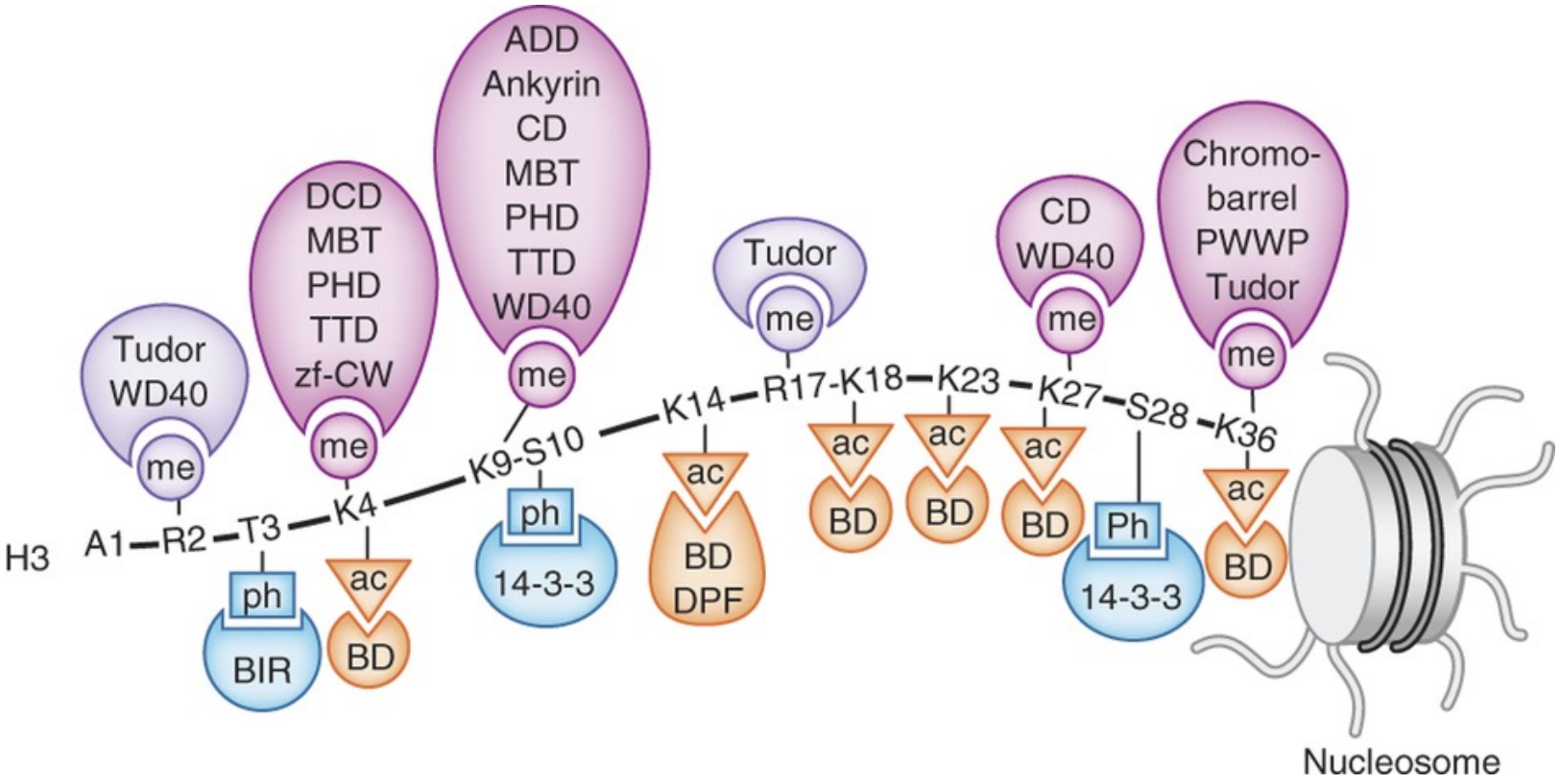
- CH_3(**Me**thyl)- recognition domain
- Targets to **Me-lys**

