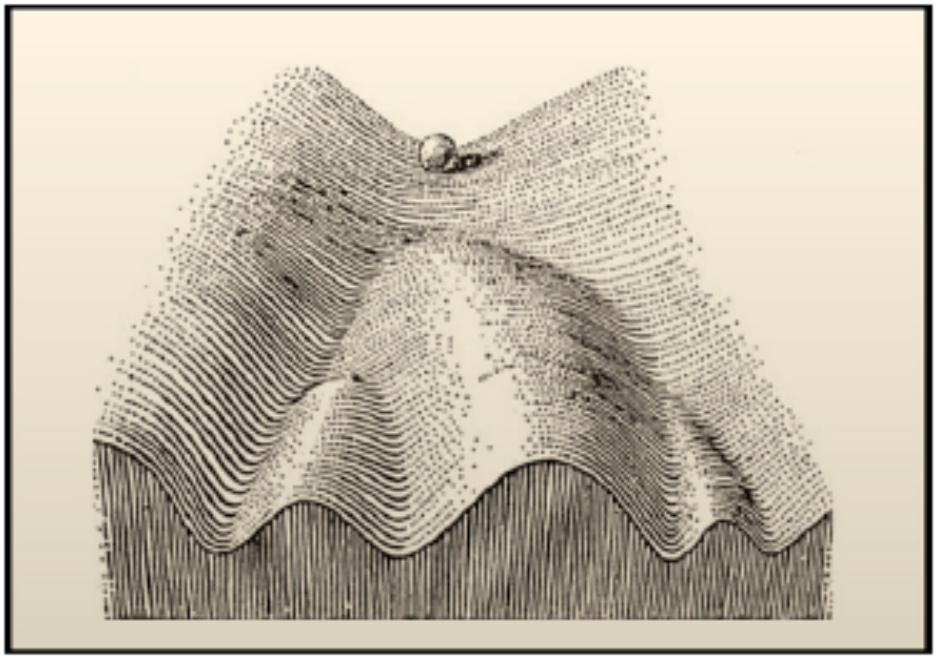
Epigenetics- Interaction of the genes with their environment

Heritable changes in phenotype or gene expression caused by mechanisms other than changes in the underlying DNA sequence. These changes may remain through cell divisions for the remainder of the cell's life and may also last for multiple generations.



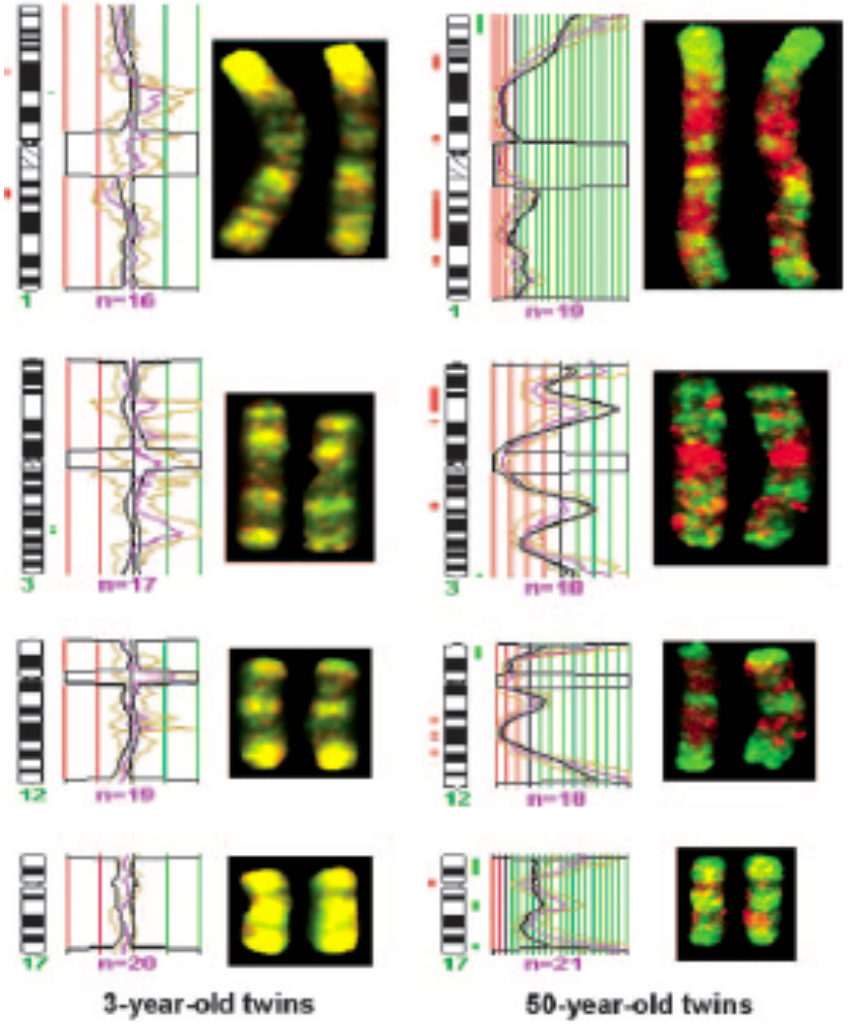
Waddington's classical epigenetic landscape

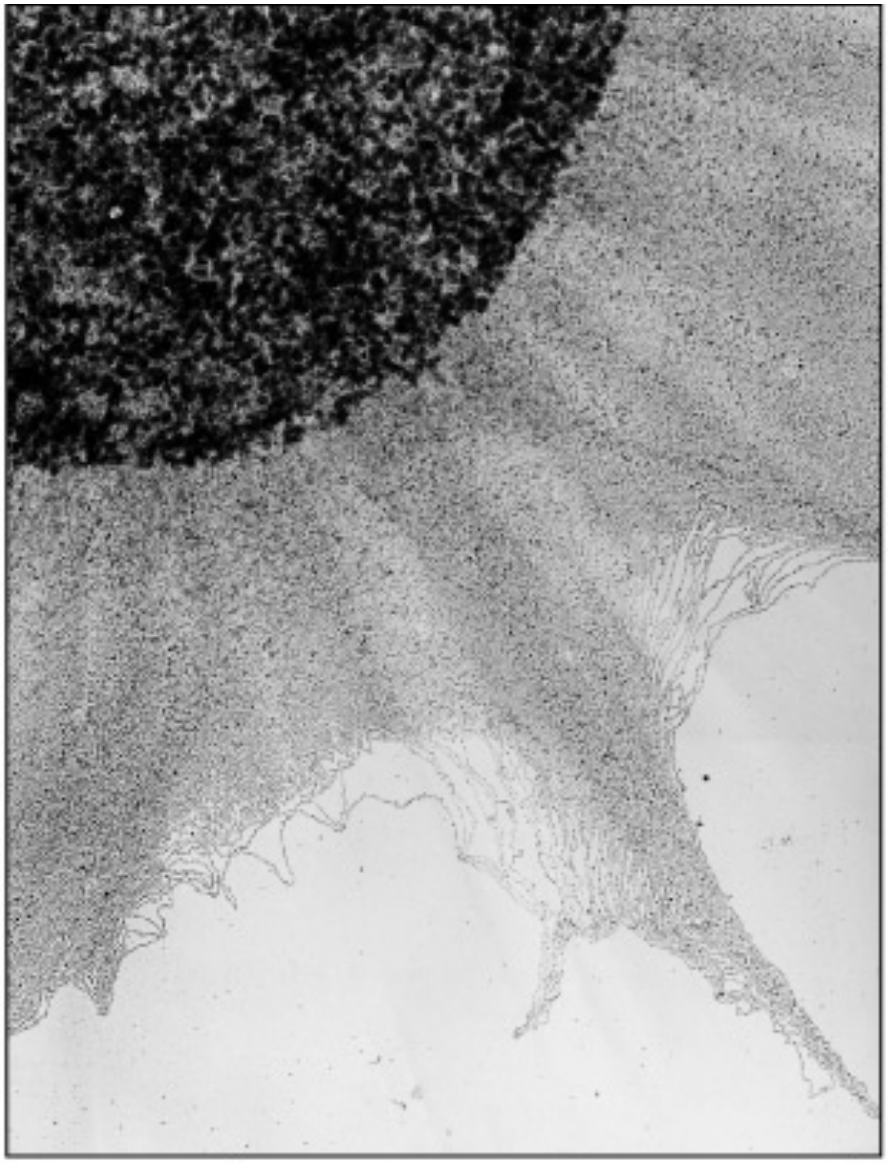
The word "epigenetics" (as in "epigenetic landscape") was coined by C. H. Waddington in 1942 as a fusion of the words "genetics" and "epigenesis". Epigenesis is an older word used to describe the differentiation of cells from a totipotent state in embryonic development. At the time Waddington first used the term "epigenetics," the physical nature of genes and their role in heredity was not known. **Epigenetics was Waddington's model of how genes within a multicellular organism interact with their surroundings to produce a phenotype.** Because all cells within an organism inherit the same DNA sequences, cellular differentiation processes crucial for epigenesis rely strongly on epigenetic rather than genetic inheritance.

Epigenetic differences arise during the lifetime of monozygotic twins

Mario F. Fraga*, Esteban Ballestar*, Maria F. Paz*, Santiago Ropero*, Fernando Setien*, Maria L. Ballestar†, Damia Heine-Suñer‡, Juan C. Cigudosa§, Miguel Urioste¶, Javier Benitez¶, Manuel Boix-Chornet†, Abel Sanchez-Aguilera†, Charlotte Ling||, Emma Carlsson||, Pernille Poulsen**, Allan Vaag**, Zarko Stephan††, Tim D. Spector††, Yue-Zhong Wu‡‡, Christoph Plass‡‡, and Manel Esteller*§§

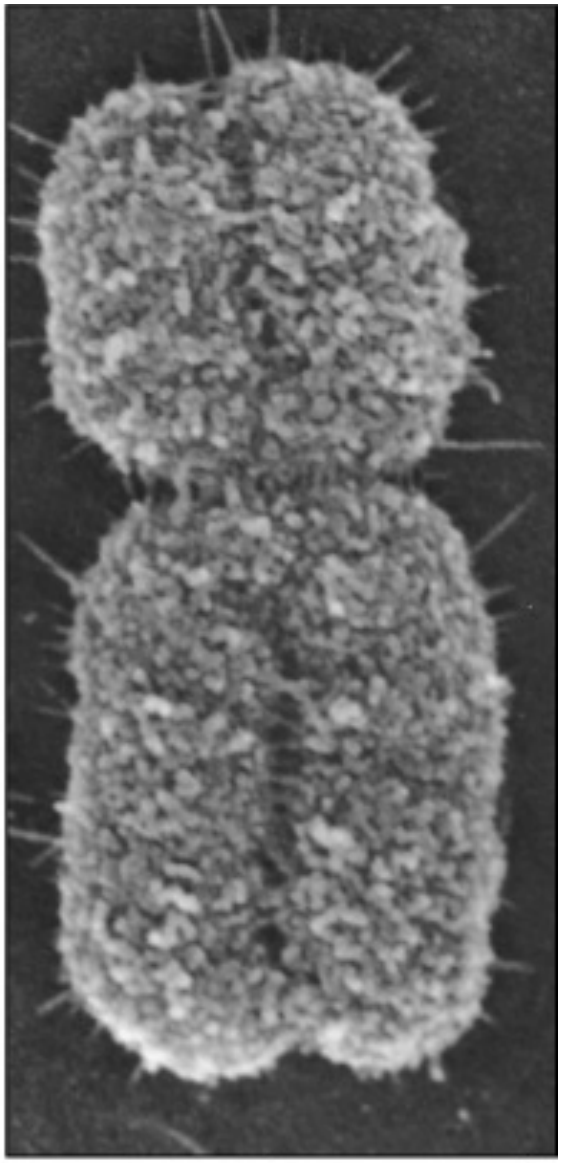
"Epigenomics is where genomics was 30 years ago, when everyone was working on part of the puzzle."
— Peter Jones





10 µm

(A)



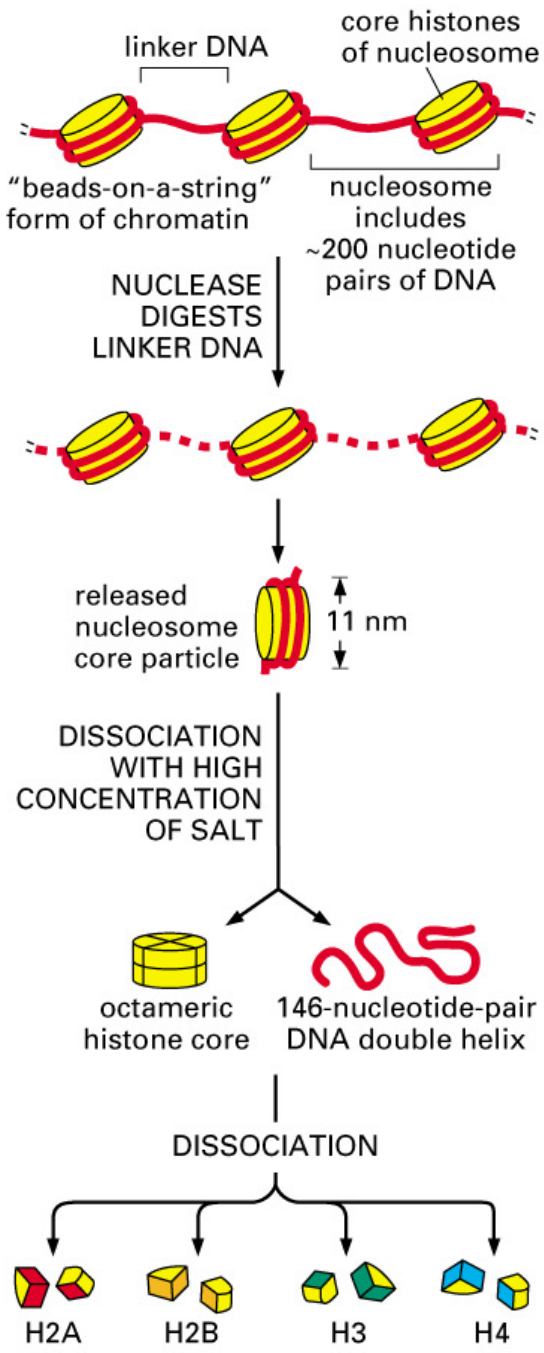
1 µm

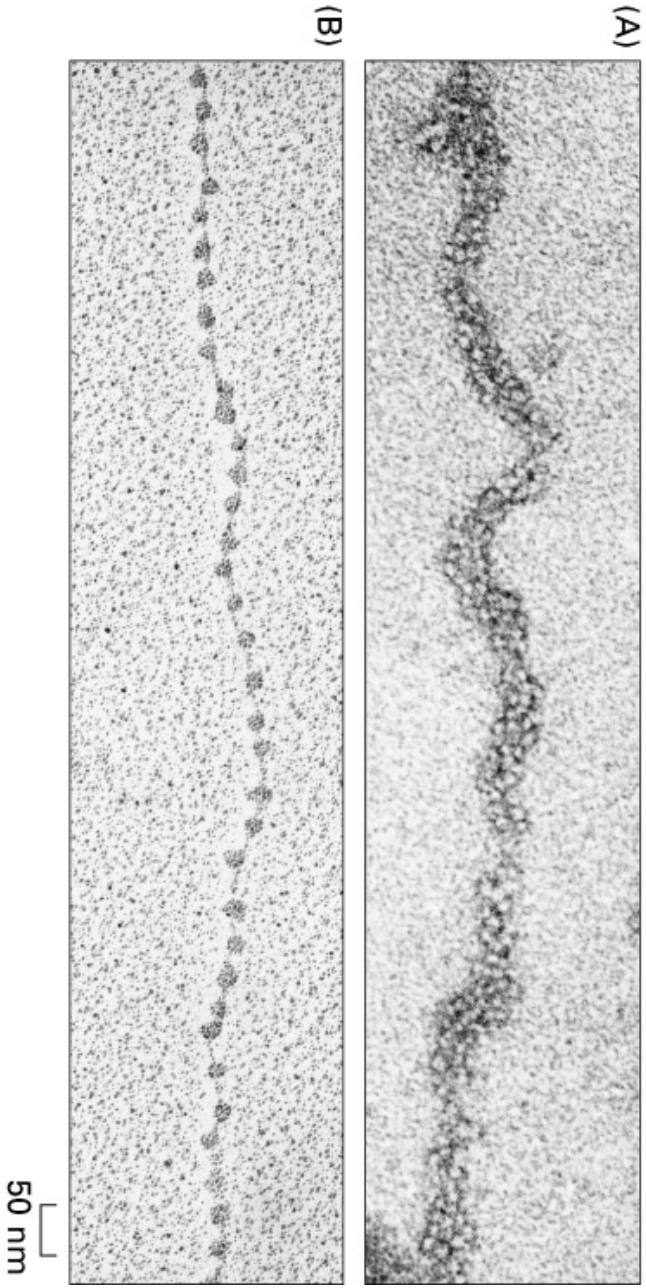
(B)

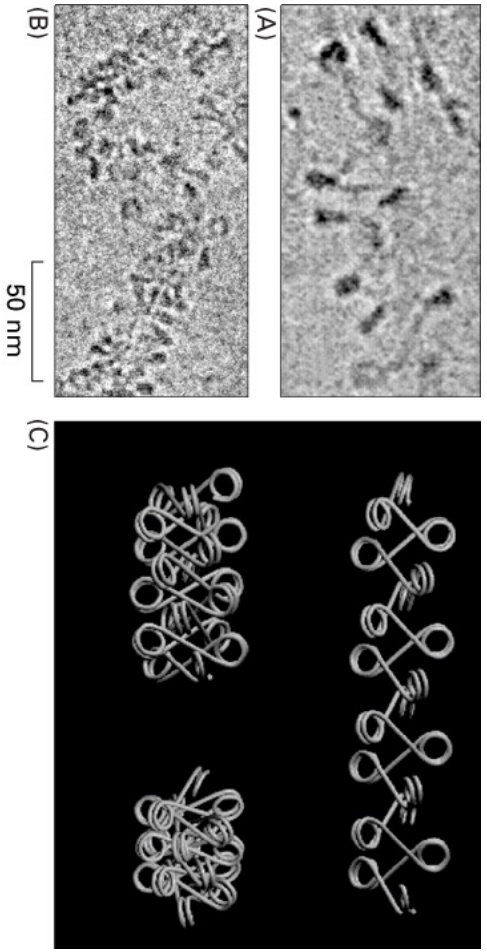
The proteins that bind to the DNA to form eucaryotic chromosomes are traditionally divided into two general classes: the **histones** and the **nonhistone chromosomal proteins**. The complex of both classes of protein with the nuclear DNA of eucaryotic cells is known as **chromatin**.

