Epigenetics
and the regulation of gene expression

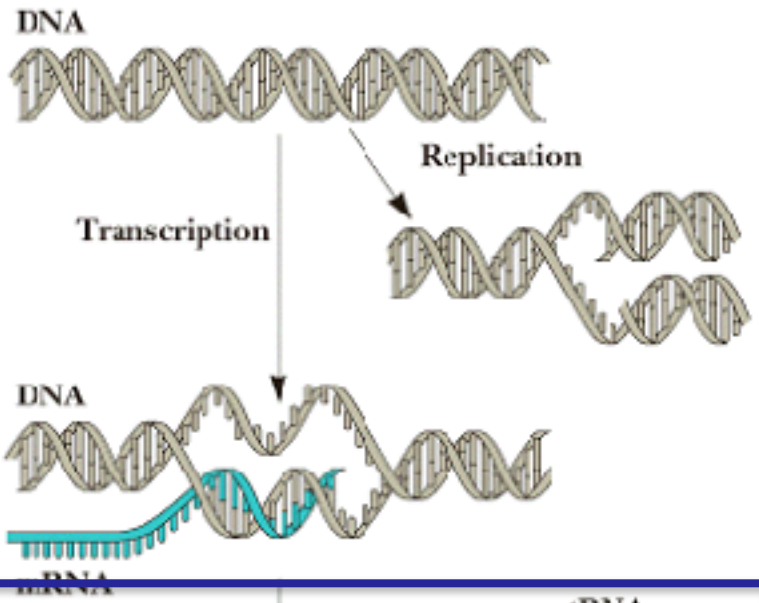
Philippe Vernier

Institut des Neurosciences Paris-Saclay, CNRS, Gif-sur-Yvette

Institut des Sciences du Vivant F. Joliot, CEA, Saclay

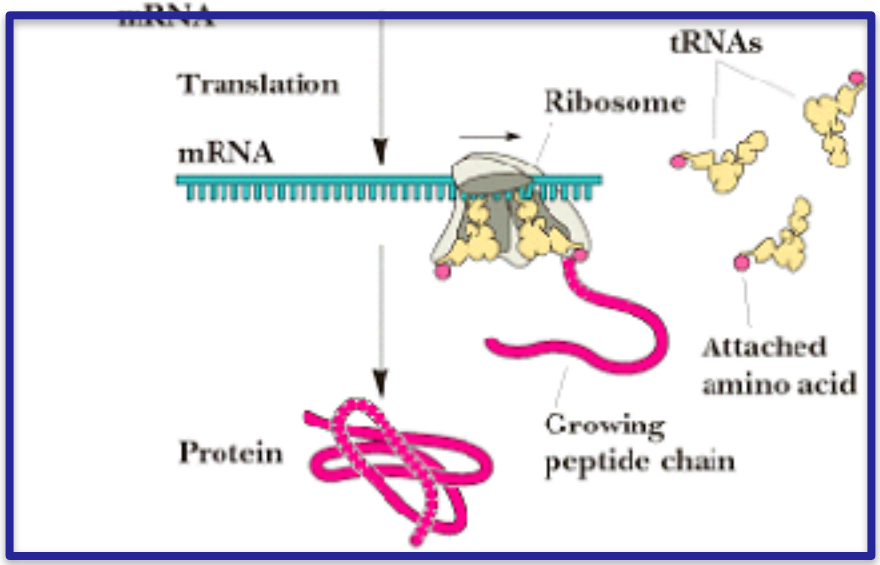
philippe.vernier@cnrs.fr

The central dogma of molecular biology



Replication
DNA replication yields two DNA molecules identical to the original one, ensuring transmission of genetic information to daughter cells with exceptional fidelity.

Transcription
The sequence of bases in DNA is recorded as a sequence of complementary bases in a single-stranded mRNA molecule.



Translation
Three-base codons on the mRNA corresponding to specific amino acids direct the sequence of building a protein. These codons are recognized by tRNAs (transfer RNAs) carrying the appropriate amino acids. Ribosomes are the "machinery" for protein synthesis.

ADN

↓

ARN

↓

Protéine

Is heredity influenced by the environment?

- selection acts on phenotypic variation via genetic variation (mutations) which is not sensitive to environmental cues.

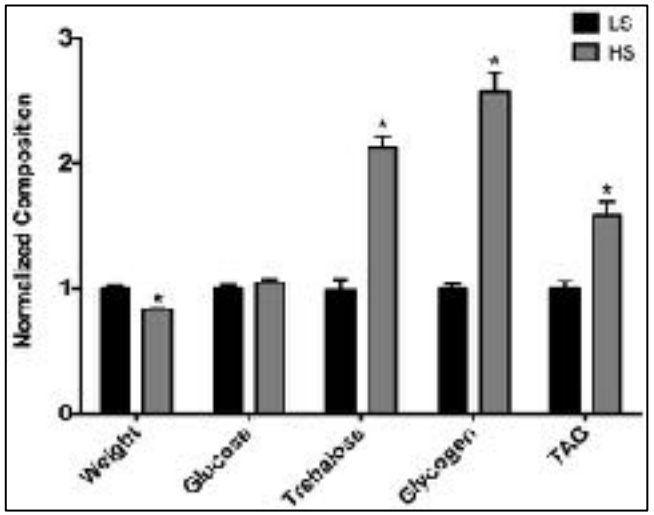
- Weismann's principle of the « germplasm » (1892): somatic cells are separated from germ cells, and thus, no mechanisms were thought to exist for germ cells to be modified by the environment.

But: environmentally challenged parents sometimes give rise to modified progeny, which cannot be accounted for by Mendelian inheritance

- Waddington observation of heat-induced fly wing structure alteration (coined the word epigenetics – 1942)

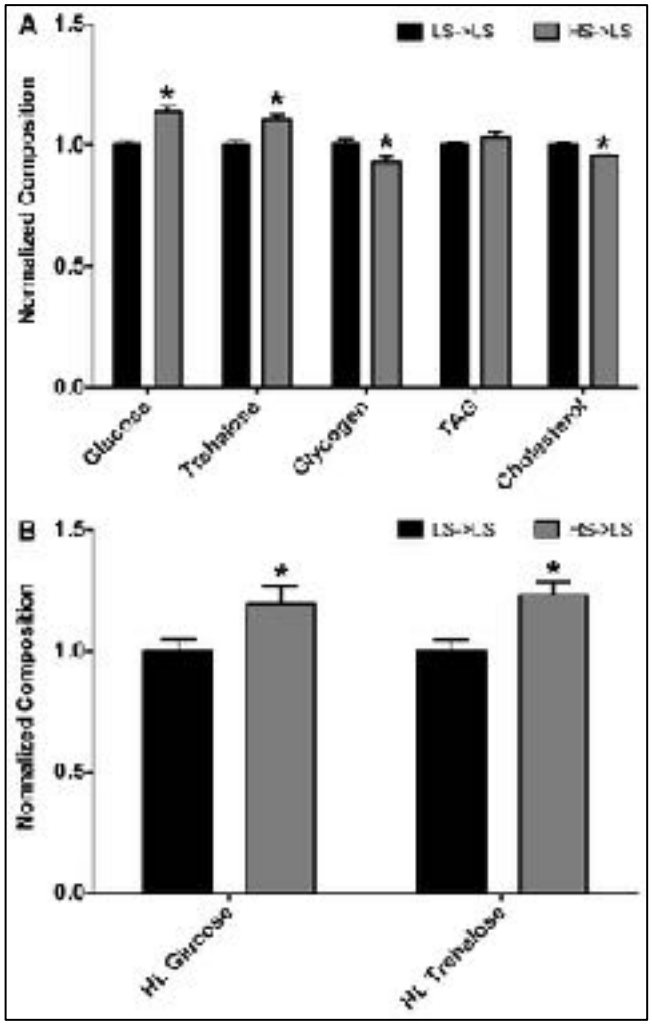
- Changes in disease rate and metabolic status in the offspring of women who experienced the Dutch Famine in 1944–1945. If individuals experienced famine during the last trimester of pregnancy or within the first few months of life, the rates of obesity significantly decreased. However, if individuals experienced famine during the first half of pregnancy, they had a significantly increased rate of obesity

Transgenerational metabolic programming in *Drosophila*



High Sucrose-fed females exhibit an obese-like phenotype

Male larval offspring from HS-fed maternal flies have altered body composition and circulating sugar levels over at least two generations



Environmentally induced epigenetic transgenerational inheritance

Environmental toxicants

Agricultural fungicides (Vinclozclin)

Agricultural pesticides (Methoxychlor)

Industrial contaminants (Dioxin/TCDD)

BPA and phthalates (Plastic compounds)

Herbicides (Atrazine and glyphosate)

Insect repellants (Permethrin and DEET)

Pesticides (DDT)

Industrial toxicants and biocides (Tributyltin)

Hydrocarbons (Jet fuel JP8)

Heavy metals (Mercury)

Other types of exposure

Nutrition (High fat or caloric restriction)

Temperature and drought (Plant health and flowering)

Smoking and alcohol

Stress and trauma (behavioral)

Trends in Endocrinology & Metabolism

Trends in Endocrinology & Metabolism 2020 31478-494DOI: (10.1016/j.tem.2020.02.009)

Epigenetics

Epigenetics can be defined as any (mitotically or meiotically) **heritable** modifications in the function of specific genes not related to modification in the DNA sequence.

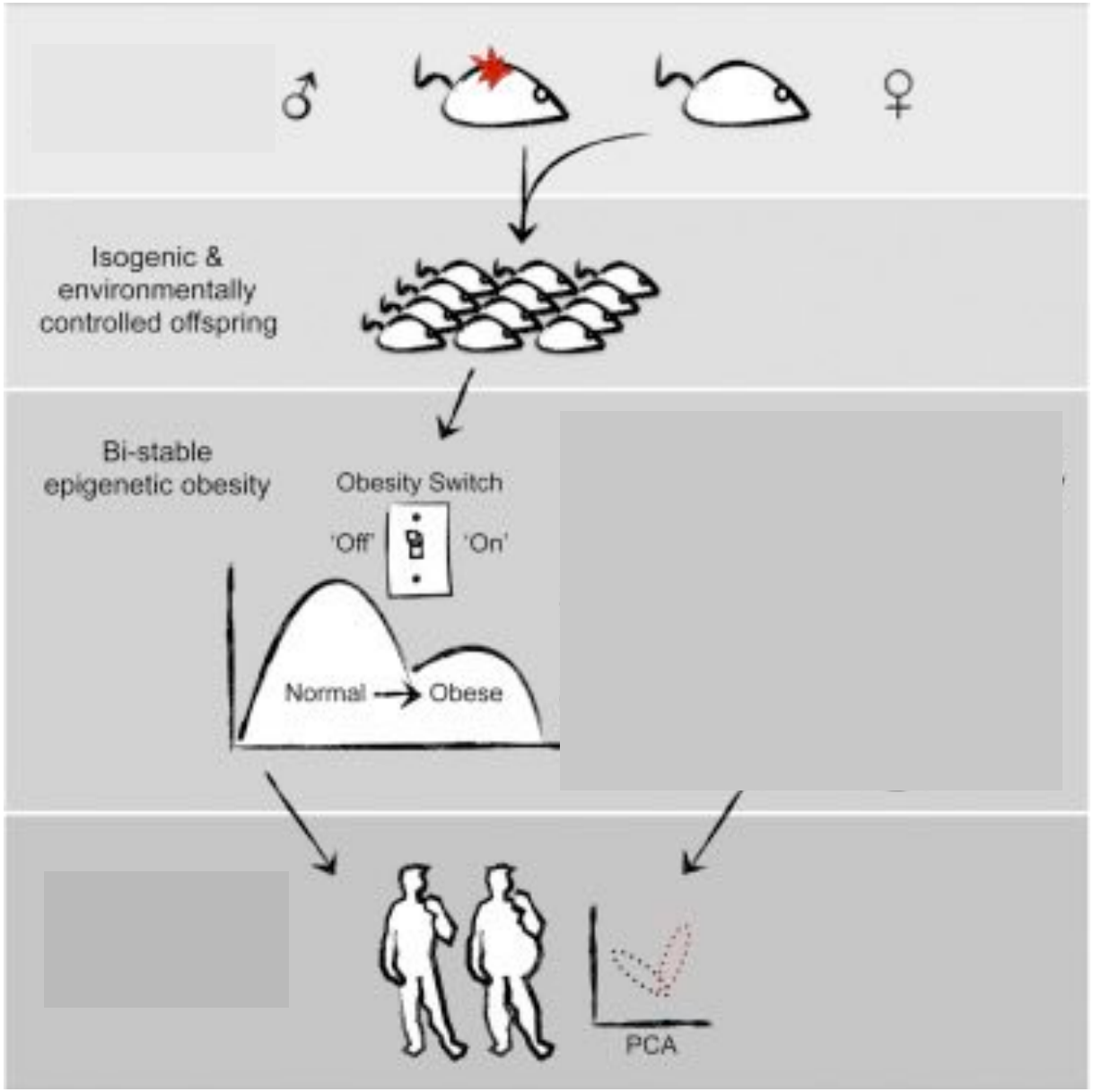
"the study of molecules and mechanisms that can perpetuate alternative gene activity states in the context of the same DNA sequence"

➢ Powerful way to turn a transient external influence / signalling event into a long-lived change in organism performance or function

Common mechanisms may include but not limited to:

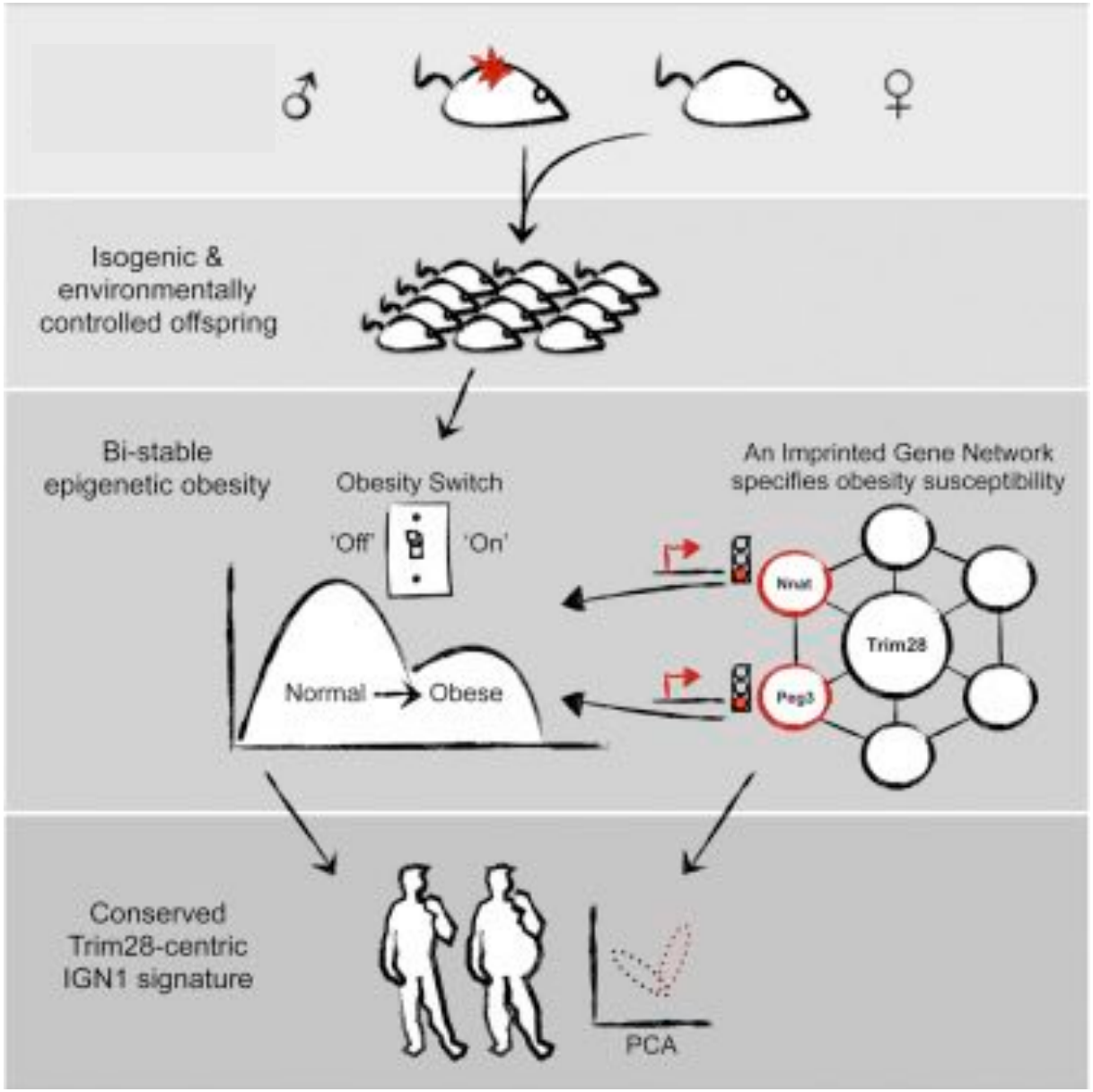
- Histone modification/histone variants
- Regulatory non-coding RNAs
- DNA methylation

Example of epigenetically transmitted characters: Obesity



Dalgaard et al., 2016, Cell

Example of epigenetically transmitted characters: Obesity



TRIM28 is a large multi-domain protein that supports heterochromatin deposition and silencing

Dalgaard et al., 2016, Cell

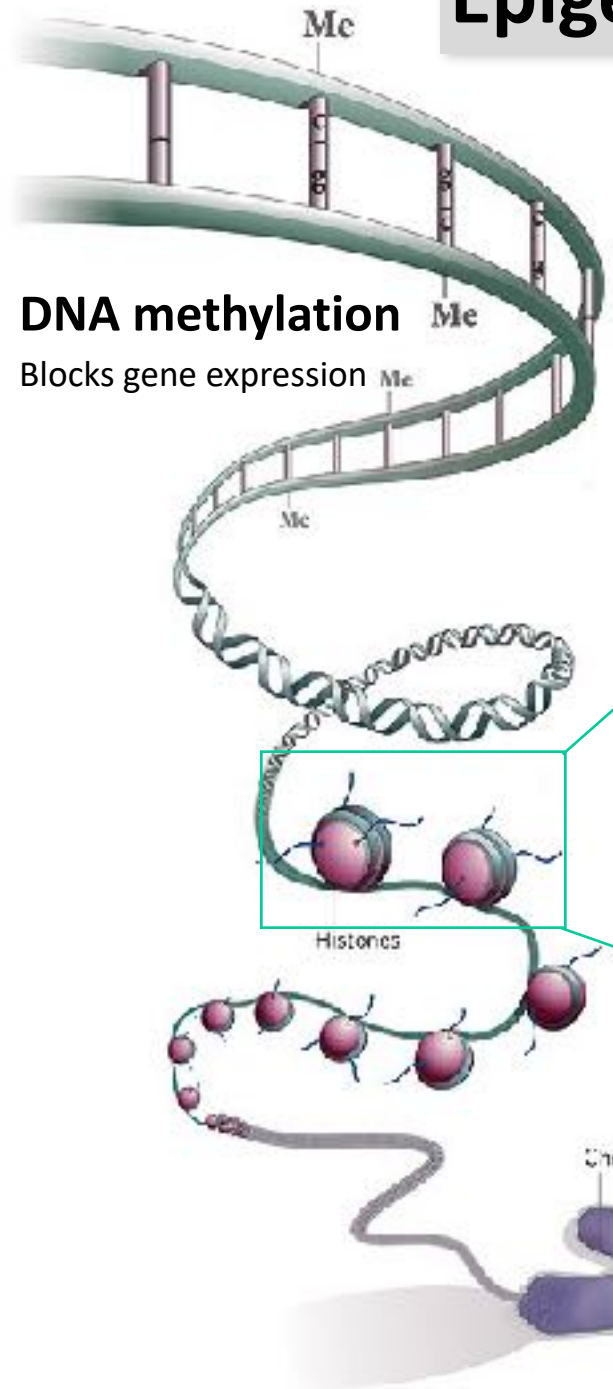
Epigenetic Control of Cell Function

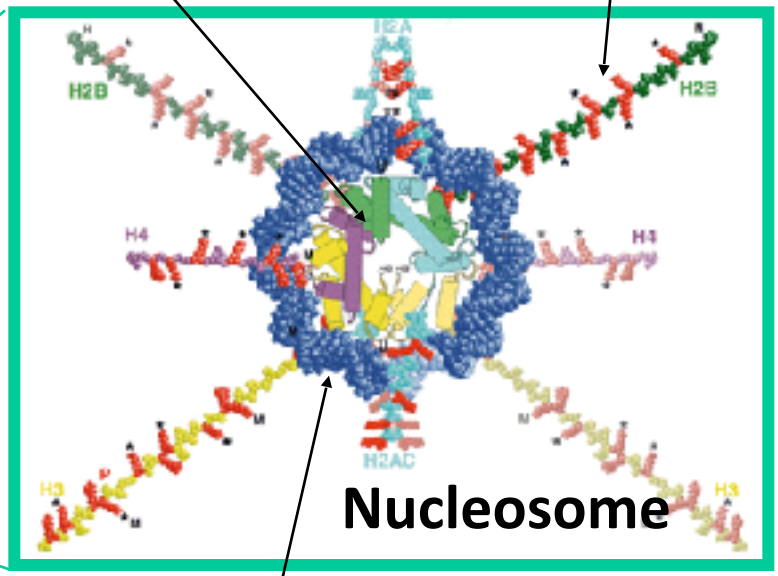
DNA methylation

Blocks gene expression

Histone Proteins (core) H2A,H2B, H3 & H4

Histone *Tails* are chemically modified by many types of enzymes: alter the interaction between DNA and histones and DNA accessibility to transcription.



