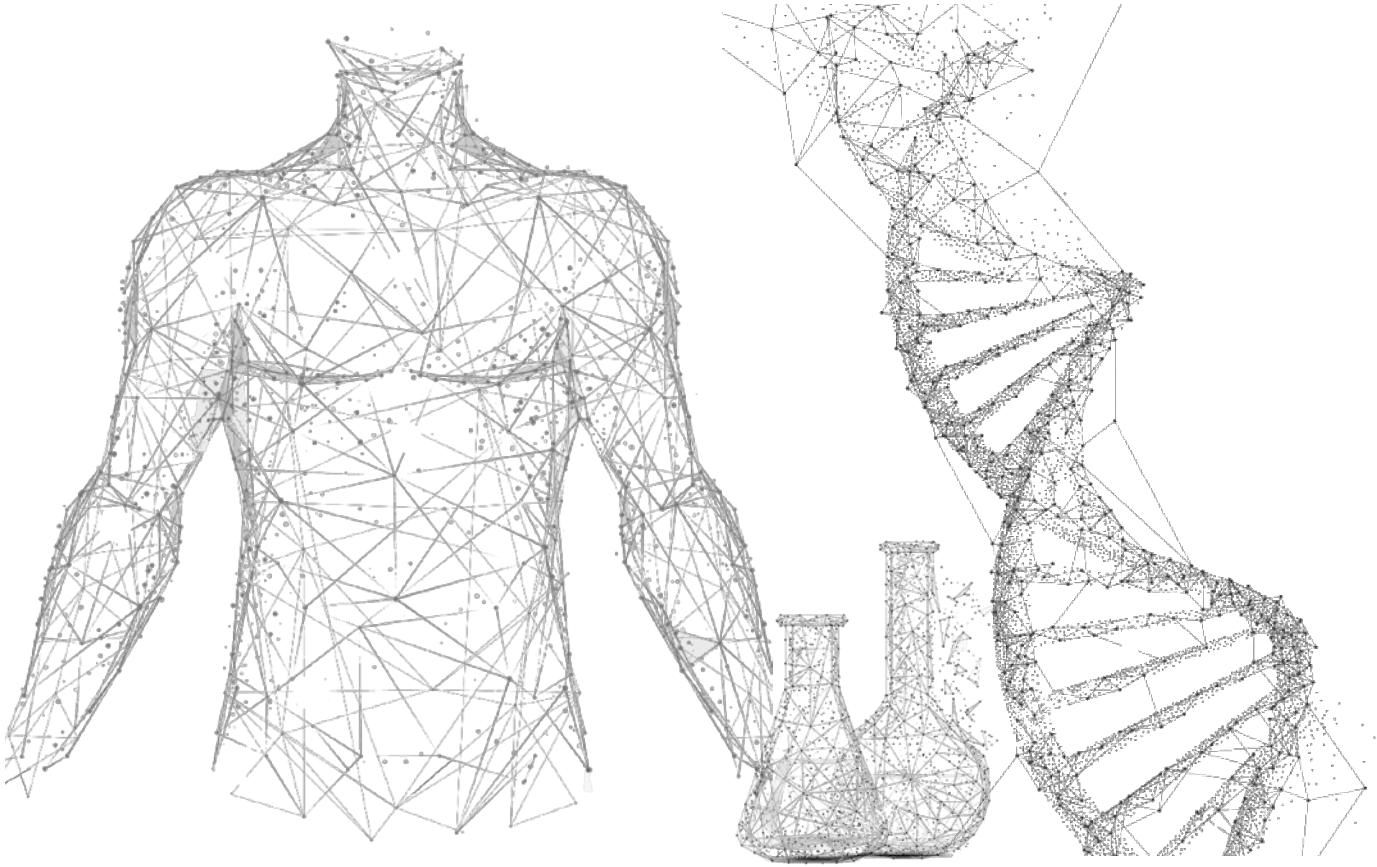


Genome Condensation and Epigenetics

San Luis Potosi State University (UASLP) Mexico
Molecular Biology Course, Faculty of Medicine graduate program

Dr. Christian A. García-Sepúlveda

Viral & Human Genomics BSL-3 Laboratory

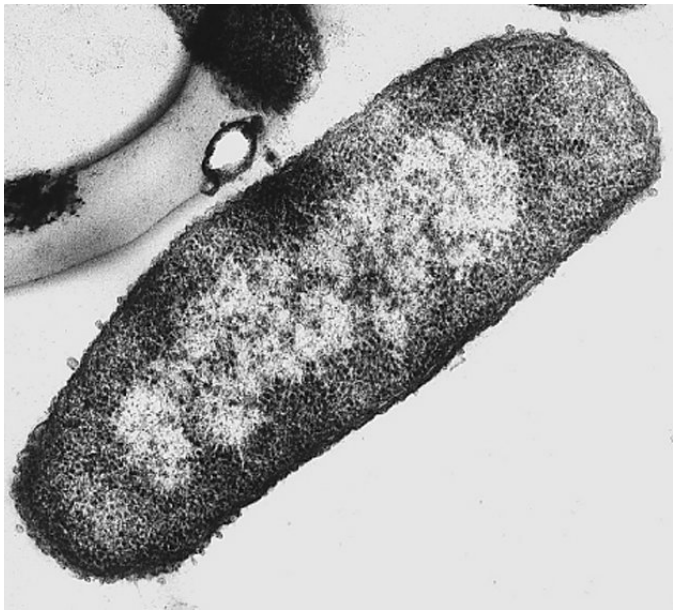
Last updated January 22, 2025 v3

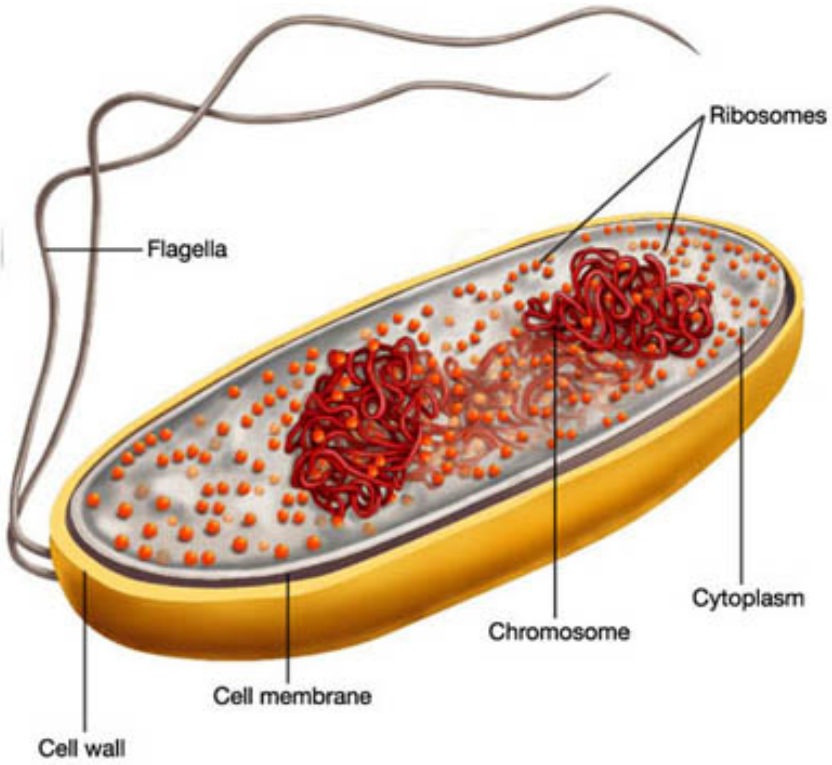
Prokaryote genophore

Although bacteria do not have morphological structures similar to ours, their DNA is organized in a circular chromosome.

It usually forms two or three diffuse agglomerates, occupying one third of the cell volume (nucleoid).

Genophore = Bacterial chromosome without chromatin





Prokaryote genoma condensation

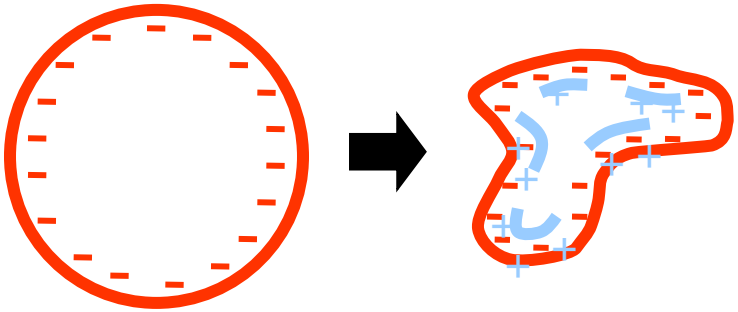
Circular bacterial chromosomes are compacted with polyamines (such as spermine).

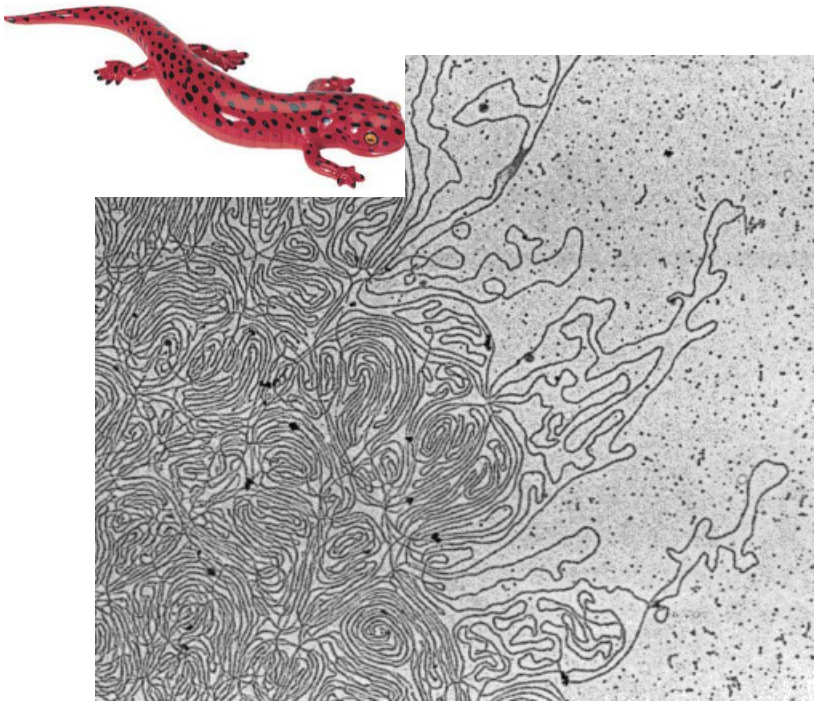
Bind to the negative charges of DNA to neutralize its charge and prevent repulsion.

Spermine is synthesized from its spermidine precursor by adding an aminopropyl group.

Spermine interacts with negatively charged molecules like DNA, RNA, and phospholipids, helping to stabilize their structure and function.

permine acts as a scavenger of reactive oxygen species, providing a protective role against oxidative stress.





The human genome

23 chromosome pairs

24 distinct human chromosomes:
- 22 autosomal
- sex-determining X and Y.

Chromosomes numbered in order of size (karyotype/gram)

63,494 total genes

19,969 protein-coding genes (40% of genome).

Haploid genome size of 3,234,000,000 bp

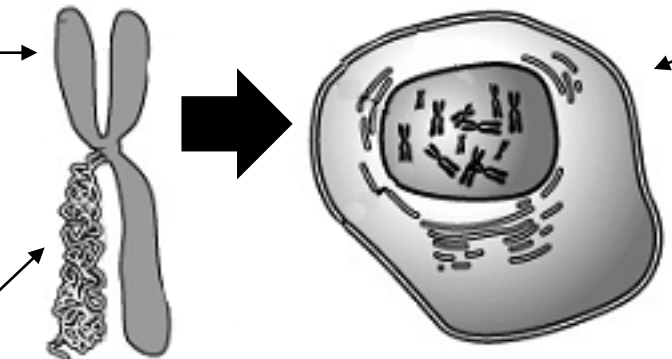
Diploid 6,469,000,000 bp

Between 1.1 and 2 mts DNA (x2)

cell type	average volume (μm^3)
sperm cell	30
red blood cell	100
lymphocyte	130
neutrophil	300
beta cell	1,000
enterocyte	1,400
fibroblast	2,000
HeLa, cervix	3,000
hair cell (ear)	4,000
osteoblast	4,000
alveolar macrophage	5,000
cardiomyocyte	15,000
megakaryocyte	30,000
fat cell	600,000
oocyte	4,000,000

Bianconi E, et al., An estimation of the number of cells in the human body. Ann Hum Biol. 2013 Nov-Dec40(6):463-71.

Eukaryote genome compaction



Chromosomes
(300 nm)
47 Mbp #21
247 MBP #1

Human cell
(4 to 120 µm)
6,469,000,000 bp

Chromatids
(765 nm)
500,000 bp

10 nm fiber
"pearls-on-a-string"
(6x11 nm)
200 bp

Domains or loops
(300 nm)
75,000 bp

DNA duplex
(2 x 3.4 nm)
10 bp

30 nm chromatin fiber
(30 nm)
1200 bp

Compaction hierarchies

Nucleosome

Histones are positively charged proteins and act as bacterial polyamines.

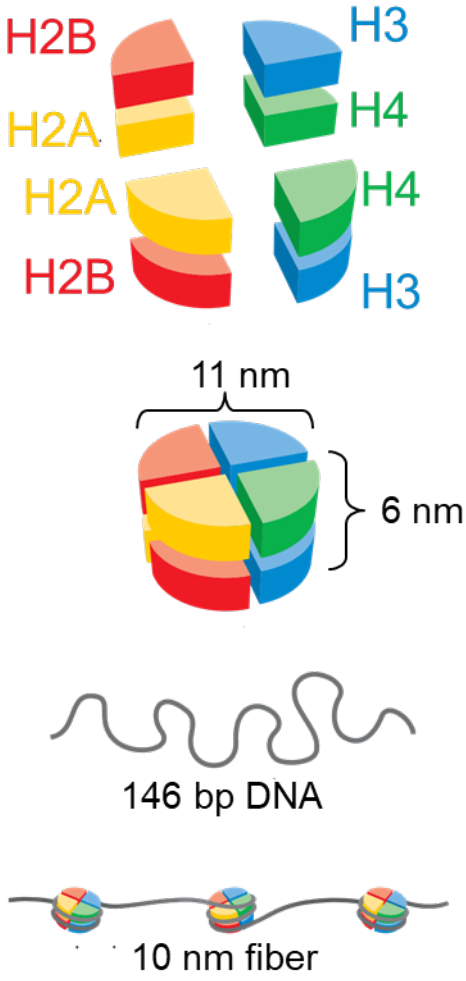
Nucleosome cores are composed of eight histones forming a 6x 11 nm disk.