Histones are present in such enormous quantities in the cell (about 60 million molecules of each type per human cell) that their total mass in chromatin is about equal to that of the DNA.

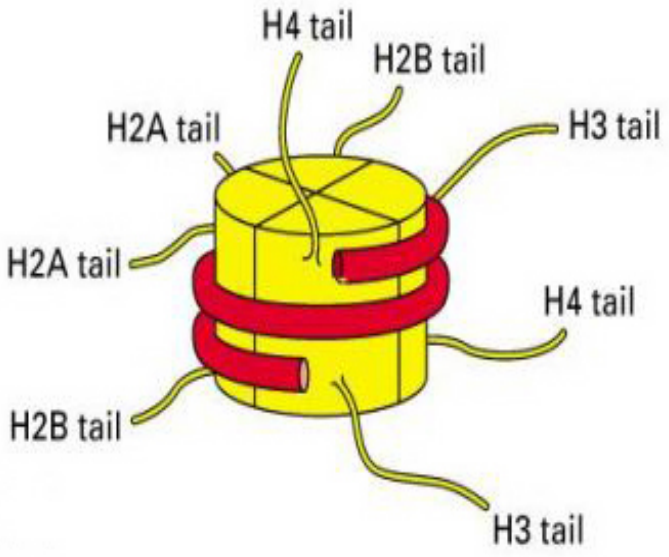
Histones are responsible for the first and most basic level of chromosome organization, the nucleosome


linker DNA
core histones of nucleosome
"beads-on-a-string" form of chromatin
nucleosome includes ~200 nucleotide pairs of DNA
NUCLEASE DIGESTS LINKER DNA
released nucleosome core particle
11 nm
DISSOCIATION WITH HIGH CONCENTRATION OF SALT
octameric histone core
146-nucleotide-pair DNA double helix
DISSOCIATION
H2A
H2B
H3
H4


(B)
(A)
50 nm

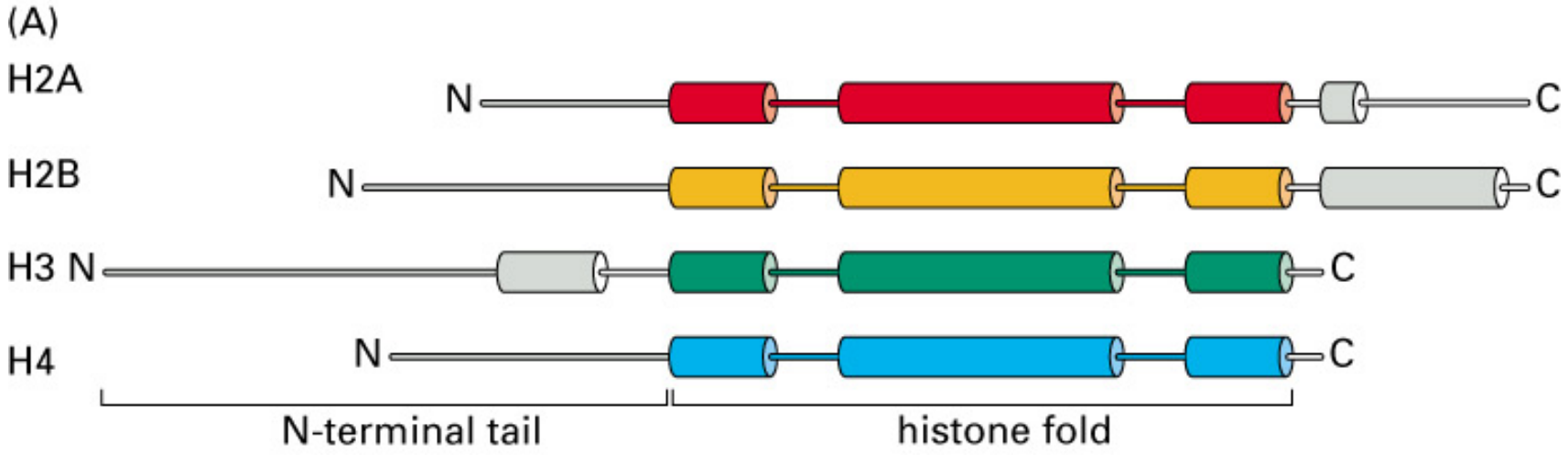

(B)
(A)
50 nm
(C)

NUCLEOSOME - THE FUNDAMENTAL UNIT OF CHROMATIN

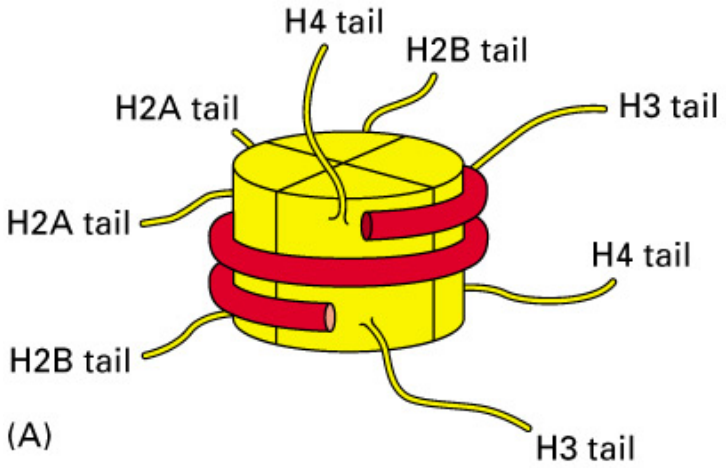


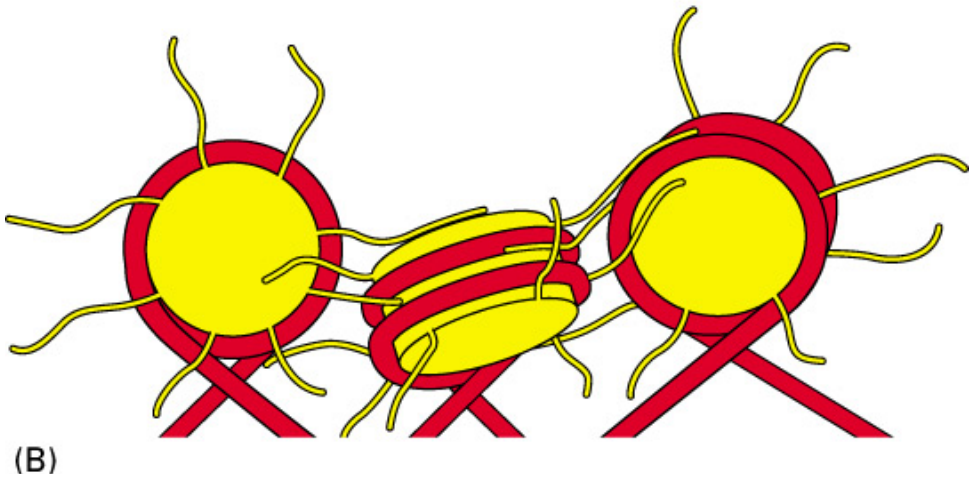
Histone modifications occur at the N-terminal tails of histones and are highly dynamic processes

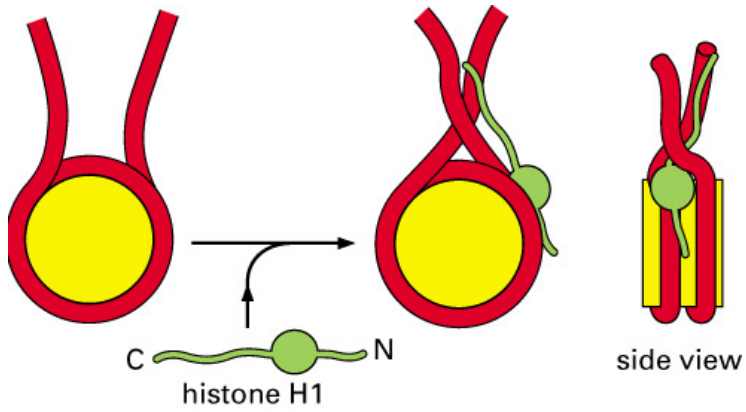
- Octamer of four core histones (H3 H4 H2A and H2B) with 147 base pairs of DNA wrapped around
- Core histones are predominantly globular except for their N-terminal tails which are unstructured