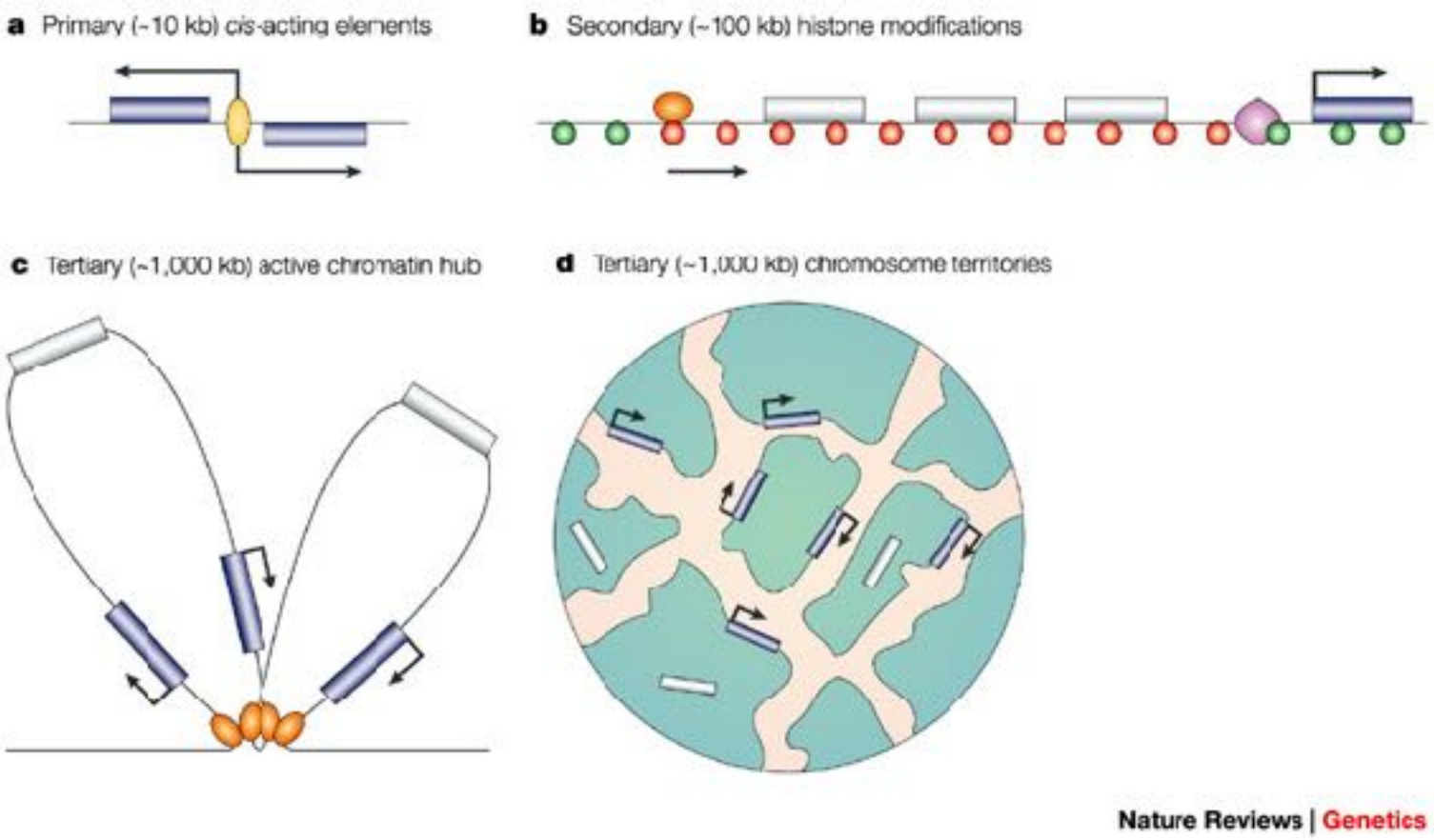


Nucleosome

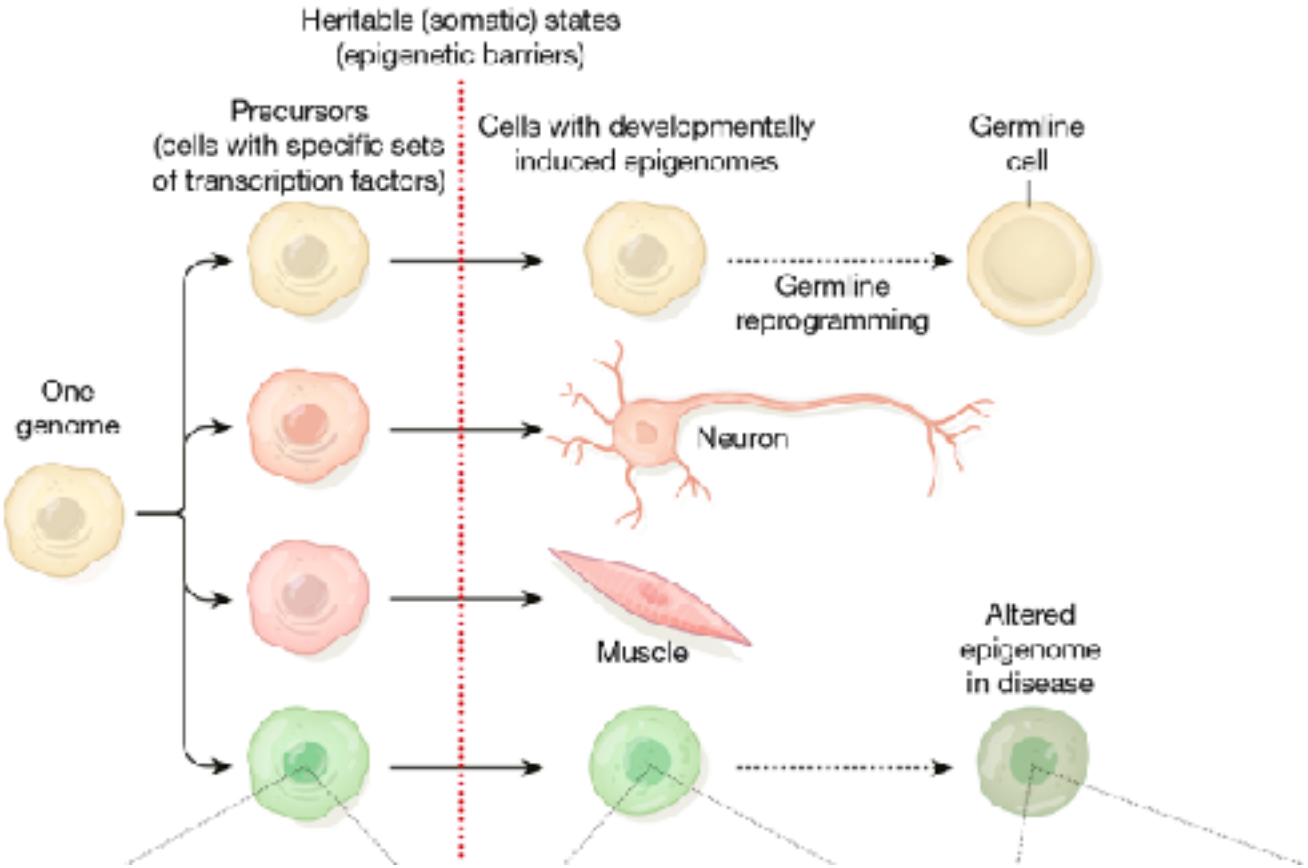
DNA

Heterochromatin vs. Euchromatin

Nuclear architecture and histone code



Epigenetics and maintenance of cell states



Heard and Cavalli, Nature 2019

Two main questions for chromatin inheritance

- **First:**

How are chromatin state specified?

- **Second**:

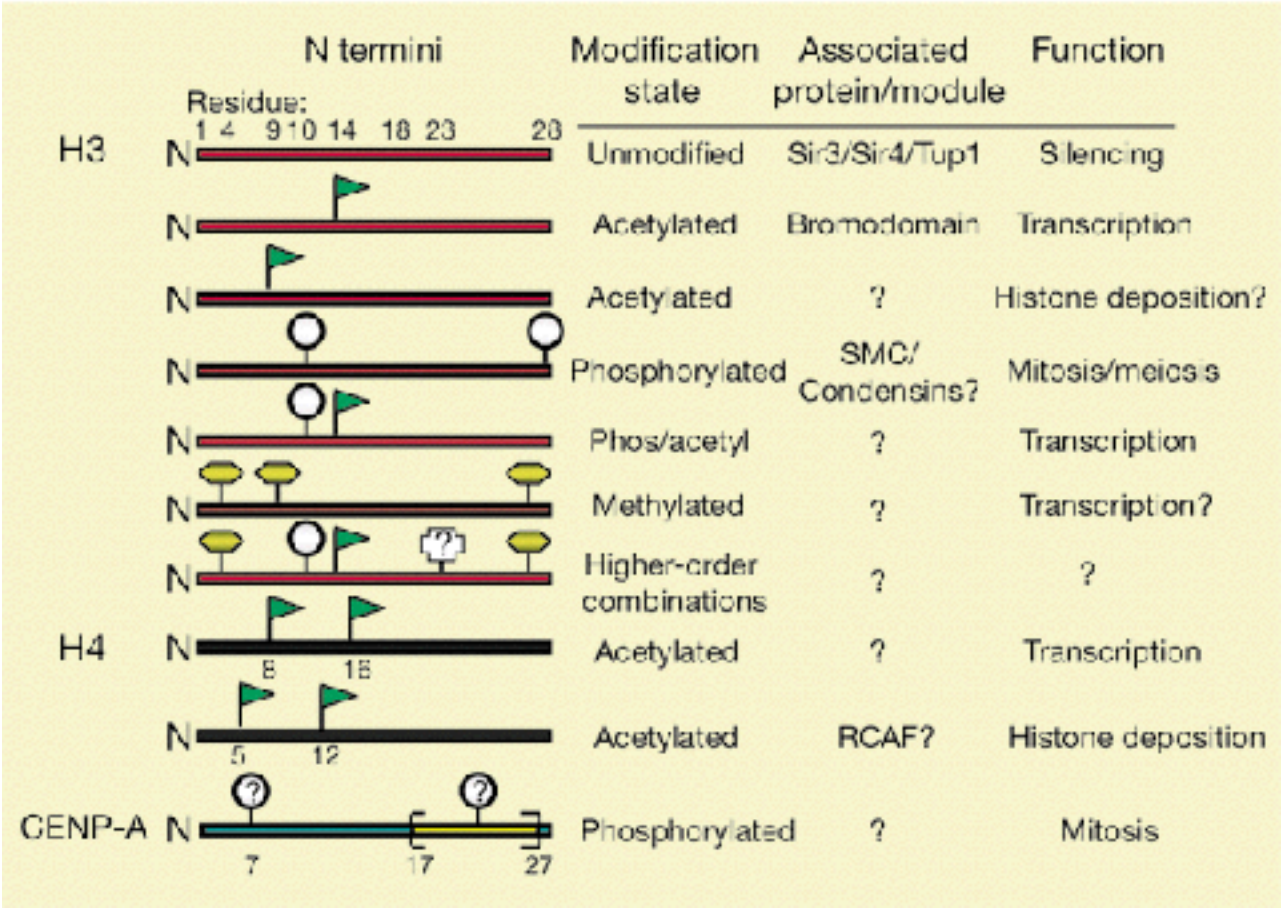
How are these states transmitted with high fidelity?

Major carriers of epigenetic information

- Heterochromatin components: megabasesized repetitive DNA domains coated in a specific histone H3K9 trimethylation mark. Heterochromatin components can both write and read the H3K9me3 mark and compact their target chromatin.
- Polycomb proteins: Polycomb (PcG) and Trithorax:two antagonistic groups that maintain the memory of spatial patterns of expression of genes throughout development, (maintenance of developmentally or environmentally programmed expression states)
- Noncoding RNAs: Many different classes and function. They are also involved in the regulation of chromatin architecture.
- DNA methylation: involve specific proteins that recognize CpG hemimethylated DNA and thereby redeposit DNA methylation on newly replicated DNA.

Histone Code Hypothesis

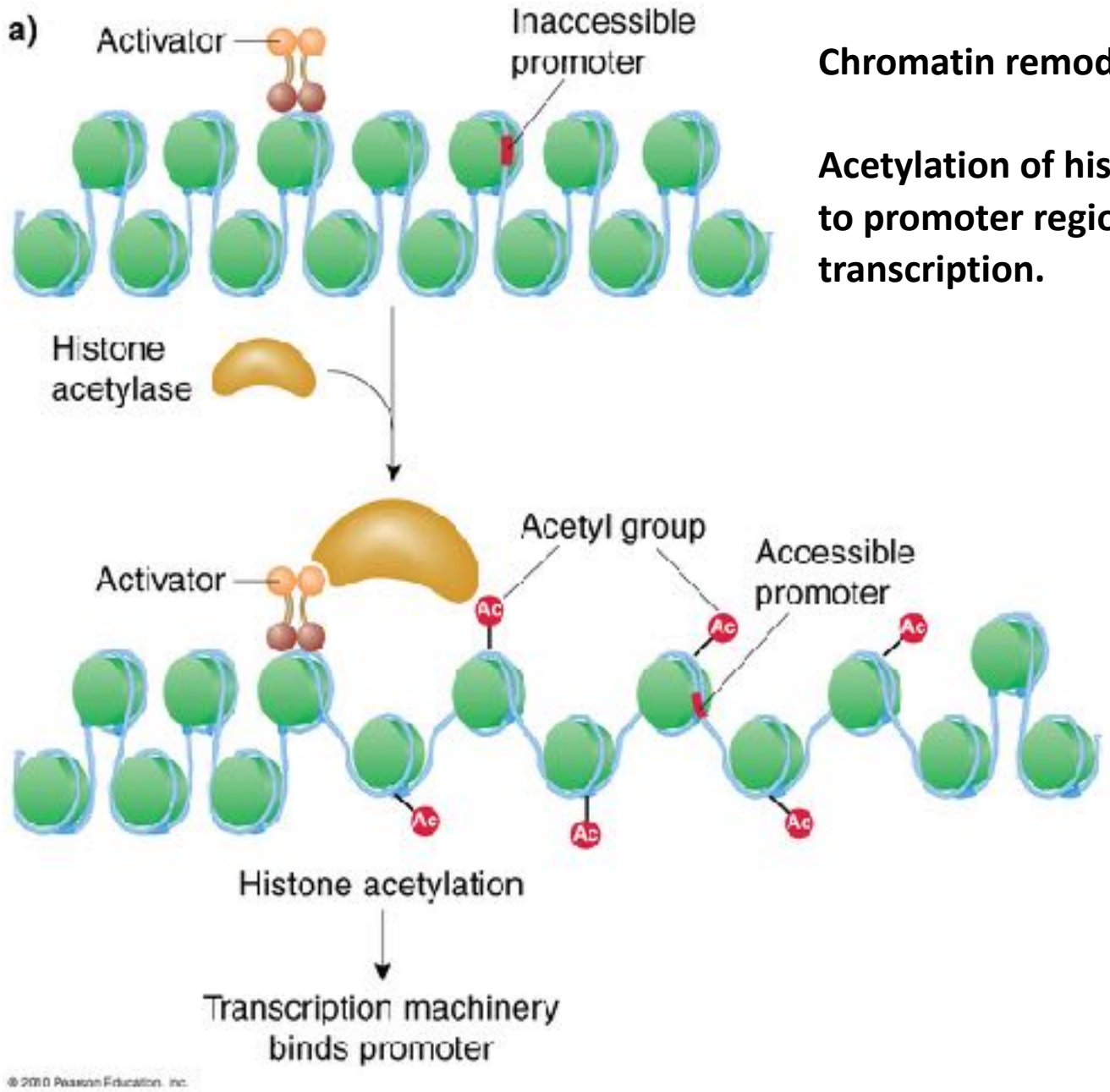
As proposed by Allis and Strahl: “that multiple histone modifications, acting in a combinatorial or sequential fashion on one or multiple histone tails, specify unique downstream functions”