Bromodomain

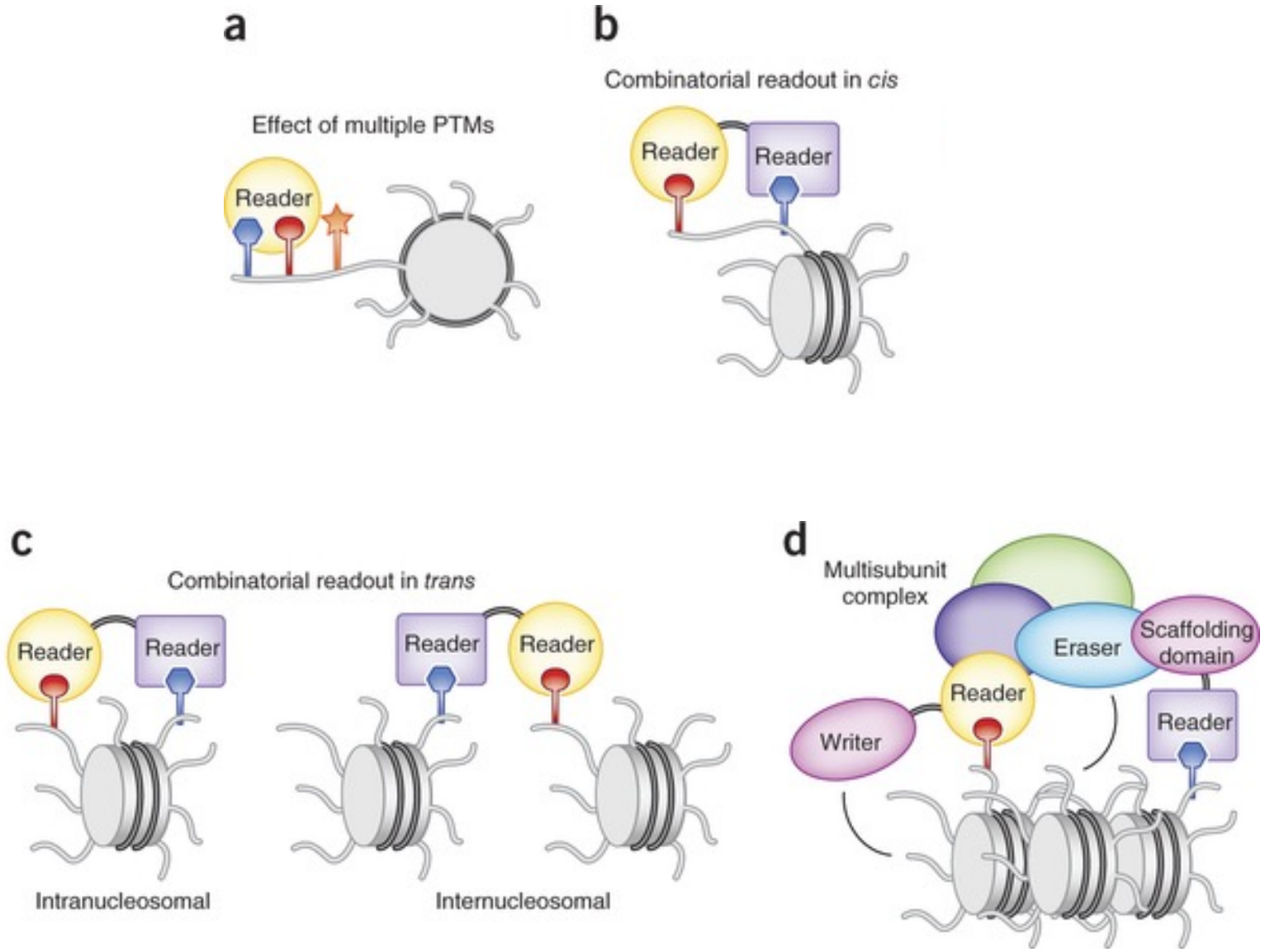
- Binds to acetylated lysines

The list of newly identified histone readers has grown rapidly

Recognition of the methylated (me) lysine, methylated (me) arginine, acetylated (ac) lysine and phosphorylated (ph) serine and threonine residues of the N-terminal histone H3 tail by indicated **readers**.