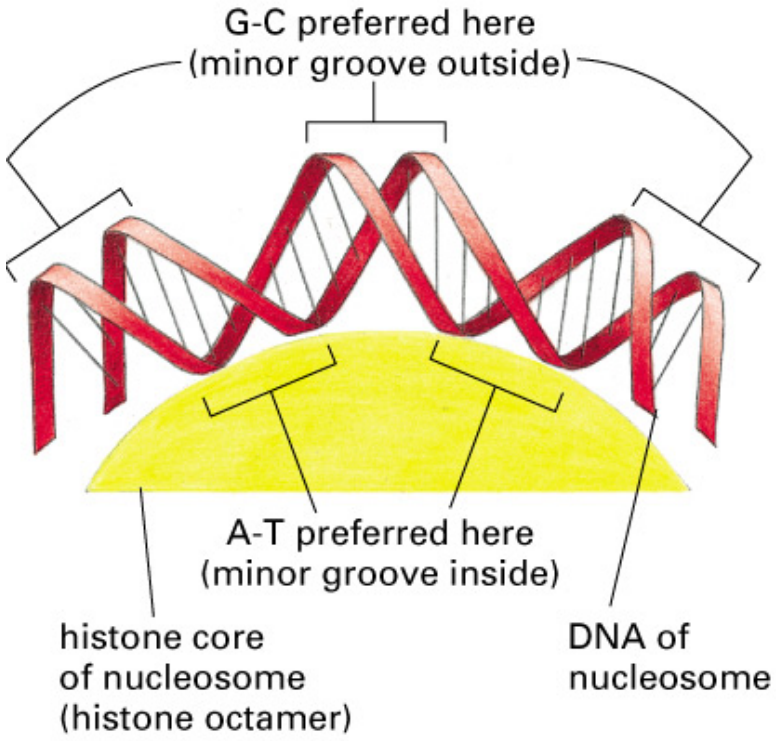
H1 hystone



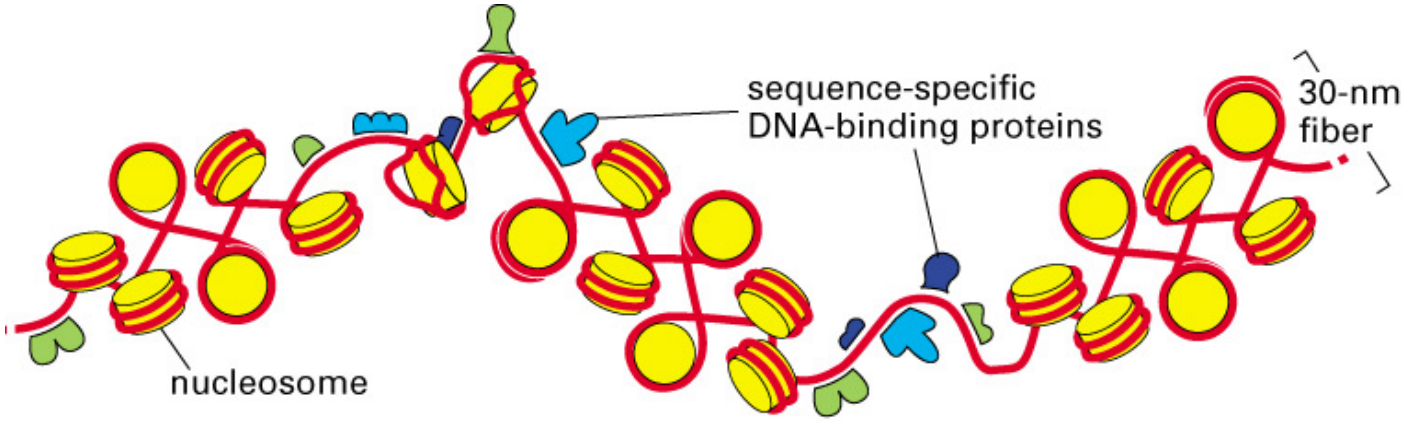


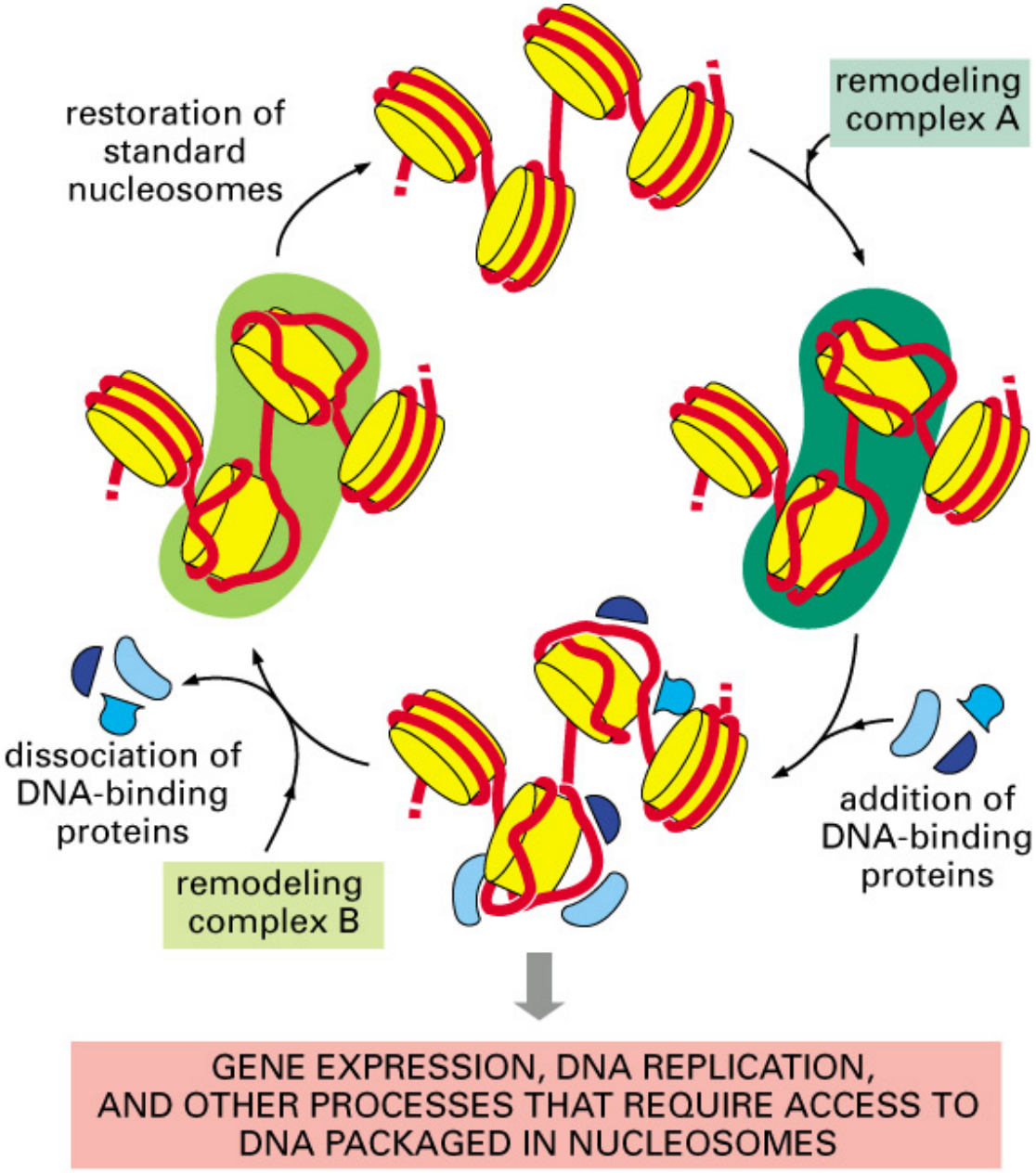


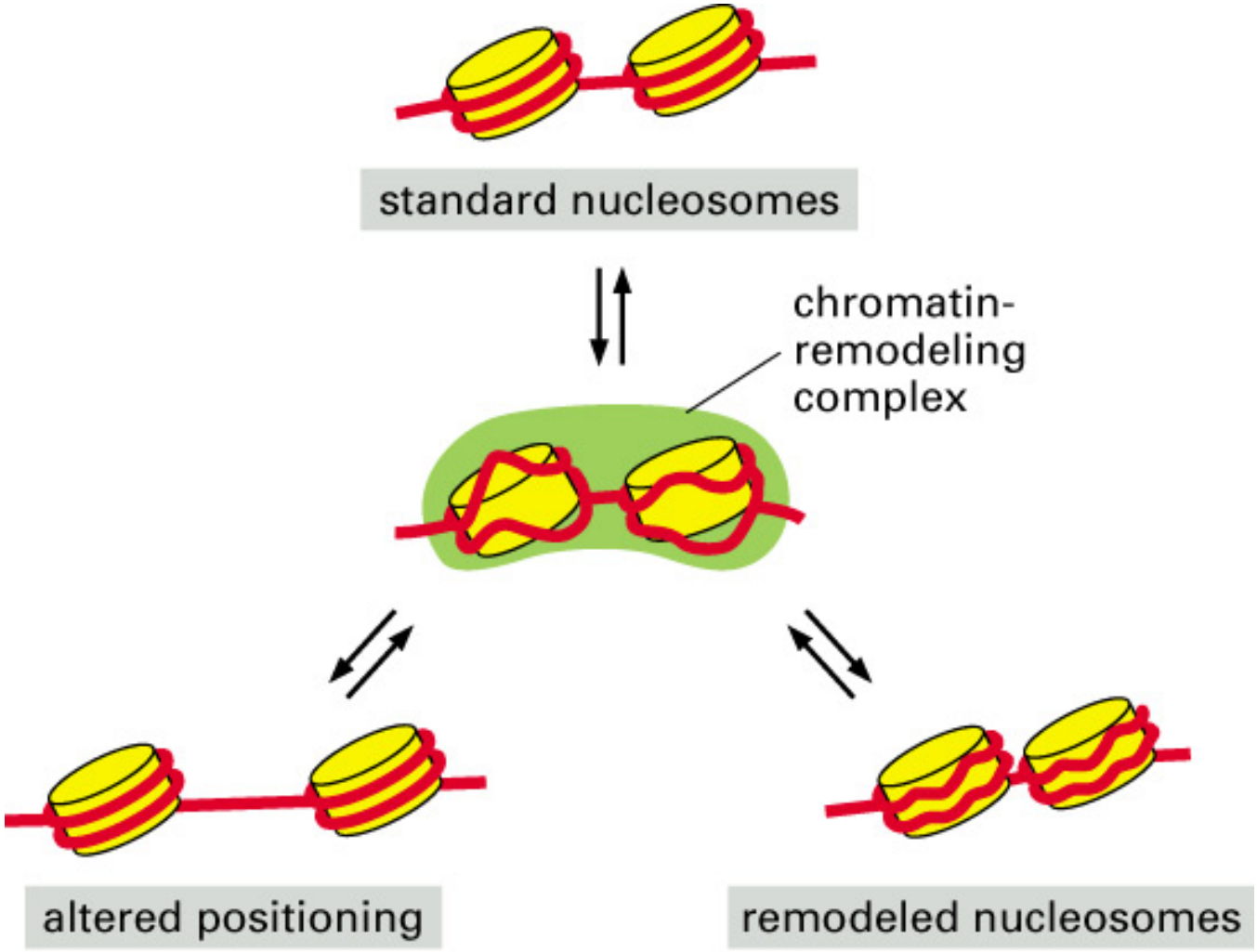
Histone H1 consists of a globular core and two extended tails.
Part of the effect of H1 on the **compaction** of nucleosome organization may result from *charge neutralization*: like the core histones, H1 is positively charged (especially its C-terminal tail), and this helps to compact the negatively charged DNA.
Unlike the core histones, H1 **does not seem to be essential** for cell viability; in one ciliated protozoan the nucleus expands nearly twofold in the absence of H1, but the cells otherwise appear normal



Each nucleosome core particle is separated from the next by a region of linker DNA, which can vary in length : on average, therefore, nucleosomes repeat at intervals of about 200 nucleotide pairs.

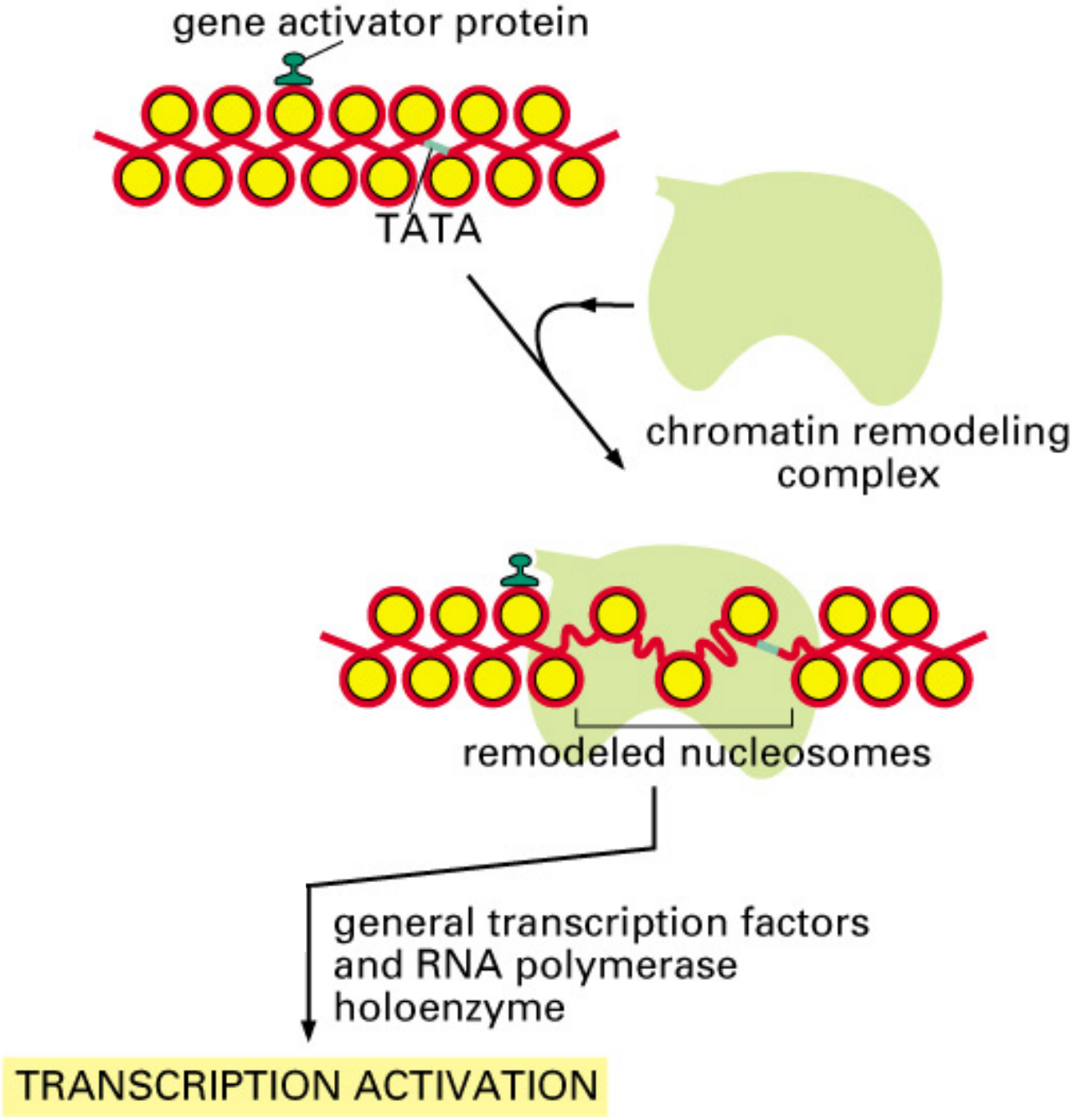


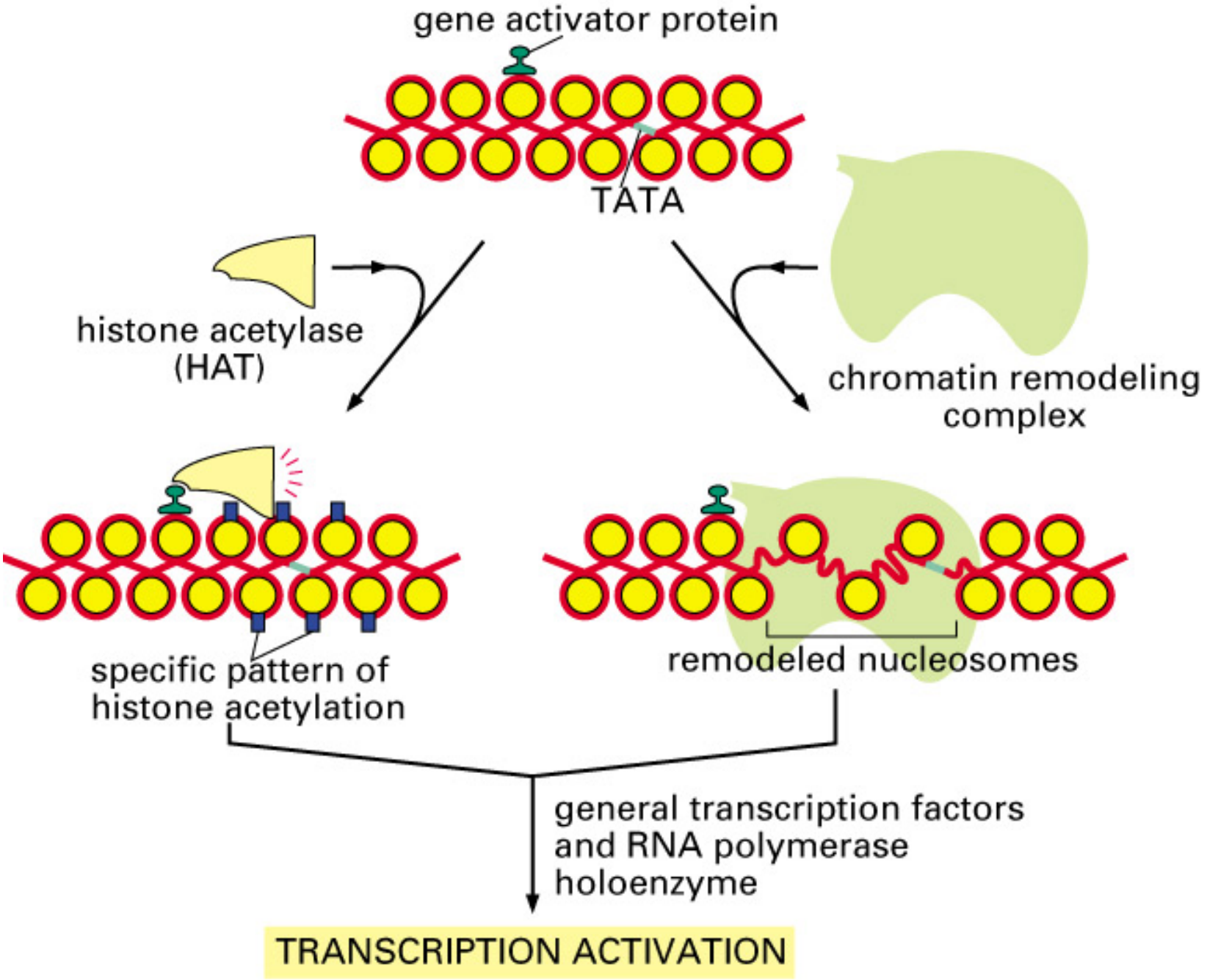

restoration of standard nucleosomes
remodeling complex A
addition of DNA-binding proteins
remodeling complex B
dissociation of DNA-binding proteins
GENE EXPRESSION, DNA REPLICATION, AND OTHER PROCESSES THAT REQUIRE ACCESS TO DNA PACKAGED IN NUCLEOSOMES



Chromatin remodelling complexes:

protein machines that use the energy of ATP hydrolysis to change the structure of nucleosomes temporarily so that DNA becomes less tightly bound to the histone core


gene activator protein
TATA
chromatin remodeling complex
remodeled nucleosomes
general transcription factors and RNA polymerase holoenzyme
TRANSCRIPTION ACTIVATION

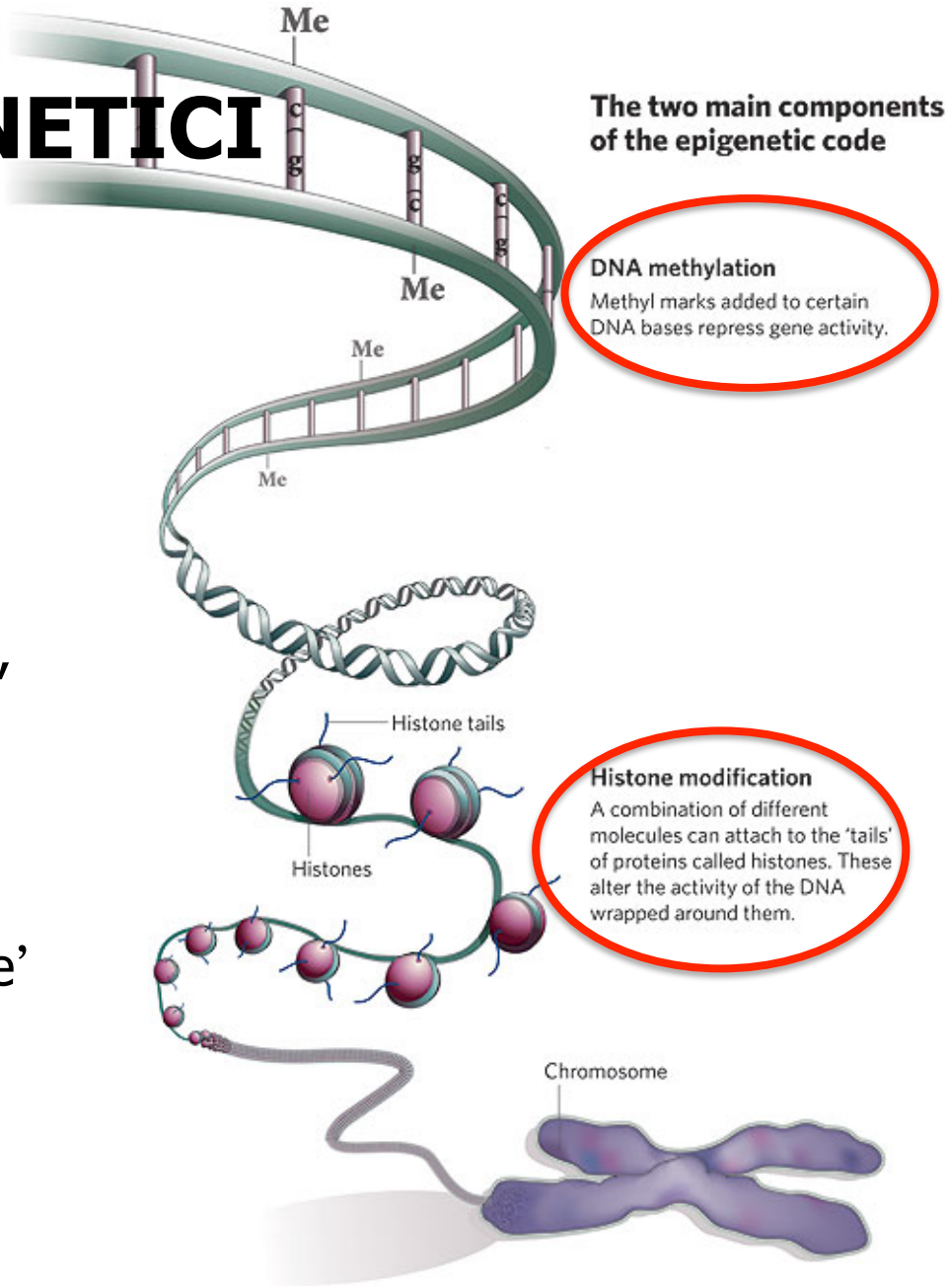

gene activator protein
TATA
histone acetylase
(HAT)
chromatin remodeling
complex
specific pattern of
histone acetylation
remodeled nucleosomes
general transcription factors
and RNA polymerase
holoenzyme
TRANSCRIPTION ACTIVATION

MECCANISMI EPIGENETICI

Fattori che vengono trasmessi alla progenie, ma che non sono direttamente attribuibili a sequenze di DNA.

<u>MODIFICAZIONI DEGLI ISTONI</u>
Acetilazioni, fosforilazioni e metilazioni, responsabili dei cambiamenti conformazionali della cromatina.

<u>METILAZIONE DEL DNA</u>
Nelle cells eucariotiche la metilazione e' a carico della G. Solo il 3% delle C e' metilato; in genere e' bersaglio della metilazione la C delle doppiette CpG.