Approximately 146 bp of DNA are wound around the core.

Core nucleosomes are separated from each other by 54 bp (varies from 8 to 114 bp) of spacer DNA wrapped around H1 histone.

Total amount of 200 bp of DNA.

Non-condensed nucleosomes without H1 resemble "beads on a string of DNA"



Nucleosome

Nucleosome cores are composed of eight histones forming a 6x 11 nm disk.

Peptides between 10 and 25 kDa.

Positively charged (Arg & Lys) to interact with negative DNA.

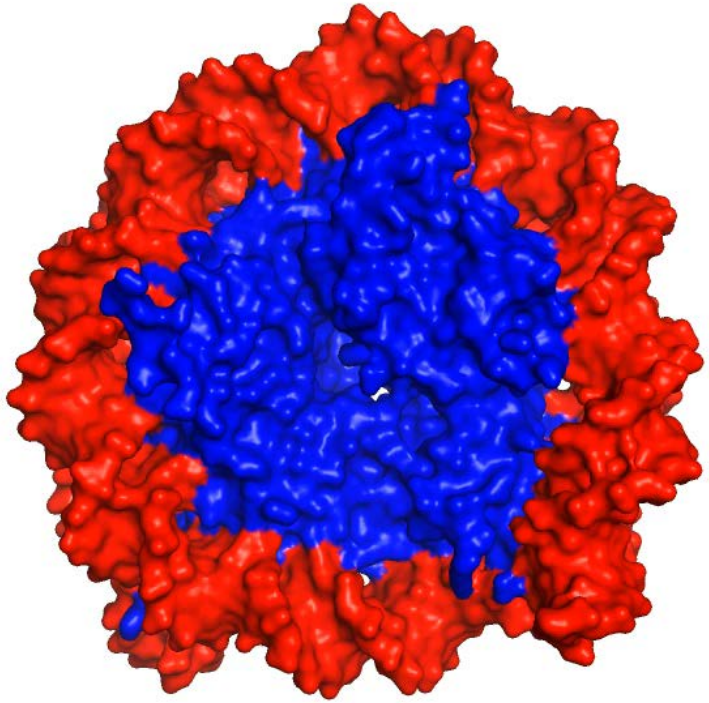
Histones interfere with the binding of other proteins to DNA.

The winding of DNA around histones is a dynamic and regulated process.

1.67 turns of DNA per nucleosome.

> 120 direct protein-DNA interactions.

Non-sequence-specific DNA-binding.



H3 dark blue
H4 light blue
H2A Orange
H2B Tan

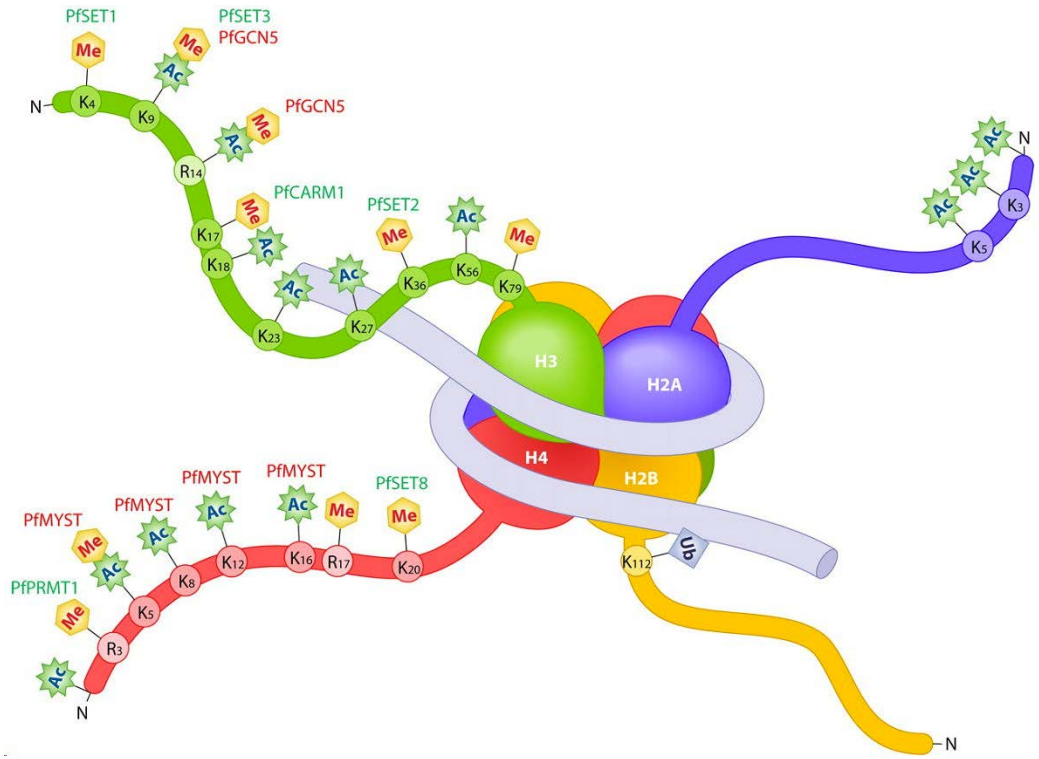
Histone tails and epigenetic information

Most post-translational histone modifications are concentrated in the tails.

Modifications affect histone-DNA[and histone-histone interactions in the core.

Acetylation or phosphorylation lower charge of the histone core and "loosen" core-DNA interactions.

Lysine (Lys, K):	Acetylation Methylation Ubiquitination
Arginine (Arg, R):	Methylation
Serine (Ser, S):	Phosphorylation



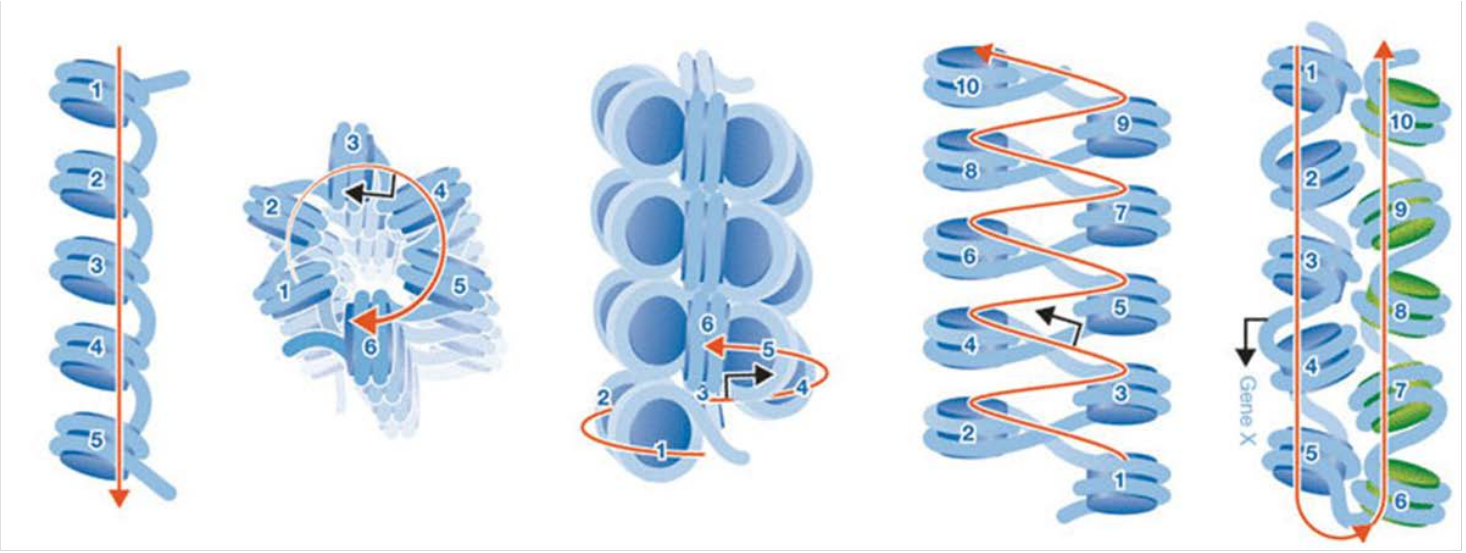
Higher-order nucleosome structure (30 nm chromatin fiber)

Not visible in crystal structures due to their high flexibility.

N-terminal tails of H3 & H2B pass through a channel formed by two DNA strands.

N-terminal tail of H4, interacts with H2A-H2B of nearby nucleosome, and is responsible for higher-order structure (nucleosome kernel).

H1 histone seals DNA around nucleosome and stabilizes formation of kernel.



10 nm fiber "pearls-on-a- string" | Nucleosomal kernel or solenoid | 30 nm fiber | Zig-zag model of the 30 nm fibre | Boustrophedon model of the 30 nm fiber

Topologically Associating Domains or loops (300 nm loops)

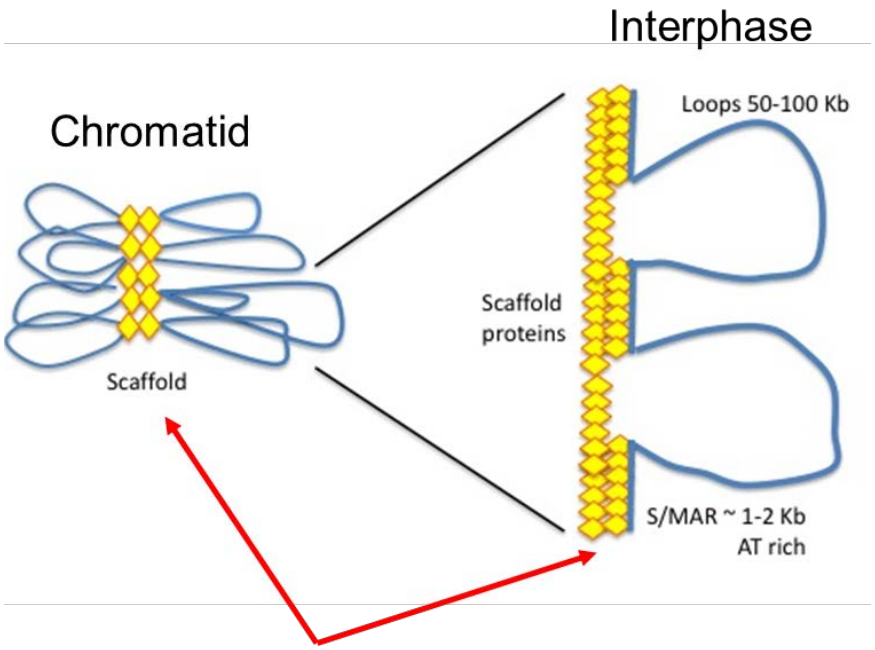
Topologically Associating Domains (TADs) also called Chromatin Loops

Domains vary in scale and complexity, ranging from local chromatin loops to large chromosome territories.

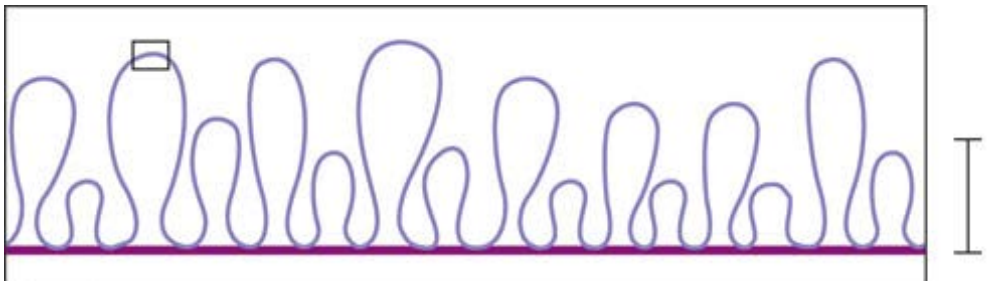
Regions of the genome where DNA sequences interact more frequently with each other.

Help organize the genome into functional units, enabling interactions between genes and regulatory elements like enhancers and promoters.

TAD boundaries are often defined by proteins such as CTCF and cohesin.



Each TAD anchored at specific sites on a **nuclear scaffold** or proteinaceous axis involving scaffold/matrix attachment regions (SARs/MARs).



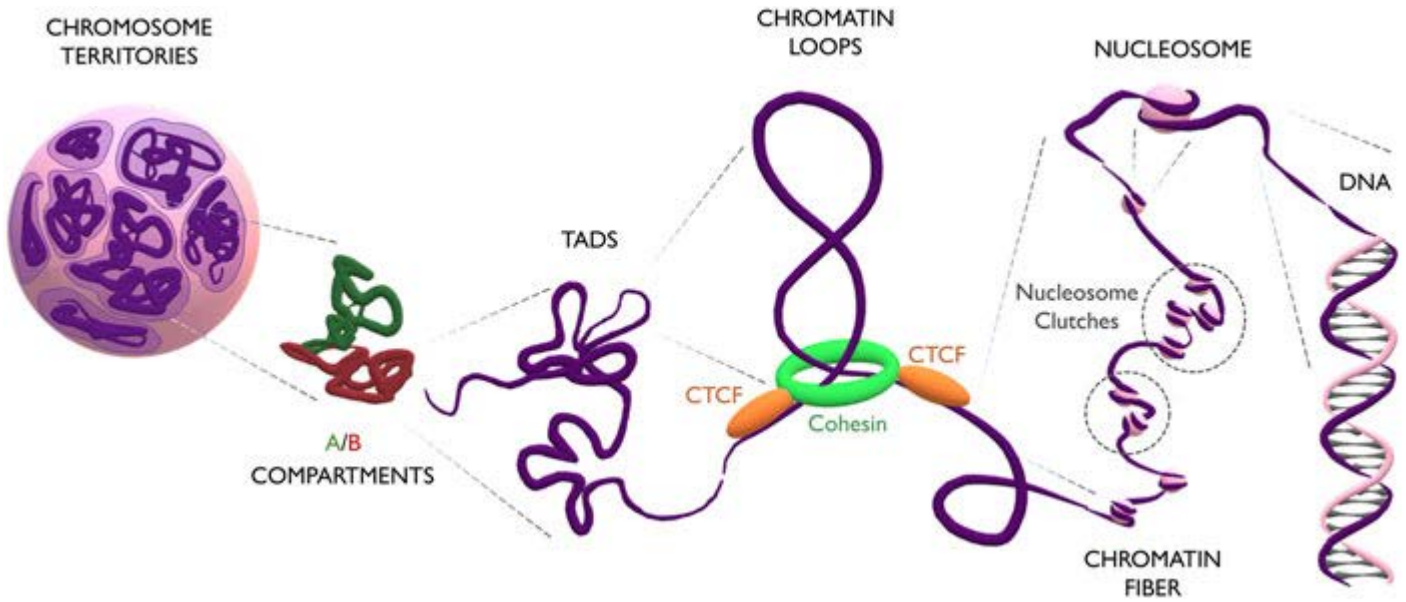
Topologically Associating Domains or loops (300 nm loops)

Topoisomerases play a critical role in maintaining DNA domain organization.