Combinatorial readout of PTMs



Crosstalk between chromatin marks

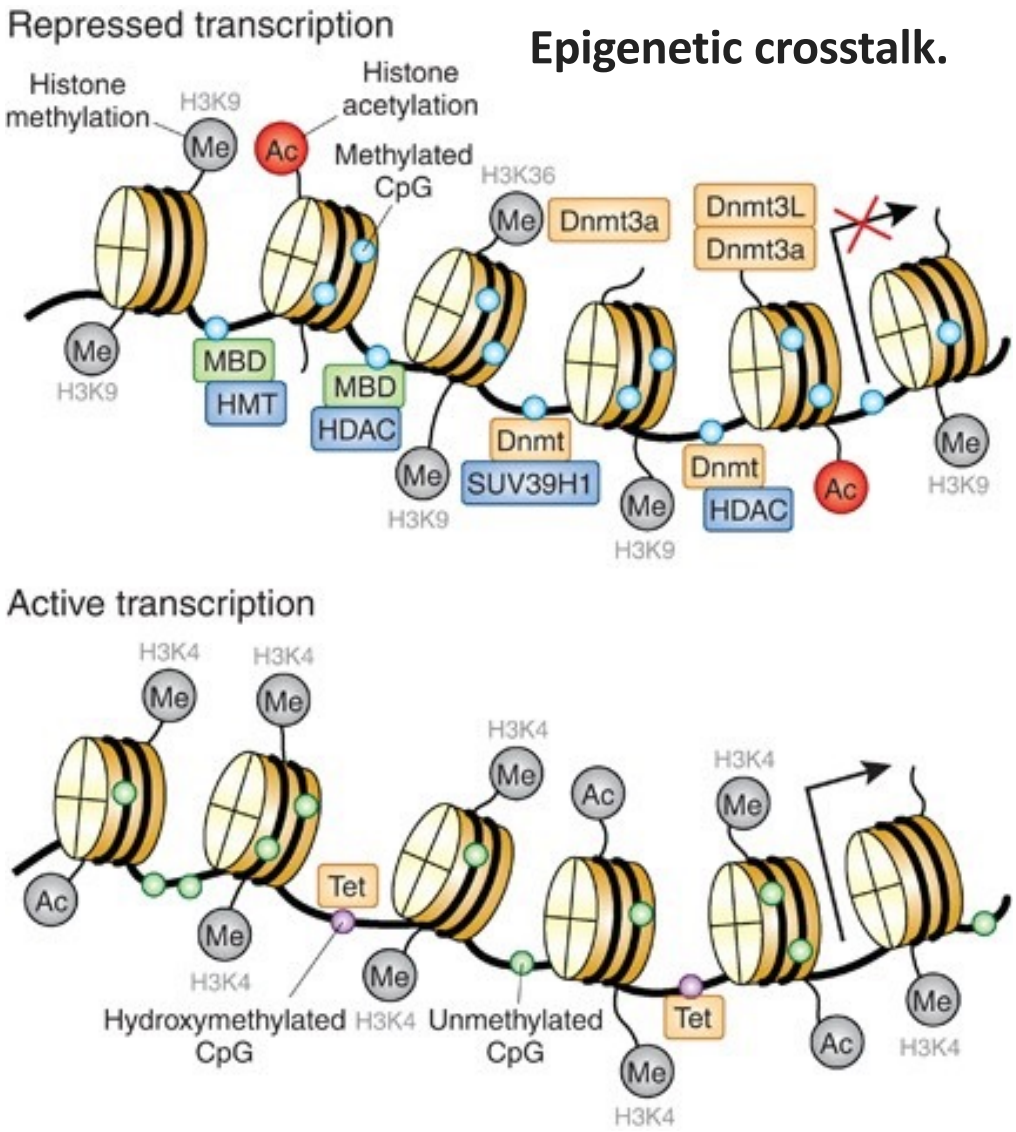
- Chromatin is typically marked by multiple modifications, and it is essential that these marks work in concert to achieve a coordinated cellular response. In some cases, the presence or absence of one modification can stimulate or inhibit deposition of another.

- Several chromatin regulators have important catalytically independent roles in setting chromatin modifications, often by recruiting other enzymes.

- Histone methylations are also highly coordinated with DNA methylation through multiple feedback loops .

Crosstalk of DNA Methylation and Other Epigenetic Mechanisms

Transcription is ultimately regulated by the interaction of multiple epigenetic mechanisms that cooperate to activate or silence gene expression.



Epigenetic crosstalk.

To suppress gene expression, Dnmts target CpG sites and actively methylate DNA. DNA methylation is recognized by methyl-binding proteins such as that along with Dnmts recruit enzymes that modify the histone tails including histone deacetylases (HDACs), which remove acetylation (Ac), and histone methyltransferases (HMTs), which methylate histones (Me) and in conjunction with DNA methylation serve to further repress gene expression.

In regions of DNA with activate transcription, Tet removes DNA methylation, and histone tails in this region often contain H3K4me^3 that inhibits Dnmt binding to unmethylated CpG sites and maintains a permissive environment for transcription.