Strahl, B.D. and Allis, C.D., *Nature*. 2000

Histone code hypothesis

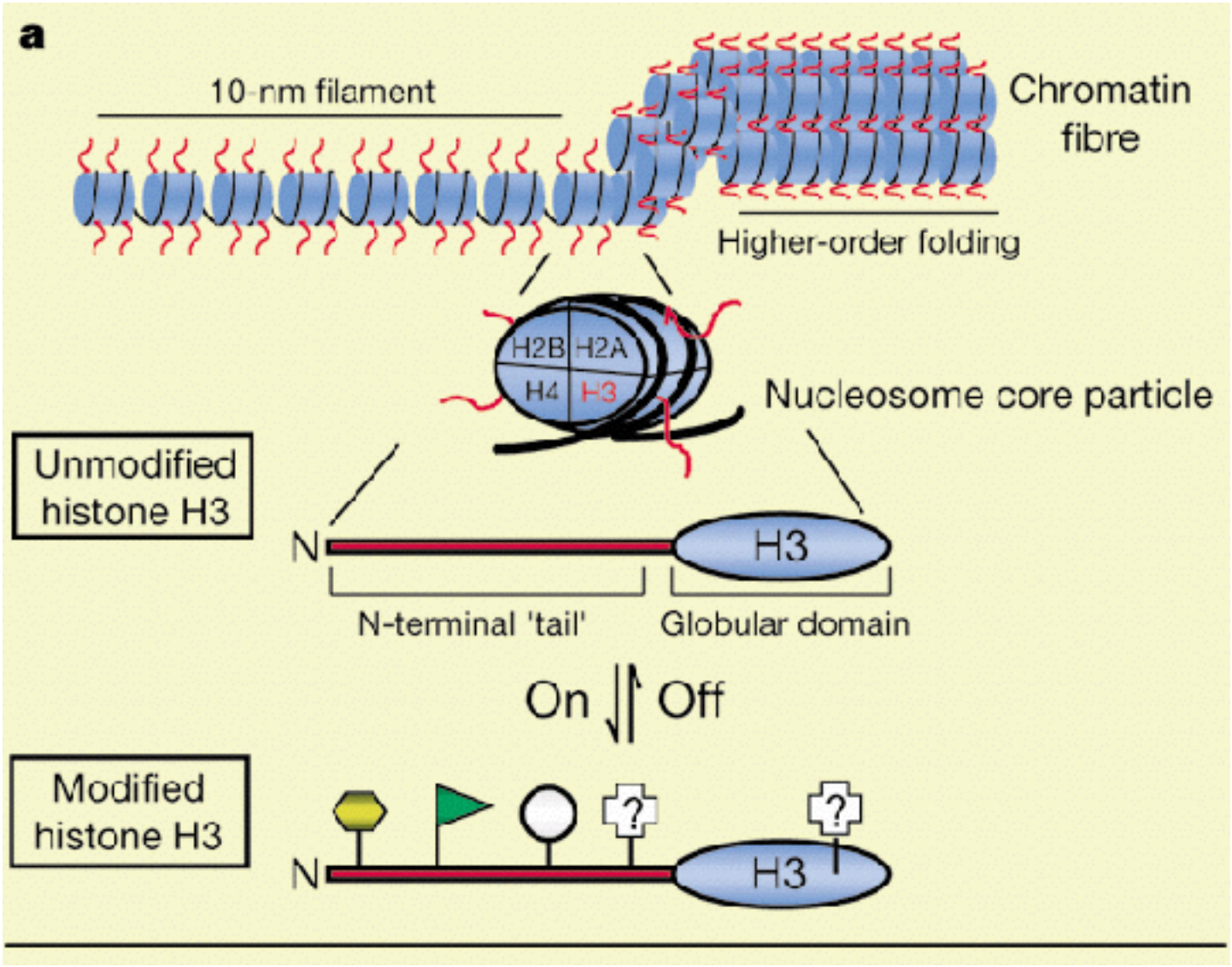
- **Post-translational modifications** are made on a specific histone residue may regulate modification of the same or different residues within the same or a different histone
- Different types or combinations of modifications are **read** by chromatin-modulating proteins, resulting in regulation of chromatin structure and, hence, transcription



Chromatin remodeling

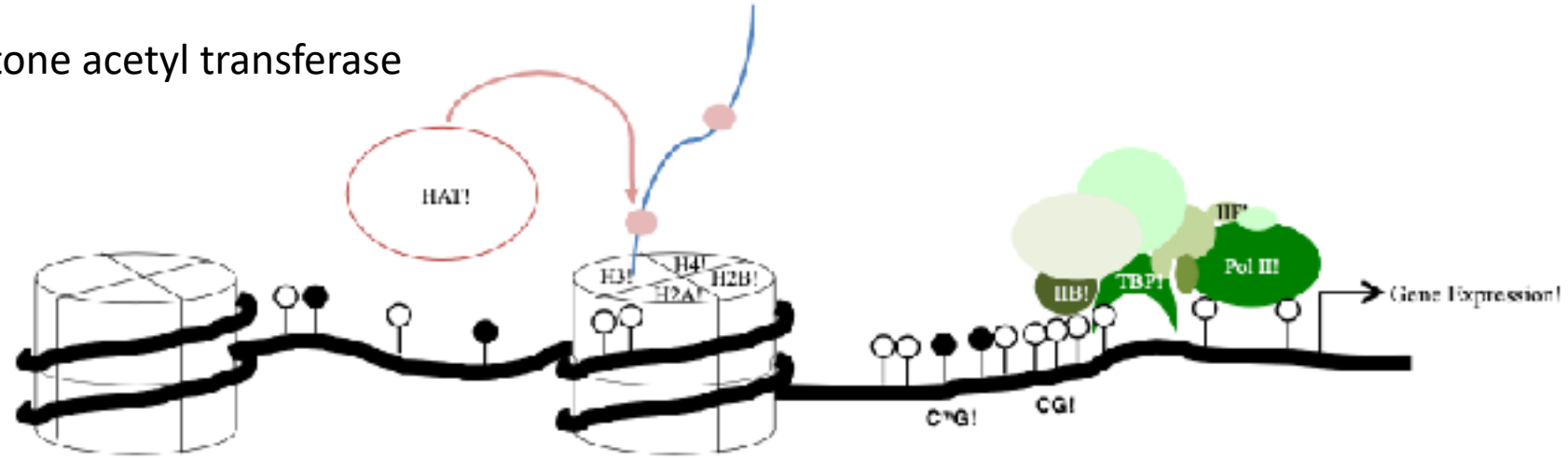
Acetylation of histones enhances access to promoter region and facilitates transcription.

Histone Modifications Alter Chromatin Structure and Gene Activation

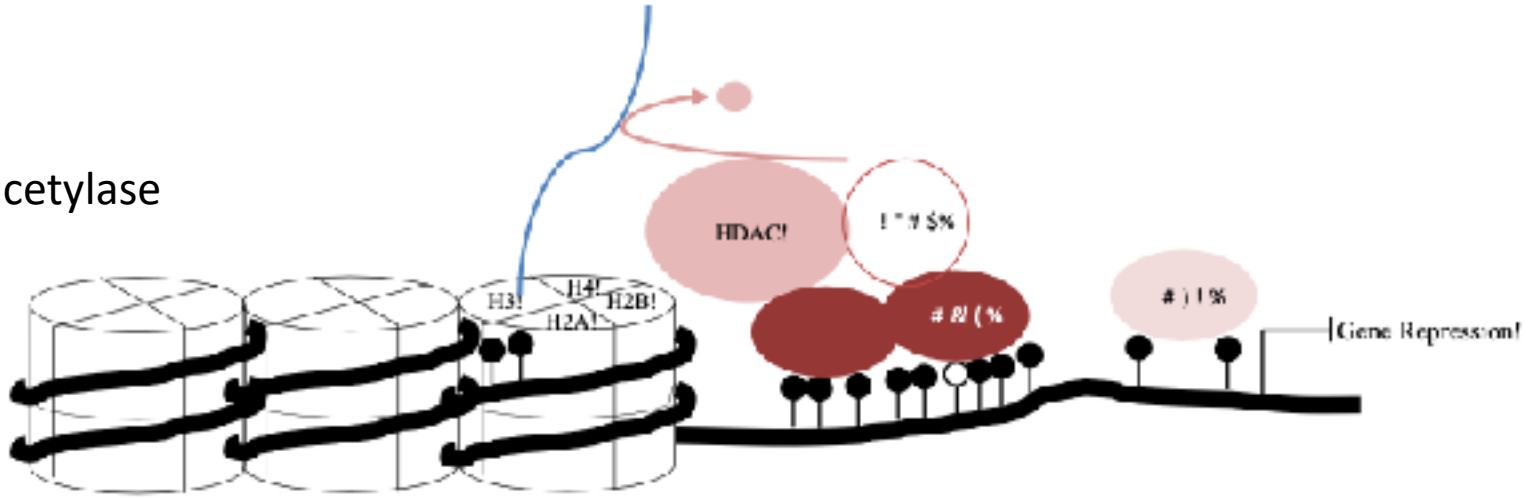


Organization of the Epigenome

Histone acetyl transferase

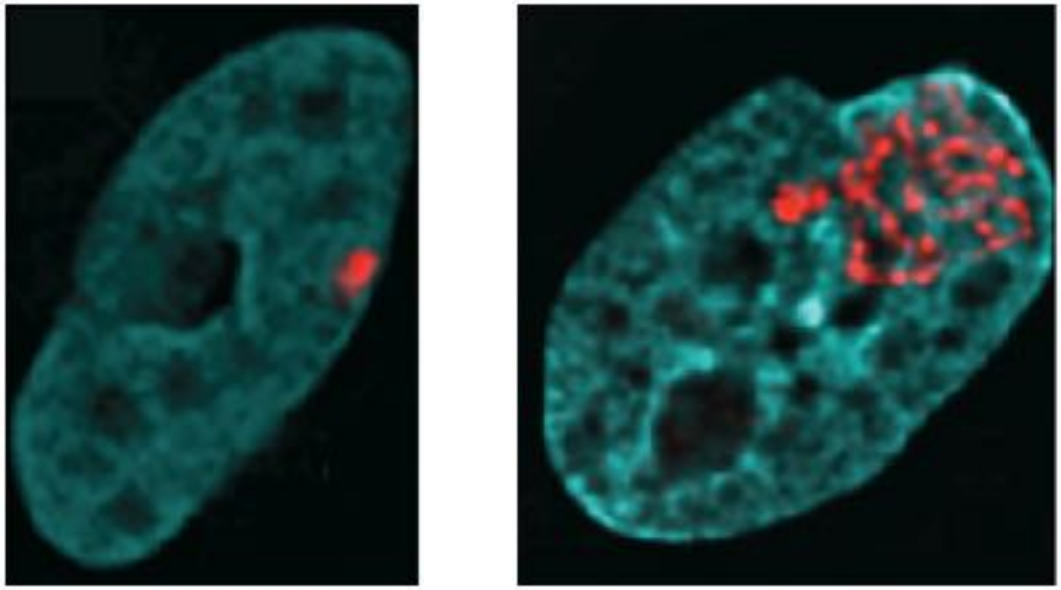


Histone deacetylase



Experimental demonstration of chromatin decondensation in vivo

An experiment demonstrating that some TF activation domains regulate chromatin condensation is shown. In this experiment, DNA consisting of a tandemly repeated *lac* operator sequence was incorporated into a yeast chromosome. When a fluorescently tagged wild-type Lac repressor is introduced into the cell, the DNA is shown to be confined to a small region of the nucleus (*left*). However, when the Lac repressor is fused to a yeast activation domain that interacts with a histone acetylase, staining spreads throughout a larger volume of the nucleus indicating the DNA has been decondensed (*right*).



Histone modifications

- "**Histone tails**" = N-termini of Histones
  - e.g. Histones H3 and H4 N-termini
  - e.g. Histones H2A, H2B and H1 N- and C-termini
- Post-translational modifications:
  - Acetylation : K-ac
  - Methylation : K-me1, K-me2, K-me3 and R-me1, R-me2a, R-me2as
  - Phosphorylation : S-Ph and T-ph
  - Ubiquitination : K-ub
  - Sumoylation : K-su
  - ADP-ribosylation : E-ar
  - Deimination : R > Cit
  - Proline isomerization : P-cis > P-trans

Acetylation, methylation, phosphorylation, and deimination can appear and disappear on chromatin within minutes of stimulus arriving at the cell surface

Histone modifications

- **Histone acetylation**
  - **Histone acetyl transferases (HATs)**
    - Adds acetyl groups to histone tails
    - Reduces interaction of histones with DNA
    - Facilitates transcription
  - **Histone de-acetylases (HDACs)**
    - Removes acetyl groups from histone tails
    - Increases interaction of DNA and histones
    - Represses transcription (usually)
  - May involve the same Lys residues as targeted for methylation

Histone modifications

- **Histone methylation**
  - **Histone methyl transferases (HMTs):**
    - **Histone lysine methyl transferases (HKMTs)**
      - Methylate lys (K) residues
    - **Protein arginine methyl transferases (PRMTs)**
      - Methylate arg (R) residues
  - Varying number of methyl groups:
    - Lys – mono- di- or tri-methylated (on e-amino group)
    - Arg – mono- or di-methylated (symmetric or asymmetric) (on guanidino-e-amino groups)
  - Methylation can result in repression or activation of expression

Histone modifications

- **Histone phosphorylation**
  - E.g. by aurora AIR2–Ipl1 kinase family
    - Required for chromosome condensation and cell cycle progression
  - E.g. by MSK1 and 2 or IKKa kinase
    - Required for signal transduction leading to gene activation
    - Can prevent nearby histone methylation due to (i) steric hindrance or (ii) facilitation of competing acetylation
  - Reversed by phosphatases like PP1 or PP2
  - Alters recruitment of binding proteins; e.g.-
    - If phospho-acceptor precedes methylated residue → activates transcription
    - If phospho-acceptor follows methylated residue → silences transcription

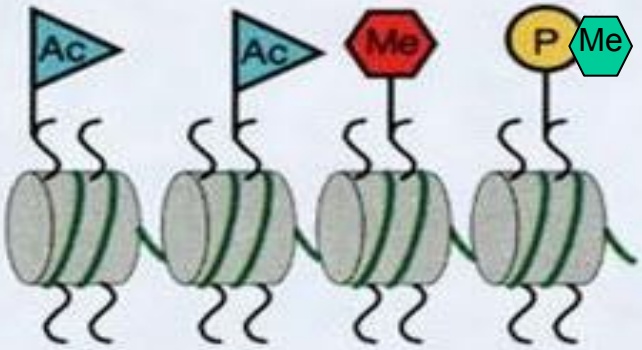
Histone modifications

- **Histone ubiquitination**
  - Mono-ubiquitination (by Rad6) and recruitment of proteasomal ATPases (Rpt4 and Rpt6)
    - alters chromatin structure
    - regulates H3 methylation
  - De-ubiquitination (by SAGA-associated Ubp8)
    - regulates mono- vs tri-methylation

Structure & Epigenetics of Euchromatin versus Heterochromatin

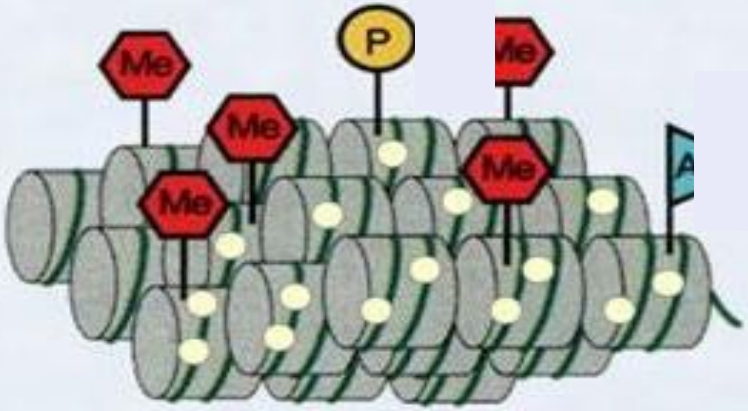
DNA methylation and histone modifications help to compartmentalize the genome into domains of different transcriptional potentials

Euchromatin



- High histone acetylation
- Low DNA methylation
- H3-K4 methylation

Heterochromatin

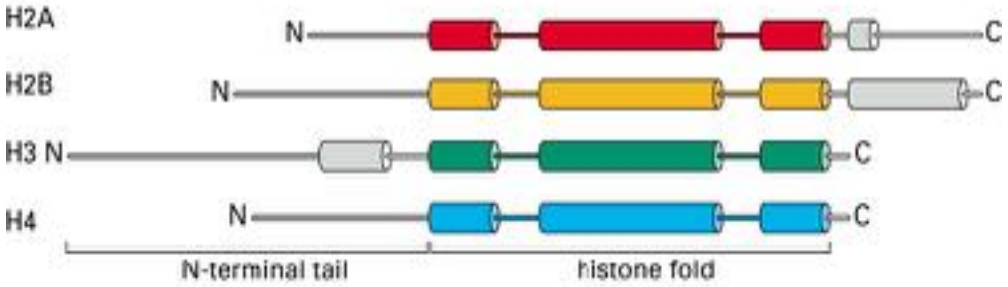


- Low histone acetylation
- Dense DNA methylation
- H3-K9 methylation

The histone fold

Ribbon drawing

Simple.
Conserved.
Adopted by all 4 “core” histones (H2A, H2B, H3 and H4).

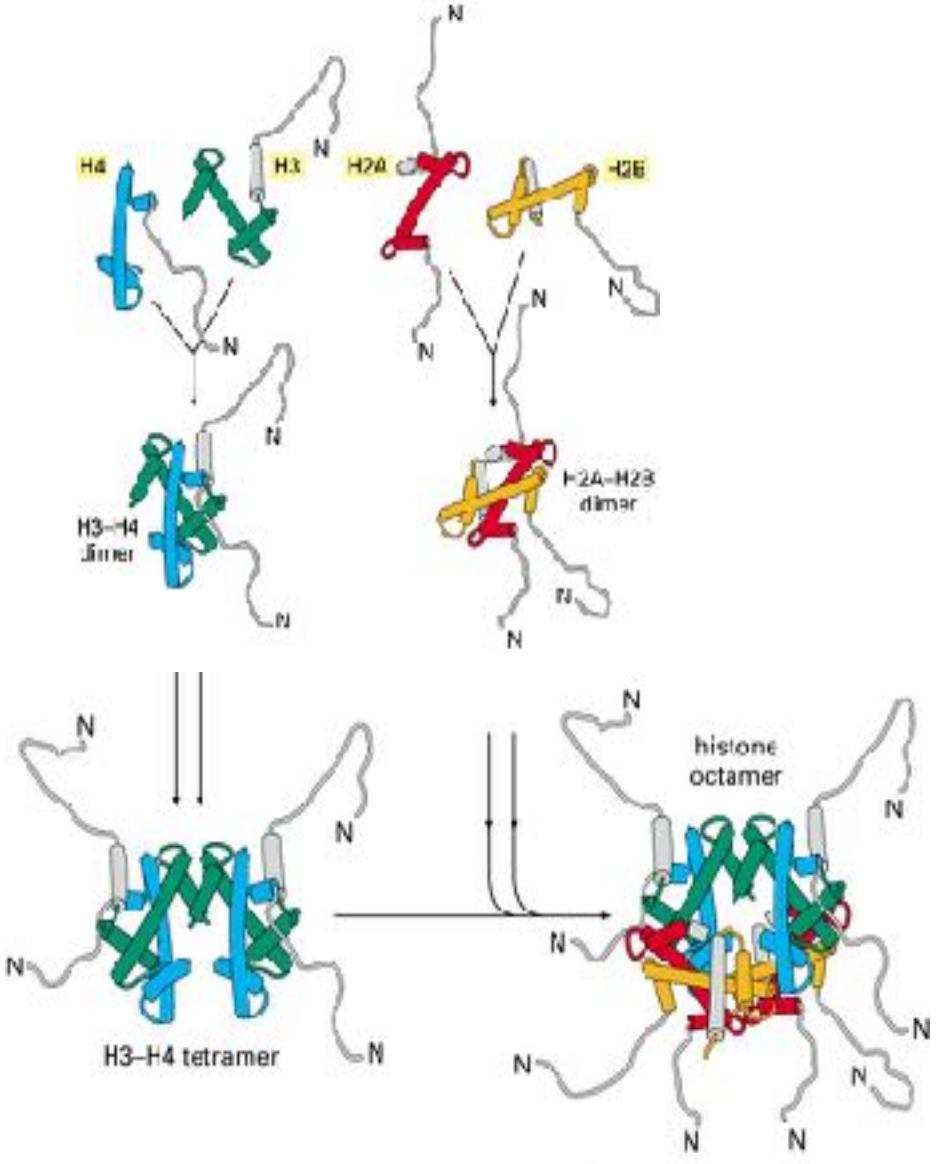


H3-H4 tetramer binds two H2A-H2B dimers to form the histone octamer

Monomers

Dimers

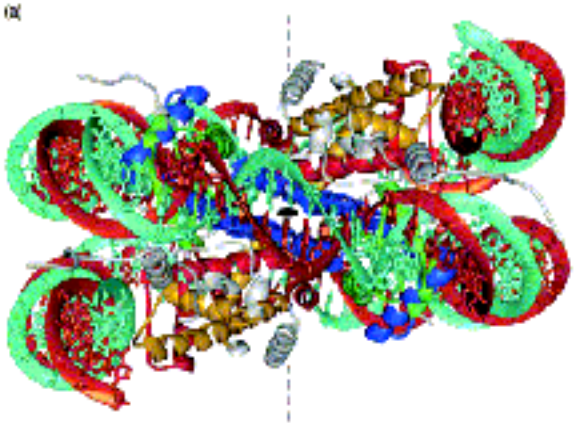
Tetramer



Octamer

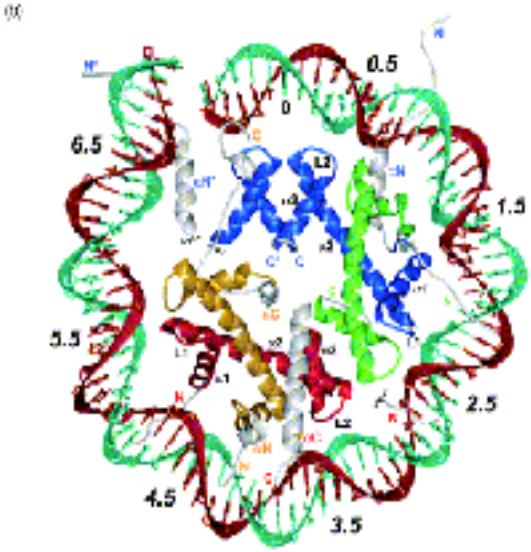
Crystal structure of the nucleosome

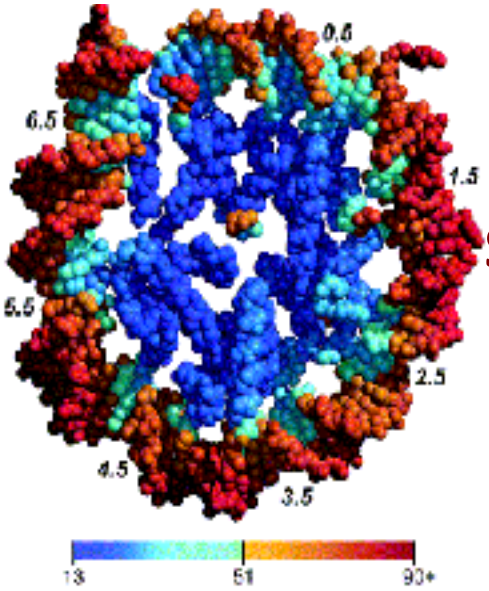
Ribbon drawings



147 bp of DNA wrapped almost twice around 8 core histones: (H2A, H2B, H3, H4)$_2$