DNA domains subject to supercoiling due to replicational & transcriptional torsional strain.

Topoisomerases I and II relieve strain by cutting one or both strands of the DNA, allowing it to unwind and relax, and then resealing the breaks.

Proteins such as cohesin and CTCF, which define TAD boundaries, depend on the resolution of topological stress to maintain stable chromatin interactions.



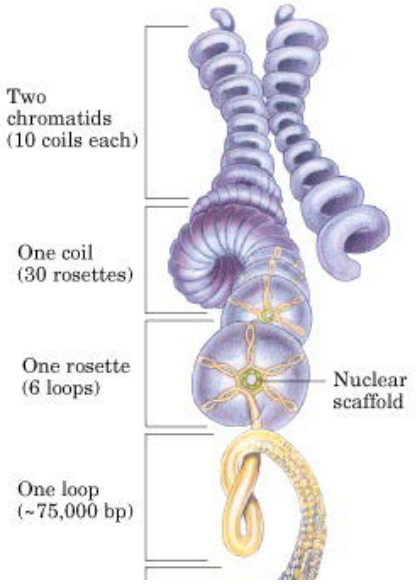
Chromatin Loop Rosettes

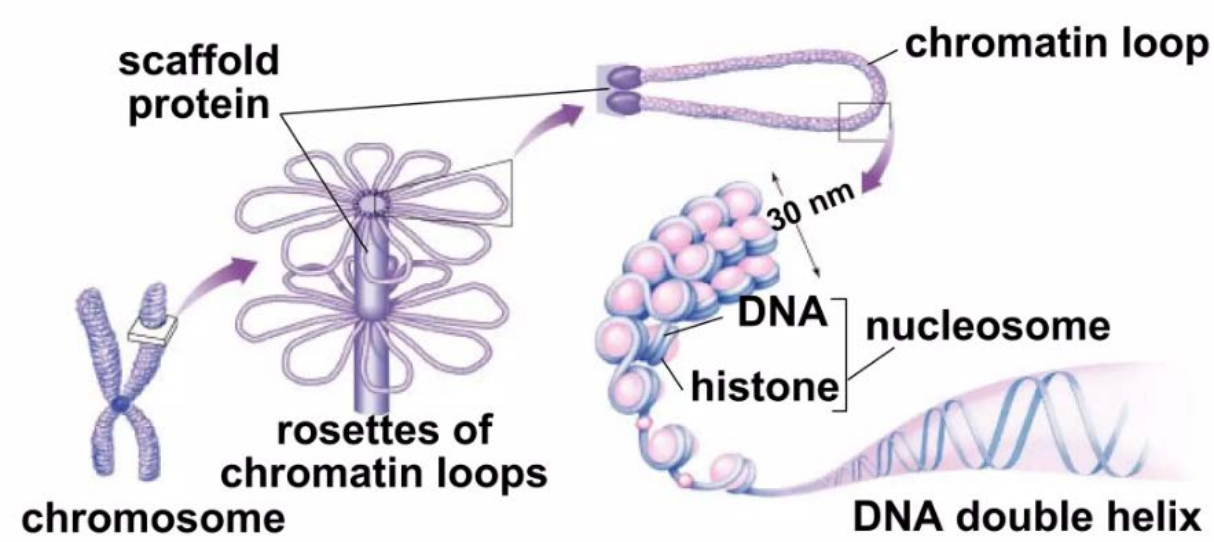
Conceptualized as clusters of **chromatin loops** arranged around a central axis, resembling the petals of a flower.

Consist of 6–12 loops, depending on the organism, cell type, and level of chromatin compaction.

Also stabilized by cohesin and CTCF.

Serves as an intermediate level of compaction between TADs and fully condensed metaphase chromosomes.





Chromosome coils

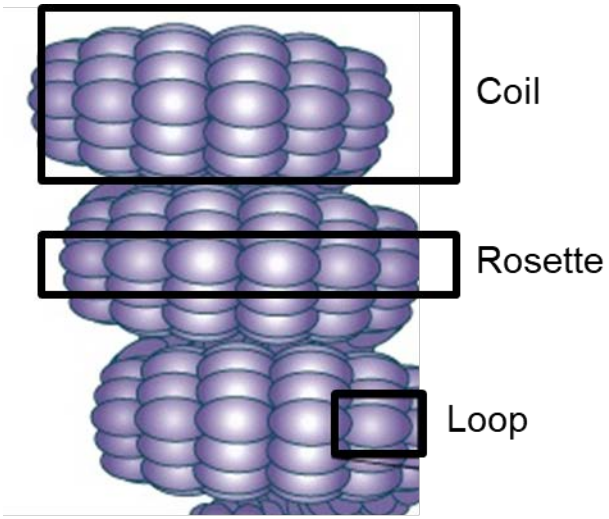
Superposition of several chromatin rosettes leads to Higher-Order Coiling.

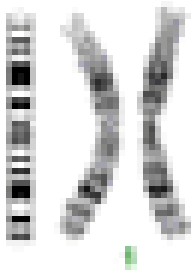
Only seen during mitosis, chromatin coils further into the highly condensed X-shaped metaphase chromosomes, ensuring efficient segregation during cell division.

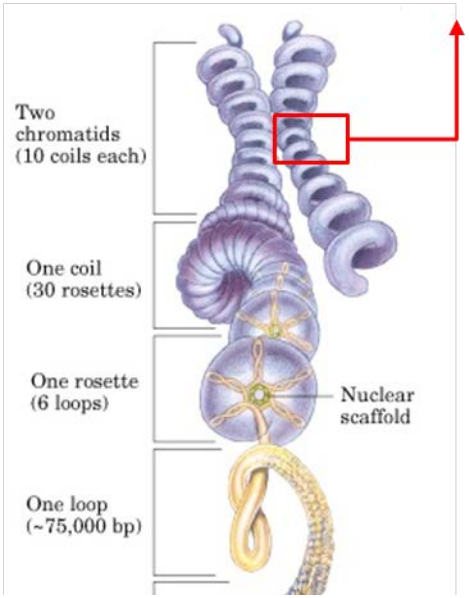
Chromosomal bands, are visible patterns observed on optical microscope during metaphase.

Giemsa (G)-banding dark bands (G-positive) correspond to gene-poor, tightly packed heterochromatin.

Light bands (G-negative) correspond to gene-rich, loosely packed euchromatin.







Chromosomal centromere and telomere

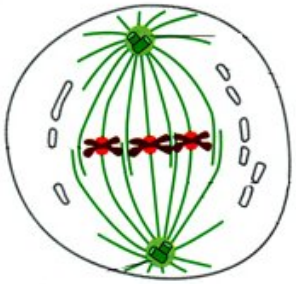
Centromere formed by heterochromatin used in attachment to the mitotic/meiotic spindle (microtubules).

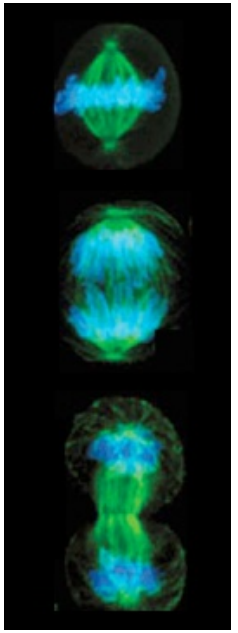
Telomere formed by repetitive heterochromatin intended to protect chromosome ends.

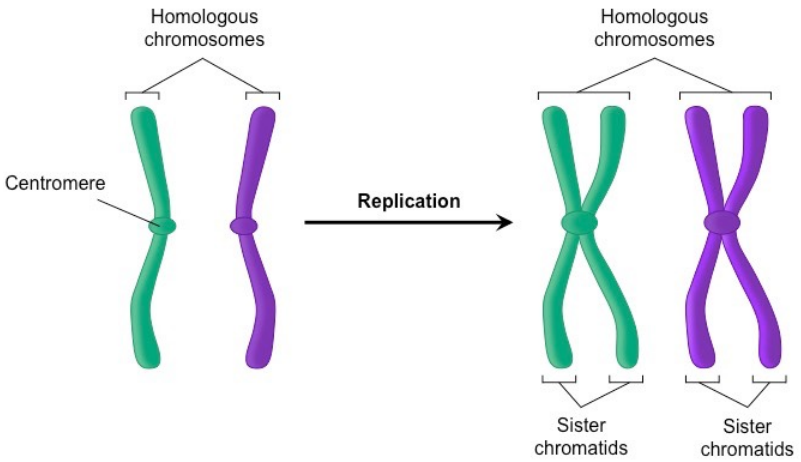
Centromeres and telomeres contain abundant amounts of simple sequence repetitive (SSR) DNA.

SSR: Tandem multiple copies of short sequences).

SSRs play structural roles in telomeres and centromeres.







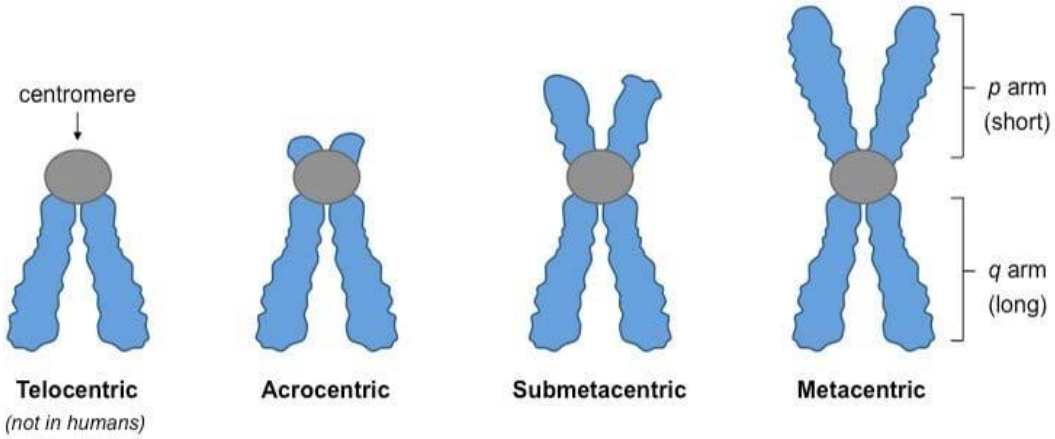
Human centromeres have a 170 bp sequence repeated between 500 and 3,000 times!

Same telomere sequence for all vertebrates!
TTAGGG

Chromosome classification

Position of centromere divides chromosome into short (p) and long (q) arms.

- Metacentric: Centromere located in the middle, two arms of equal length (Chr 1).
- Submetacentric: Centromere slightly off-center (Chr 4).
- Acrocentric: Centromere located near one end (Chr 13, 14, 15, 21, and 22).
- Telocentric: Centromere at the very end of the chromosome (Not in humans).



Condensed chromosomes

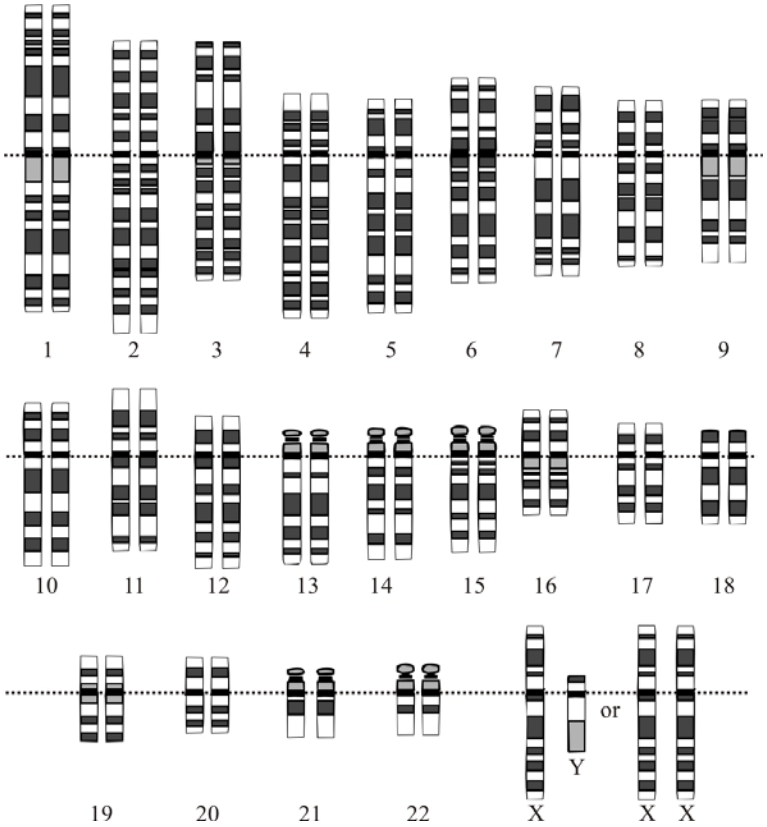
Normally 23 pairs of chromosomes (euploidy)

Poorly tolerated autosomal aneuploidies (abortions).

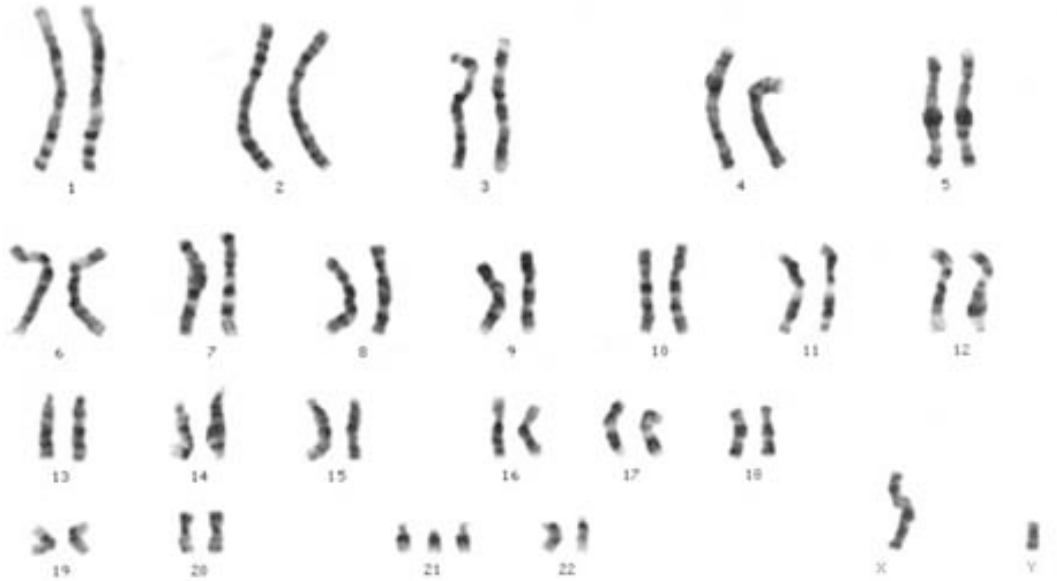
In general, aneuploidies contribute to infertility, stillbirths and deaths, congenital malformations, mental retardation, abnormal sexual development and cancers.