Regulation of transcription by chromatin

1. Opening of chromatin through directed **modication of histone tails** (e.g. acetylation and methylation)

2. Histone variants

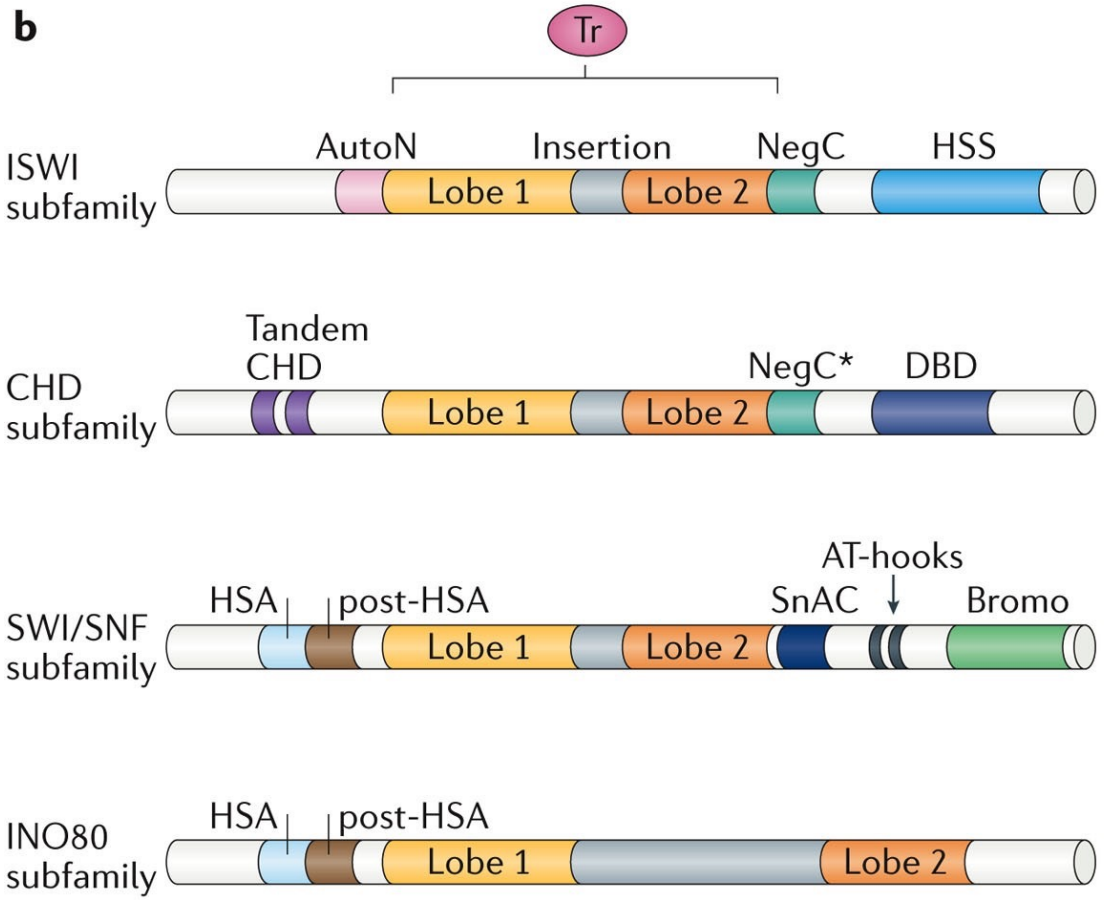
3. Opening of chromatin through directed **nucleosome mobilization** by **remodeling complexes**

4. **Positioning of nucleosomes** creates promoters with different requirement for remodeling

Three families of ATP-driven remodeling factors

Chromatin remodellers can be classified into four subfamilies.
Each subfamily is specialized to preferentially achieve particular chromatin outcomes: assembly, access or editing.