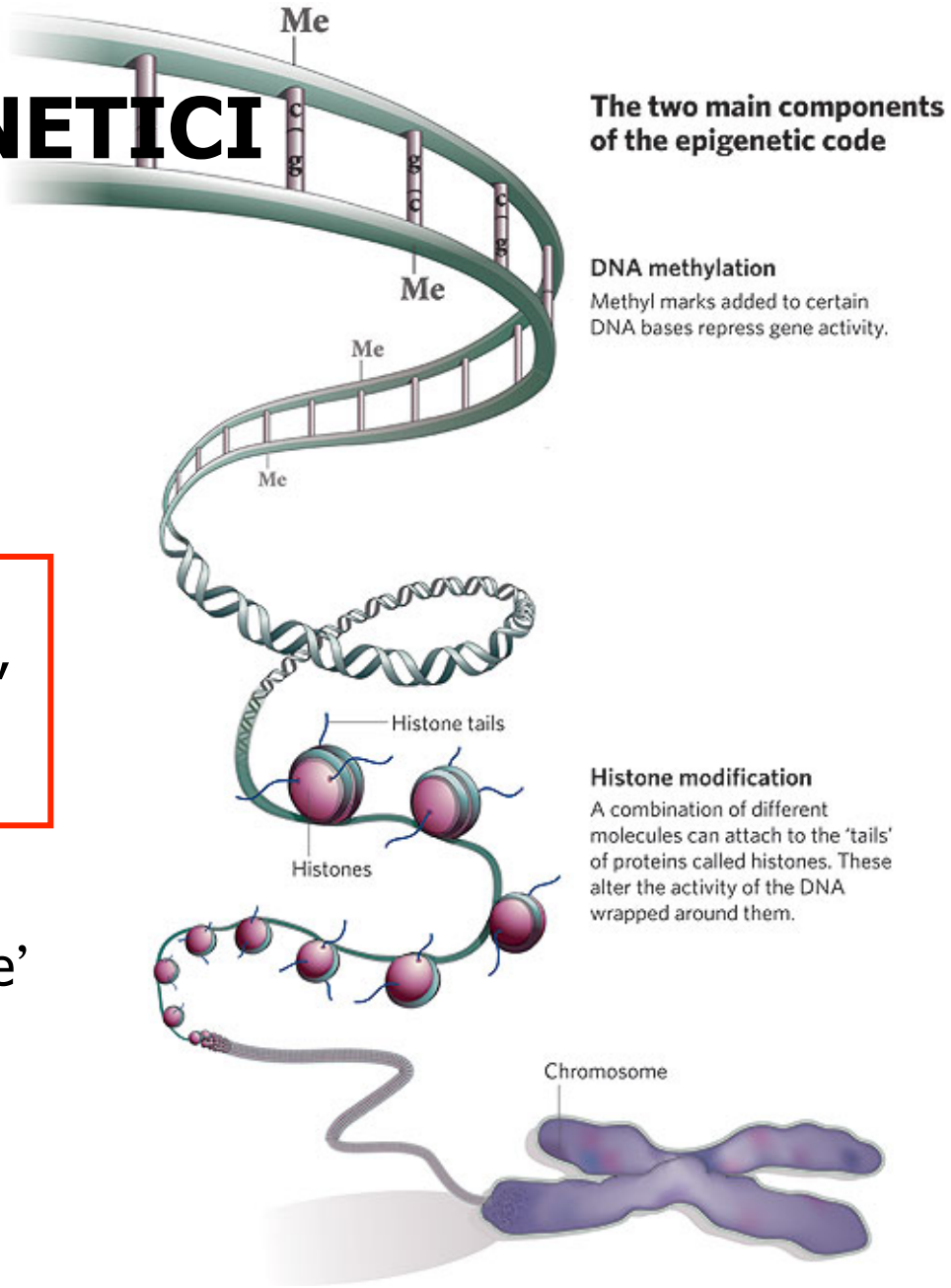


MECCANISMI EPIGENETICI

Fattori che vengono trasmessi alla progenie, ma che non sono direttamente attribuibili a sequenze di DNA.

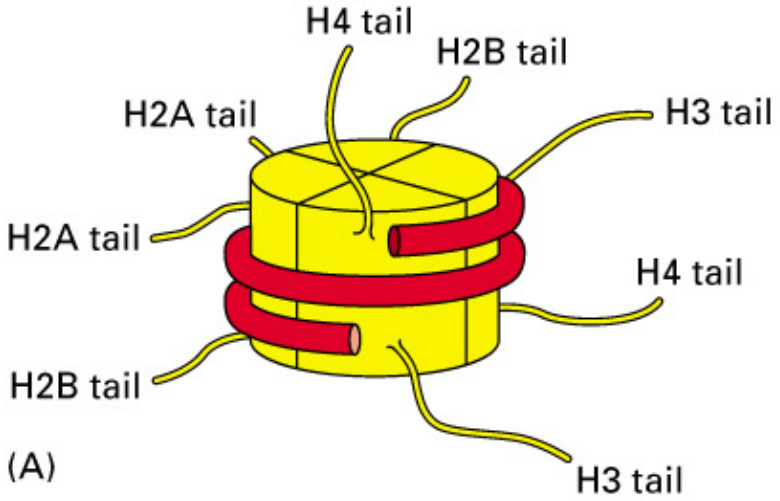
MODIFICAZIONI DEGLI ISTONI
Acetilazioni, fosforilazioni e metilazioni, responsabili dei cambiamenti conformazionali della cromatina.

METILAZIONE DEL DNA
Nelle cells eucariotiche la metilazione e' a carico della G. Solo il 3% delle C e' metilato; in genere e' bersaglio della metilazione la C delle doppiette CpG.

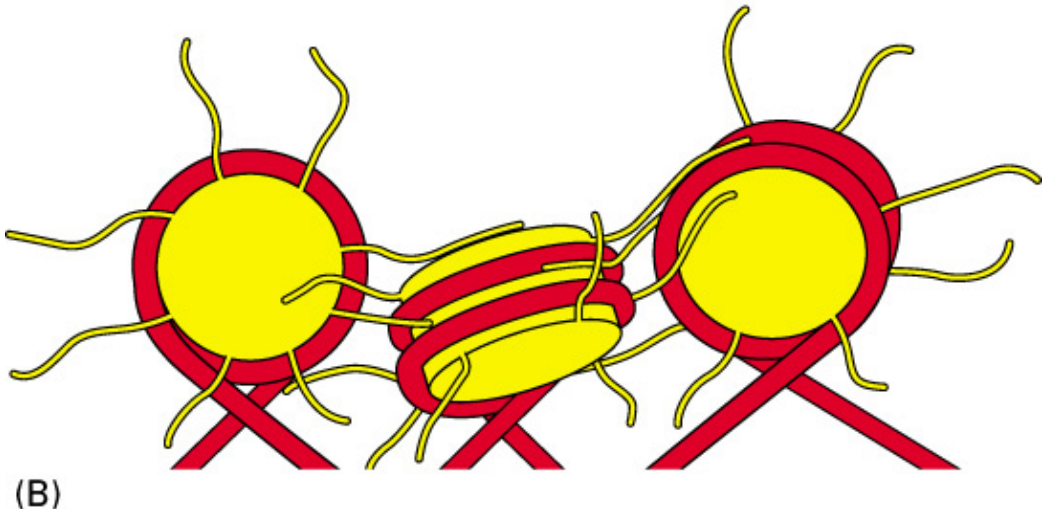


Le **code N-terminali** degli istoni sporgono dal nucleo dell'ottamero

Le modificazioni chimiche degli istoni forniscono siti di legame per proteine che possono cambiare lo stato della cromatina in attivo o inespresso

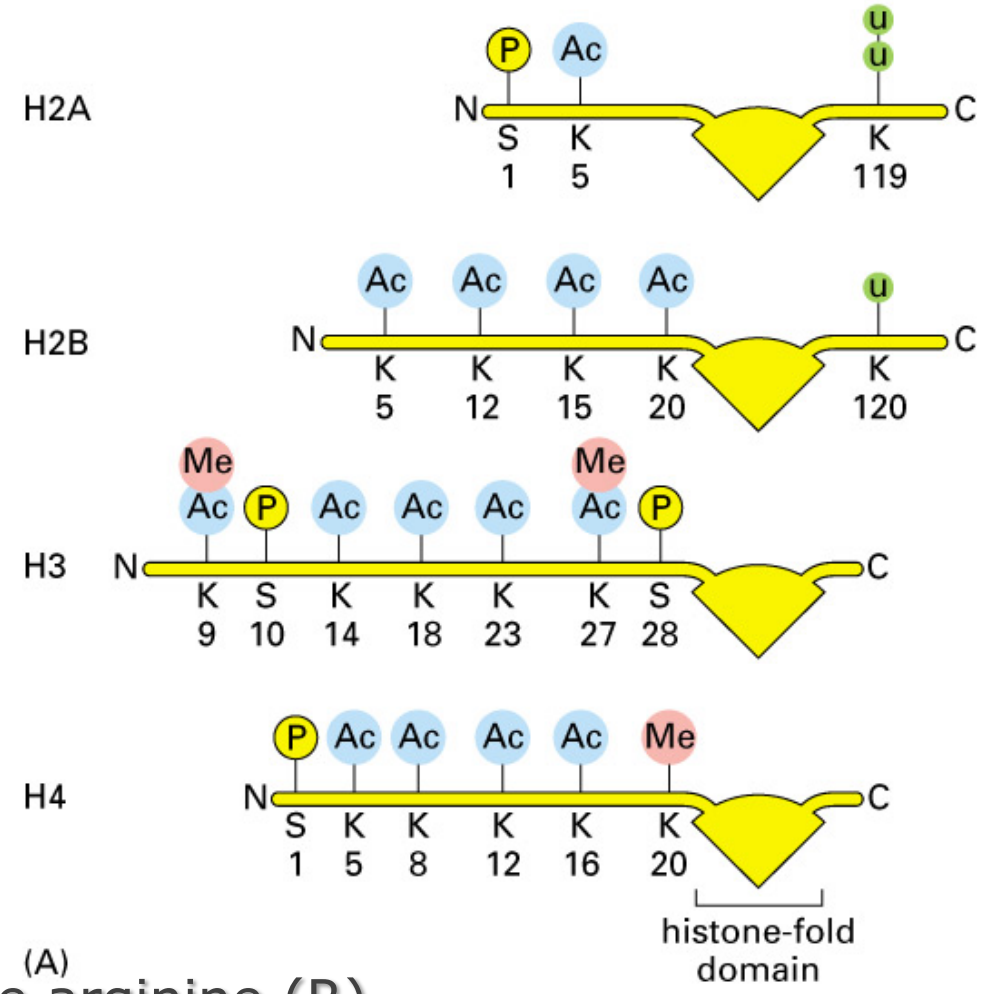


Una particolare combinazione di tali modificazioni ha un significato biologico (CODICE ISTONICO)



The histone code

L'ipotesi del **codice istonico** propone che modificazioni covalenti post-traduzionali delle code degli istoni vengano "lette" dalla cellula portando ad un risultato trascrizionale combinatorio complesso



Modificazioni possibili:

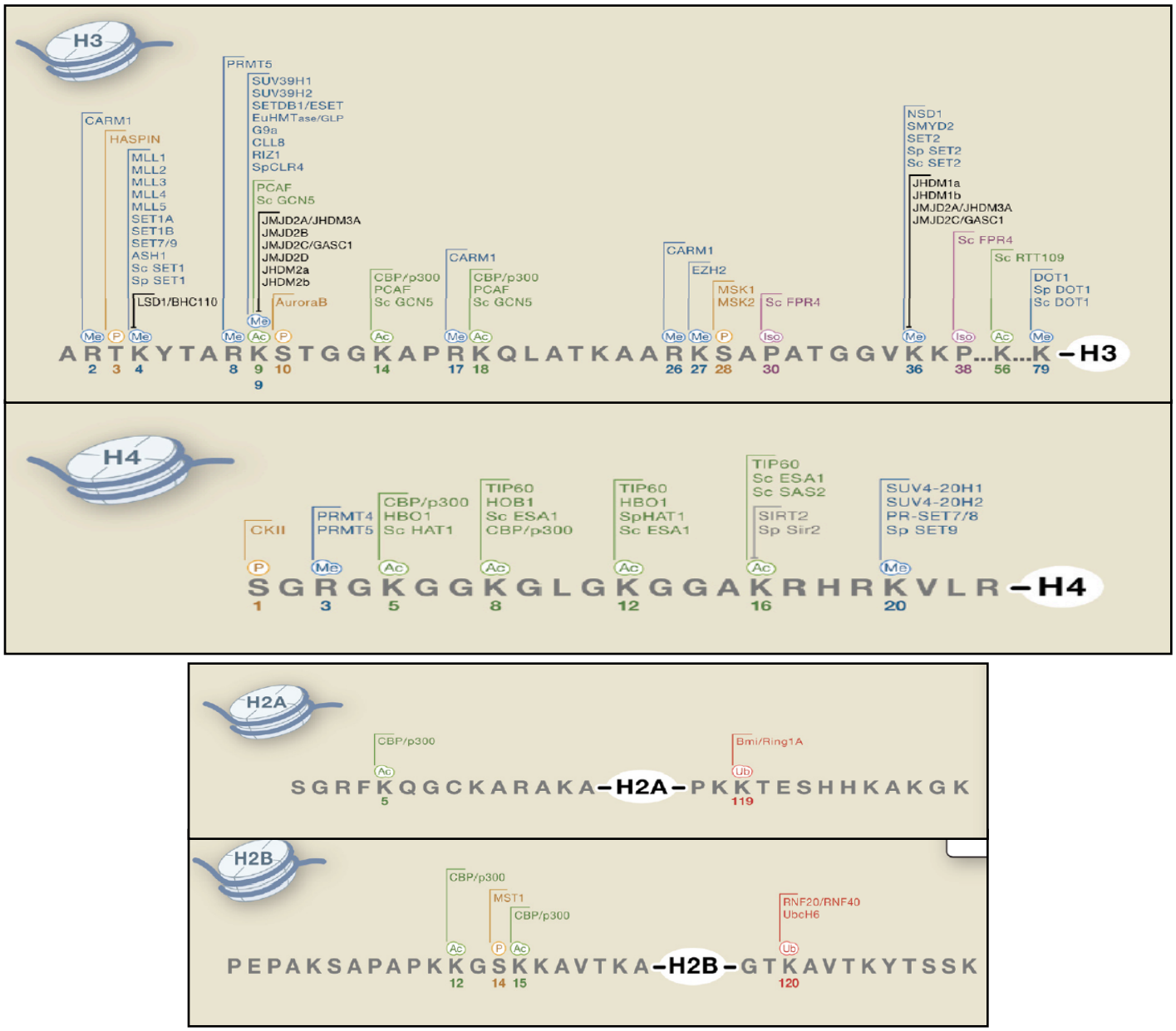
A = Acetilazione di lisine (K)

M = Metilazione di lisine (K) e arginine (R)

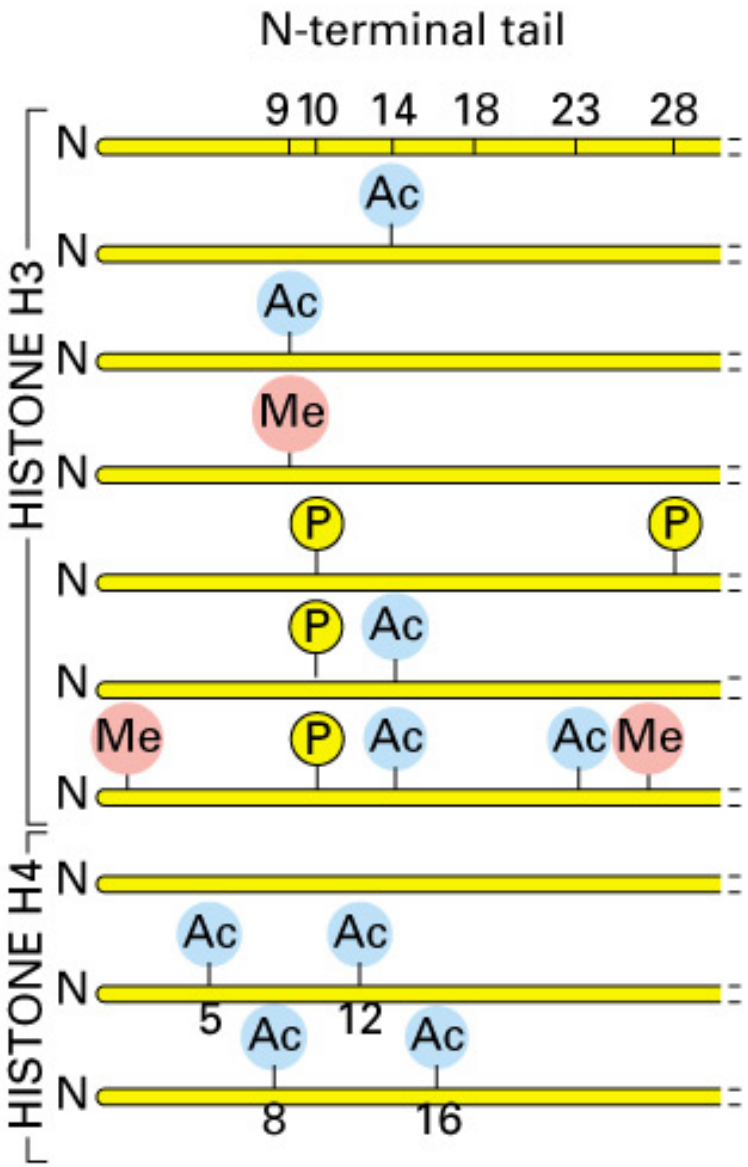
P = Fosforilazione di serine e treonine (S/T)

U = Ubiquitinazione di lisine (K)

60 different residues on histones can be modified



Cell snapshot on histone modifications 2007



modification state	"meaning"
unmodified	gene silencing?
acetylated	gene expression
acetylated	histone deposition
methylated	gene silencing/ heterochromatin
phosphorylated	mitosis/meiosis
phosphorylated/ acetylated	gene expression
higher-order combinations	?
unmodified	gene silencing?
acetylated	histone deposition
acetylated	gene expression

MODELS OF THE FUNCTIONS OF HISTONE MODIFICATIONS

3 POPULAR MODELS that attempt to explain the function of post-translational histone modifications in gene regulation:

1) **Charge neutralization** - specific modifications of histone acetylation and histone phosphorylation change the overall charge of the chromatin structure. The acetylation neutralizes positive charge on DNA and phosphorylation adds a negative charge. According to this model, these modifications can lead to a general decondensation of the chromatin fiber

2) **Histone code** was originally introduced to explain how multiple histone modifications occurring in the same region could control gene regulation and it hypothesizes that multiple modifications can function combinatorially or sequentially to regulate downstream functions.

3) **Signalling pathway** model postulates that histone modifications serve as signalling platforms to facilitate binding of enzymes for their function on chromatin. It is more general than the histone code model, and suggests that multiple histone modifications provide stability, robustness and specificity through feedback loops, redundancy and combination.