Sexual aneuploidies relatively well tolerated.

Approximately 20% of oocytes/spermatocytes are aneuploid

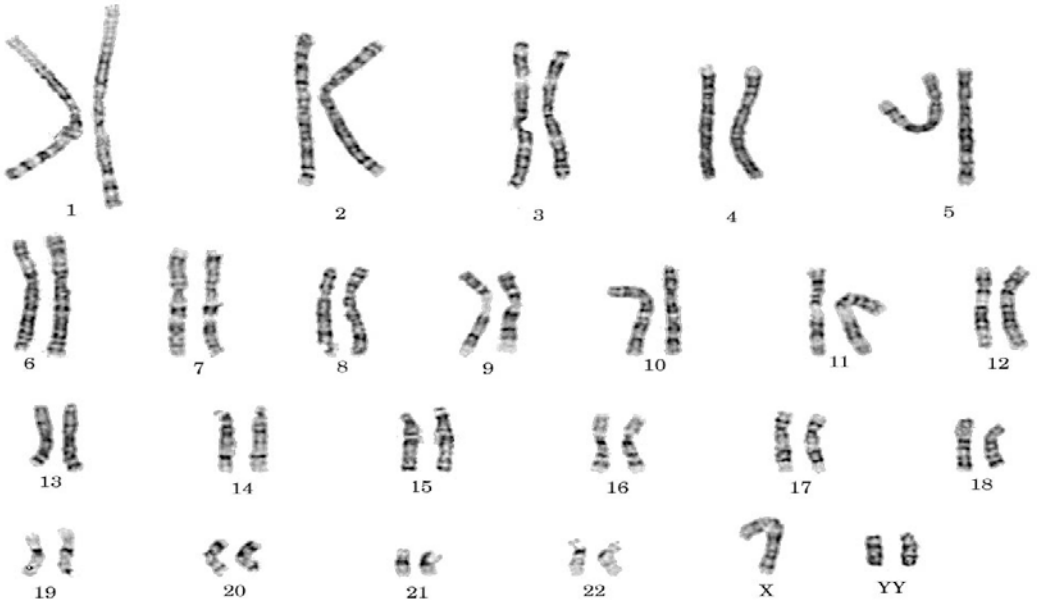


Aneuploidy (Trisomy 21, Down's Syndrome, 47XY or 47XX +21)



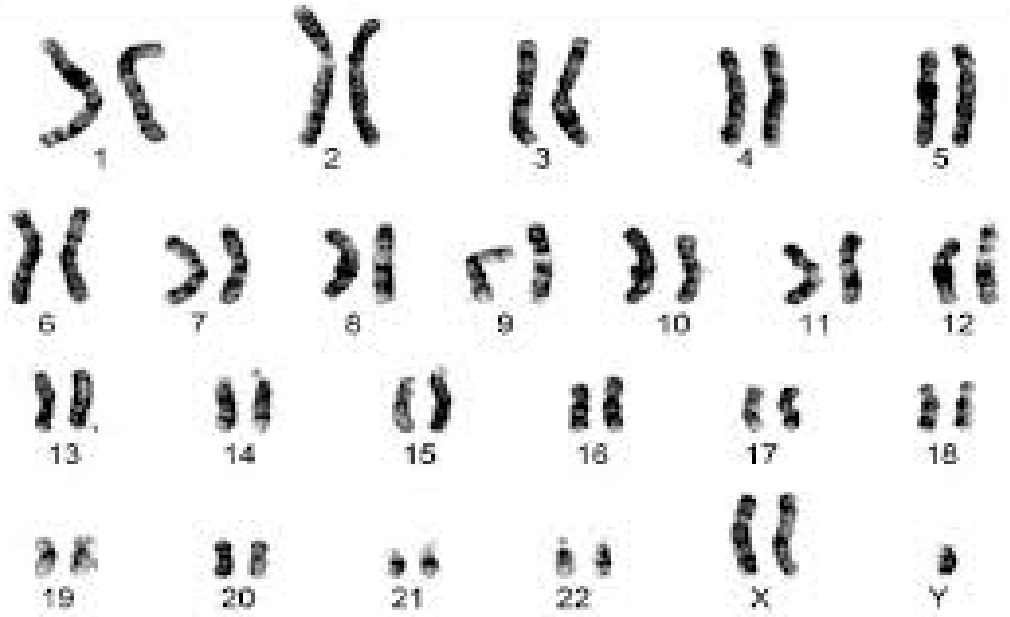
- **Cause**: Extra copy of chromosome 21, usually due to nondisjunction during meiosis.
- **Prevalence**: 1 in 700 live births worldwide.
- **Types**: Includes full trisomy (95% of cases), mosaic trisomy, and translocation trisomy.
- **Symptoms**: Macroglossia, Epicanthal folds, Small, low-set ears, simian crease, Short stature, Delayed developmental milestones and mild to moderate Intellectual disability.
- **Health Risks**: Congenital heart defects, thyroid dysfunction, and Alzheimer's disease.
- **Lifespan**: Improved medical care has extended life expectancy to 60+ years in many cases.

Aneuploidy (Poly Y Syndrome 47,XYY)



- **Cause**: Results from nondisjunction during sperm meiosis, leading to an extra Y chromosome.
- **Prevalence**: Affects 1 in 1,000 male births, often undiagnosed due to mild symptoms.
- **Physical Features**: Typically normal male sexual development; taller-than-average height is common.
- **Cognition**: Possible learning difficulties, especially in language and speech.
- **Behavioral Traits**: Increased risk of ADHD, impulsivity, and mild social or emotional challenges.
- **Fertility**: Normal fertility in most cases, with no major reproductive issues.
- **Health Risks**: Rarely associated to skeletal anomalies, tremors, or autism spectrum disorder (ASD).
- **Prognosis**: Most individuals lead normal, healthy lives with typical life expectancy.
- **Misconceptions**: Early myths linking to aggression or criminality unsupported by modern research.

Aneuploidy (Klinefelter Syndrome 47,XXY)



- **Cause**: Results from an extra X chromosome in males.
- **Prevalence**: Occurs in 1 in 500 to 1,000 male births.
- **Physical Features**: Tall stature, reduced muscle mass, broader hips, and gynecomastia.
- **Reproductive Effects**: Small testes, low testosterone levels, infertility, and reduced facial/body hair.
- **Cognitive Traits**: Learning difficulties, especially language and reading, normal/slightly reduced intelligence.
- **Behavioral Traits**: Increased risk of social anxiety, emotional immaturity, and low self-esteem.
- **Health Risks**: Higher risk of osteoporosis, diabetes, cardiovascular disease, and autoimmune disorders.
- **Treatment**: Testosterone replacement therapy.
- **Prognosis**: Can lead normal, healthy lives.
- **Variants**: Mosaicism (46,XY/47,XXY) can result in milder symptoms.

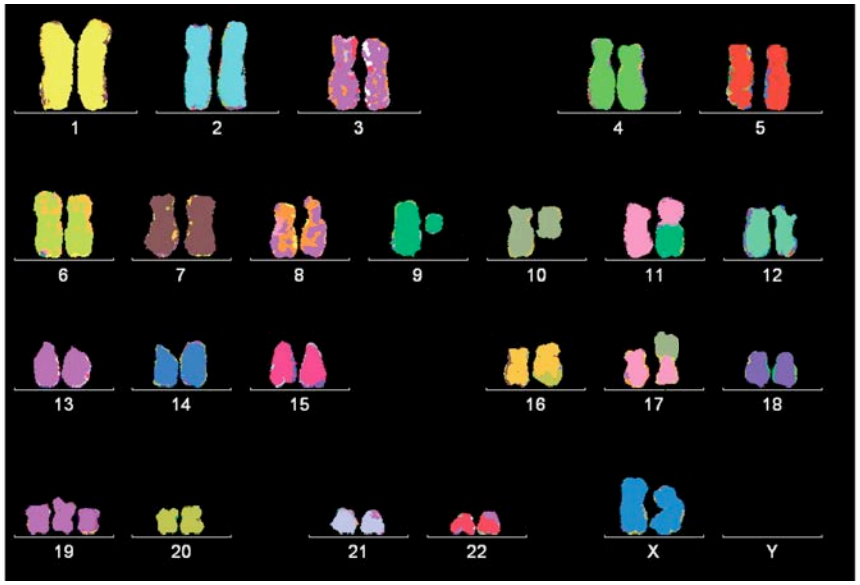
Fluorescence In Situ Hybridization (FISH)

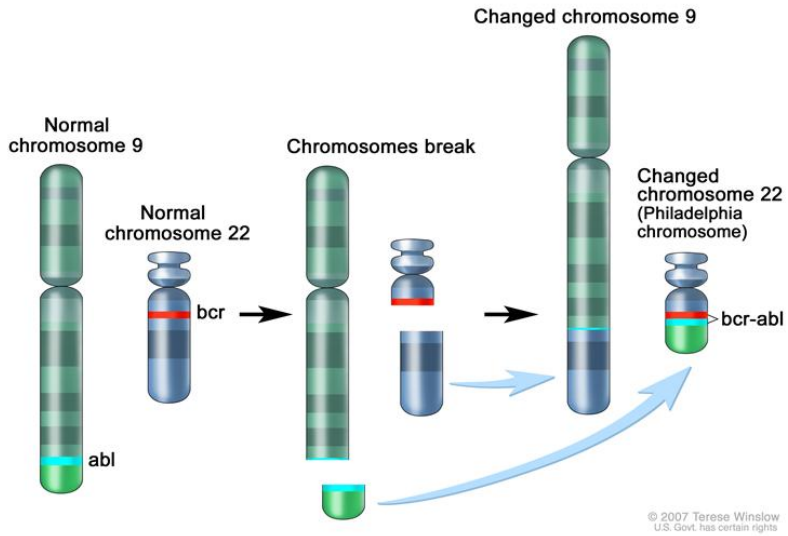
Uses fluorescently labeled DNA probes that bind to specific DNA sequences on chromosomes.

Provides higher resolution than traditional karyotyping, detecting microdeletions or duplications.

Identifies structural abnormalities, such as translocations, deletions, duplications, and inversions.

Detecting cancer-associated chromosomal abnormalities (e.g., BCR-ABL Philadelphia chromosome in CML).





Chromosome territories

During interphase, each chromosome occupies a specific region in the nucleus, known as a chromosome territory.

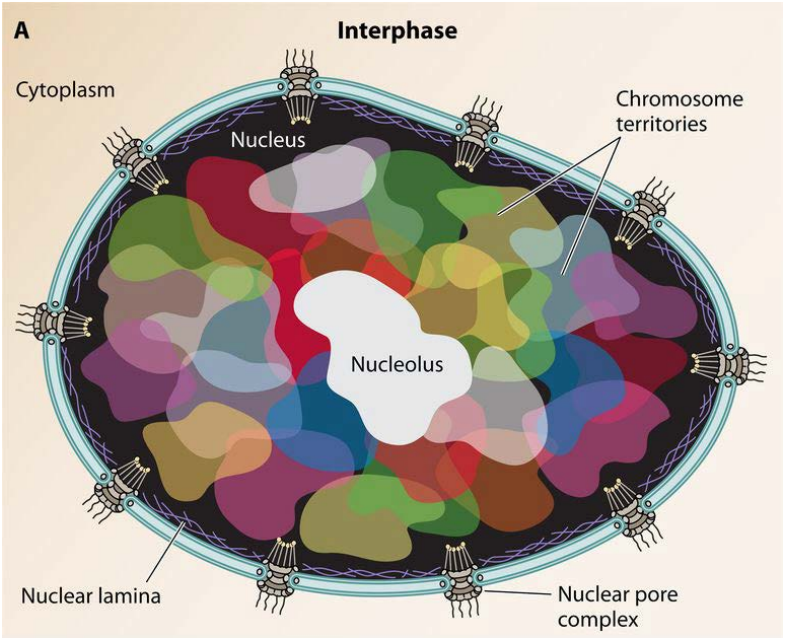
Discovered through FISH.

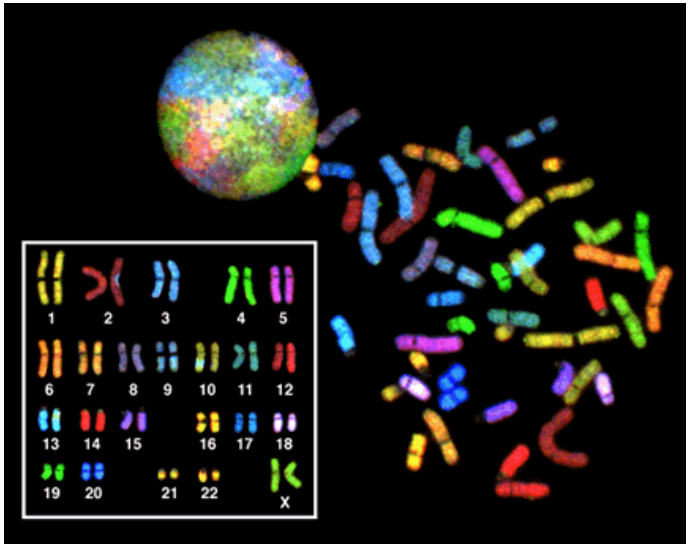
Prevents entanglement of chromosomes and ensures proper gene expression regulation.

Larger, gene-poor chromosomes (Chr 1) located in nuclear periphery.

Smaller, gene-rich chromosomes (Chr 19) localize near the nuclear centre.

Disruptions in territories linked to **translocations** (e.g., Philadelphia chromosome in leukemia), aneuploidies, aging, developmental disorders, and neurodegenerative diseases.