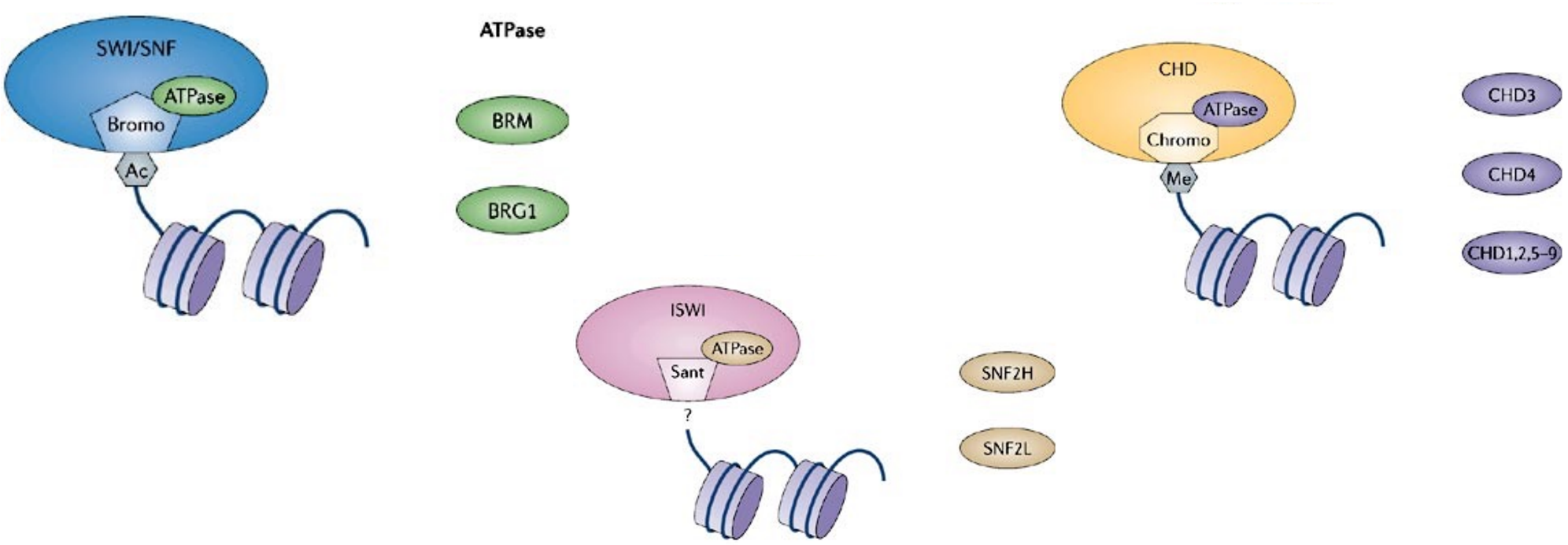
- **Swi/Snf**
- **ISWI**
- **Mi-2/CHD**
- **INO80**



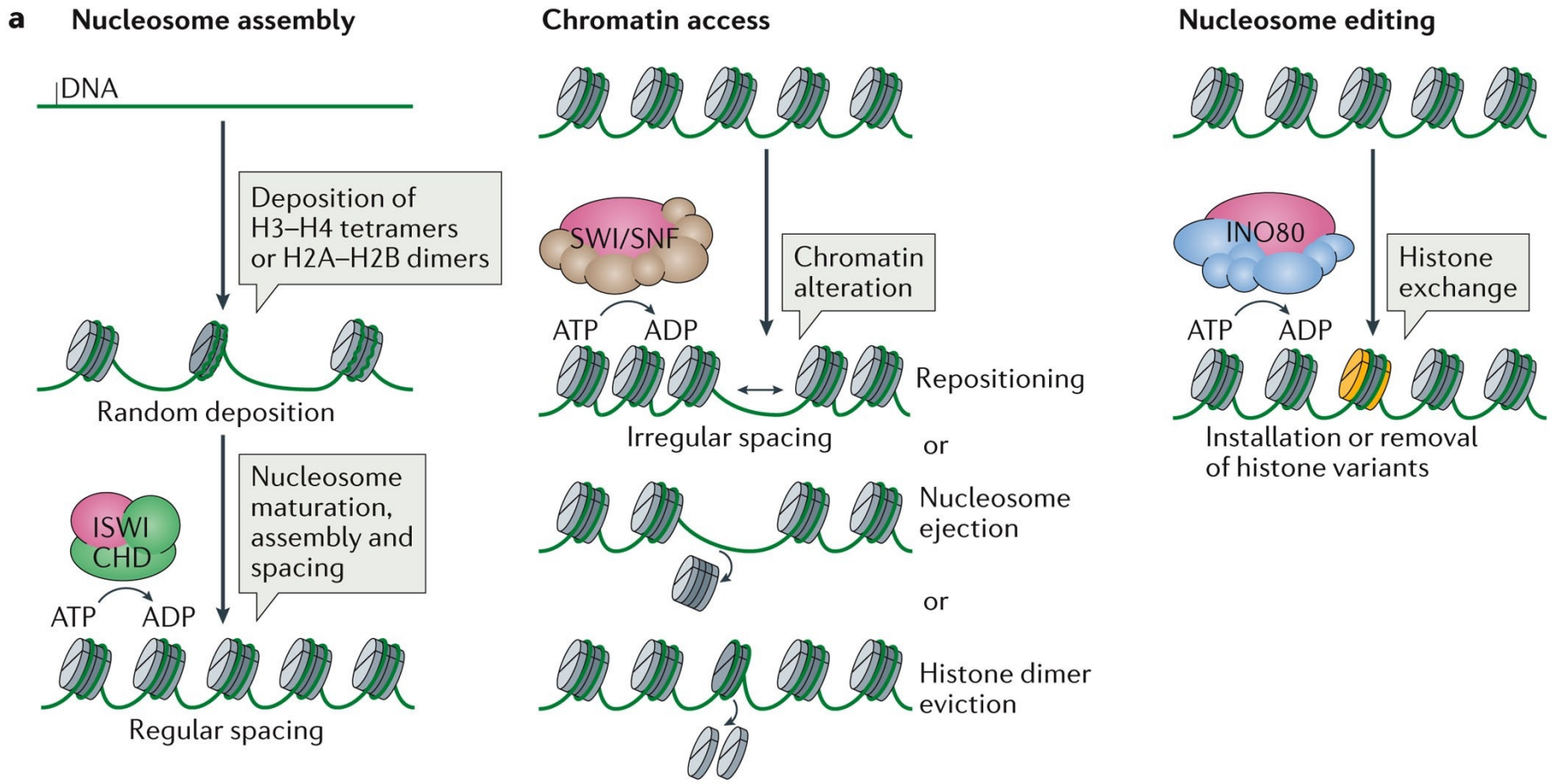
Clapier et sl., 2017 Nature Reviews Molecular Cell Biology