"Top" view





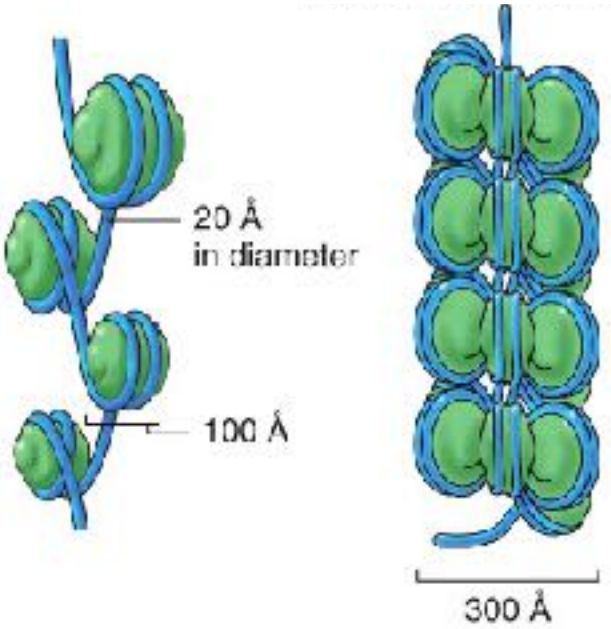
Space-filling drawing

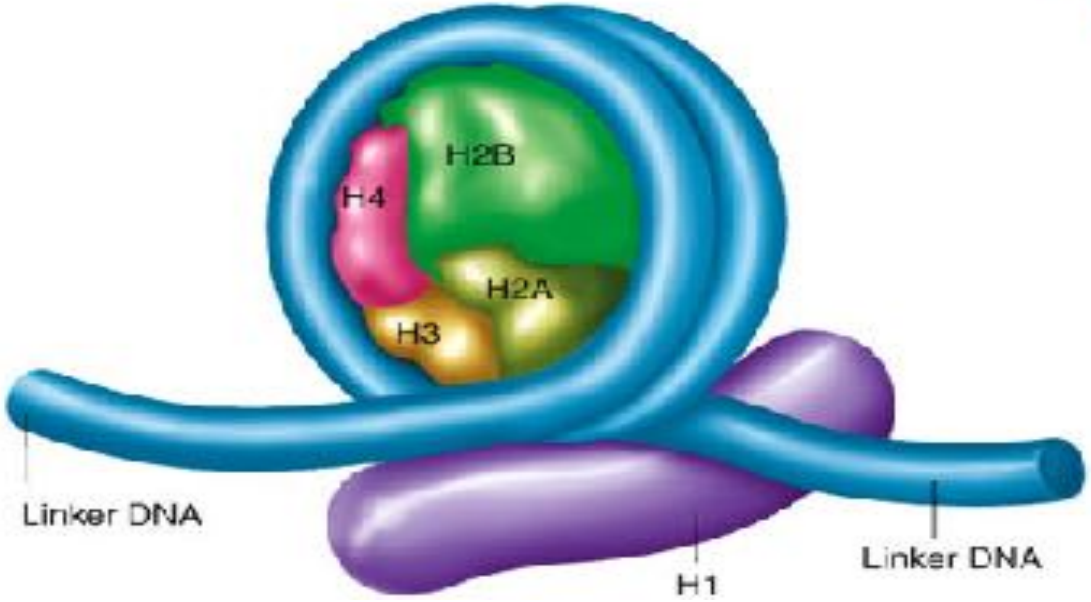
"Side" view of 1/2 nucleosome

Carolyn Luger and Tim Richmond
Cur. Op. Struct. Biol. 1998

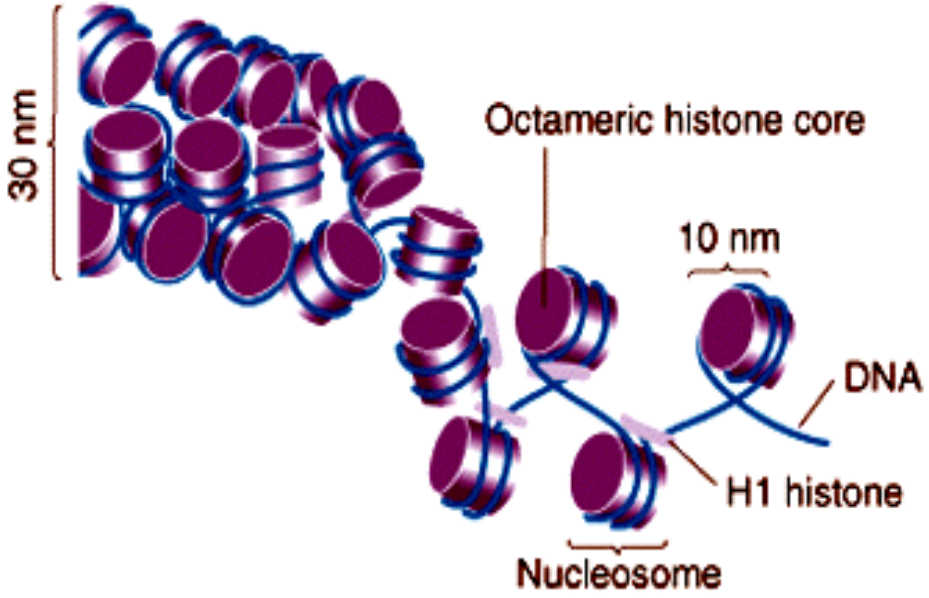
Chromosome Structure

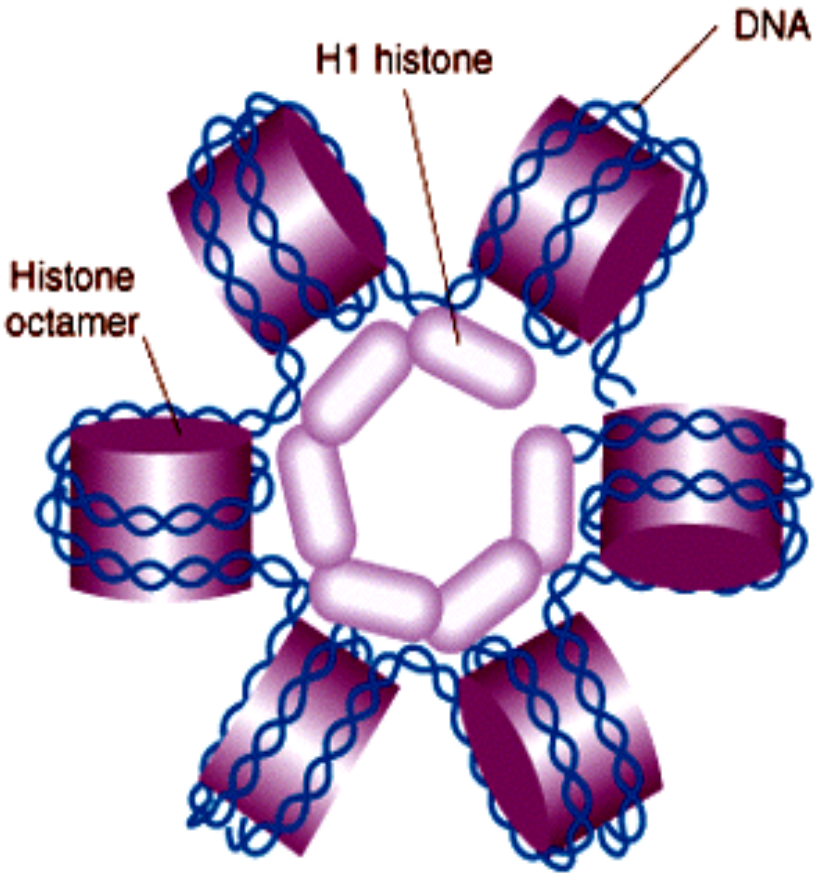
- **Nucleosome**
  - fundamental unit of chromatin
  - 147bp DNA wound 1.75 turns around histone core (octamer)
    - $2(H2A/H2B) + (H3/H4)_2$
  - 11 nm fiber ("beads on a string")



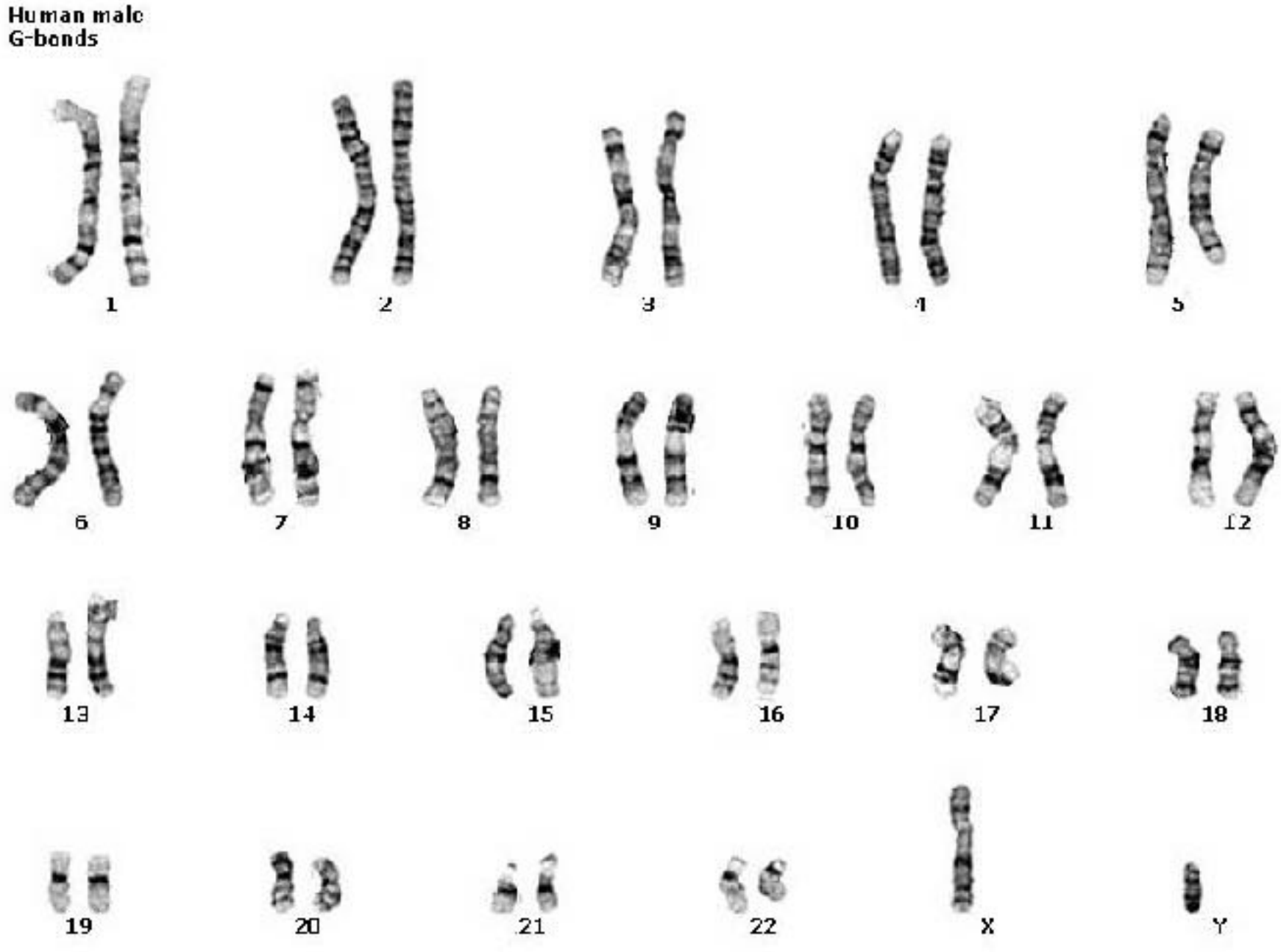


Chromatine packaging
Nucleosomes / 30nm Solenoid

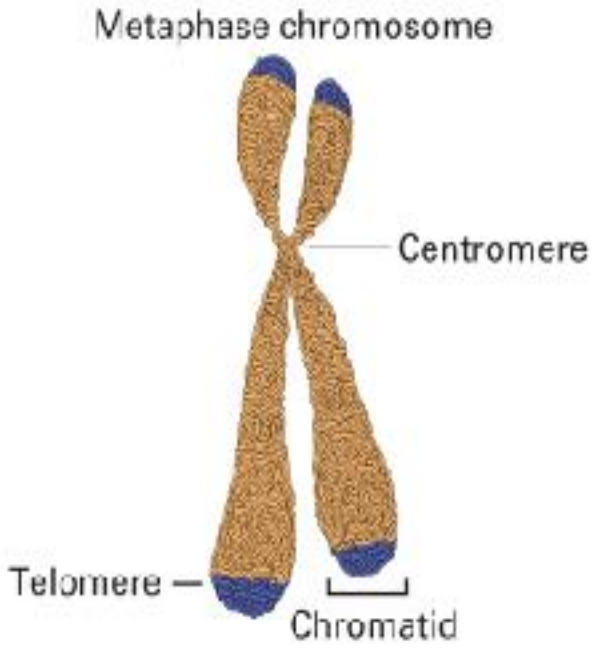


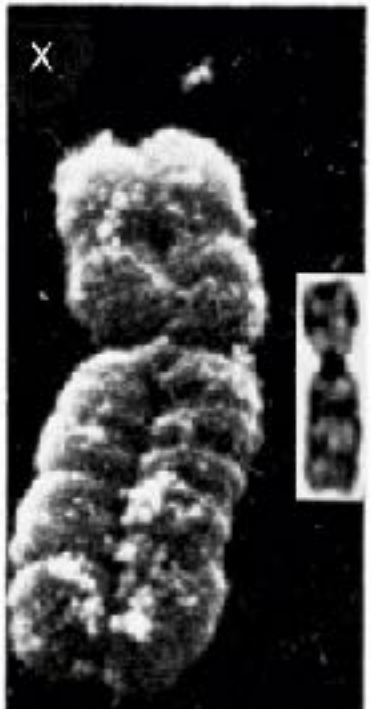


*Set of 46 **homologous chromosomes** of the human male*



Functional DNA sites – Telomere & Centromere

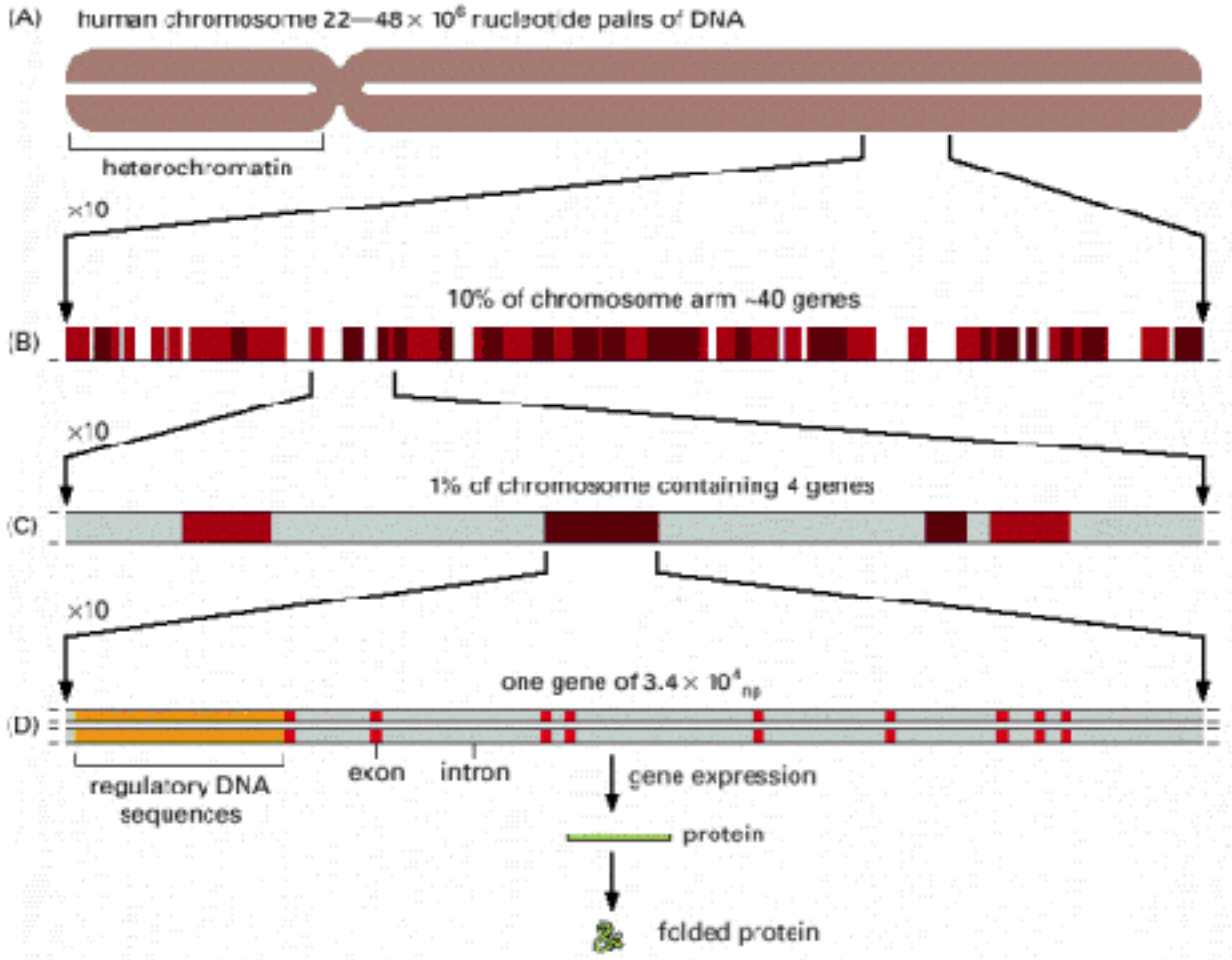




Schematic and electron micrograph of X chromosome.