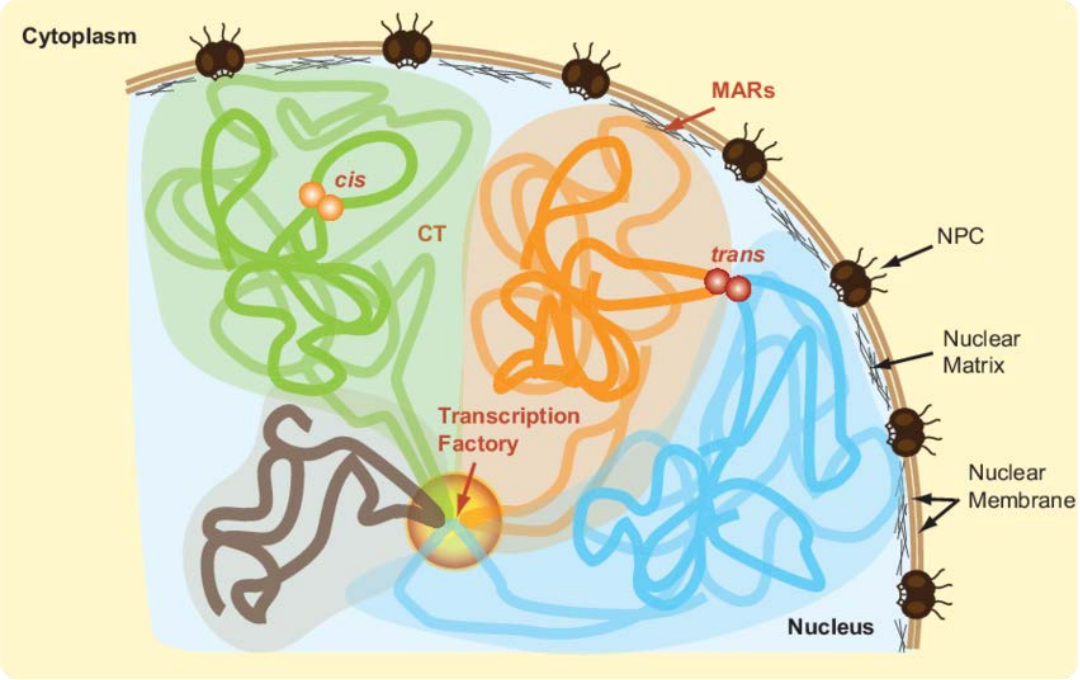
Chromosome territories

Facilitates gene regulation by clustering euchromatin and separating inactive chromatin.

Prevents entanglement of chromosomes, aiding in replication, transcription, and repair.

Enables spatial clustering of co-regulated genes and enhancer-promoter interactions.

Are dynamic and can reorganize in response to cellular signals, differentiation, and stress.



Chromatin

The degree of **DNA compaction must be dynamic** and modifiable to allow its function **(transcription and replication).**

It is necessary to provide (occasionally) access to other proteins to the DNA.

Chromatin mass: ca 30% nucleosomes (histones)
ca 30% non-histone
ca 30% DNA
<3% RNA

All chromatin-associated proteins that are not histones.

Much more variable and diverse group of proteins.

Involved in the control of gene expression and maintenance of superior compaction.
E.g. RNA Pol

A/B Compartmens and Euchromatin/Heterochromatin

The genome classified into two broad compartments: Active (A compartment) and inactive (B compartment) chromatin.

The A compartment is gene-rich, transcriptionally active, and associated with euchromatin.

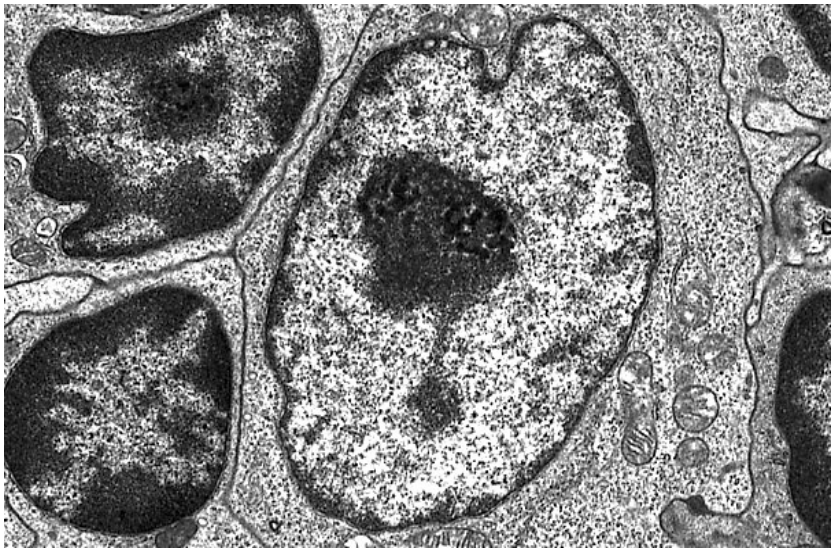
The B compartment is gene-poor, transcriptionally silent, and associated with heterochromatin.

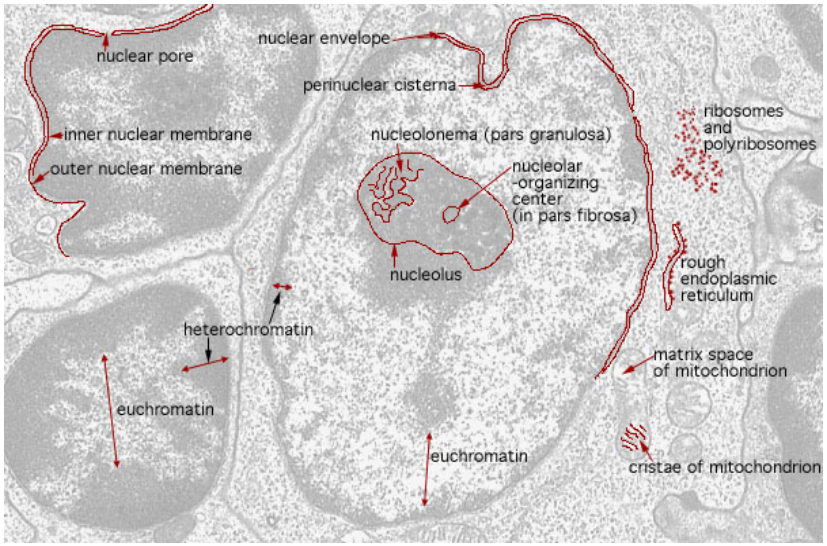
Euchromatin
Low uptake of Feulgen stain, active gene expresión.

Heterochromatin
High uptake of Feulgen stain, inactive.

Constitutive heterochromatin Inactivation at the organic level (whole organism).

Facultative heterochromatin tissue-specific inactivation.





Heterochromatin

Heterochromatin can be:

Constitutive: Specific non-coding sequences

satellite DNA

Centromeres

Facultative: Whole chromosomes that are condensed (inactivated)

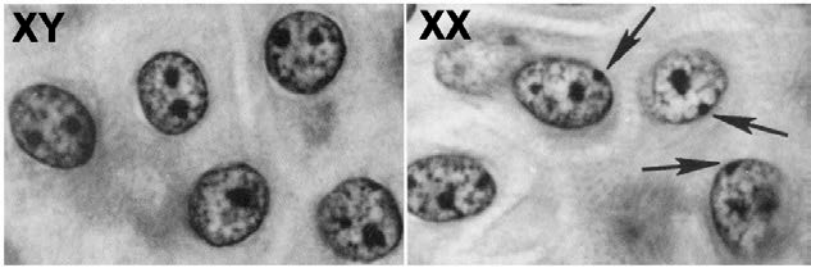
It depends on the cell lineage or stage of development.

Classic example: X chromosome and Barr body.

Inactivation of the X chromosome is **a way to balance the amount of genetic material available** (between both sexes).

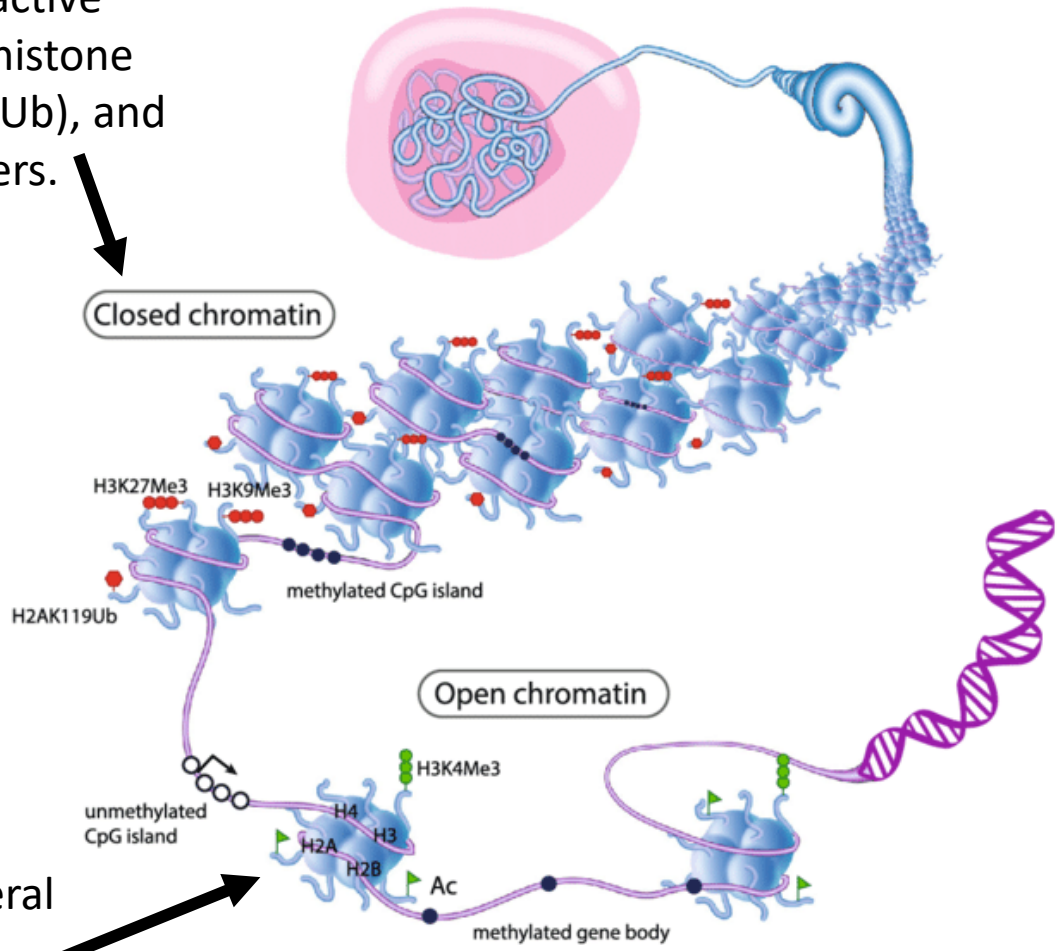
While one X chromosome remains active (euchromatin) the other does not (hetero).

Barr body corresponds to an inactivated X chromosome, only present in women or individuals with X polyploidy.



Epigenetic regulation of chromatin structure

The majority of DNA is packaged into inactive heterochromatin possessing repressive histone marks (H3K27Me3, H3K9Me3, H2AK119Ub), and methylated CpG islands in gene promoters.



Regions of open chromatin possess active histone marks (H3K4Me3), general histone acetylation and unmethylated CpG islands in gene promoters.

Heterochromatinization

Nucleation

Histone H3 lysine 9 trimethylation (H3K9me3)

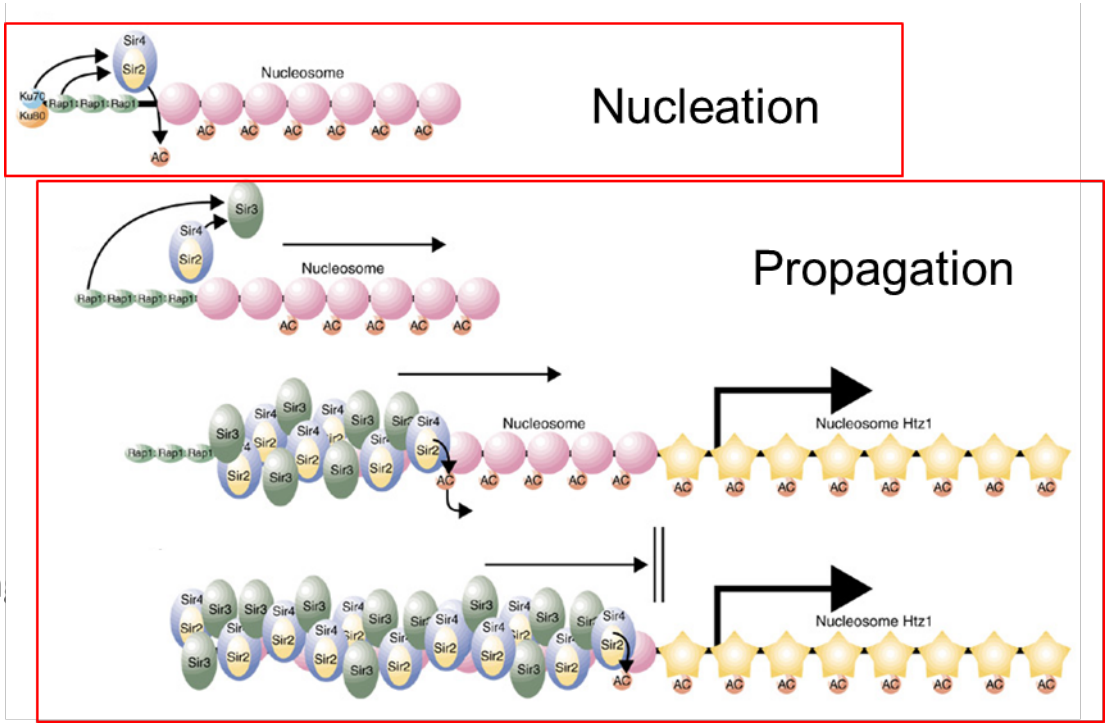
Histone H4 lysine 20 methylation (H4K20me3)

5-methylcytosine methylation in CpG islands reinforces heterochromatin nucleation.

HP1 (Heterochromatin Protein 1) Silencing Protein: Recognizes H3K9me3 marks and propagates heterochromatin.

Non-coding RNAs and RNA interference (RNAi) guide heterochromatinization.

Satellite DNA and transposable elements often serve as nucleation sites for heterochromatin.



Propagation and Spreading

Heterochromatin spreads to adjacent regions via a self-reinforcing loop of histone modifications and protein recruitment.

The extent of spreading is regulated by boundary elements to prevent inappropriate gene silencing (Htz1).

Subtelomeres

Highly variable 500 kb long domain adjacent to telomere.

Located between telomere and a gene-enriched chromosome body.

Has a relatively small number of genes.