Recruitment

Each remodeliling complex has a unique domain (bromo, chromo and sant), and these are known or thought to interact with specific chromatin substrates. Moreover, they interact only on a subset of genes (Genome-analysis in yeast shows that <5% of all gener require Swi/Snf). Diversity in the protein composition of remodellers enables their specific interaction with particular transcription activators, repressors and histone modifications, which together specify targeting.



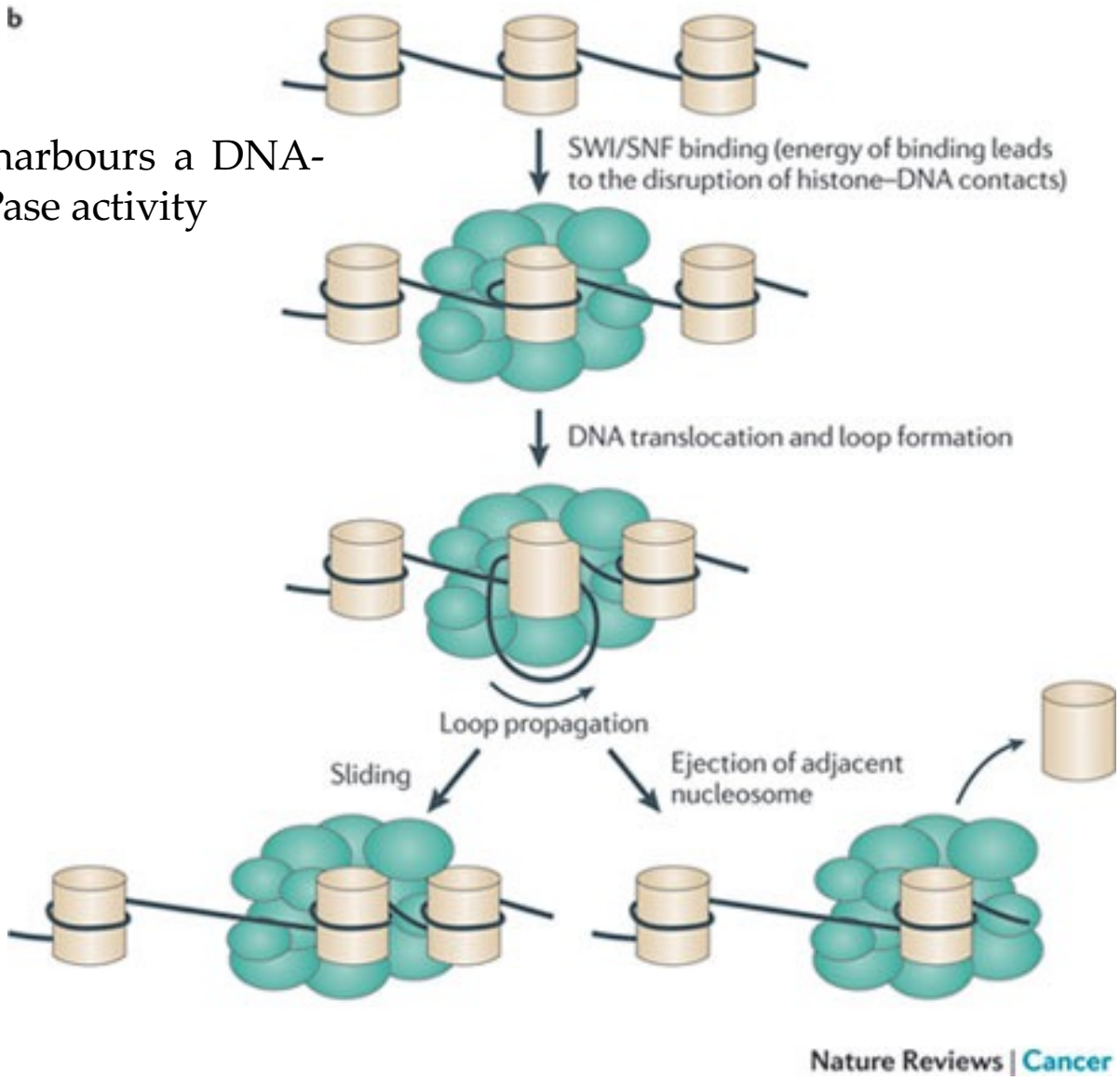
Functions of chromatin remodellers



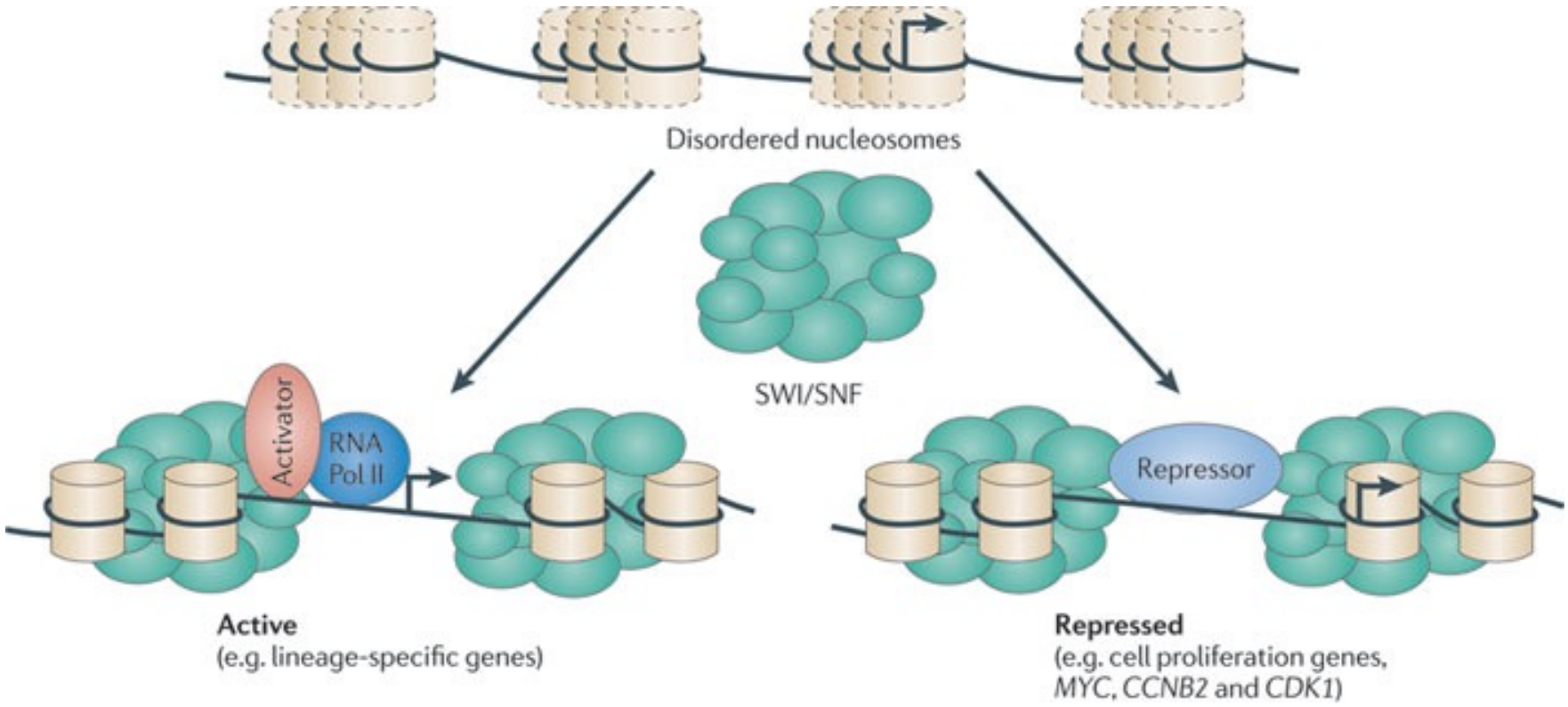
Clapier et sl., 2017 Nature Reviews Molecular Cell Biology