- **Telomeres protect the ends.**
- **Centromere is at the primary constriction. It mediates chromosome cohesion, spindle attachment and chromosome segregation.**

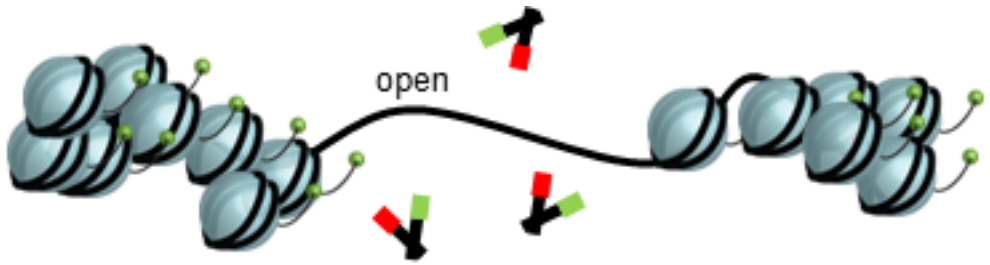
The organization of genes on a human chromosome.



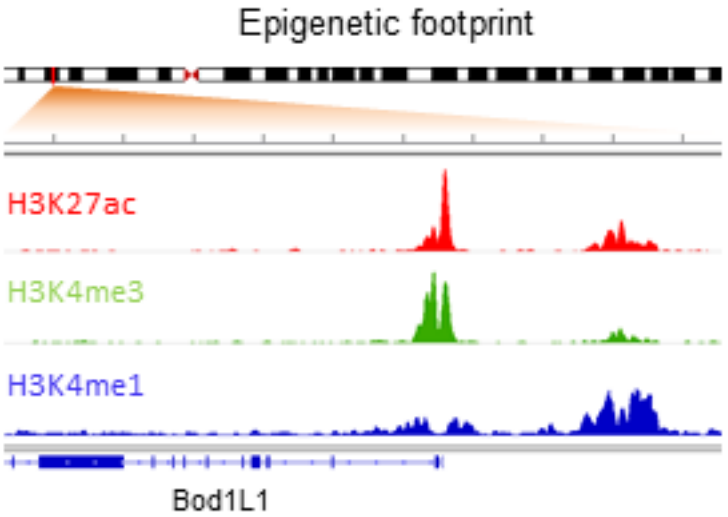
With the publication of the “first draft” of the entire human genome in 2001 and the “finished DNA sequence” in 2004.

Analysis of transcriptional rate in open chromatin in human cells (mesure of RNA produced at a given transcription site on the chromosome).

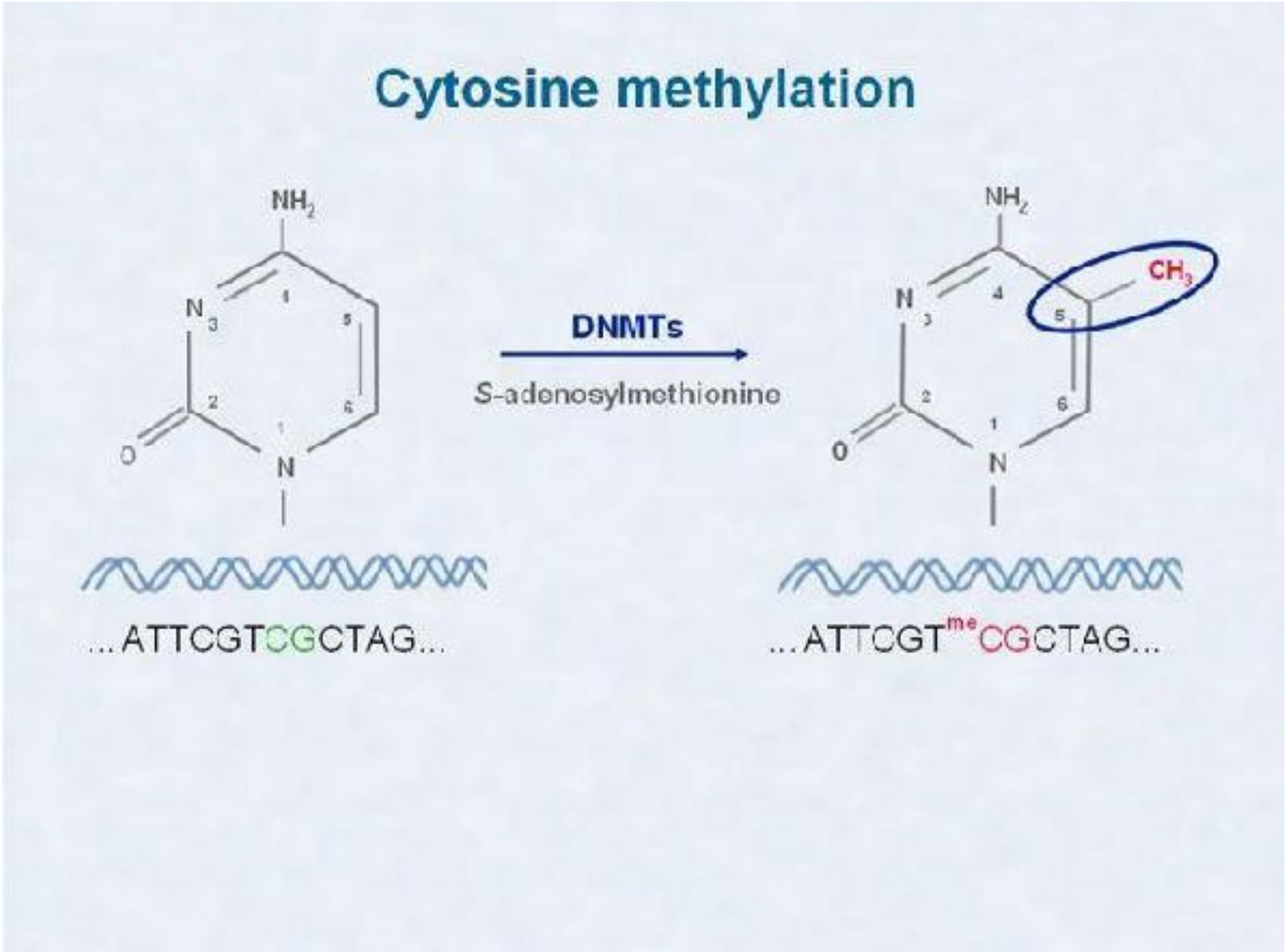
H3K27ac



1 → Enrich for open chromatin using ATAC-Seq

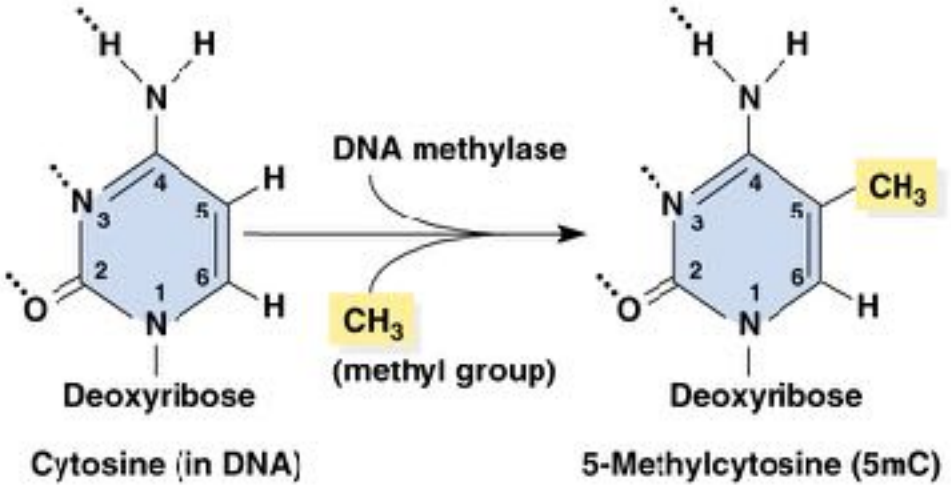


Methylation of Cytosine in DNA

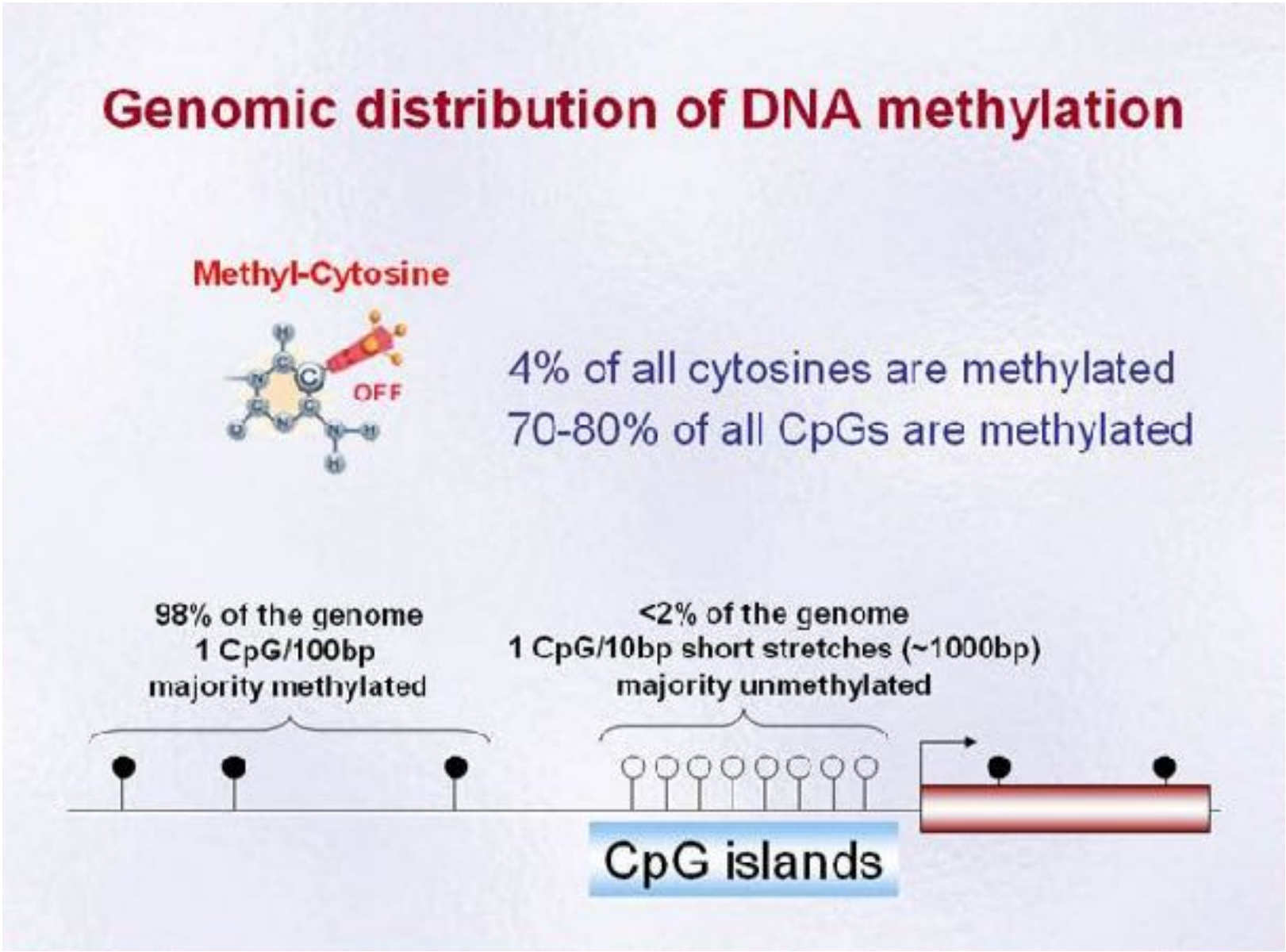


DNA methylation and transcription control:

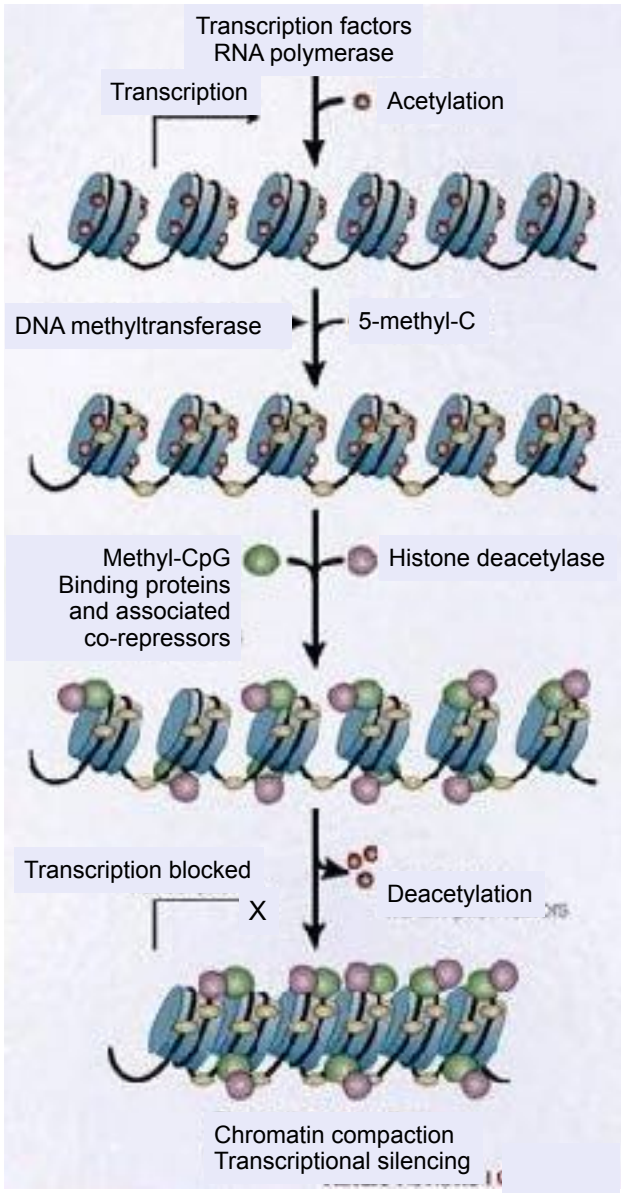
- Small percentages of newly synthesized DNAs (~3% in mammals) are chemically modified by methylation.
- Methylation occurs most often in symmetrical CG sequences.
- Transcriptionally active genes possess significantly lower levels of methylated DNA than inactive genes.
    - A gene for methylation is essential for development in mice (turning off a gene also can be important).
    - Methylation results in a human disease called fragile X syndrome; FMR-1 gene is silenced by methylation.

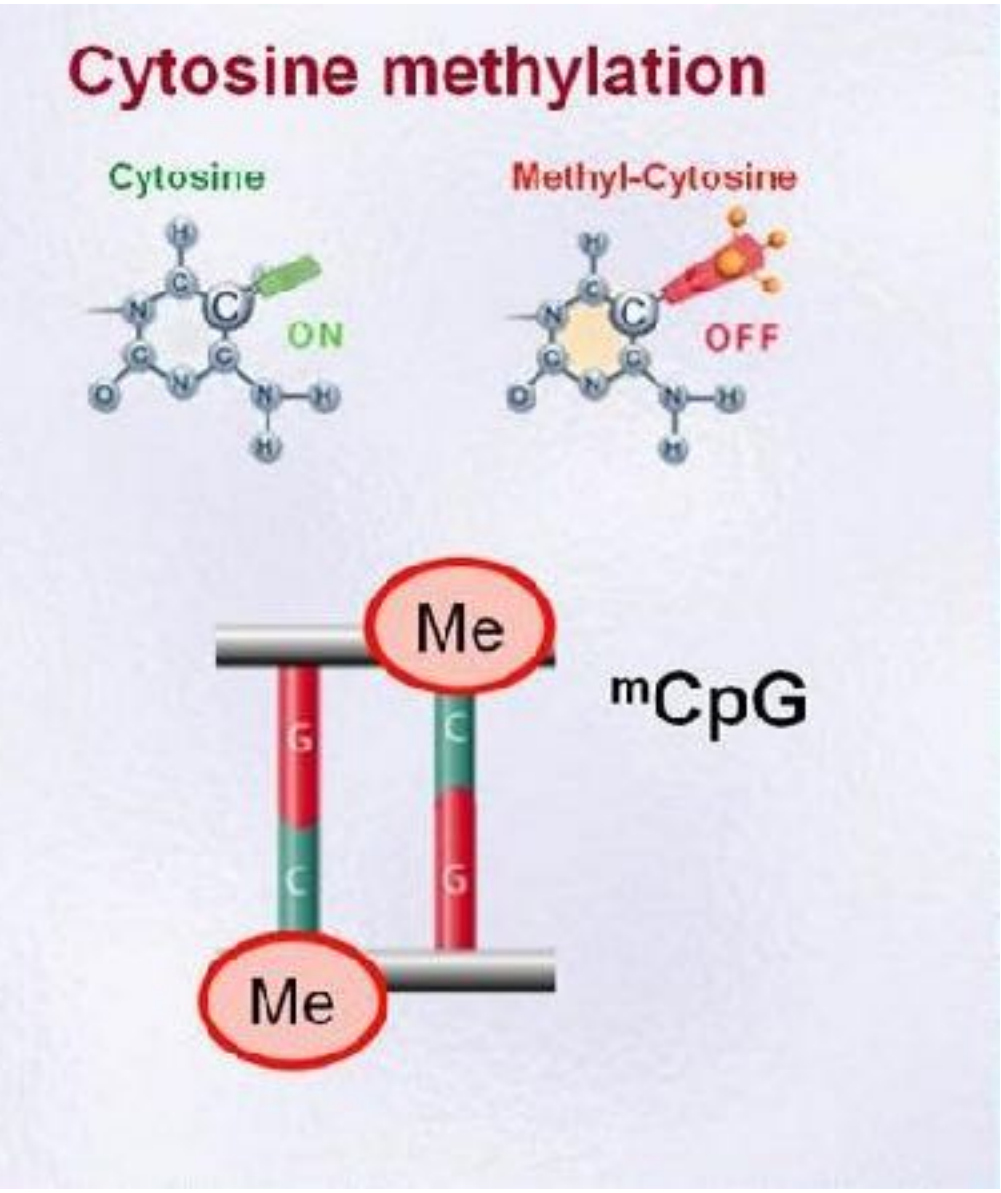


Critical CpG Sequences in CpG Islands Near Promoters



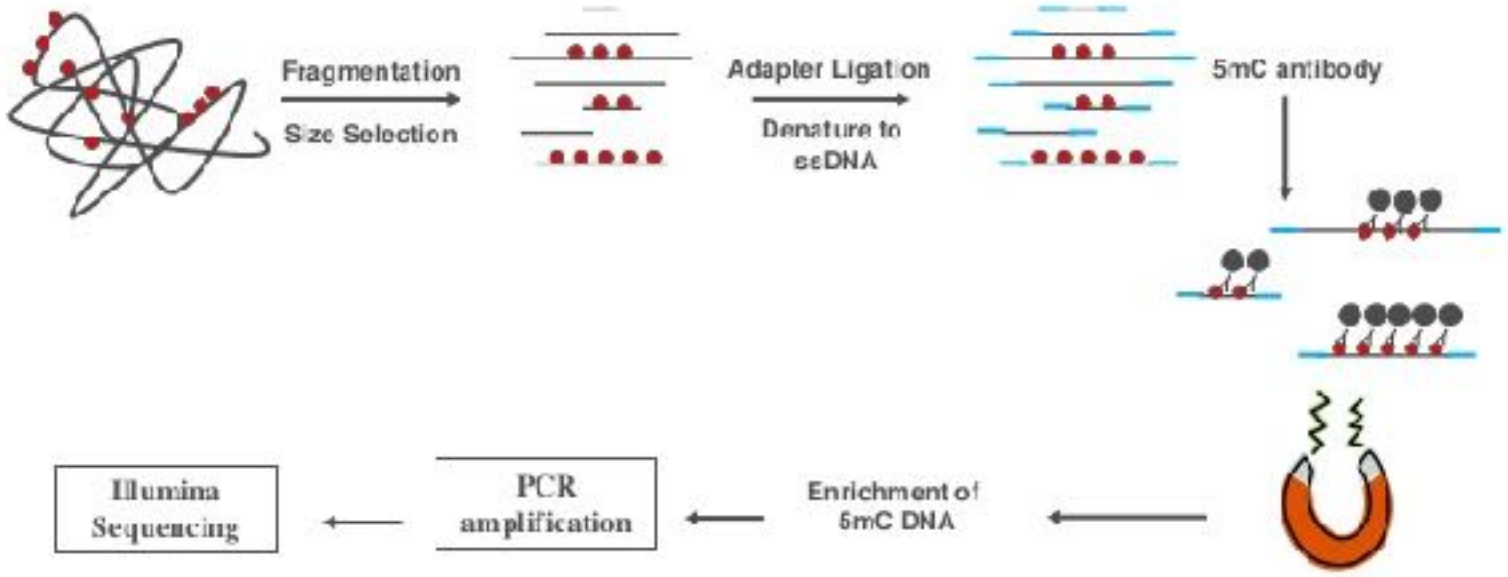
Cytosine Methylation Maintains Inactive-Condensed Chromatin State