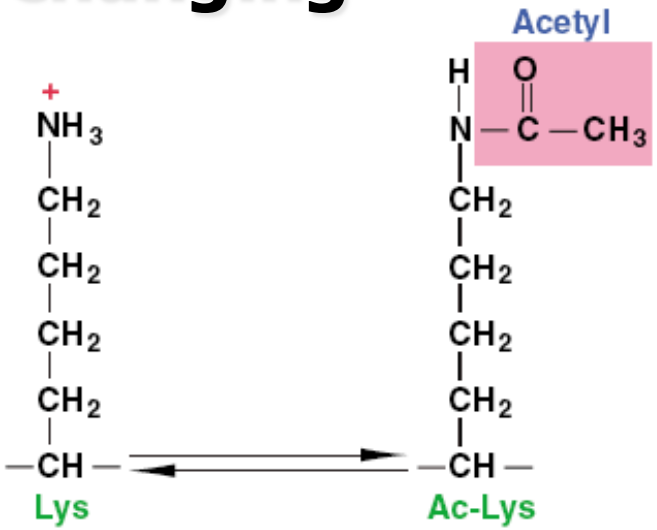
Modificazioni possibili:
A = Acetilazione di lisine (K)
M = Metilazione di lisine (K) e arginine (R)
P = Fosforilazione di serine e treonine (S/T)
U = Ubiquitinazione di lisine (U)

CHI AGISCE?

COMPLESSI DI MODIFICAZIONE DELLA CROMATINA:

HAT, HDAC
ISTONE METILTRANSFERASI (HMT) E DEMETILASI
CHINASI
ENZIMI CHE CONIUGANO UBIQUITINA

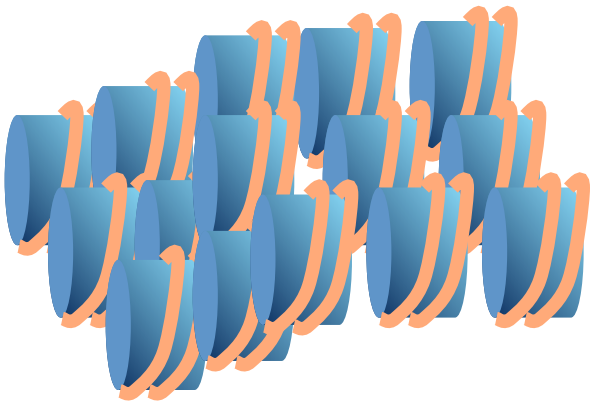
Acetylation is very dynamic and rapidly changing



HAT catalyzes the transfer of an acetyl group from AcCoA to the ε– amino group of the lysine residue, releasing its positive charge and therefore lowering its affinity for DNA

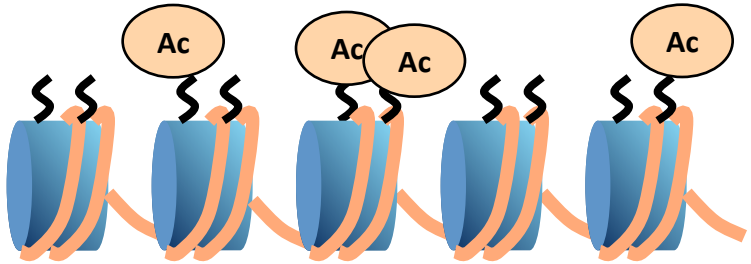
HDAC promotes the removal of the acetyl group from the acetyl-lysine regenerating the ε– amino group and releasing the acetate molecule

heterochromatin
(transcriptionally inactive/condensed)



HAT

euchromatin
(transcriptionally active/accessible)



Writing, erasing and reading histone lysine methylations

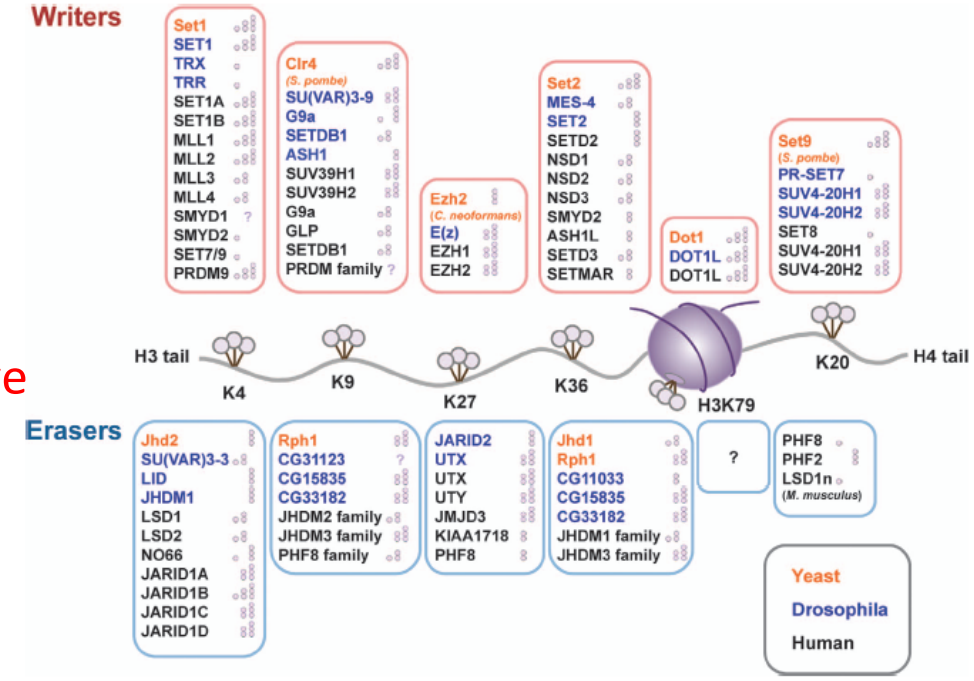
Experimental & Molecular Medicine (2017) 49, e324; doi:10.1038/emm.2017.11

www.nature.com/emm

Kwangbeom Hyun, Jongcheol Jeon, Kihyun Park and Jaehoon Kim

Histone lysine methylations confer active or repressive transcription depending on their positions and methylation states.

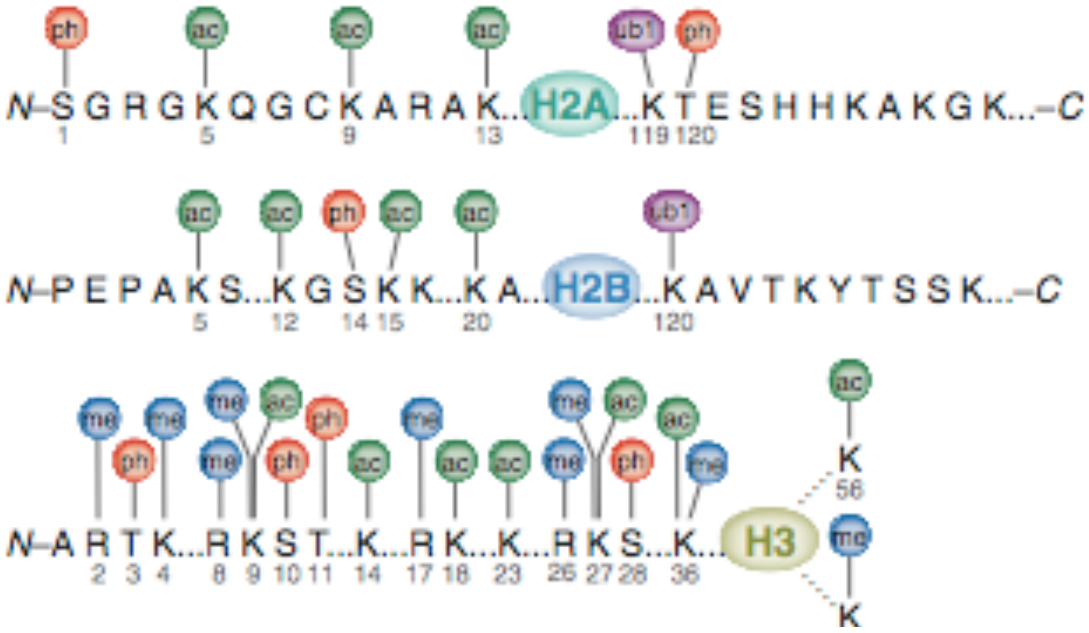
Generally, H3K4, H3K36 and H3K79 methylations are considered to mark active transcription, whereas H3K9, H3K27 and H4K20 methylations are associated with silenced chromatin states.



Histone lysine methylation functions are exerted by effector molecules that specifically recognize the methylated site. These '**reader**' proteins contain methyl-lysine-binding motifs the ability to distinguish target methyl-lysines and surrounding amino-acid sequence.

Histone ubiquitination

Histone ubiquitination occurs on histones H2A, H2B and H3



Histone ubiquitination is formed as isopeptide bond between the carboxy-terminal glycine of ubiquitin and a lysine residue on histones; this bond is formed with the catalytic actions of E1, E2 and E3 ligases.

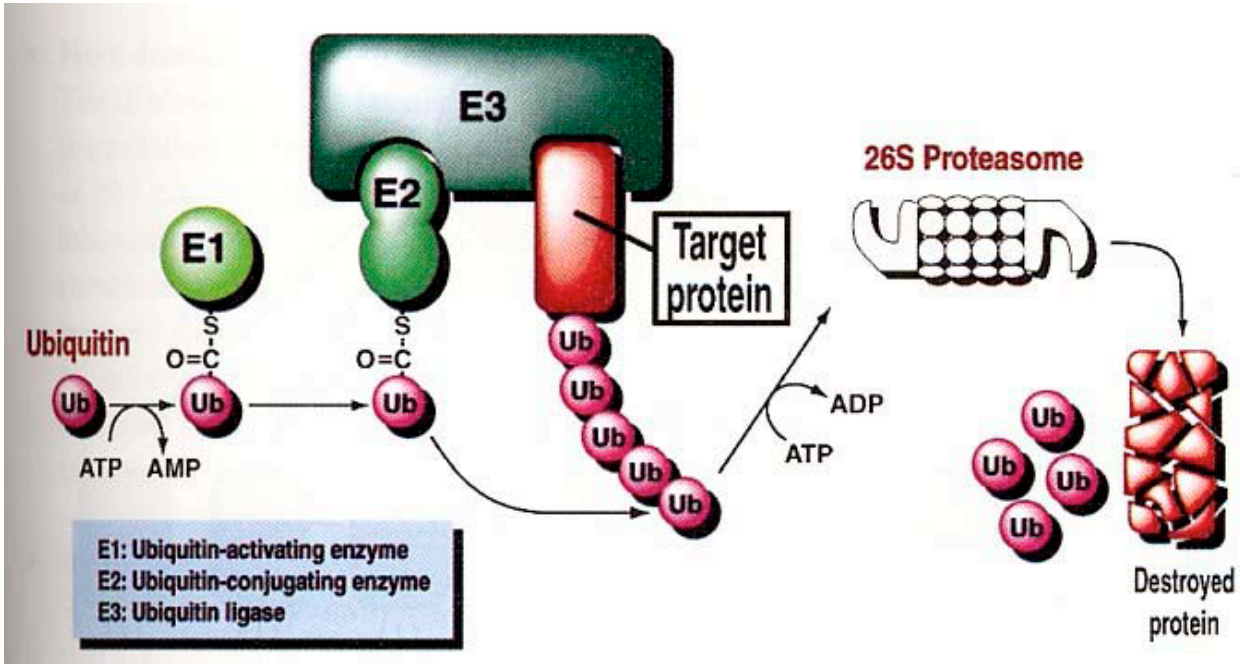
Histone ubiquitination can be reversed by deubiquitinases

Ubiquitin is a small, highly-conserved regulatory protein that is *ubiquitously* expressed in eukaryotes.

Ubiquitination (or ubiquitylation) refers to the post-translational modification of a protein by the covalent attachment (via an isopeptide bond) of one or more ubiquitin monomers.

The most prominent function of ubiquitin is **labeling proteins for proteasomal degradation**.

Besides this function, MONO-ubiquitination also controls the stability, function, and intracellular localization of a wide variety of proteins.

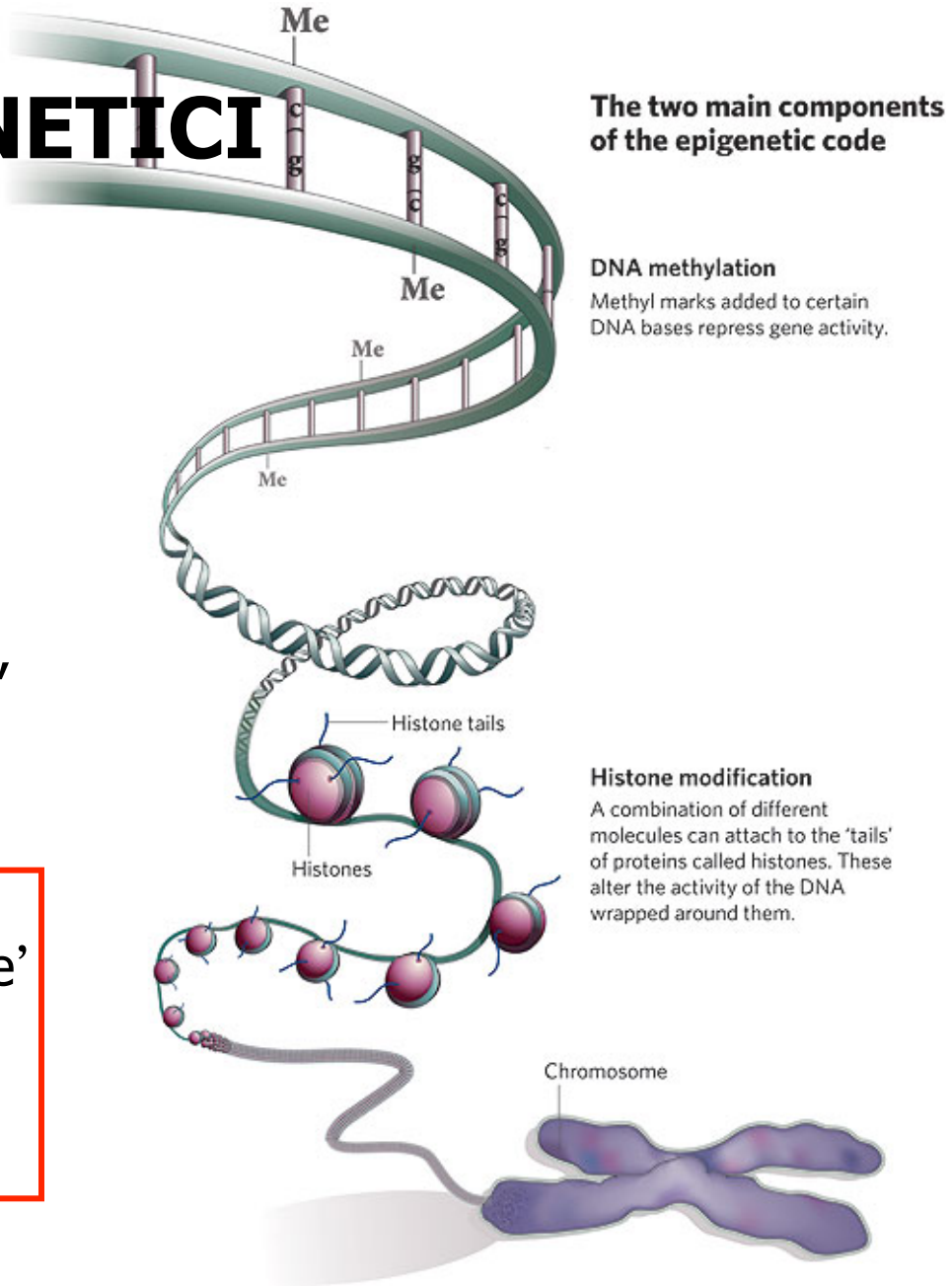


MECCANISMI EPIGENETICI

Fattori che vengono trasmessi alla progenie, ma che non sono direttamente attribuibili a sequenze di DNA.

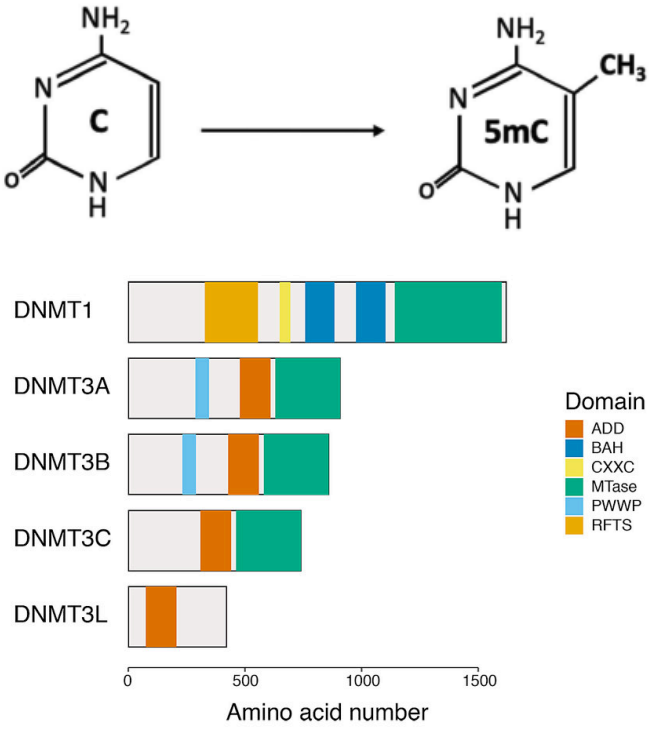
MODIFICAZIONI DEGLI ISTONI
Acetilazioni, fosforilazioni e metilazioni, responsabili dei cambiamenti conformazionali della cromatina.

METILAZIONE DEL DNA
Nelle cells eucariotiche la metilazione e' a carico della G. Solo il 3% delle C e' metilato; in genere e' bersaglio della metilazione la C delle doppiette CpG.