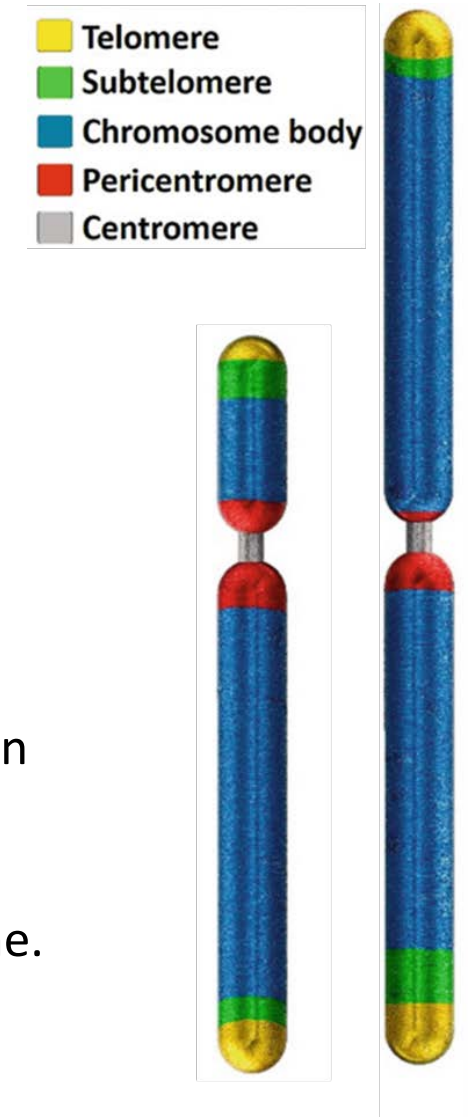
Has many repeats and segmental duplications.

Participates in intra- and inter-subtelomeric recombinations (genetic diversity of species).

Play a role both in development and evolution of mammalian karyotypes, due to chromosomal reorganizations.

Subtelomeres: most rapidly evolving regions of HoSa genome.

Human telomeres are quite short (5–15 kb)



Subtelomeric silencing

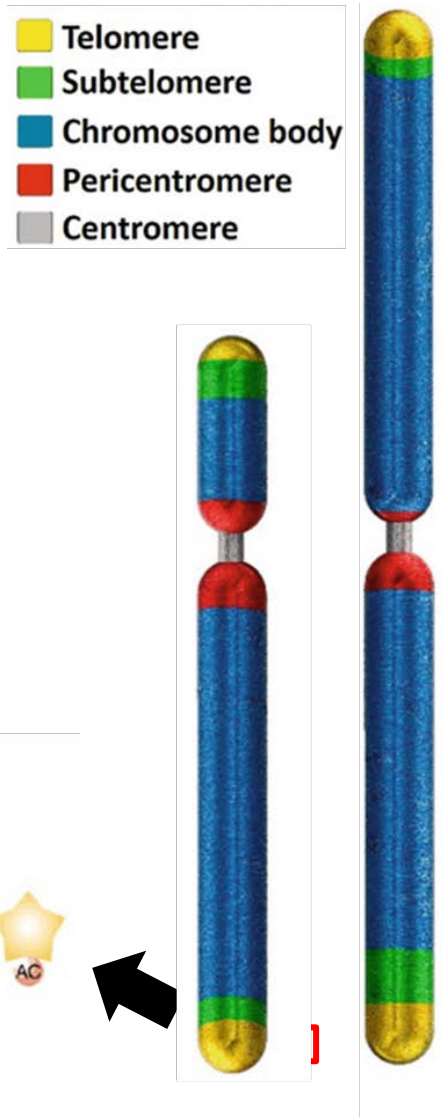
Subtelomeric silencing is an epigenetic mechanism.

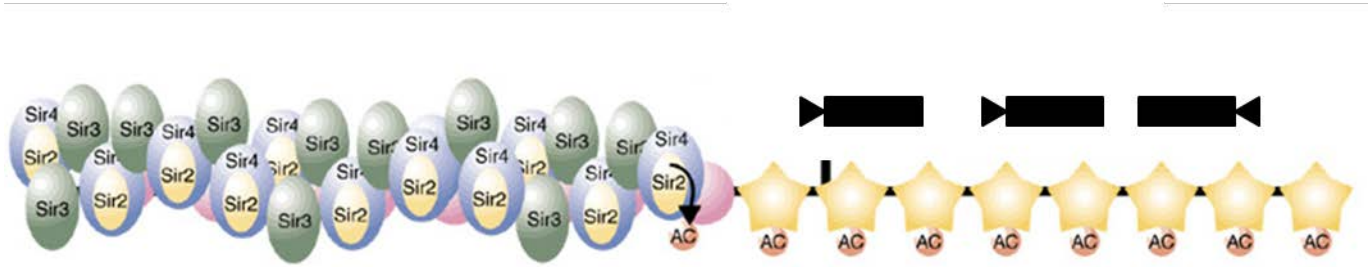
Genes in the subtelomeric regions are silenced.

Consequence of spread of telomeric heterochromatin.

Histone modifications, like H3K9 or H3K27 methylation, mark these regions for repression.

The silencing can be reversible and influenced by factors such as telomere length and environmental conditions.





X chromosome innactivation

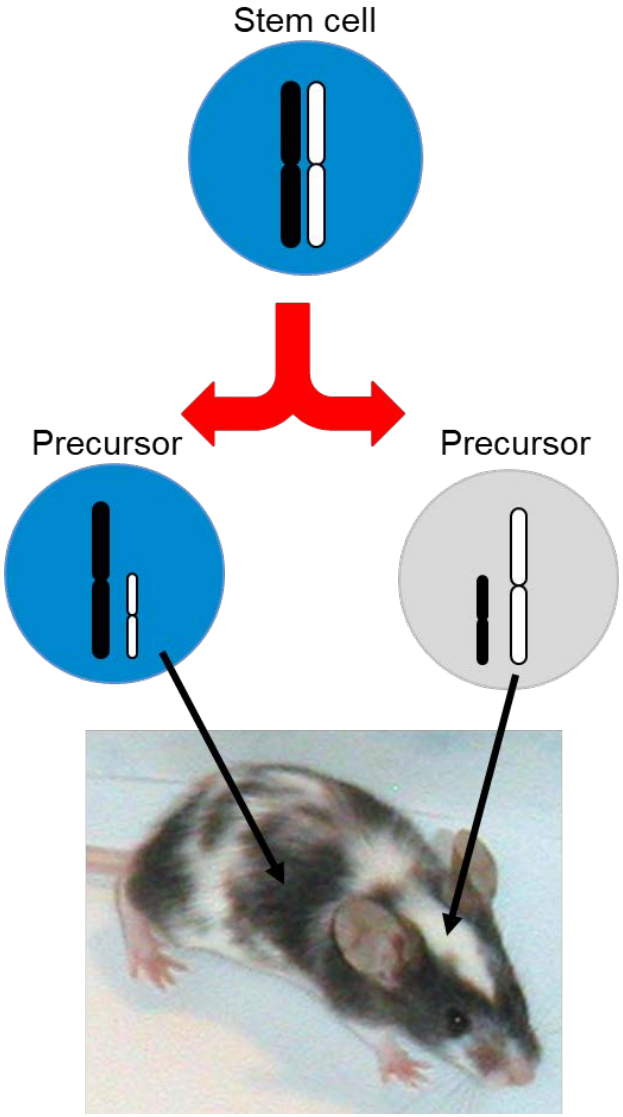
Inactivation of X chromosomes in mammals follows the n-1 rule: No matter how many X chromosomes exist, **only one can remain active,** the others will be inactivated.

Even in cases of **aneuploidies (47, XXX) it rarely shows phenotypic variations.**

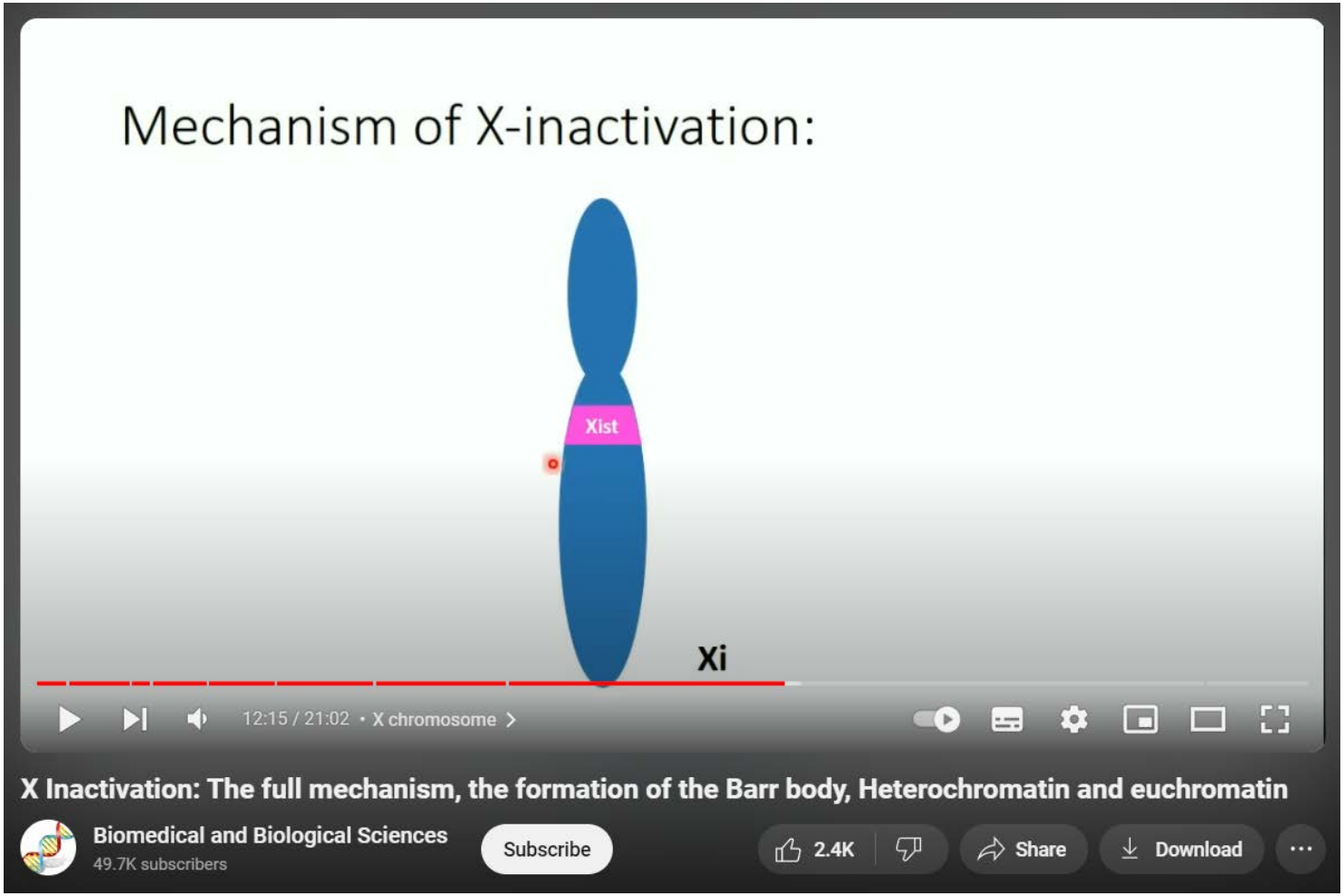
This **implies the existence of a counting mechanism** and another global inactivation mechanism.

A **locus** (ca 450 kbp) on the X chromosome identified as responsible for inactivation: **XIC (X inactivation center)** contains the **Xist** gene (RNA without ORF, open reading frame).

When this sequence is inserted into an autosome, the neosome participates in X counting and is subject to stochastic inactivation...however, the degree of inactivation is usually imperfect.



X inactivation summary video



X chromosome innactivation

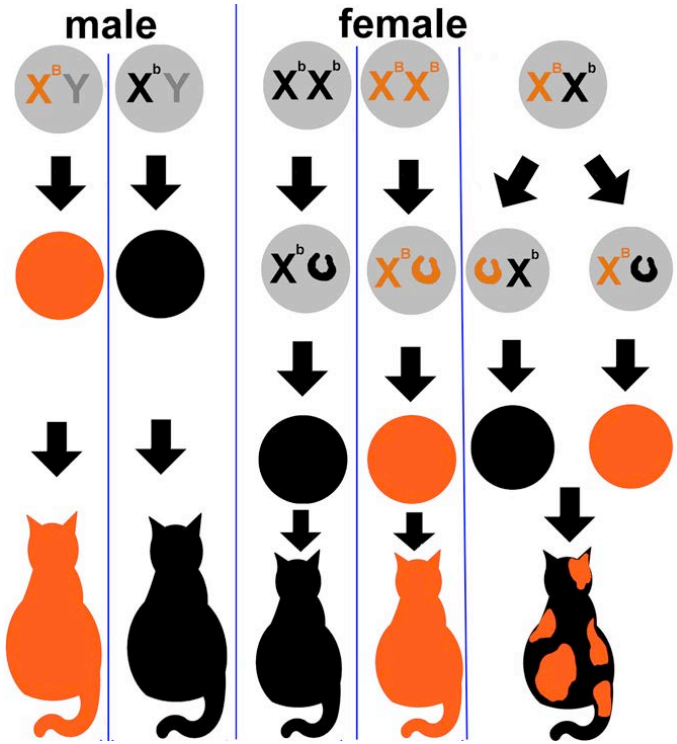
The calico cat is a classic example of X-chromosome inactivation in mammals.

Only seen in female cats (have two X chromosomes).

One X chromosomes is randomly inactivated during early embryonic development.

One allele might code for orange fur, and the other for black fur.

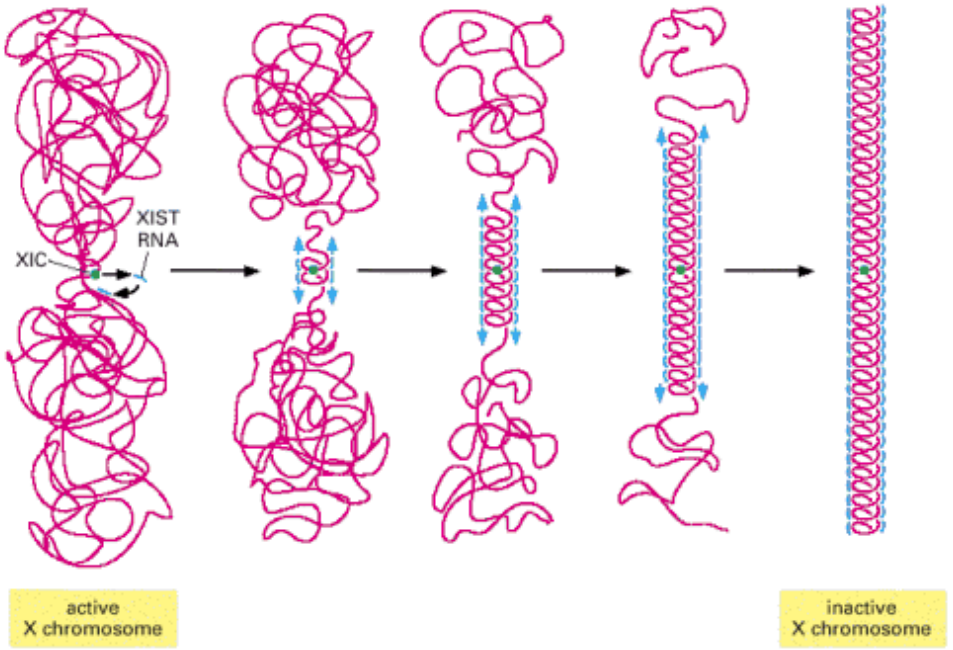
Characteristic patchwork of orange and black fur seen in female calico cats.