5mC mapping technique

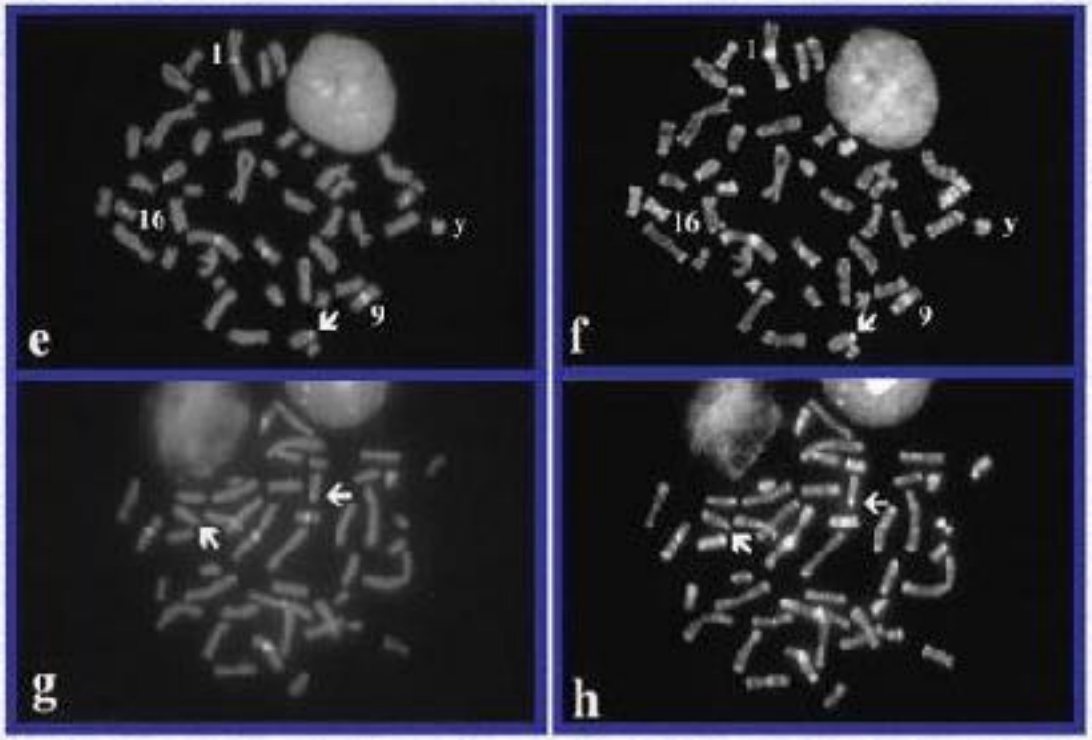
Me-DIP-Seq: Methylated DNA ImmunoPrecipitation Sequencing



5-Methyl Cytosine is Found in Heterochromatic Regions

The distribution of cytosine methylation in mammals

- Heterogeneity visible at cytogenetic scale
- Associated with heterochromatic regions

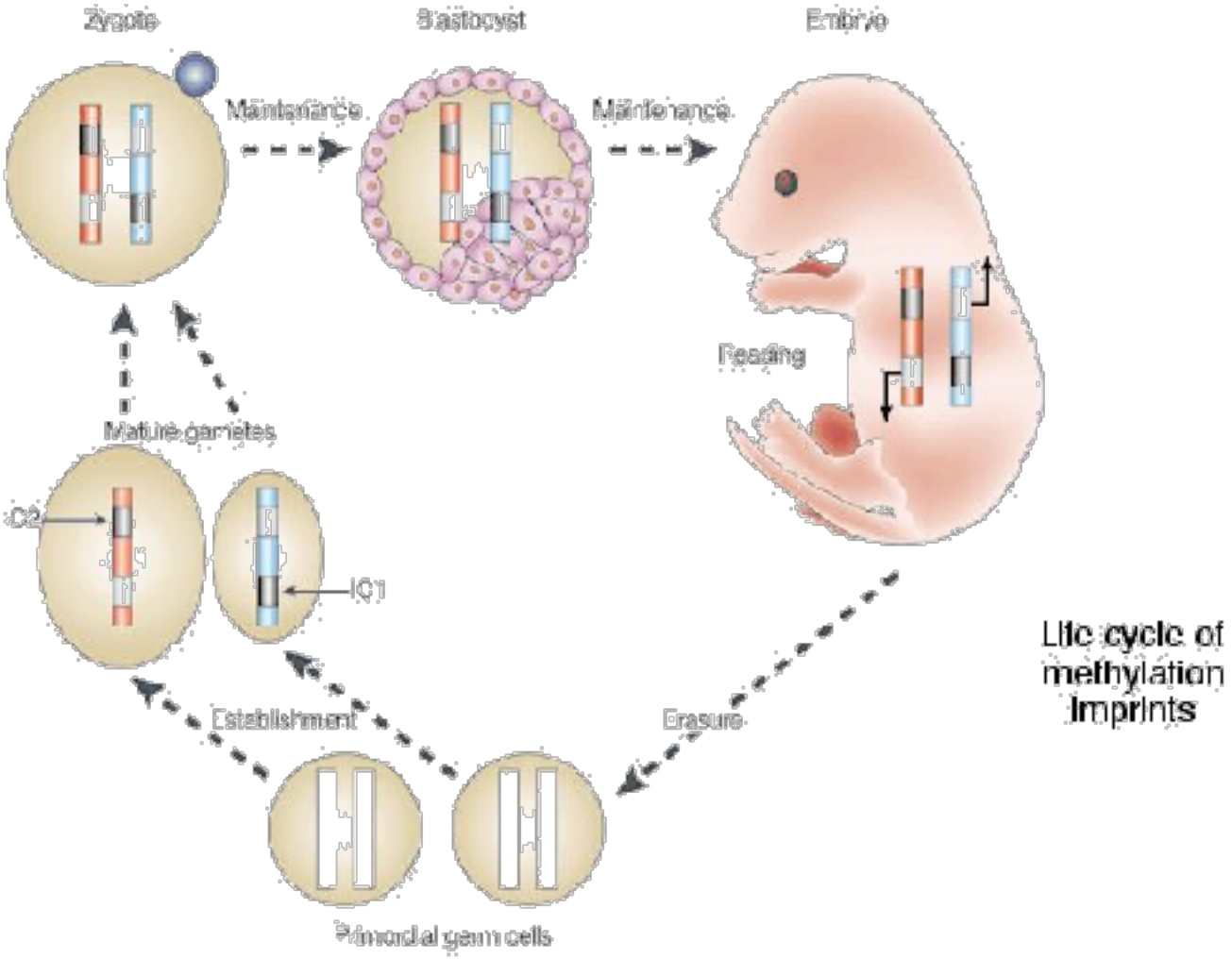


PMID: 9809658

Genomic imprinting:

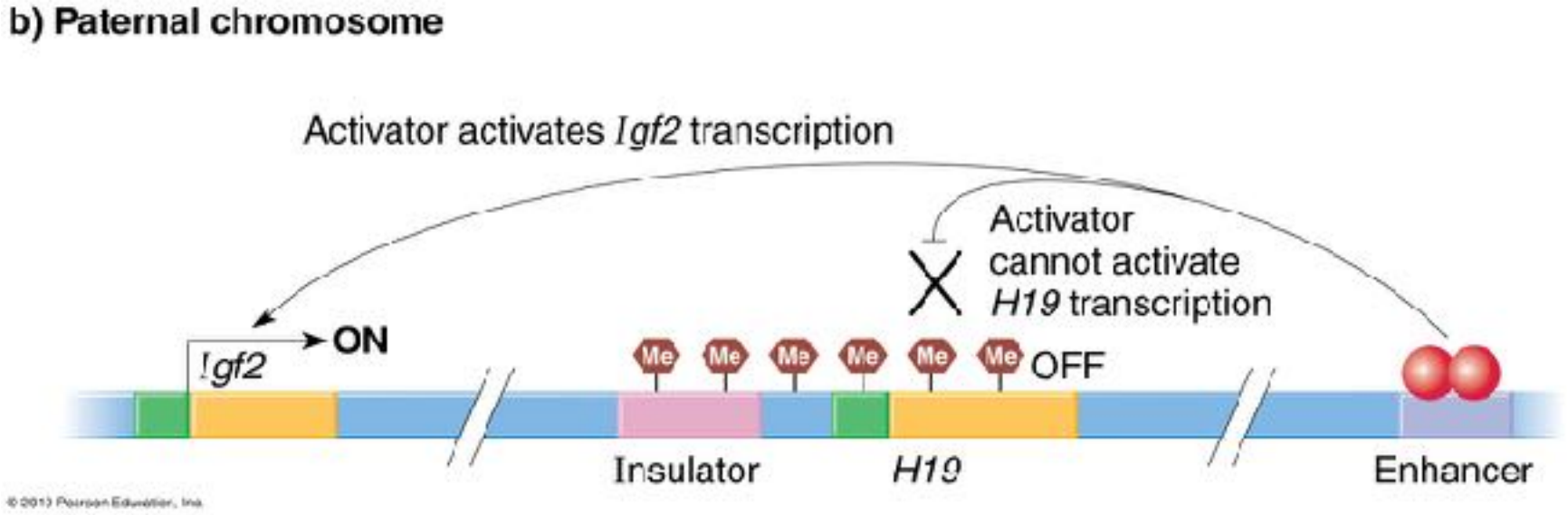
inactivation of maternal or paternal genes

Some alleles are tagged by methyl C.

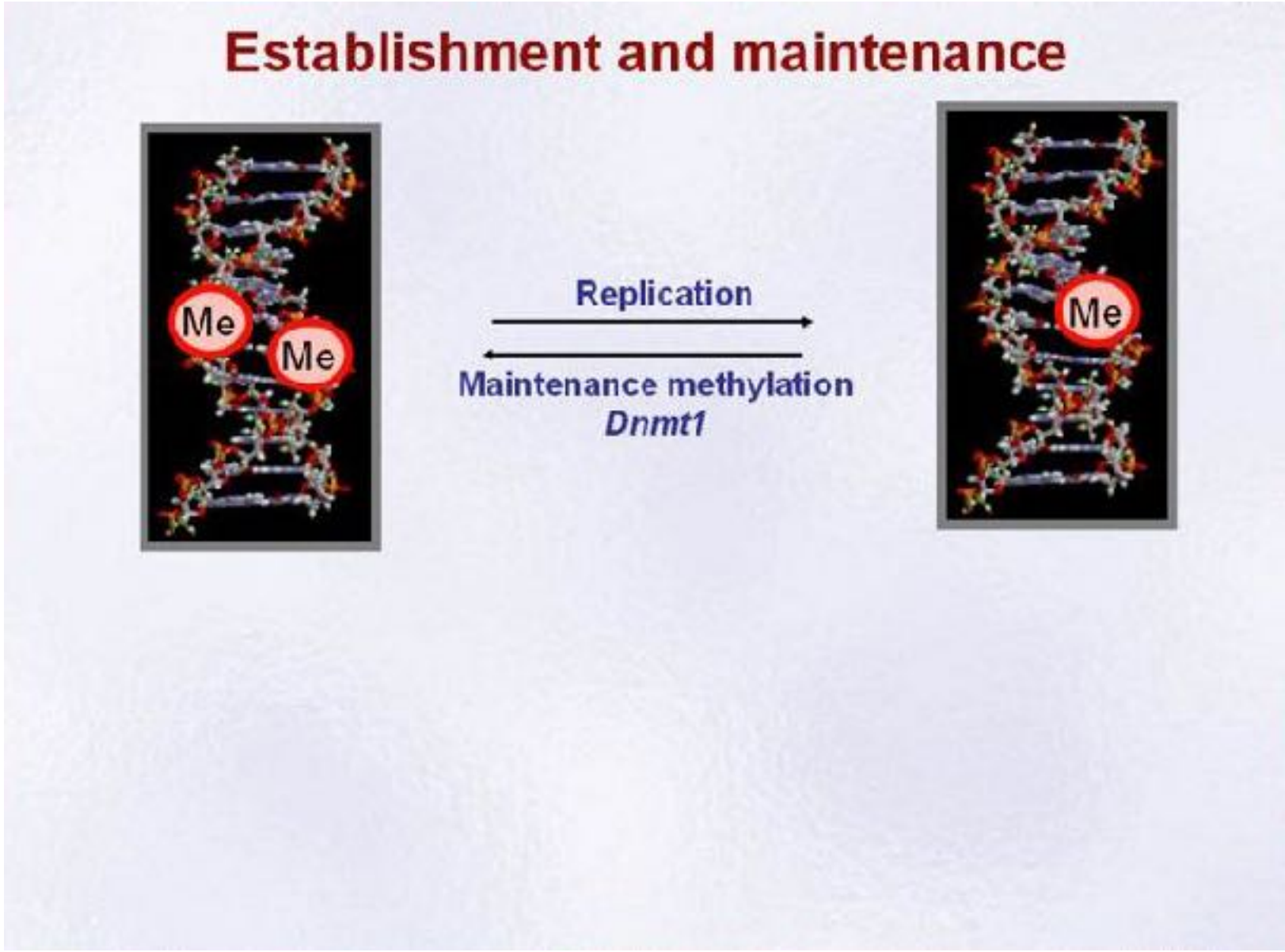


Methylation of H19 inactivates transcription

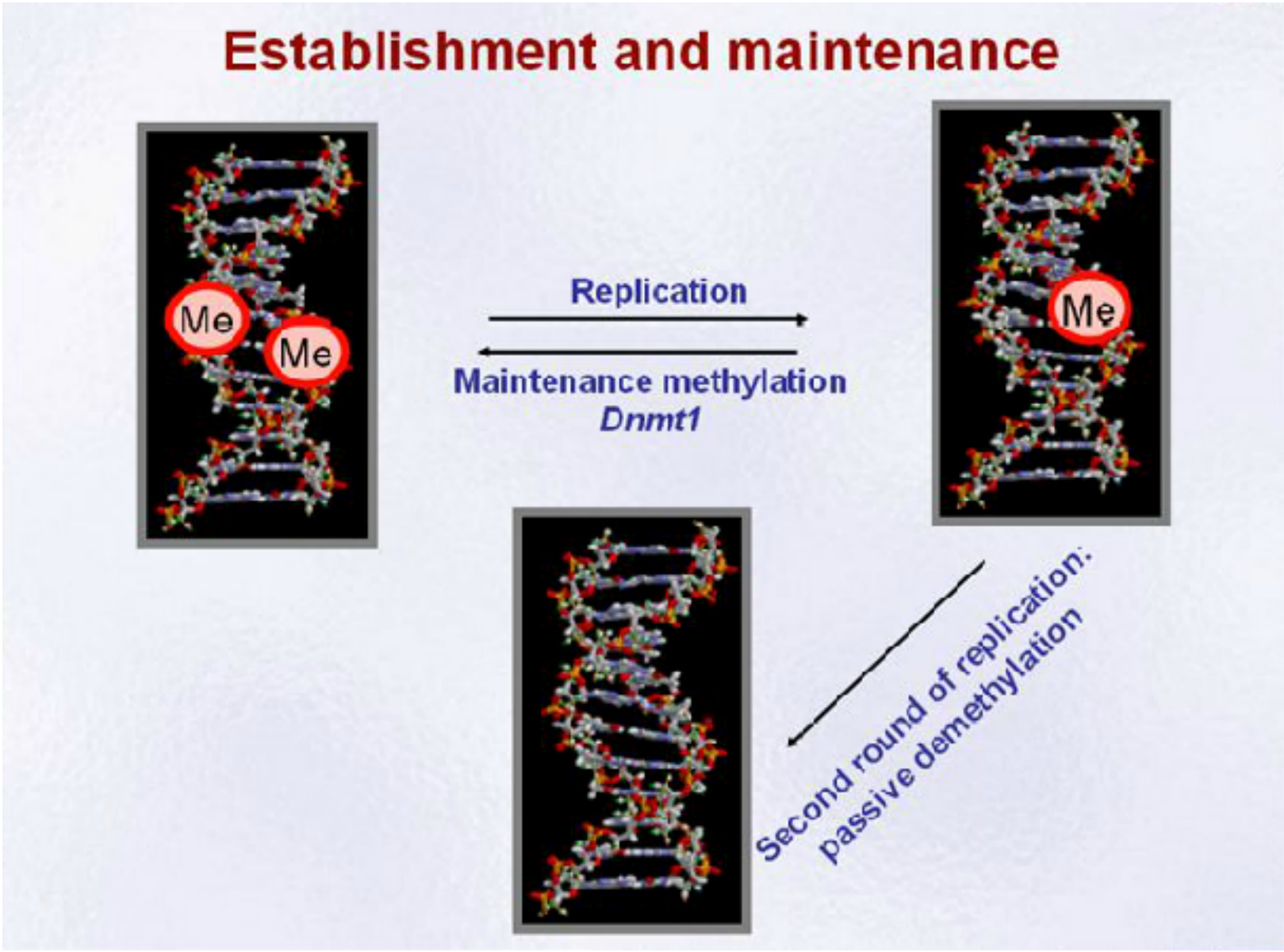
(involved in expression of insulin like growth factor)