DNA Methyl-transferases

Factor	Function	Mouse loss-of-function phenotype	Human diseases associated with genetic mutations
DNMT1	Maintenance DNA methyltransferase	• Low global DNA methylation • Derepression of IAP transposons • Early embryonic lethality	• Hereditary sensory autonomic neuropathy 1E (HSAN1E; OMIM 614116) • Autosomal-dominant cerebellar ataxia, deafness and narcolepsy (ADAC-DN; OMIM 604121)
UHRF1	DNMT1 cofactor	• Low global DNA methylation • Early embryonic lethality	
DNMT3A	De novo DNA methyltransferase	• Constitutive knockouts die ~4 weeks after birth[a] • Sterility in both males and females in germline-specific knockouts	• Microcephalic dwarfism • Tatton-Brown–Rahman syndrome (TBRS; OMIM 602729) • Acute myeloid leukaemia (AML; OMIM 601626)
DNMT3B	De novo DNA methyltransferase	Constitutive knockouts die mid-gestation[a]. More important for embryonic DNA methylation than for germline DNA methylation	Immunodeficiency, centromeric instability and facial anomalies syndrome (ICF; OMIM 602900)
DNMT3C	De novo DNA methyltransferase (Muroidea specific)	Males are infertile likely owing to defect in methylating transposon promoters during spermatogenesis	
DNMT3L	De novo DNA methyltransferase cofactor	• Male germline cells unable to undergo meiosis • Females unable to establish maternal imprinting, leading to mid-gestation lethality of progeny	



In mammals, there are 2 families of **DNA Methyl-transferases**:

- DNMT1, the maintenance methyltransferase that is responsible for the methylation of hemi-methylated CpG sites during DNA replication.
- *de novo* methyltransferases (DNMT3A and DNMT3B) that act primarily on CpG dinucleotides during the embryonic life;

DNA methylation

is mainly associated with **transcriptional repression** and plays a major role in different processes such as X chromosome inactivation (XCI), genomic imprinting, silencing of transposons, repetitive elements and germ-line specific genes.

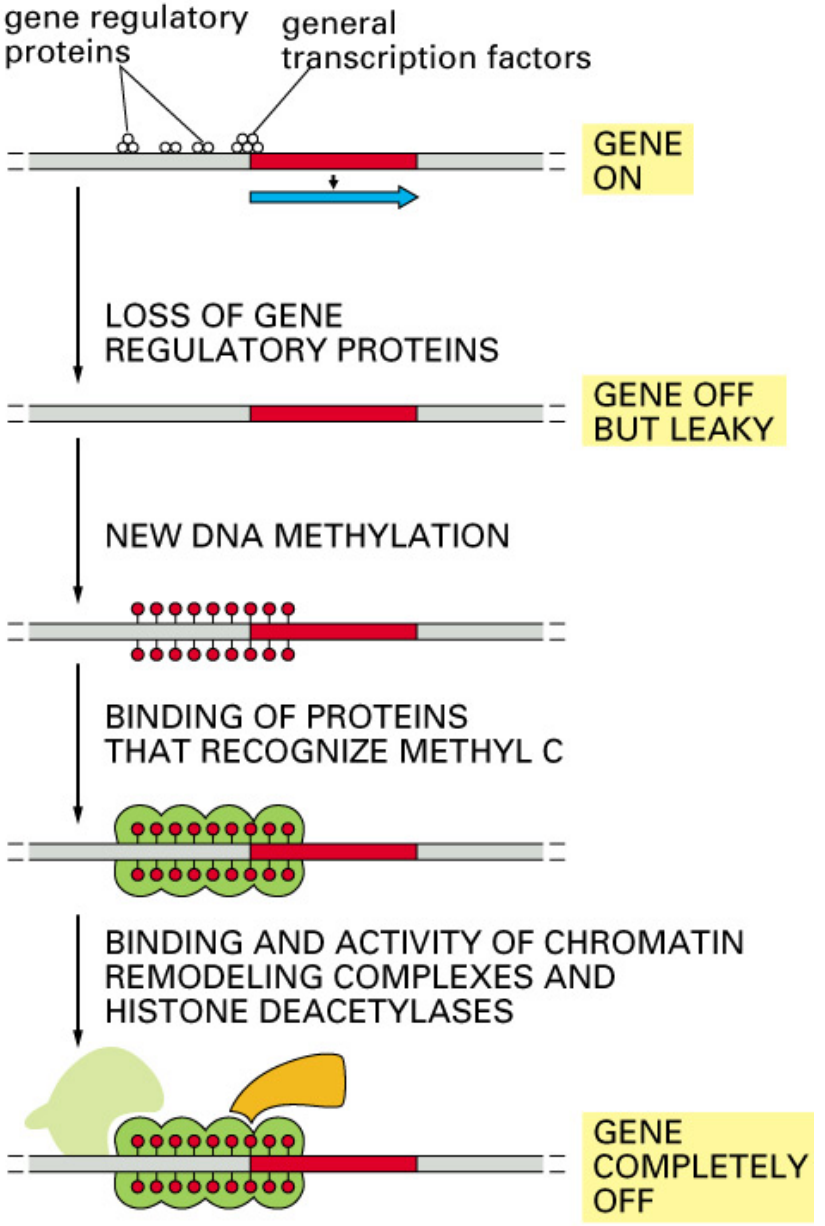
Given the robust stability of DNA methylation, ectopic promoter methylation might lead to long-term silencing of important genes. Tight regulation of the mark is of crucial importance, for proper cellular function (Greenberg, 2021).

The epigenetic memory linked to DNA methylation is robust in somatic tissues, where the levels of CpG methylation are globally stable, with 70-80% of CpG dinucleotides harboring the mark (Lee et al., 2014).

Vertebrates Use DNA Methylation to Lock Genes in a Silent State

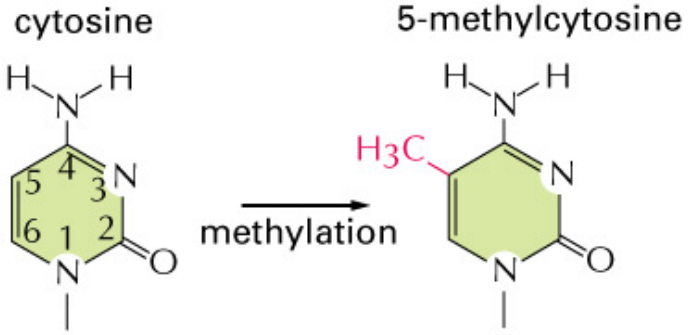
Vertebrate cells contain a family of **proteins (MeCP2) that bind methylated DNA**.

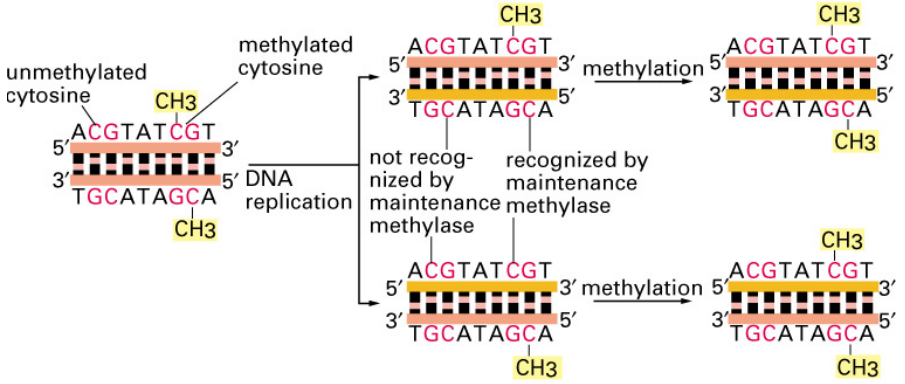
These DNA-binding proteins, in turn, interact with **chromatin remodeling complexes** and **histone deacetylases** that condense chromatin so it becomes transcriptionally inactive.



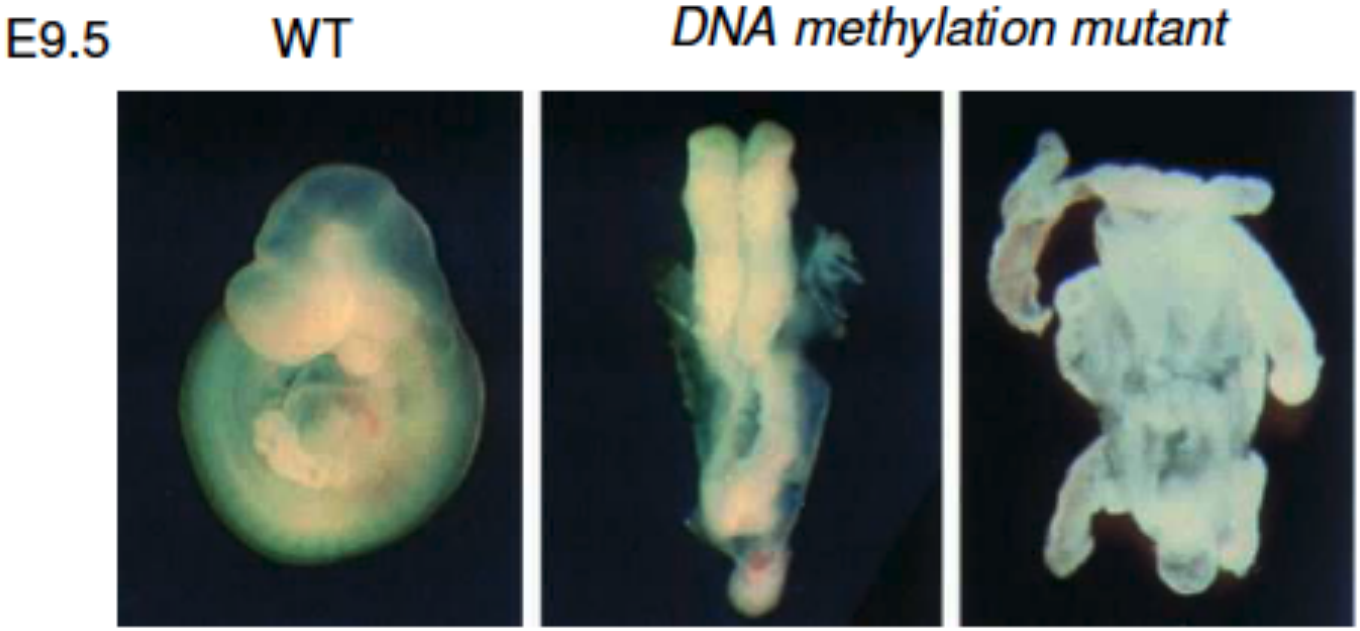
Maintenance of DNA methylation

1. **Dnmt1** maintains the methyl-CpG content of both daughter DNA duplexes following replication (higher affinity for hemimethylated mCpG DNA)
2. Dnmt1 Methyltransferase is localized to the chromosomal replication complex
3. Methylation of newly synthesized DNA takes place less than one minute following replication (chromatin assembly takes 10-20 min)

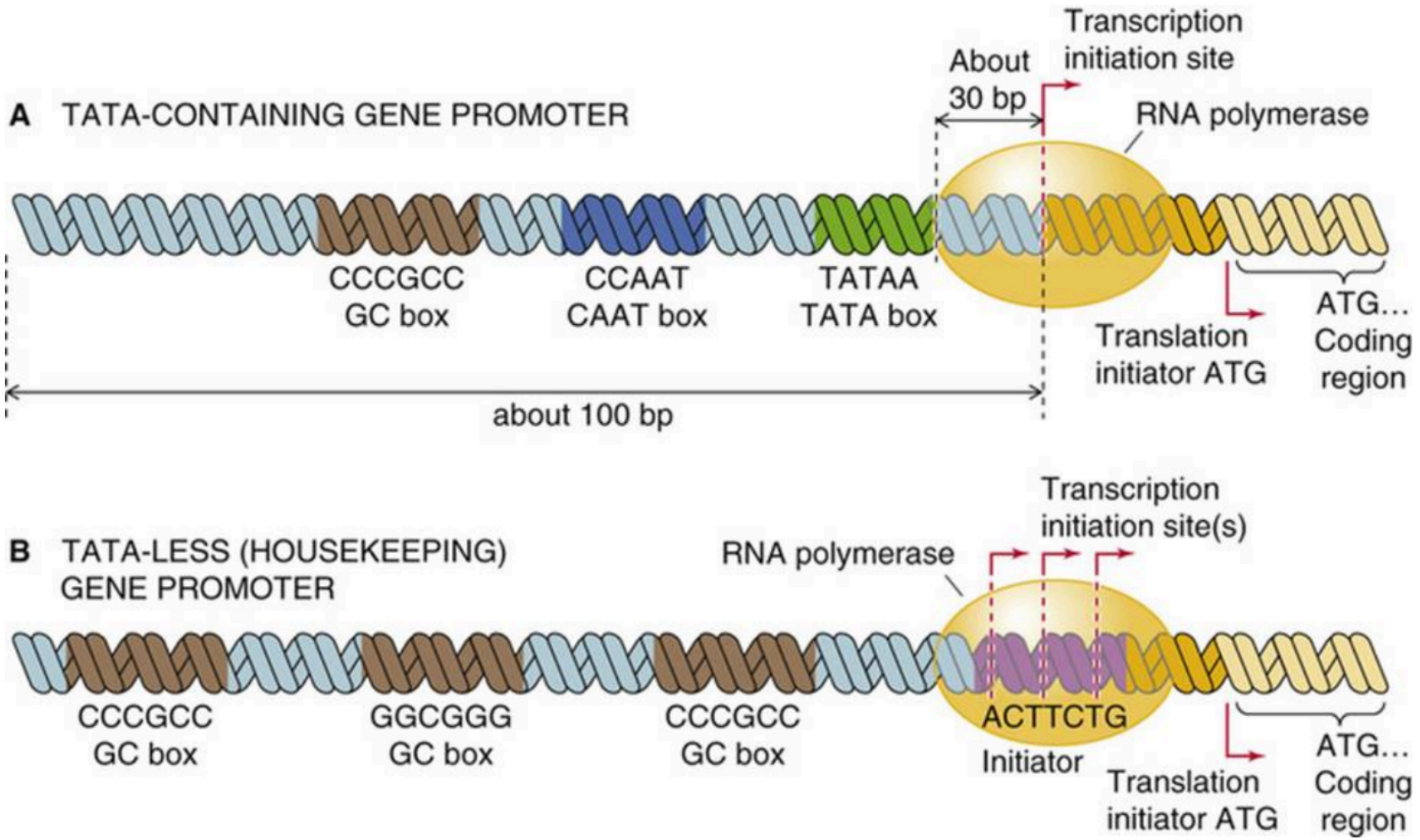




The essential role of DNA methylation for a proper differentiation is supported by the severe developmental defects and embryonic lethality exhibited in DNMT-deficient mice.



PROMOTORI PRIVI DI **TATA** box.



CpG islands (20-50 nucleotides within 100-1000 bp upstream the start site) act as promoters for **housekeeping genes**

A transcription factor called **SP1** recognizes these CG rich regions

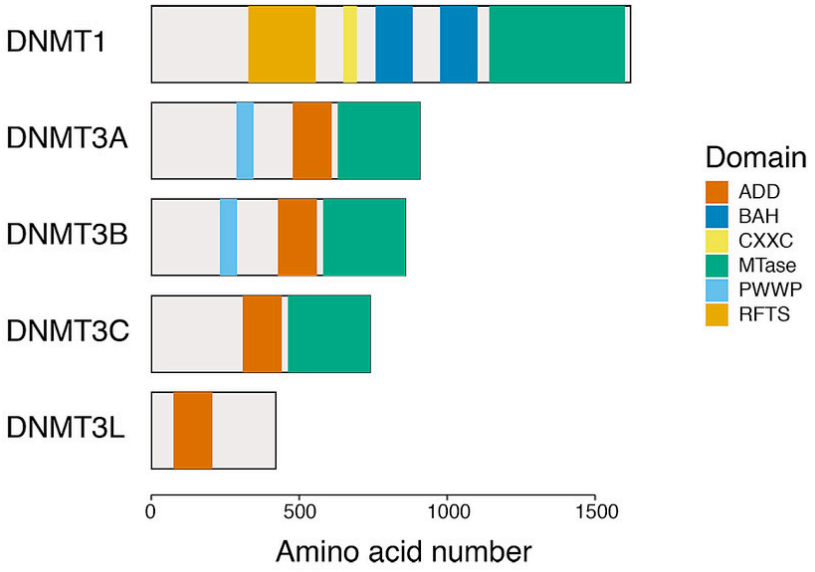
CpG island (CGI) promoters are not methylated

- Roughly two-thirds of promoters are CGIs, and comprise most housekeeping and developmental genes.
- ALL CpG islands are associated with trasnrcibed genes
- Keeping promoters free of methylation is absolutely crucial for proper cellular function.
- X-linked CpG islands become methylated upon X inasctivation; DNA methylation at CpG islands is essential fopr the maintenance of the inactive state