X chromosome innactivation

Locus **XIC** contains the **Xist** gene which is **only expressed by the inactive X chromosome (unlike the rest of the genes)** and is responsible for chromosome inactivation.

Its deletion prevents chromosome inactivation.



	♀ XX	♂ XY
	1X	2X
	½ X	1X
	X0	1X

X chromosome innactivation

Locus XIC contains the Xist gene which is initially expressed by both X chromosomes but only that of the inactive X chromosome is stabilized.

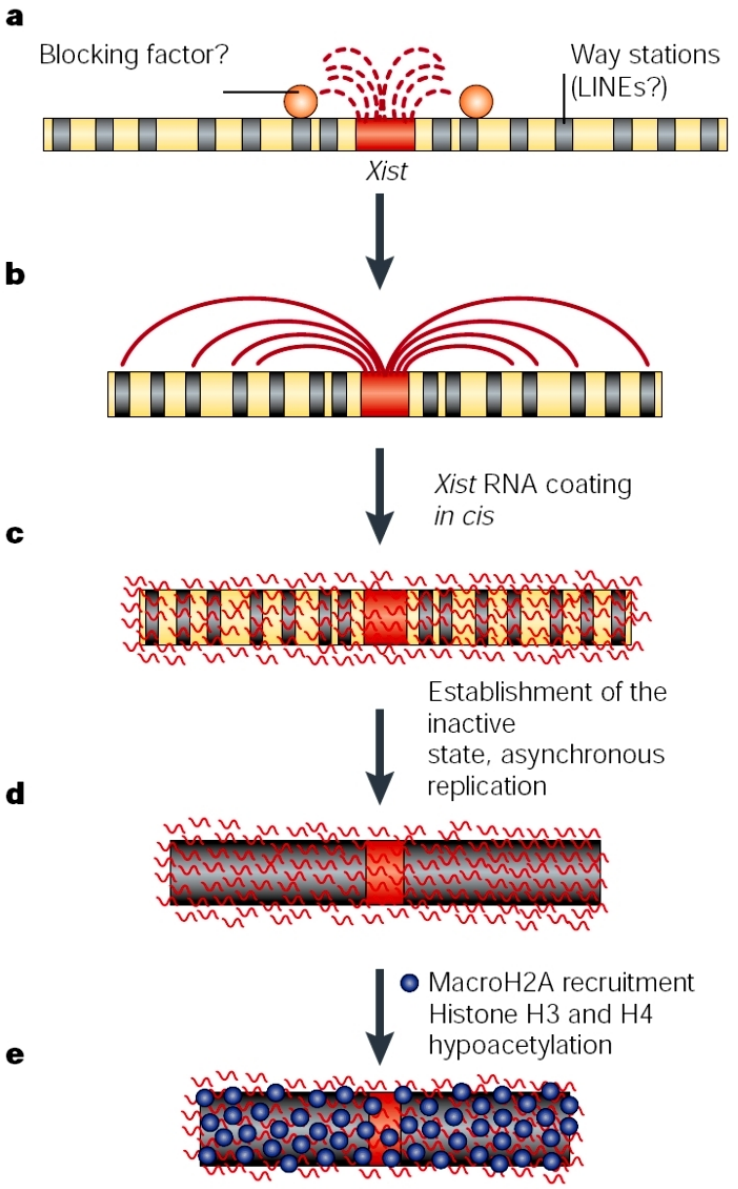
Xist encodes RNA without open reading frame (ORF).

RNA has a very short half-life (ca 2 hrs).

RNA-Xist is capable of lining the X chromosome that synthesizes it, suggesting a structural role.

It forms a punctate pattern around the inactivated X chromosome (before ME) which suggests that it interacts with certain structural proteins.

Finally, the process concludes with the deacetylation of H4 and the methylation of CpG sequences (ensuring the inactivation



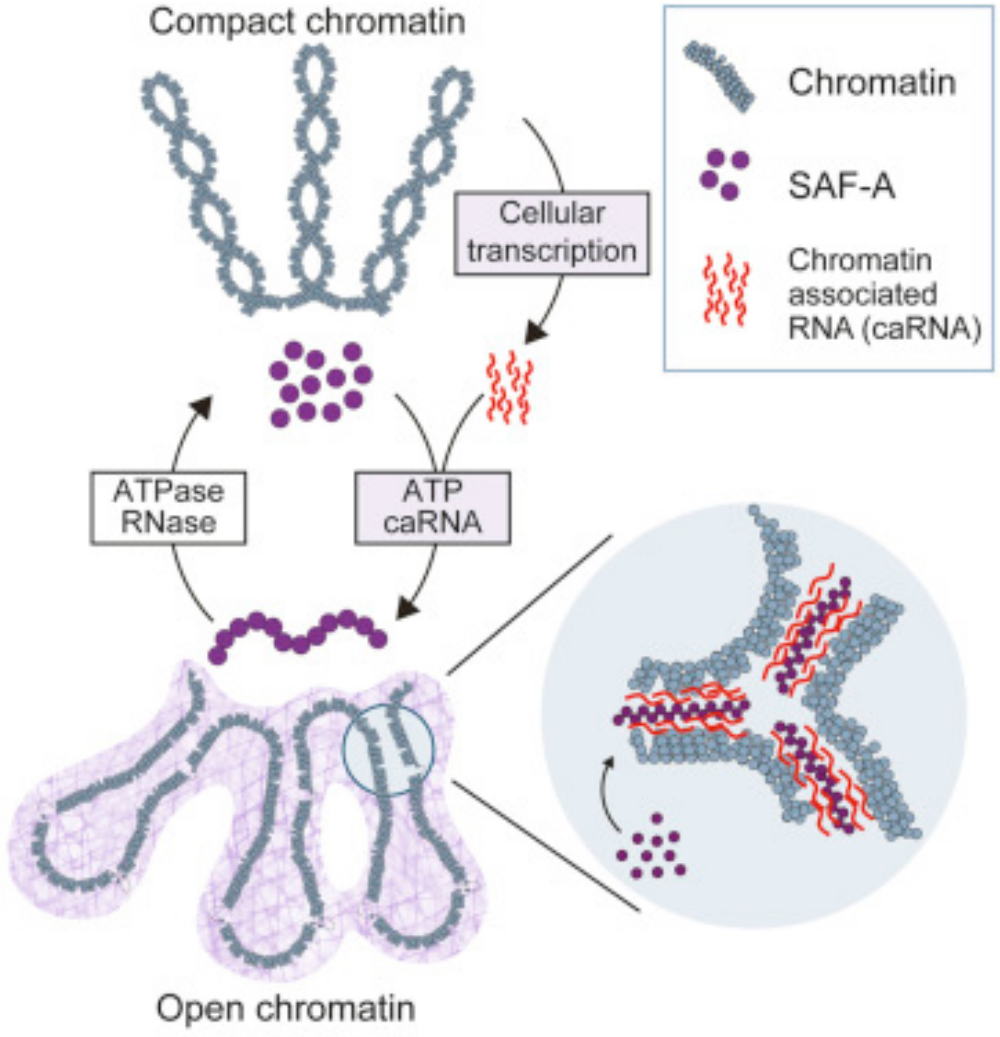
Epigenetic regulation of chromatin structure

RNAs that associate with chromatin directly or indirectly are called chromatin-associated RNAs (caRNAs).

caRNAs are involved in gene and transcriptional regulation.

They are part of non-coding RNAs class (ncRNA).

Recently developed Mapping RNA-Genome Interaction (MARGI) technology detect genome-wide RNA-chromatin interactions… including Xist.

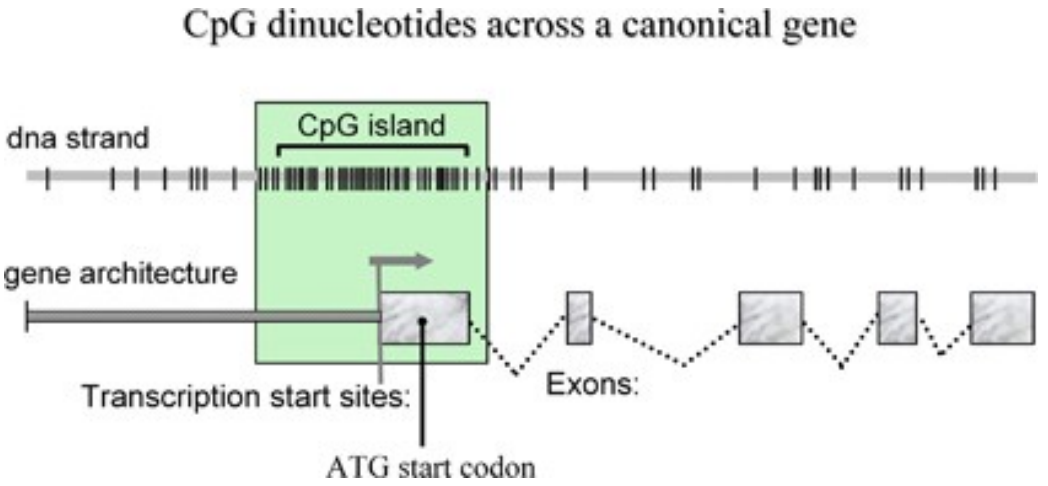


DNA CpG island methylation

CpG island hypermethylation is regulates gene expression

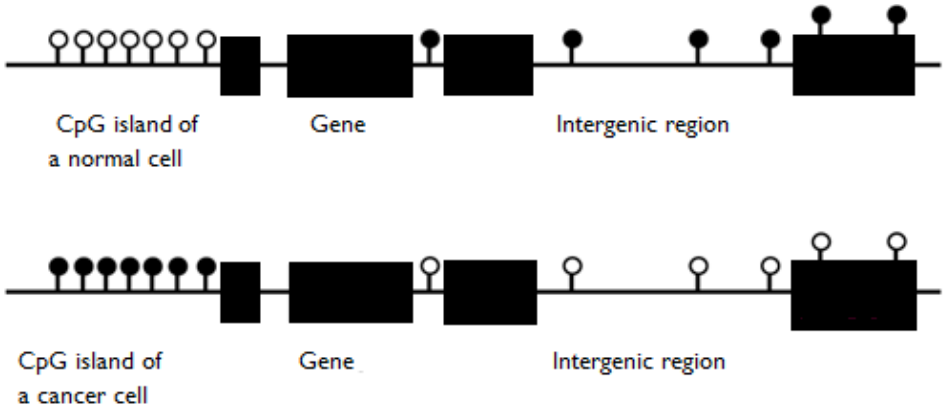
In a normal cell, the CpG island is hypomethylated while and the rest of the genome is methylated.

Can be locus-specific DNA hypermethylation or genome-wide DNA hypomethylation.



Hypermethylation of CpG islands has been described in almost every type of tumor.

CpG islands hypermethylation used for cancer diagnosis, prognosis and monitoring.



DNA CpG island methylation

There are **two types of methylases:**

De novo

Not yet characterized very well

Presumably sequence-specific

maintenance

The enzyme **responsible for the previous example**

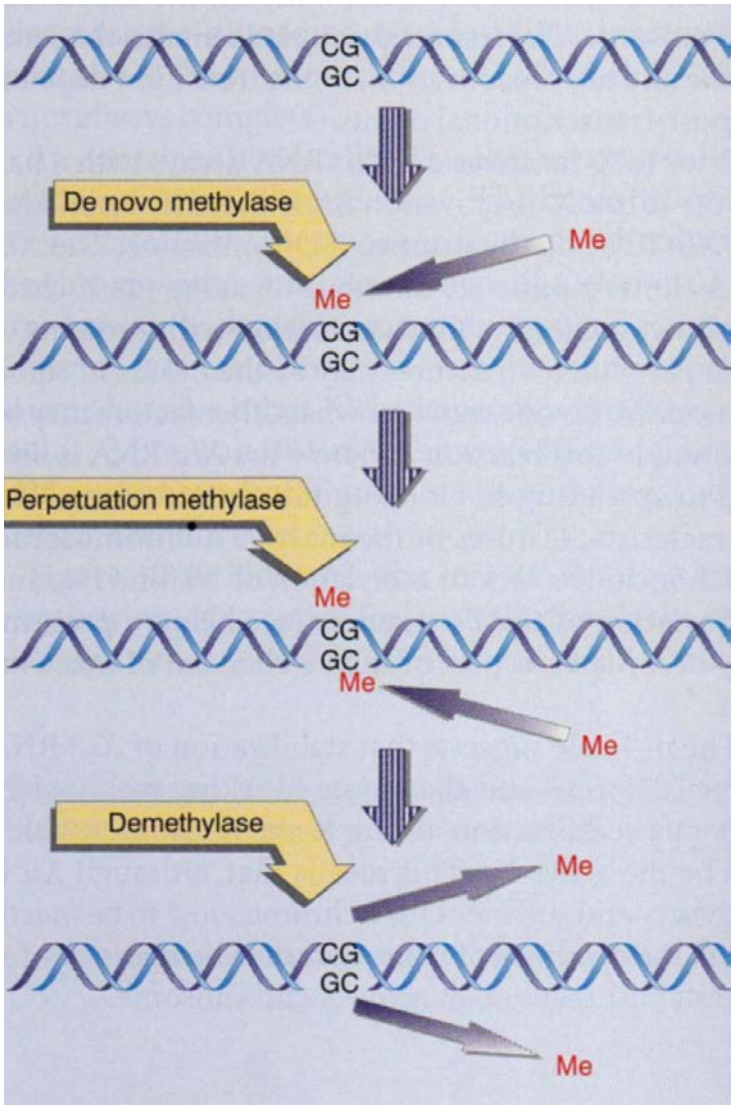
Hemimethylation to full methylation

Ubiquitous, constitutive expression

Essential for development

Deletion is embryonic-lethal

When methylated foreign DNA is inserted into a eukaryote, the methylation pattern is respected and is perpetuated indefinitely with >95% fidelity!