Possible activities of Swi/Snf

SWI2-subunit harbours a DNA-stimulated ATPase activity



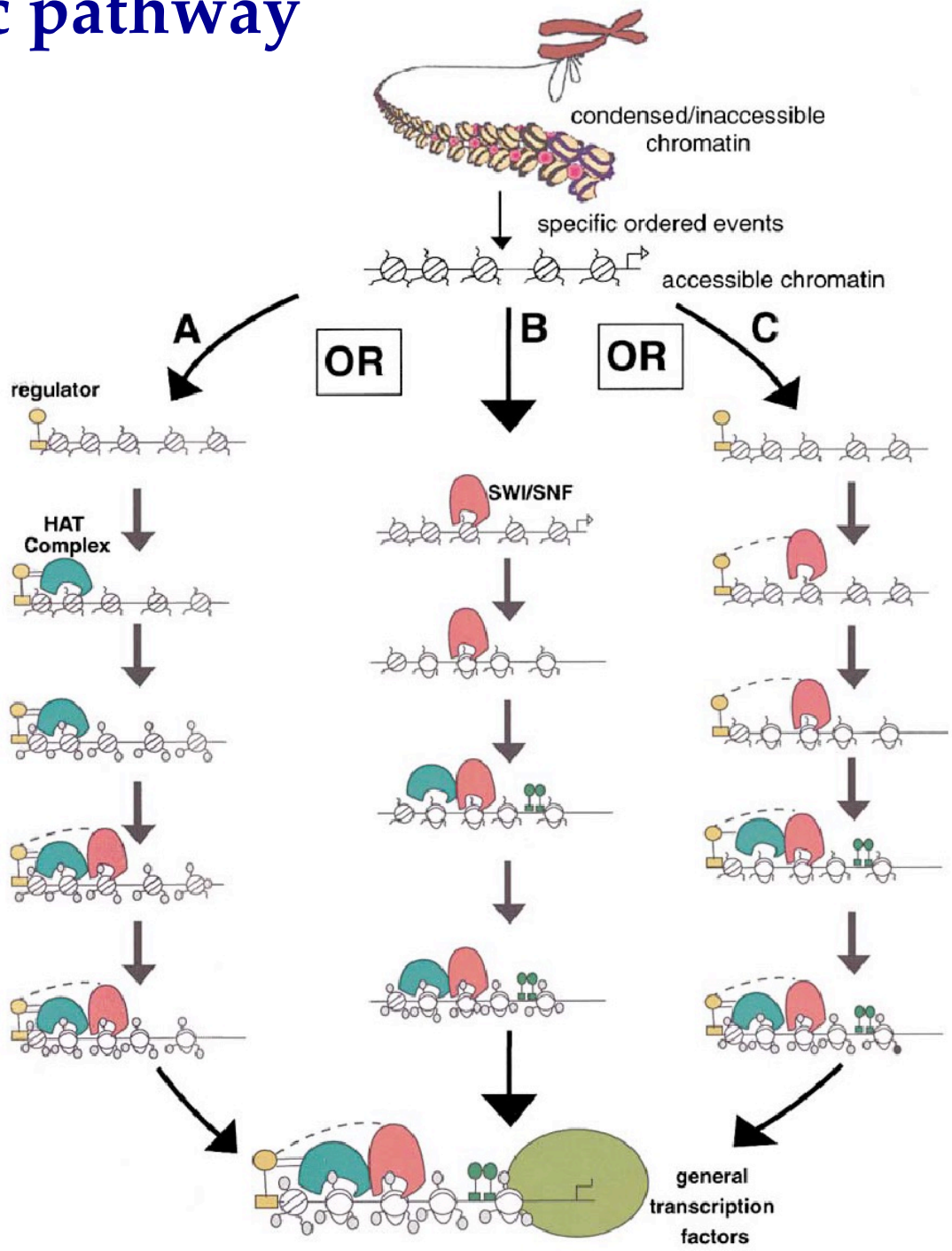
Possible activities of Swi/Snf



Each gene a specific pathway

- It appears that there is no obligate order for function of ATP-dependent remodelers and covalent modifiers that is general for all promoters.

- Rather, it seems that each individual promoter will work using a set order of action by these complexes that differs from promoter to promoter.



Mutations in regulators of the epigenome identified in cancer