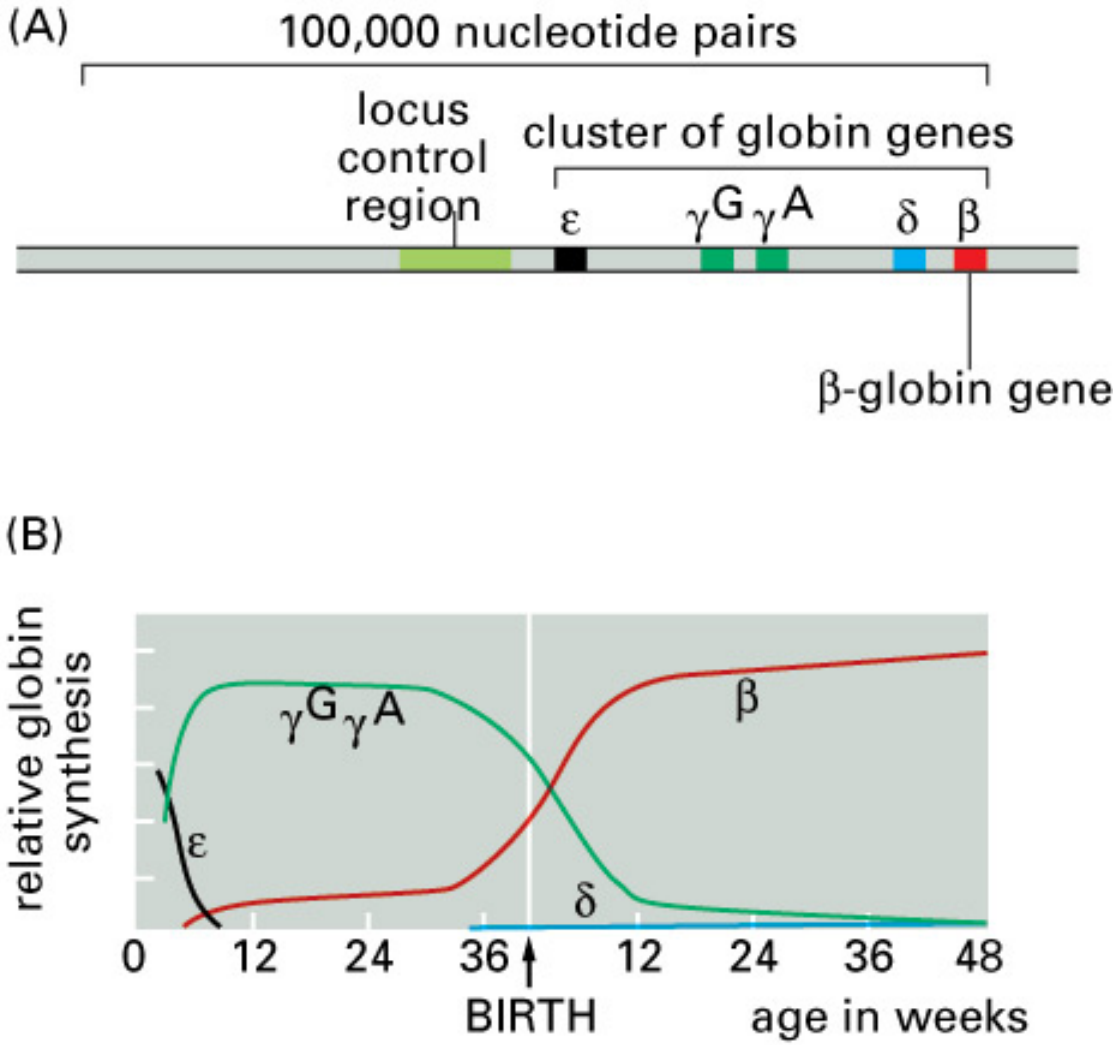
There is nothing about the sequence, per se, that should repel de novo DNA methylation.

The ADD domain, harbored by both DNMT3A and 3B enzymes , is repelled by H3K4 methylation.

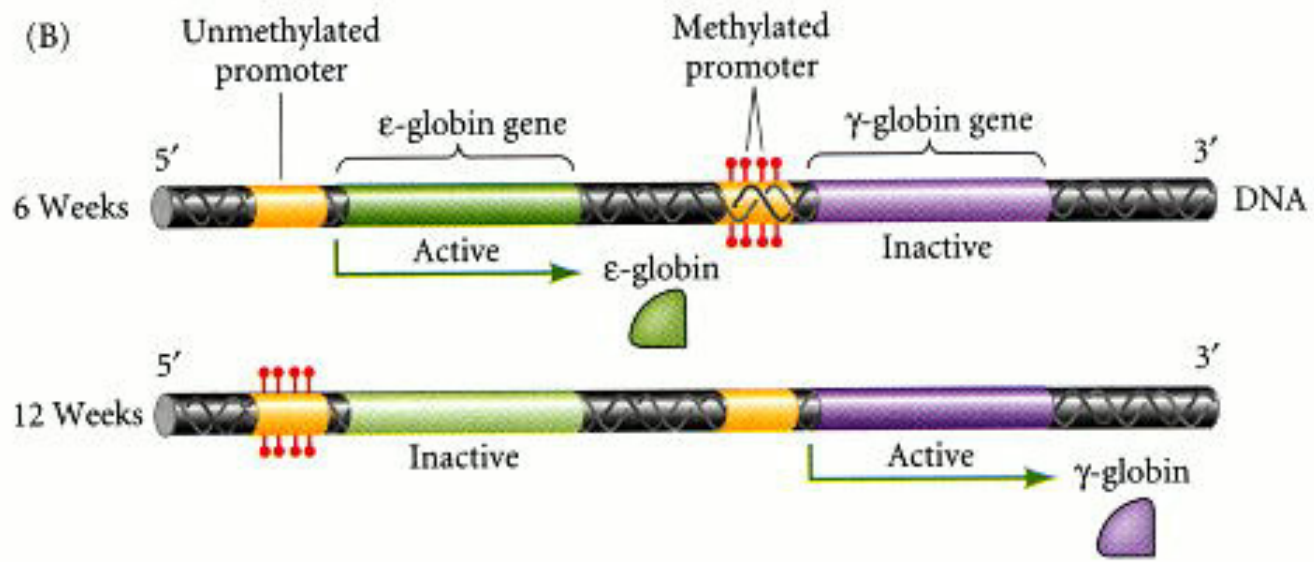
Given that H3K4me3 is strongly linked with active promoters, therein lies a simple mechanism to protect promoter sequences from DNA methylation deposition.



The β-globin gene cluster



The activity of the globin genes correlates inversely with the methylation of their promoters



In developing human and chick red blood cells, the DNA of the globin promoters is almost completely **unmethylated**, whereas the same promoters are **highly methylated** in cells that do not produce globin.

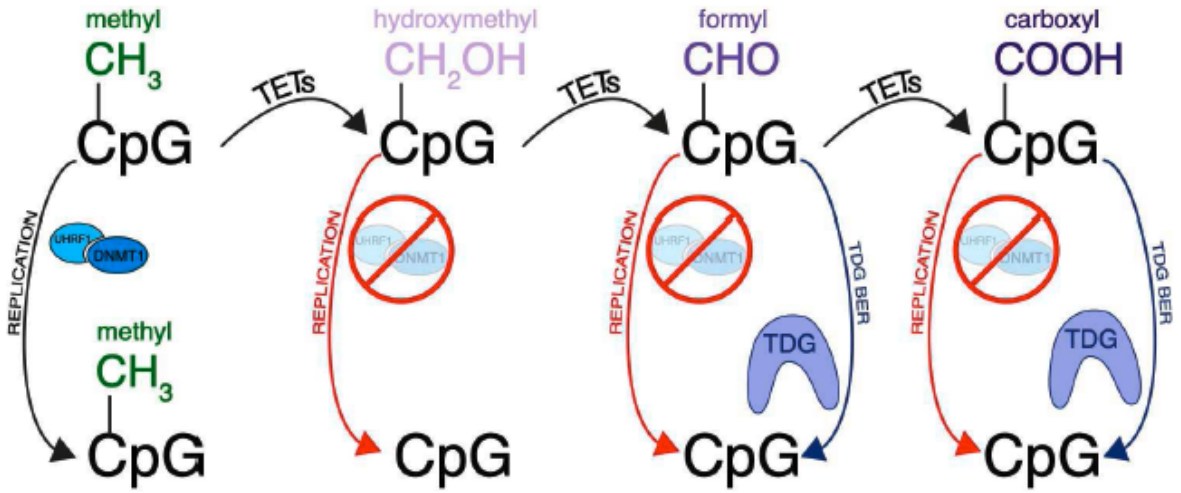
Moreover, the methylation pattern changes during development. The cells that produce hemoglobin in the human embryo have unmethylated promoters for the genes encoding the ε-globins of embryonic hemoglobin. These promoters become methylated in the fetal tissue. Similarly, when the fetal globin gives way to adult globin, the γ-globin gene promoters become methylated.

Global demethylation of the genome occurs *immediately after fertilization* and de novo remethylation follows implantation

There are different mechanisms of DNA demethylation: both **passive** and **active** processes can occur.

- Passive demethylation simply requires the impairment of maintenance DNA methylation machinery (Dnmt-1), which results in 2-fold dilution of methyl-CpGs during each round of DNA synthesis.

- Active DNA demethylation in mammals involves the action of Ten-eleven translocase (TET) family of dioxygenases.

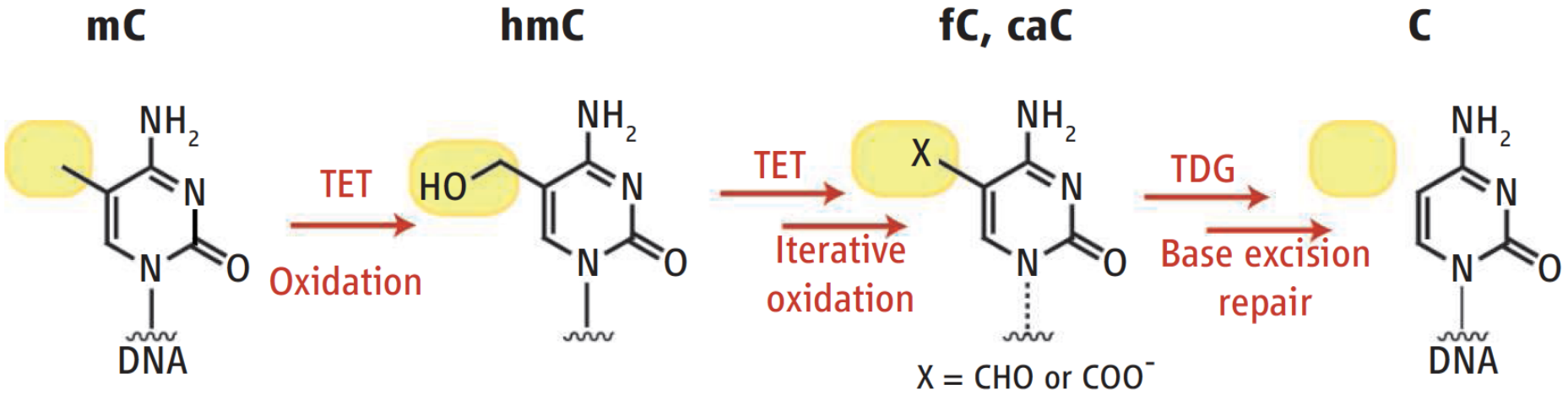


NATURE REVIEWS | GENETICS
VOLUME 13 | JANUARY 2012 | 7

Uncovering the role of 5-hydroxymethylcytosine in the epigenome

Miguel R. Branco, Gabriella Ficz and Wolf Reik

Abstract | Just over 2 years ago, TET1 was found to catalyse the oxidation of 5-methylcytosine, a well-known epigenetic mark, into 5-hydroxymethylcytosine in mammalian DNA. The exciting prospect of a novel epigenetic modification that may dynamically regulate DNA methylation has led to the rapid accumulation of publications from a wide array of fields, from biochemistry to stem cell biology. Although we have only started to scratch the surface, interesting clues on the role of 5-hydroxymethylcytosine are quickly emerging.



DNA demethylation. TET enzymes are proposed to oxidize 5-methylcytosine (mC) to 5-hydroxymethylcytosine (hmC) and subsequently to generate the higher oxidation substituents 5-formylcytosine (fC) and 5-carboxylcytosine (caC) (shown as the structure with the 5-X substituent). Unmodified cytosine (C) is on the far right. Base excision repair, initiated by thymine-DNA glycosylase (TDG), releases and replaces the entire modified oxidized base with unmodified C.

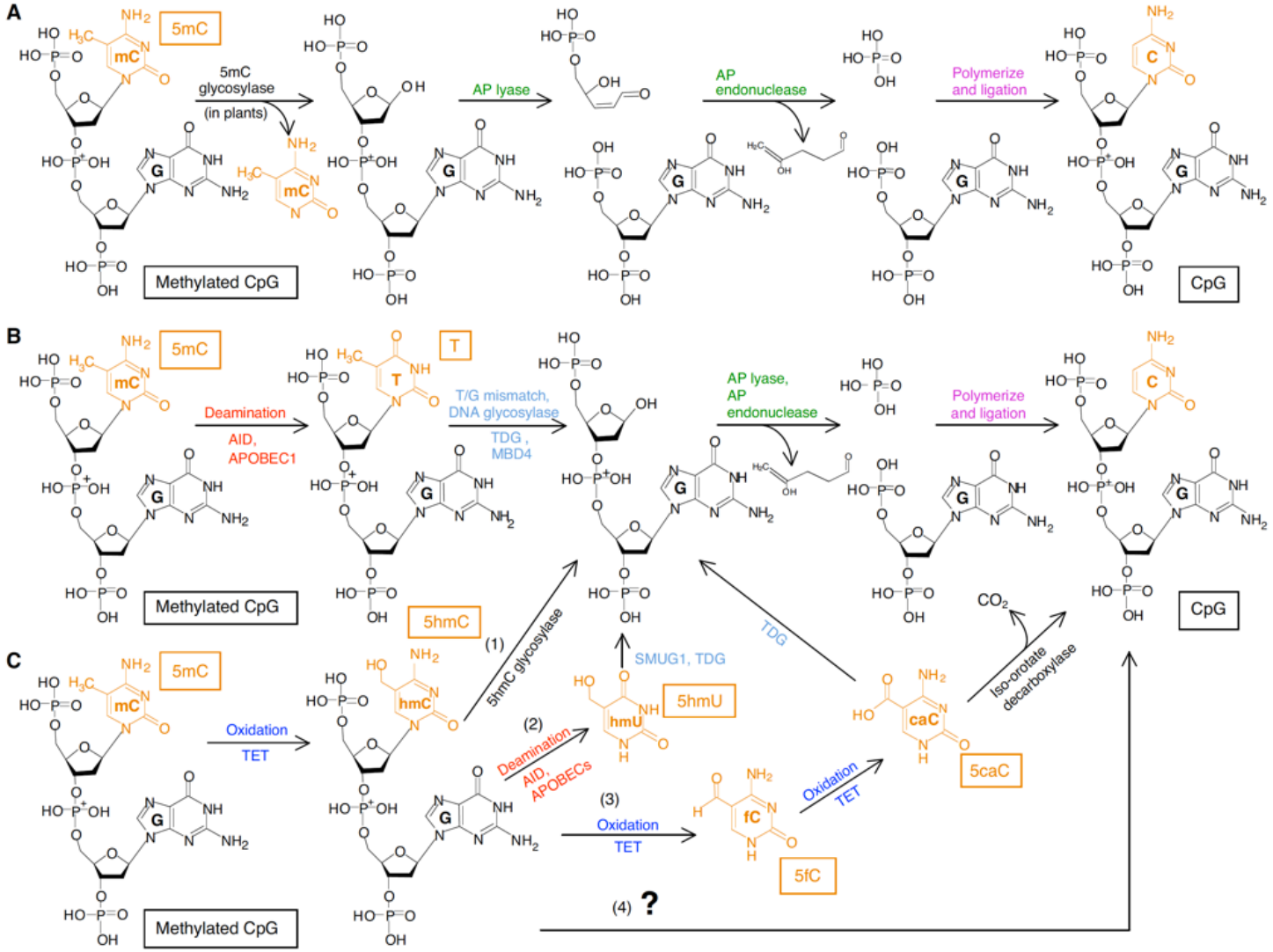


Fig. 4. Potential chemical pathways for active DNA demethylation. (**A**) Direct excision of 5mC (orange) by a 5mC glycosylase followed by repair via the base excision repair (BER) pathway (green and pink), as occurs in plants. (**B**) Cytosine deamination by AID/APOBEC1 (red), followed by base excision mismatch repair, involving the TDG/MBD4 (pale blue) and BER pathways. (**C**) Hydroxylation by TET (blue) initiates four potential pathways leading to demethylated cytosine: (1) removal of 5hmC by an unidentified 5hmC glycosylase, followed by BER; (2) deamination of 5hmC by AID or APOBECs creates 5hmU, which is removed by SMUG1 (single-strand selective monofunctional uracil DNA glycosylase) or TDG, followed by BER; (3) further oxidization of 5hmC to 5fC and then to 5caC, which then may be converted to C by a decarboxylase or by TDG followed by BER; and (4) direct conversion of 5hmC to 5mC by an unidentified enzyme (?). 5caC, 5-carboxylcytosine; 5fC, 5-formylcytosine; 5hmC, 5-hydroxymethylcytosine; 5hmU, 5-hydroxymethyluracil;5mC, 5-methylcytosine; AP, apurinic/apyrimidinic; AID, activation-induced deaminase; APOBEC1, apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1; C, cytosine; G, guanine; MBD4, methyl CpG binding domain protein 4; SMUG1, single-strand selective monofunctional uracil DNA glycosylase; T, thymidine; TDG, thymine DNA glycosylase; TET, ten-eleven translocation.