Genomic imprinting

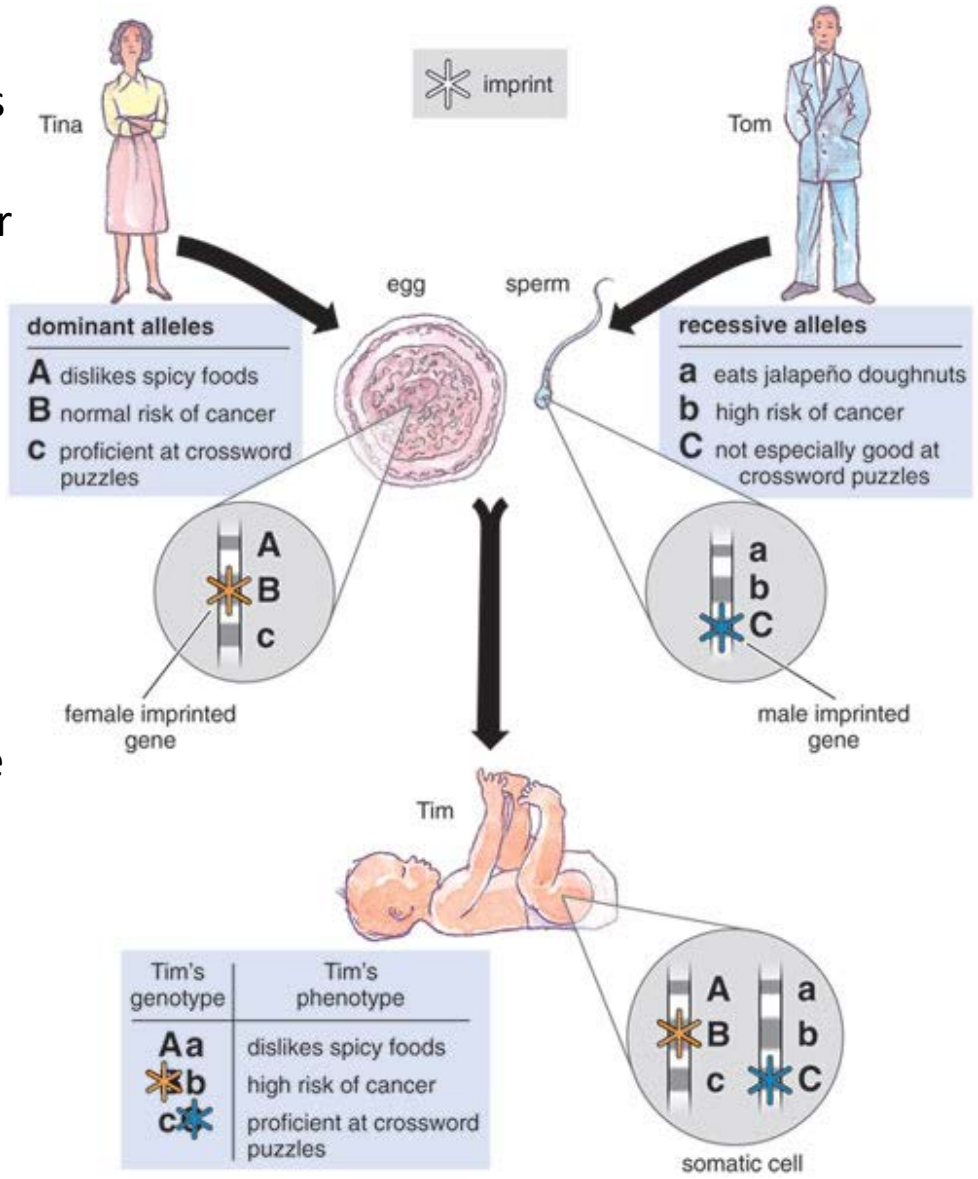
An epigenetic phenomenon that causes genes to be expressed or not, depending on whether they are inherited from the female or male parent.

Inheritance process independent of the classical Mendelian inheritance.

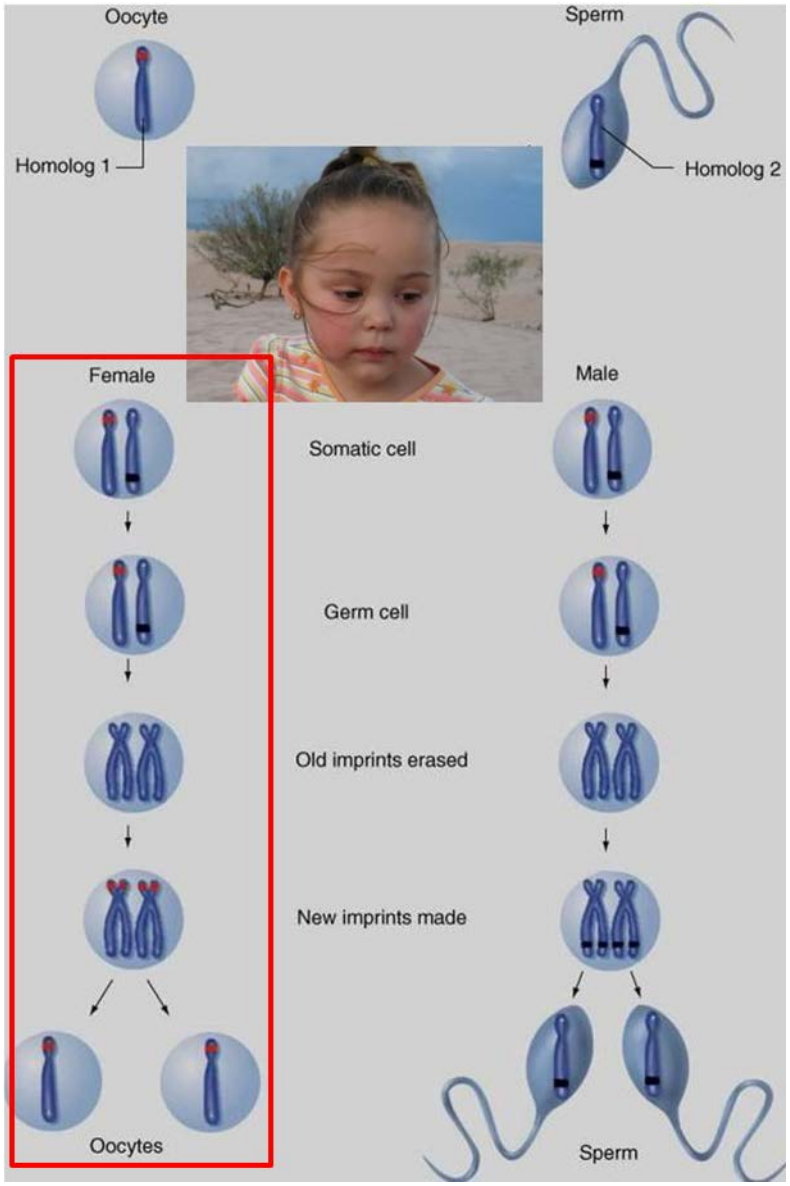
Involves DNA & histone methylation without altering the genetic sequence.

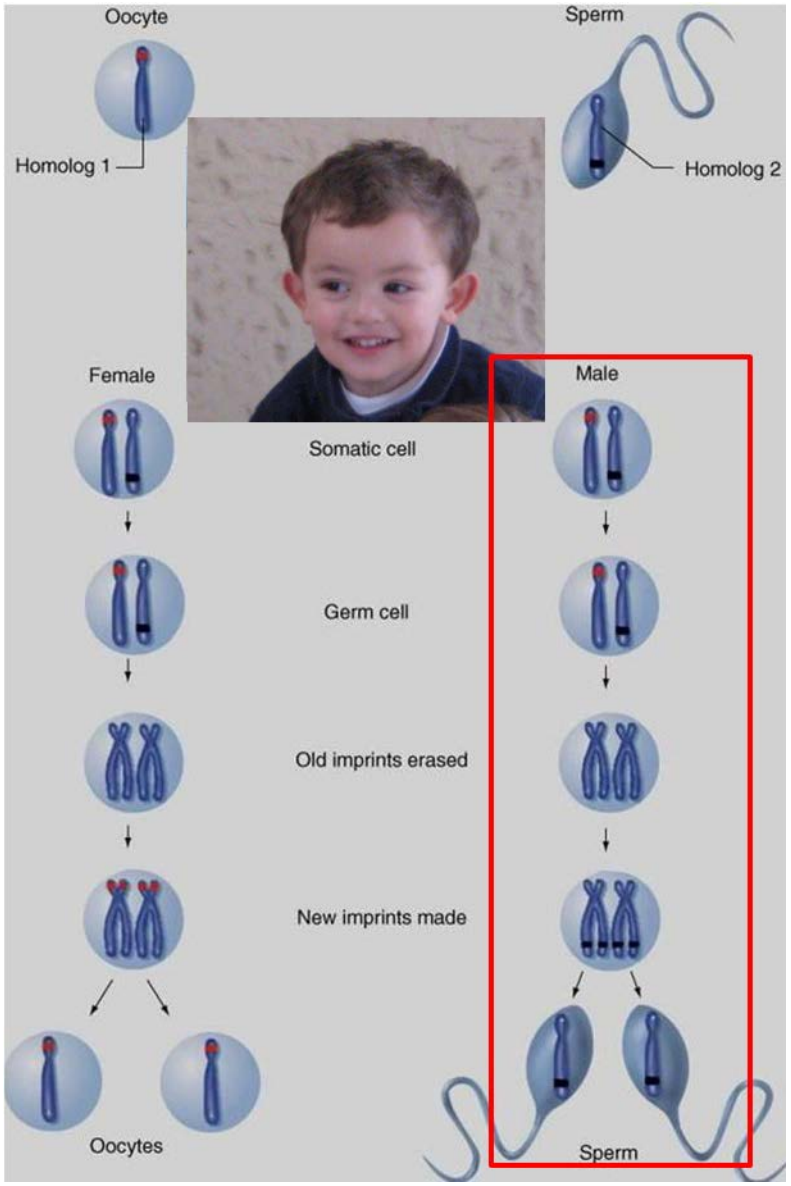
These epigenetic marks are "imprinted" in the germline (sperm or egg cells) of the parents and maintained through mitotic cell divisions in somatic cells.

As of 2019, there were about 228 imprinted genes known in humans.

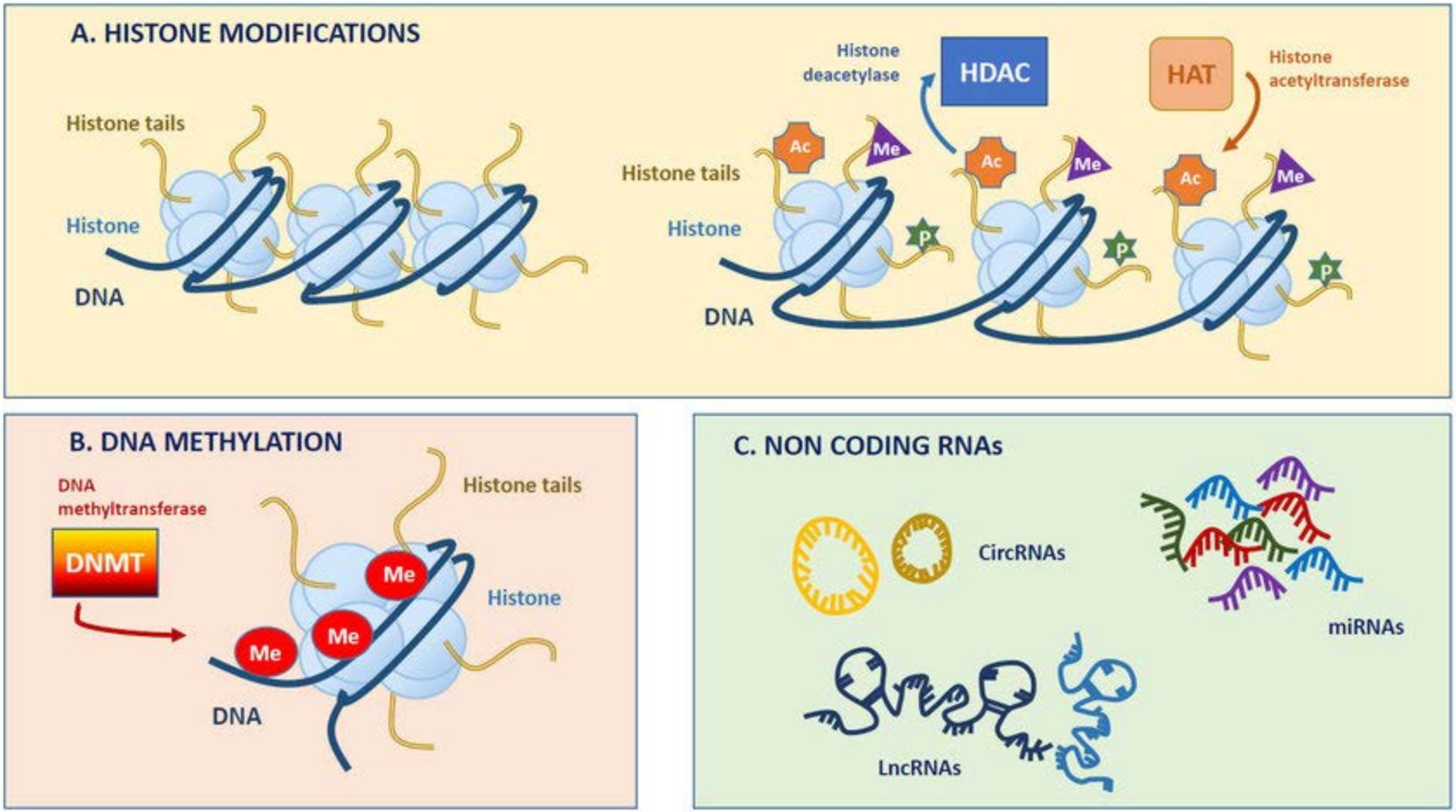


Genomic imprinting





Epigenetic information



Laboratorio de Genómica Viral y Humana

Instalaciones de Alta Contención Biológica Nivel de Bioseguridad 3 (BSL-3) CDC-certificadas

Facultad de Medicina UASLP
San Luis Potosí, México

www.genomica.uaslp.mx

GenomicaUASLP

YouTube GenomicaUASLP

Content copyright and license

The Viral and Human Genomics Laboratory is committed to promoting the human rights of free access to knowledge and to receiving the benefits of scientific progress and its applications by providing universal access to all the resources and publications it produces. This is in agreement with article 15 of the United Nations International Covenant on Economic, Social and Cultural Rights published on April 30, 2020.

(Last updated: August 23, © 2024.)