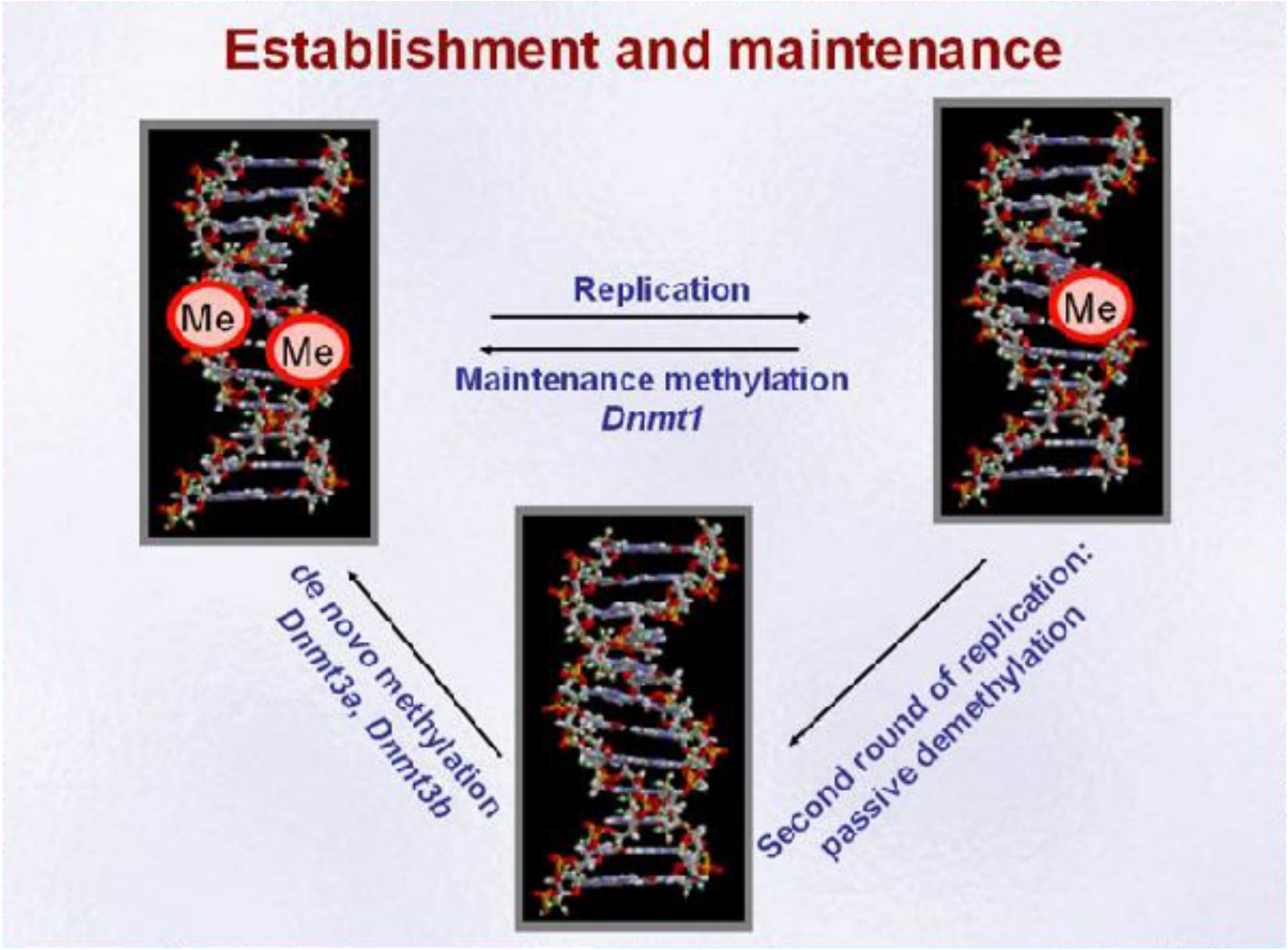
Maintenance of Cytosine Methylation



Passive Demethylation of 5-Methyl-Cytosine



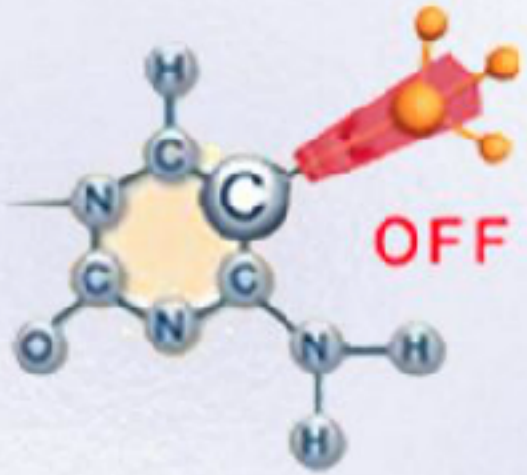
Establishment and Maintenance of Cytosine Methylation



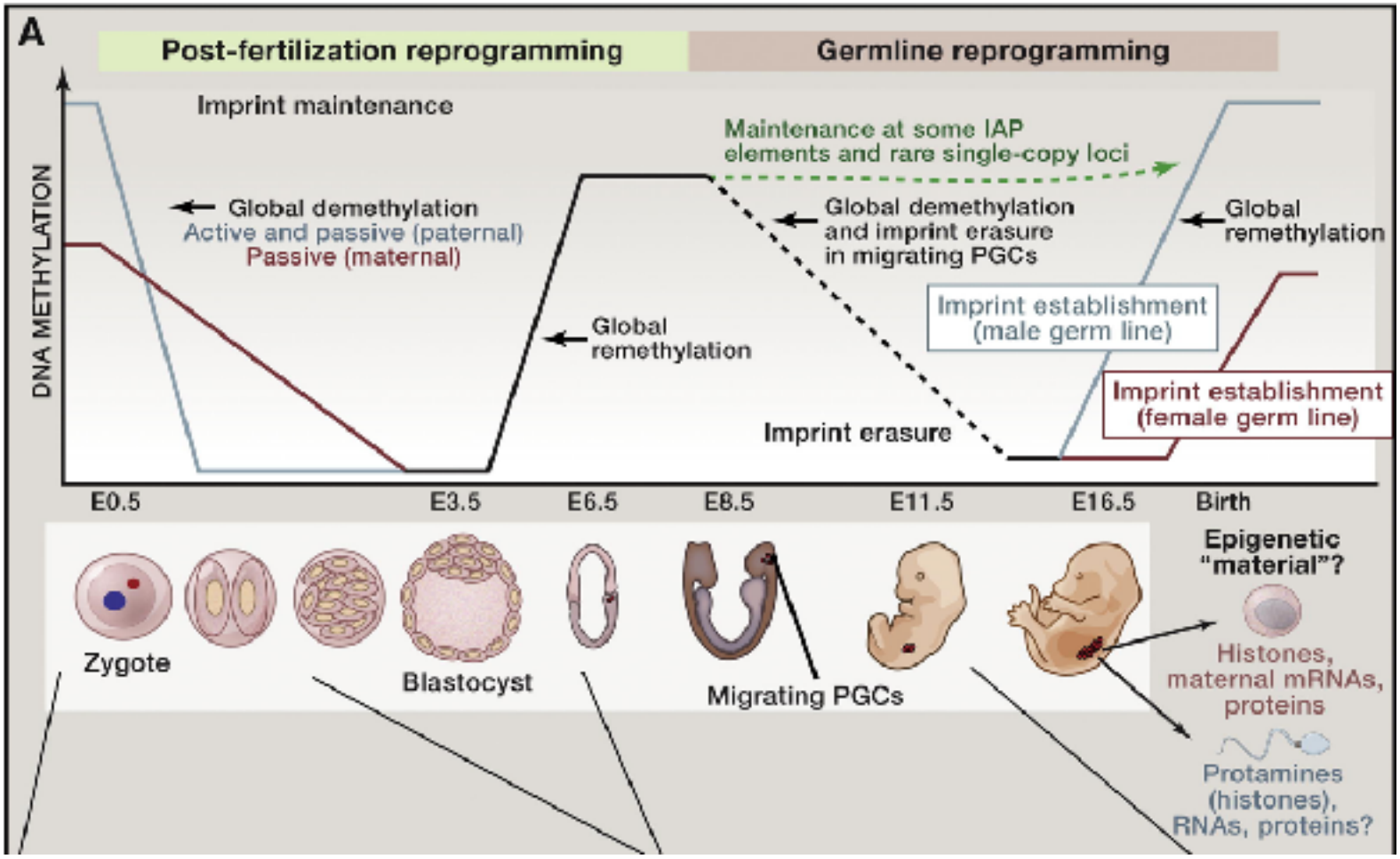
Some DNA Methyl Transferases are Essential

Cytosine methylation in mammals

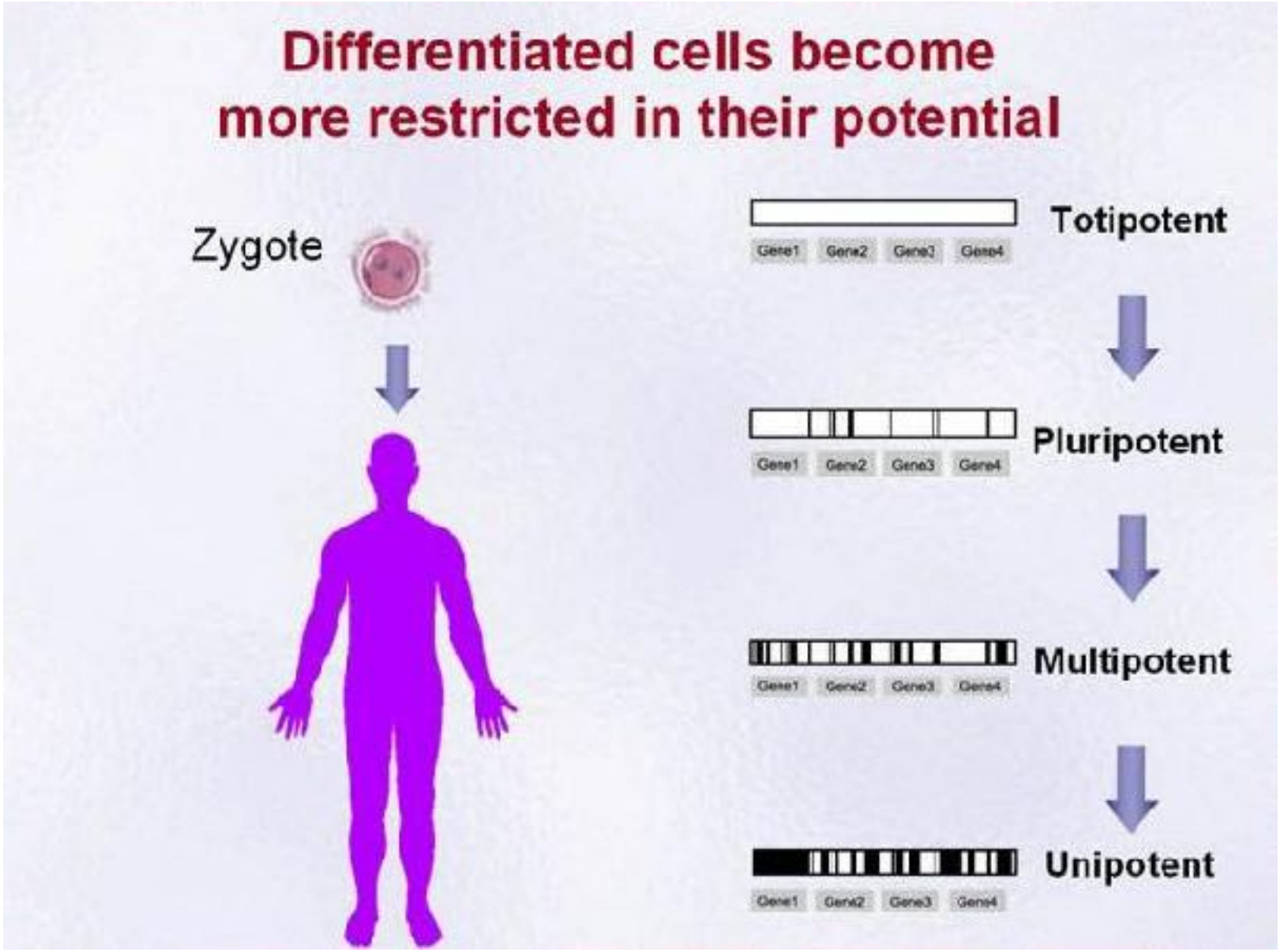
- Gene expression
- Chromosomal stability
- Cell differentiation
- Imprinting
- X-Inactivation
- Carcinogenesis
- Aging



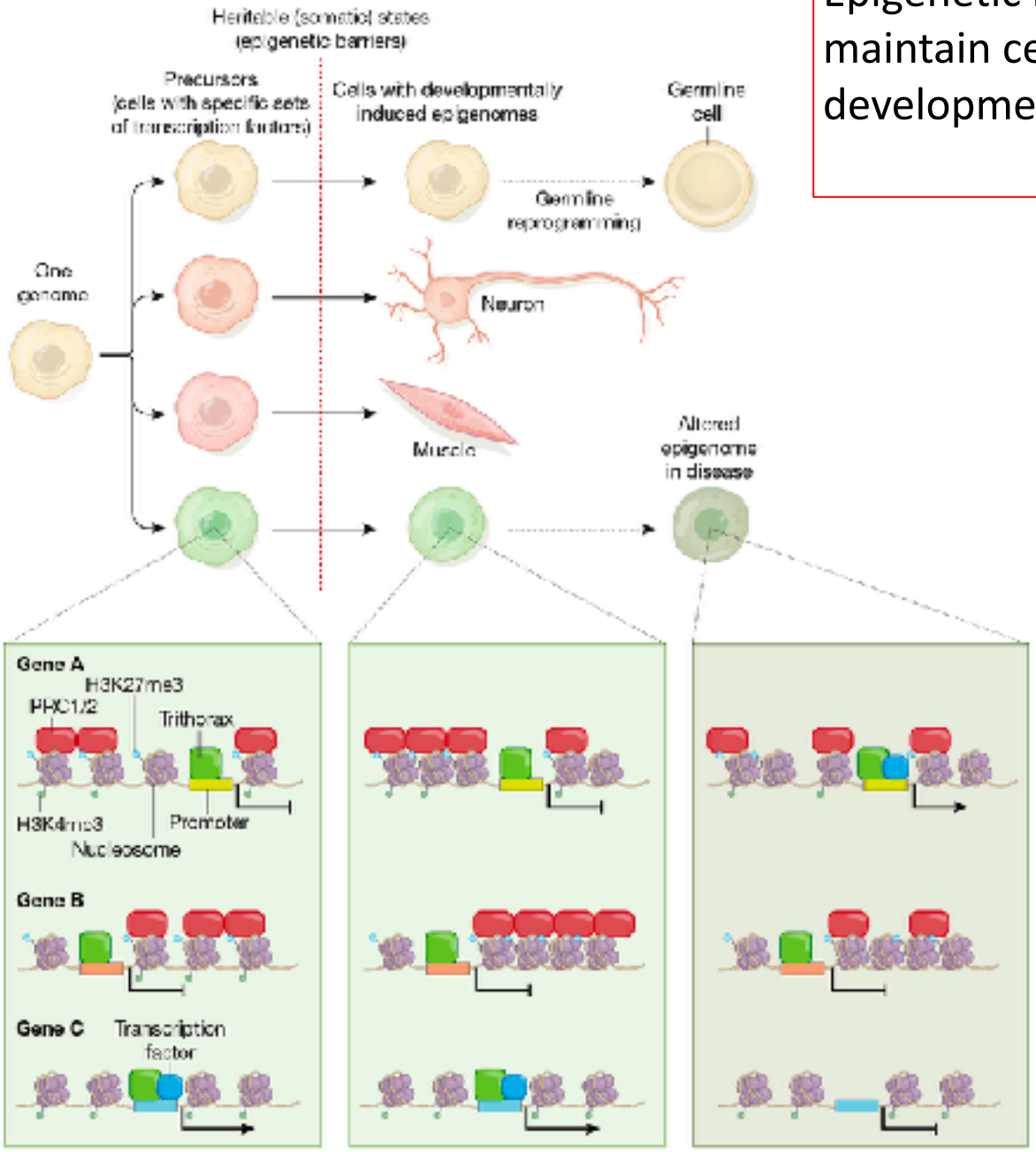
Mechanisms of germline reprogramming



Methylated DNA from Zygote to Adult

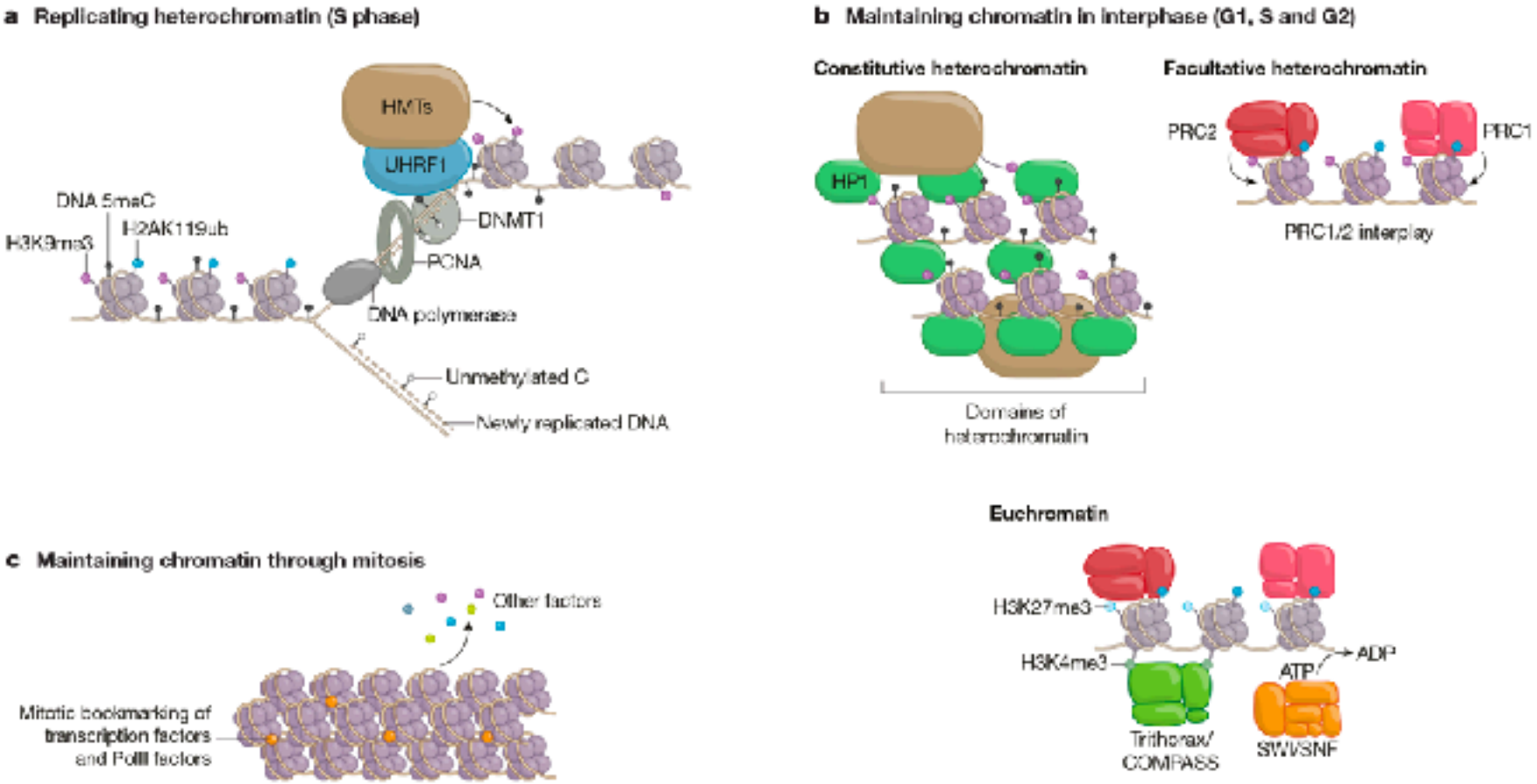


Epigenetic mechanisms that maintain cell identities during development and throughout life