The amazing demethylase

Howard Cedar and Gregory L. Verdine

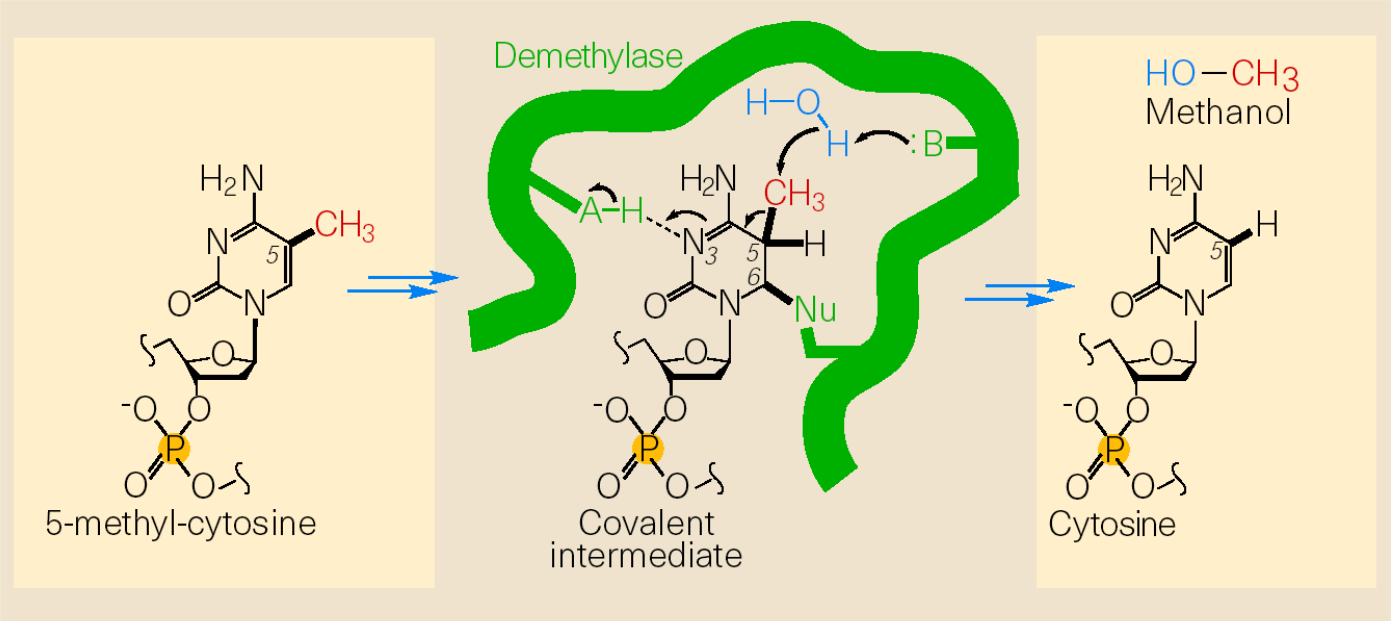


Figure 1 **Mechanism for the enzymatic demethylation of 5-methyl-cytosine. The demethylase (green) is envisaged to form a covalent intermediate by addition of an enzymatic nucleophile (Nu–H) across the 5,6 double bond, assisted by proton shuffling at N3. This intermediate is poised to attack the hydroxide ion, which is generated by *in situ* activation of water. Double arrows indicate two reaction steps, with the intermediates not shown. In the case of enzymatic methylation, an analogous covalent intermediate is formed, but is further processed by cleavage of the C5–H bond as opposed to the C5–CH_3 bond. The 3′-phosphate labelled with ^{32}P in the tracer studies of Bhattacharya *et al.*[1] is in yellow.**

Mammalian genome methylation

Mammals exhibit two rounds of dramatic DNA methylation reprogramming during embryonic development:

1. immediately after fertilization
2. in the germline (Monk et al., 1987).

The embryo must level the high DNA methylation asymmetry exhibited by the paternal and maternal gametic genomes that arrive in the zygote (Wang et al., 2014), thus mitigating dosage discrepancies between alleles.

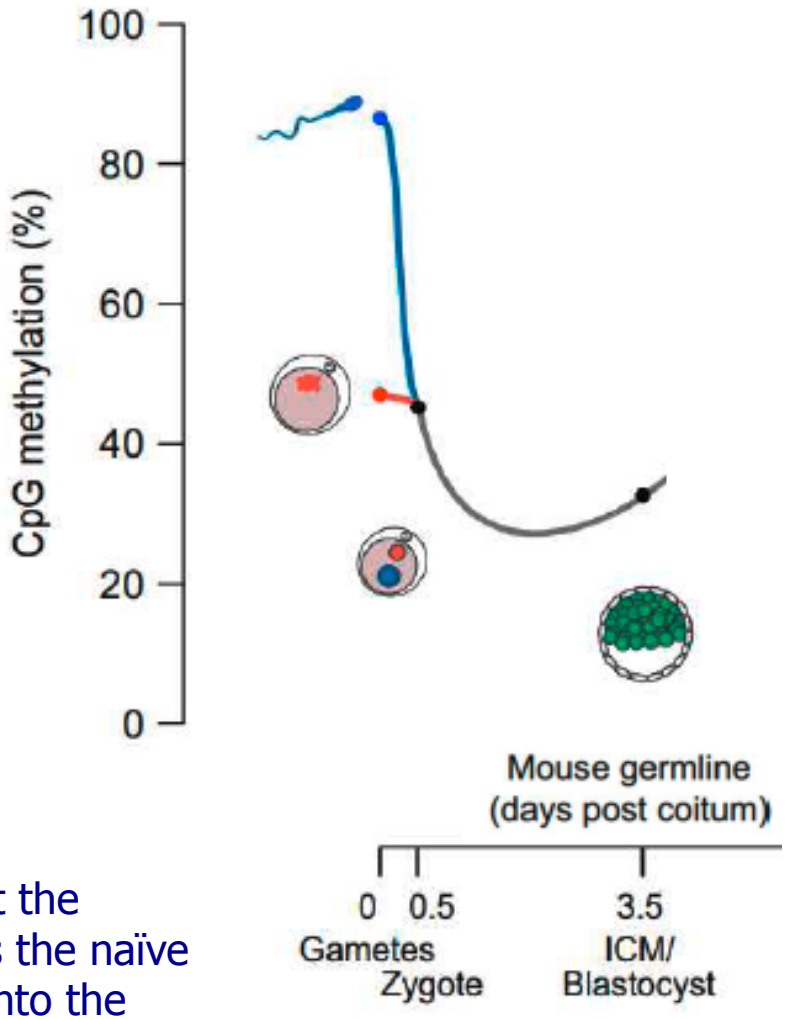
As the embryo implants in the uterus, the de novo DNA methyltransferases, DNMT3A and DNMT3B, rapidly remethylate the genome, establishing a pattern that is globally maintained in somatic tissue-types

A dramatic genome-wide reprogramming of DNA methylation occurs during **embryogenesis**

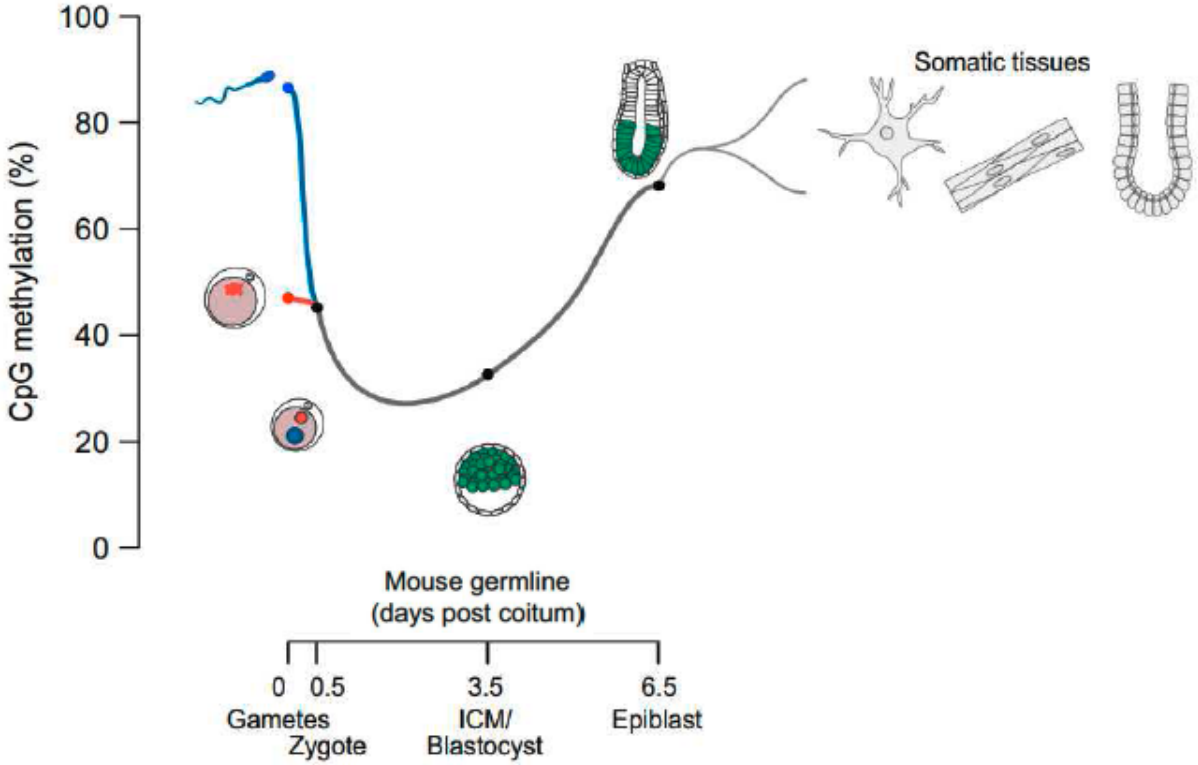
Upon fertilization, both the maternal and paternal genomes are subjected to a progressive loss of DNA methylation

Such reprogramming is found exclusively in mammals.

A reprogramming of DNA methylation at this stage of development is needed to revert cellular memory towards an undifferentiated state in order to allow for naïve pluripotency

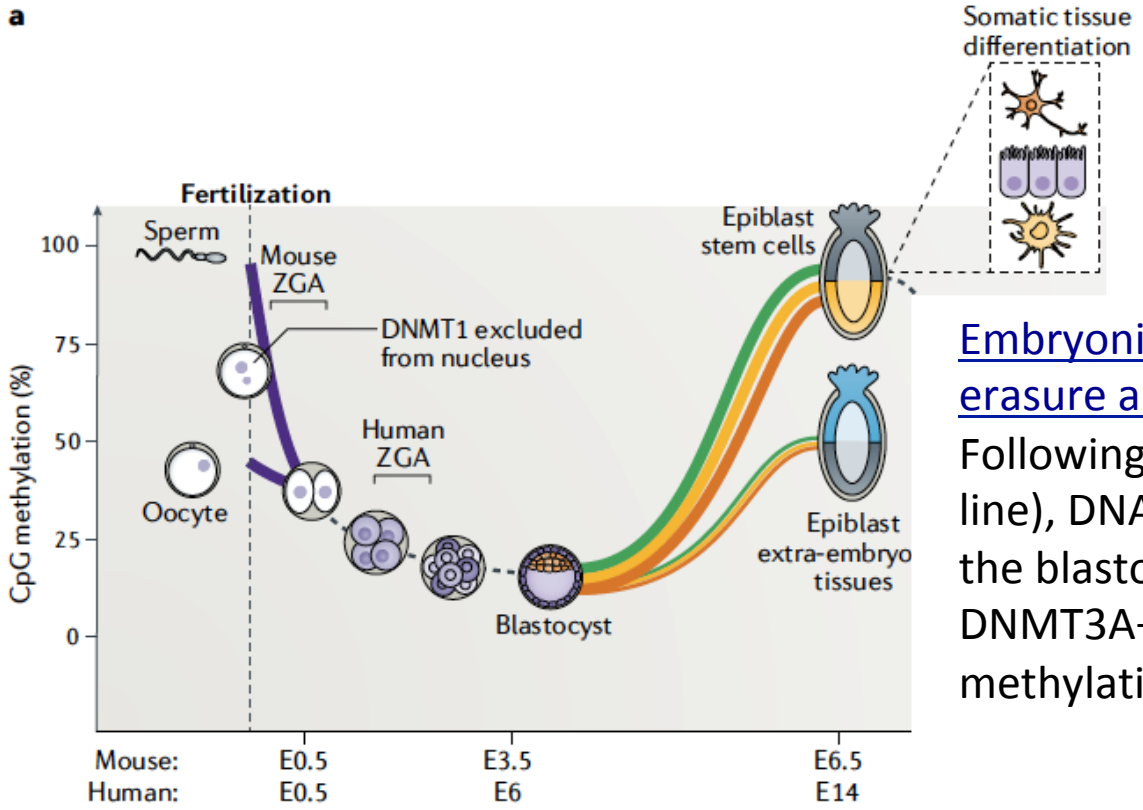


In mice, in the inner cell mass (ICM) at the blastocyst stage (E3.5), which contains the naïve pluripotent cells that will differentiate into the embryo, only 20% of CpGs remain methylated.



Following this dramatic global decrease of the mark, the *de novo* methyltransferases, **DNMT3A and DNMT3B**, rapidly re-establish high levels of methylation.
Importantly, the wave of re-methylation coincides with the transition from naïve to primed pluripotency.
By E6.5, the primed stem cells in the epiblast present a global level of CpG methylation similar to that of somatic tissues.
These cells will then further differentiate into the somatic lineages, which will globally maintain the pattern and levels of CpG methylation established during these early stages of development.

Methylation patterning in development

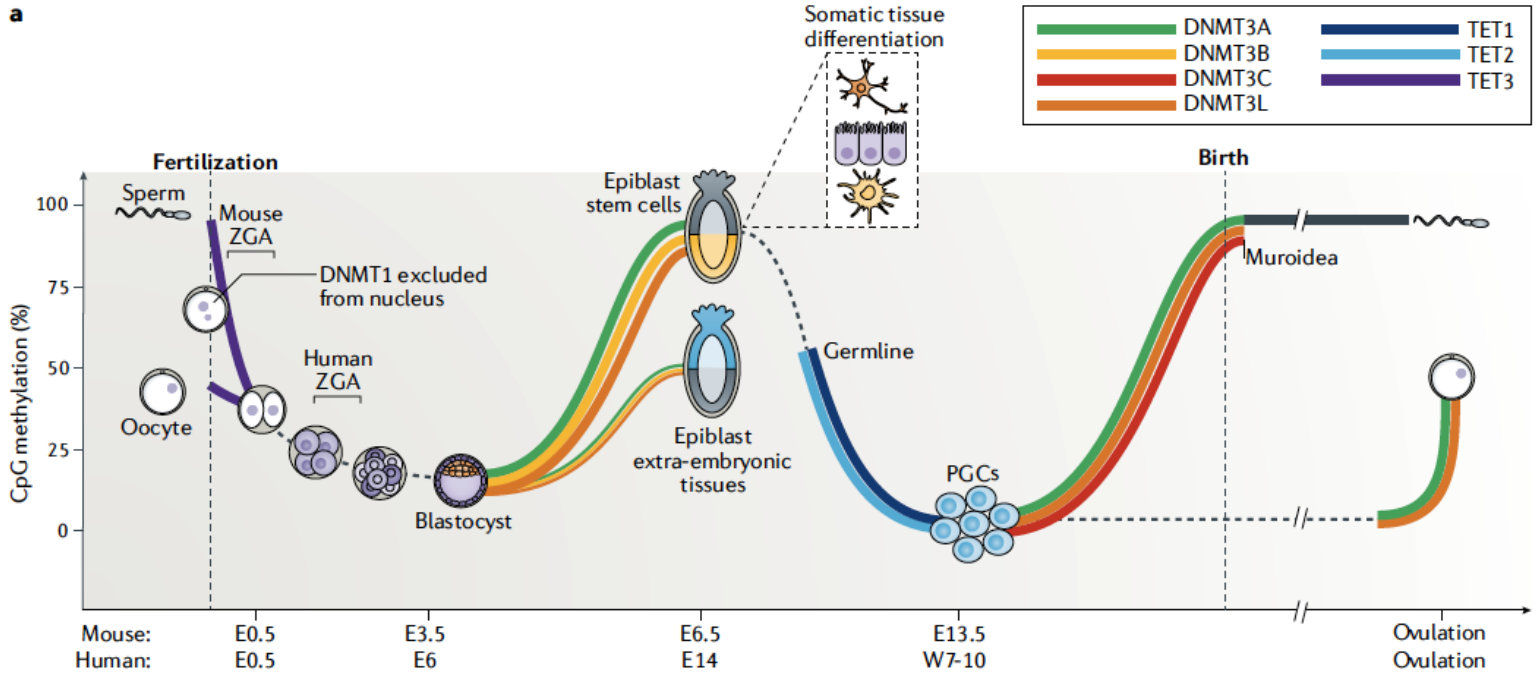


Embryonic and germline DNA methylation erasure and establishment.

Following passive demethylation (dashed line), DNA methylation reaches a low point at the blastocyst stage, which is followed by DNMT3A+B-mediated de novo DNA methylation after blastocyst implantation.

During post- fertilization reprogramming, the embryo loses gamete- specific DNA methylation patterns inherited from the oocyte and the sperm as it progresses towards pluripotency. The paternal genome is actively demethylated by TET3; the two parental genomes then undergo rounds of passive, DNA replication- dependent dilution of DNA methylation, as the maintenance enzyme DNMT1 provided by the oocyte is excluded from the nucleus during subsequent cell divisions.

Methylation patterning in development



- Demethylation occurs in developing PGCs, as a prerequisite for subsequent acquisition of sex- specific DNA methylation patterns during male and female germline differentiation.
- Post implantation, in the epiblast, a subset of stem cells is specified for the germline, where they undergo two waves of DNA demethylation: one passive and one mediated by TET1 and TET2.
- Male gametes become highly methylated before birth (E15-E16 and onwards at the prospermatogonia state preceding mitosis and meiosis) through the activity of DNMT3A and DNMT3L.
- The oocyte gains methylation after birth, after meiosis and prior to ovulation through the activity of DNMT3A and DNMT3L in mice, and likely through DNMT3A in humans.
- At the end of gametogenesis, the sperm genome exhibits ~80% CpG methylation, with a genomic distribution roughly similar to that of somatic cells.
- By contrast, the oocyte genome is only ~50% methylated, and nearly exclusively in gene bodies.