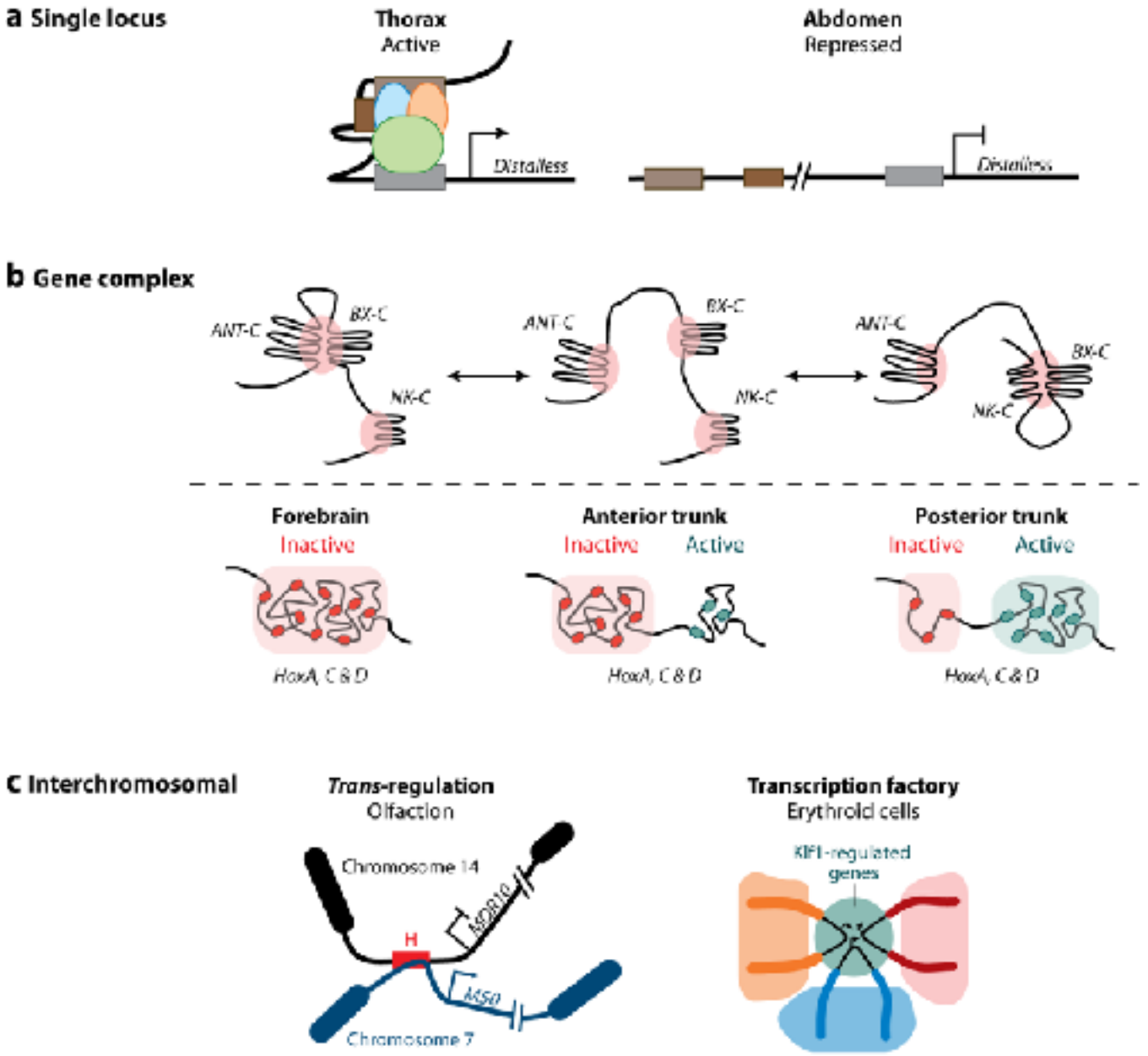


Cavalli & Heard, Nature, 2019

Maintaining chromatin states through the cell cycle.



Cavalli & Heard, Nature, 2019



Long-range interactions in transcriptional regulation

Major carriers of epigenetic information

- Heterochromatin components: megabasesized repetitive DNA domains coated in a specific histone H3K9 trimethylation mark. Heterochromatin components can both write and read the H3K9me3 mark and compact their target chromatin.
- Polycomb proteins: Polycomb (PcG) and Trithorax:two antagonistic groups that maintain the memory of spatial patterns of expression of genes throughout development, (maintenance of developmentally or environmentally programmed expression states)
- Noncoding RNAs: Many different classes and function. They are also involved in the regulation of chromatin architecture.
- DNA methylation: involve specific proteins that recognize CpG hemimethylated DNA and thereby redeposit DNA methylation on newly replicated DNA.

Non-coding RNA

- **Short interfering RNA (siRNA)**
- **Micro RNA (mir RNA)**
- **Double-stranded RNA (ds RNA)**
- **Short heterochromatic RNA (sh RNA)**
- **transcripts from repeated sequences (ALU, LTR)**

Non-coding RNA

How many different microRNA genes are there?

C. elegans ~ 40,000 pairs of hairpins
35,697 had the minimal conservation
to receive *MiRscan* score.
± 15 000 miRNAs were identified

D. melanogaster ~ 436,000 pairs of hairpins
118,000 structure with high score
±8000 miRNA genes identified in 2019

Human ~ 800,000 pairs of hairpins
15,000 have a minimal conservation
to receive *MIRscan* score (non-coding regions)
±2300 miRNA genes identified in 2019

Non-coding RNA

How is MicroRNA Activity Regulated?

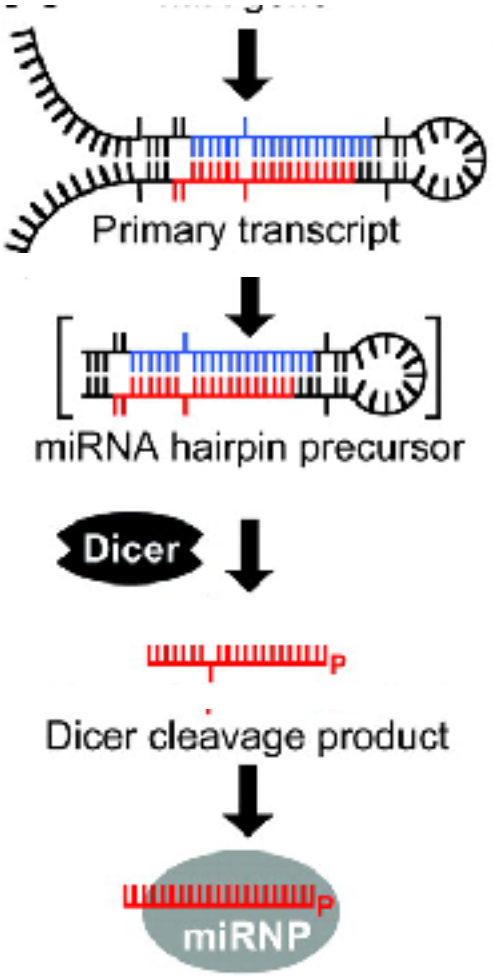
- miRNAs have diverse temporal and quantitative expression profile

- miRNA genes are known to reside in local genomic clusters with possible operon-like organization

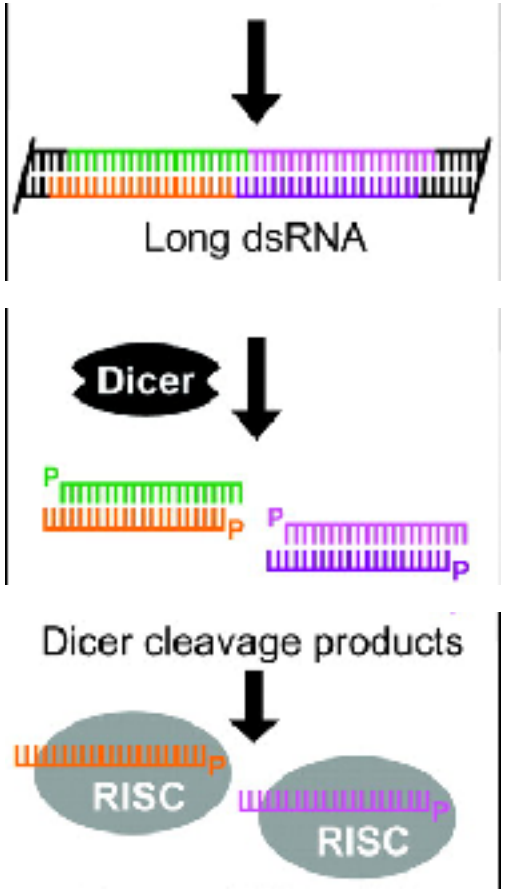
Non-coding RNA

MicroRNAs and Short Interfering RNAs Might Use the Same RNA Processing Complex.

MicroRNA gene



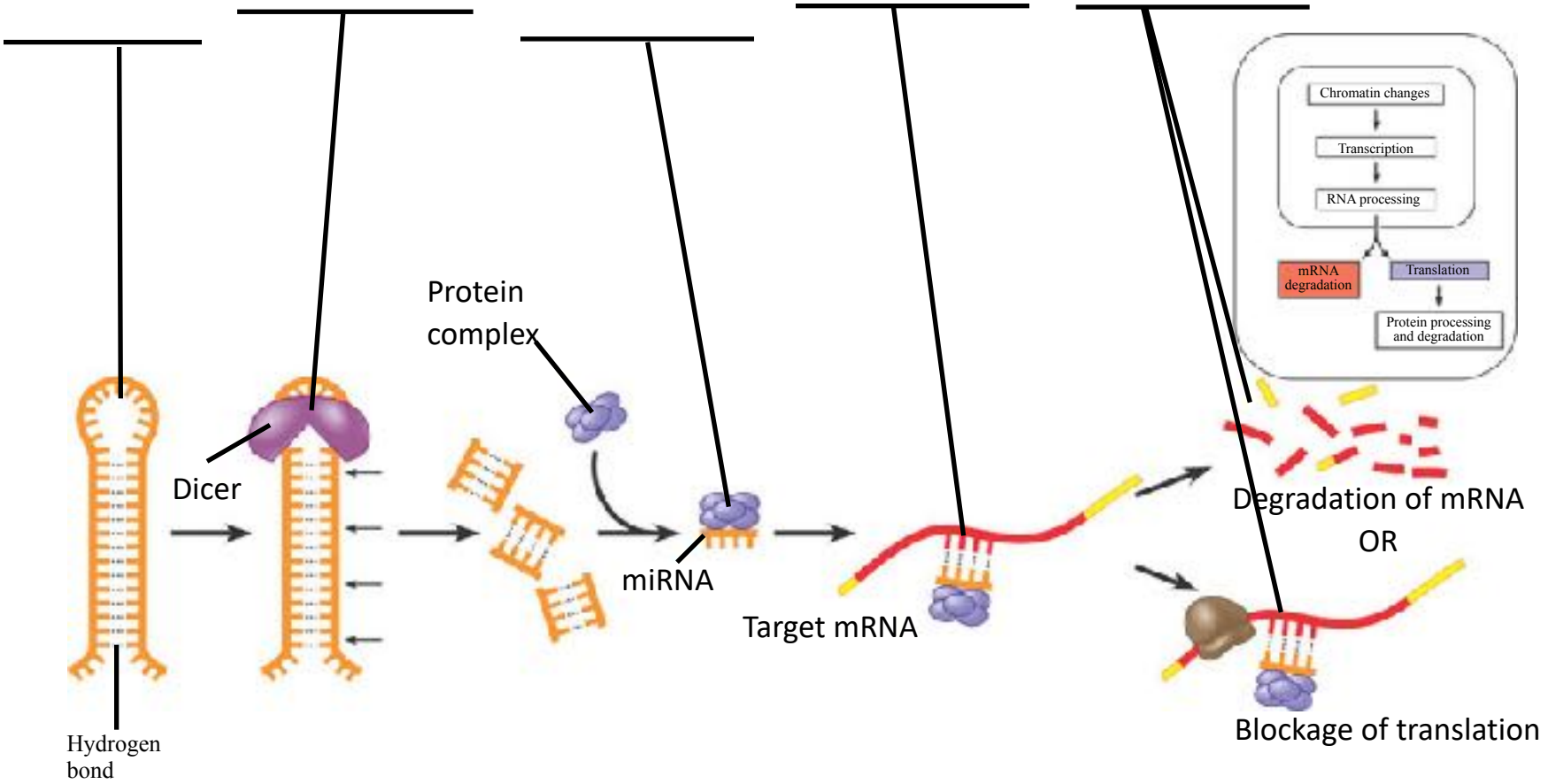
Transposons , transgenes, viruses, heterochromatic DNA…



Non-coding RNA

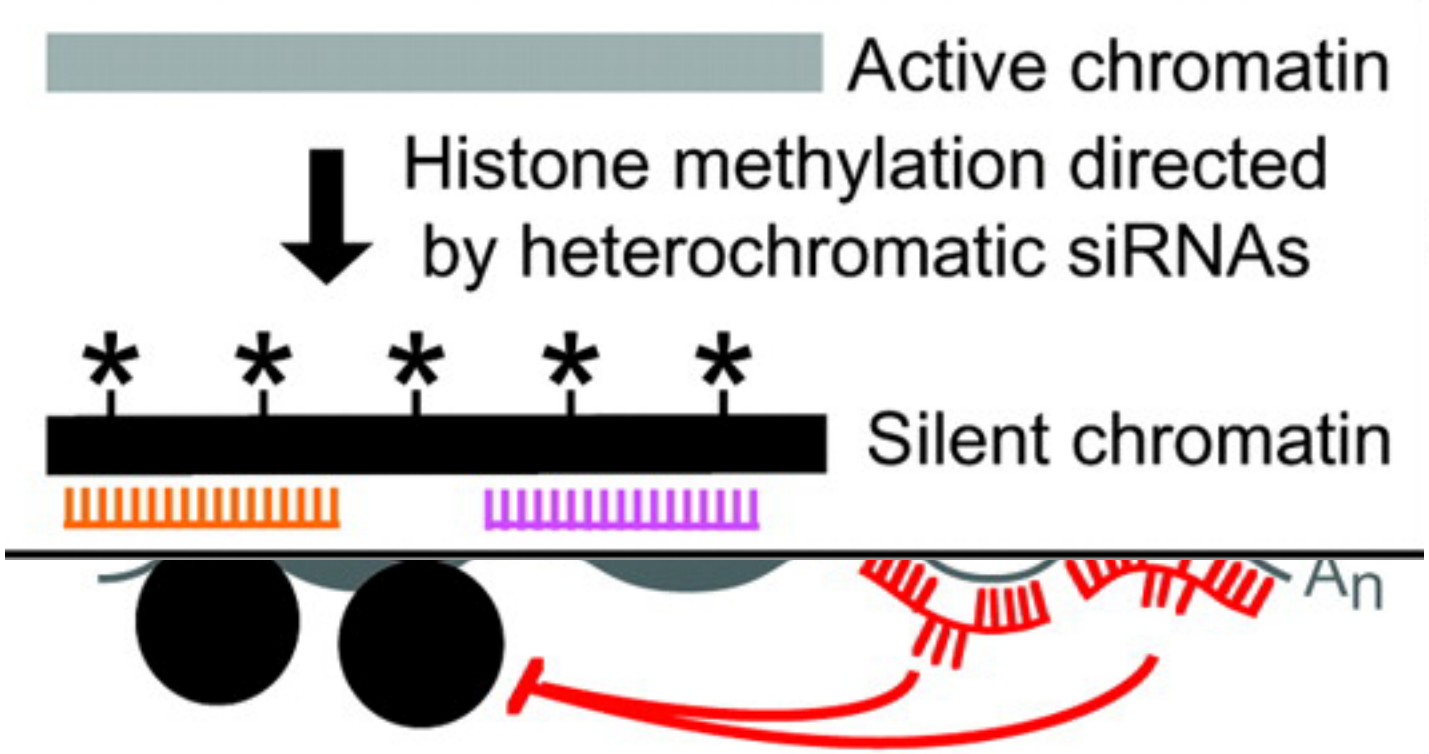
What Do MicroRNA Do ?

1. The micro-RNA (miRNA) precursor folds back on itself
2. Dicer cuts dsRNA into short segments
3. One strand of miRNA associates with protein.
4. The bound miRNA can base-pair with any complementary mRNA
5. Prevents gene expresion

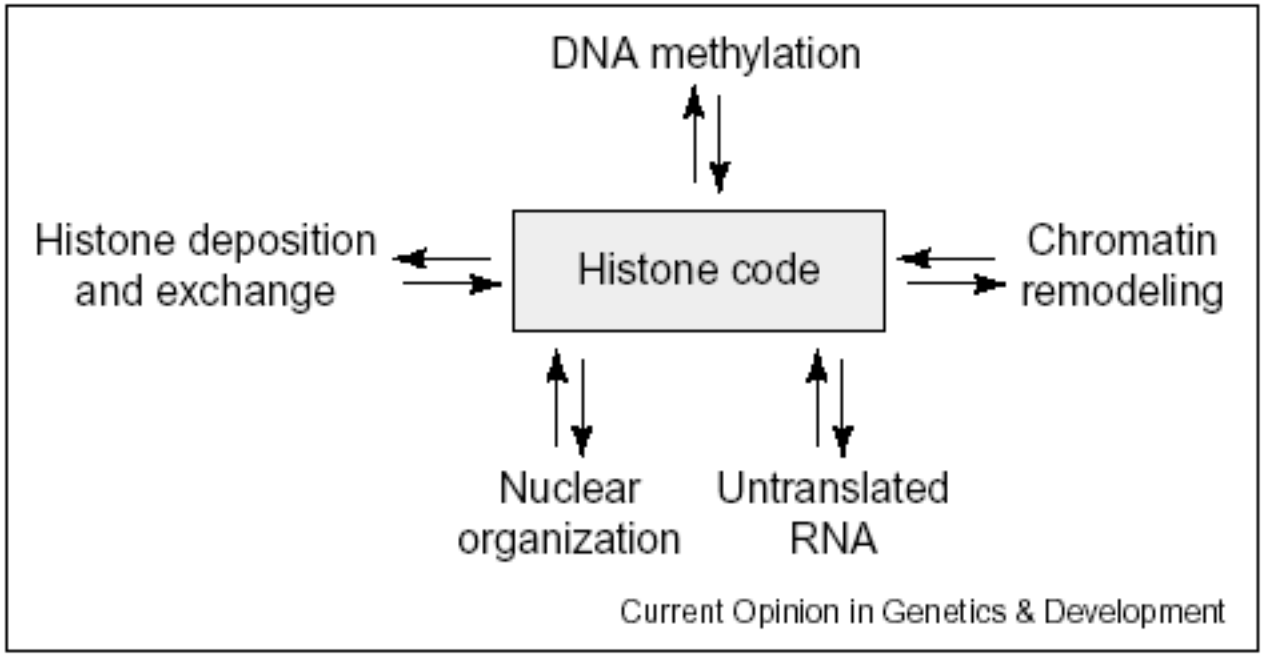


Non-coding RNA

What Do MicroRNA Do ?



Interplay between different epigenetic strategies



Margueron ,Current Opinion in Genetics & Development (2005)