Eukaryotic Gene Expression: Basics & Benefits

P N RANGARAJAN

Lecture 27

Epigenetic regulation of gene expression during development

Development of a multicellular organism is not only determined by the DNA sequence but also epigenetically through DNA methylation and post-translational modifications of histone proteins.

Epigenetic events influence the expression of several genes during embryonic development.

For ex., in embryonic stem (ES) cells, active gene expression markers are found at pluripotent genes and repressive markers are found at lineage-specific genes.

Thus, different cell types and their fates can be defined by their epigenetic and gene expression profiles.

HISTONE CODE

Acetylation of H3 or H4 = active genes

Deacetylation of histones = inactive genes

Methylation of K4 of H3 = active genes expression

Methylation of K9 of H3 = gene silencing

DNA methylation plays a key role in regulation of gene expression

5-methyl cytosine – fifth base

EPIGENTIC CODE

How does DNA methylation/demethylation regulate gene expression during development?

Genome-wide DNA demethylation of paternal and maternal nuclei.

Prior to fertilization, mammalian gametes are at different stages of the cell cycle and their genomes are organized differently.

The egg is meiotically arrested at metaphase II, resulting in a diploid genome that is packaged with histones.

Mature sperm complete meiosis, but their haploid genomes are packaged with protamines instead of histones.

When a sperm penetrates the zona pellucida to fertilize the egg, both gametes undergo rapid changes.

The egg completes its second meiosis resulting in the extrusion of one copy of the genome as the polar body;

the sperm reorganizes its genomic DNA by replacing protamines with histone proteins.

Around 4-8 hr after fertilization, after the protamine–histone exchange, the sperm-derived paternal pronucleus undergoes genome-wide DNA demethylation and this is completed before the first cell division. The paternal DNA remains demethylated after multiple rounds of cell division.

Specific cytoplasmic factors in the fertilized egg are responsible for this selective demethylation of paternal DNA.

Thus, although both the maternal and paternal DNA are exposed to the same cytoplasmic factors, only the paternal but not the maternal DNA is demethylated.

This may be due to a mechanism that protects the maternal genome from this wave of demethylation or to a putative DNA demethylase that is specifically recruited to the paternal genome.

In addition to replacement of protamines with histones, specific histone variants are also deposited in the paternal pronucleus. For example, a histone variant H3.3 is preferentially deposited in the paternal pronucleus and it was suggested such differences may be responsible for selective paternal genome-specific demethylation process.

In addition, certain histone modifications are also paternal- or maternal-specific. For example, methylation, dimethylation and trimethylation at H3 Lys27 and at Lys9 are seen in the maternal pronucleus of zygotes but not in the paternal pronucleus.

Thus, the maternal genome may use a protective mechanism against demethylation by carrying specific histone variants or modifications.

Interestingy, it has been reported that certain non-histone factors present in the oocyte might protect the maternal genome from demethylation. Zygotes lacking a maternal effect gene known as stella/*DPPA3/PGC7)* exhibited demethylation of both pronuclei, although *how* stella protects the maternal genome from demethylation remains to be determined.

Nakamura et al., Nature Cell Biology 9, 64 - 71 (2006)

PGC7/Stella protects against DNA demethylation in early embryogenesis

DNA demethylation

DNA demethylation involves removal of the methyl group from 5-methyl cytosine (active demethylation) or by the inhibition of the DNMT1 (passive demethylation).

Even though DNA methylation contributes to stable, long-term and heritable gene silencing, DNA methylation levels can change rapidly by mechanisms involving active DNA demethylation.

DNA demethylation can occur at two different levels:

Genome-wide demethylation, which occurs at specific times during early development,

Gene-specific demethylation, which occurs in somatic cells responding to specific signals.

During embryonic development, some genomic regions are resistant to demethylation.

These genomic regions include:

imprinting control regions
intracisternal A particle (IAP)
Retrotransposons
centric and pericentric hetero chromatin.

DNA methylation

Specific patterns of DNA methylation are reestablished during early development by the action of DNA methyltransferase and demethylase enzymes.

DNA methyltransferases (DNMTs)

In mammals, DNA methylation occurs predominantly in the context of CpG dinucleotides, whereas DNA methylation in plants can occur at C bases in diverse sequence contexts.

The enzymes responsible for this modification are known as DNA methyltransferases (DNMTs).

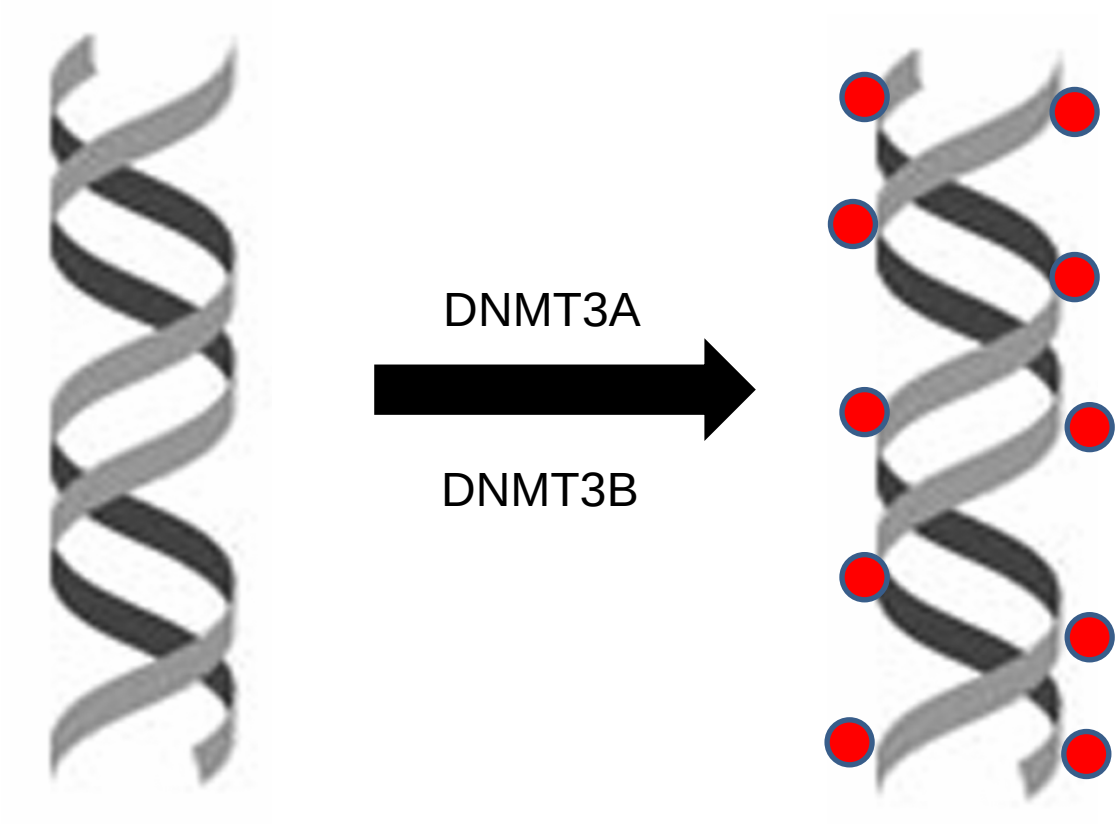
DNMTs fall under two categories: *de novo* DNMTs and maintenance DNMTs

DNA methylation patterns are initially established by the *de novo* DNA methyltransferases, DNMT3A and DNMT3B during the blastocyst stage of embryonic development.

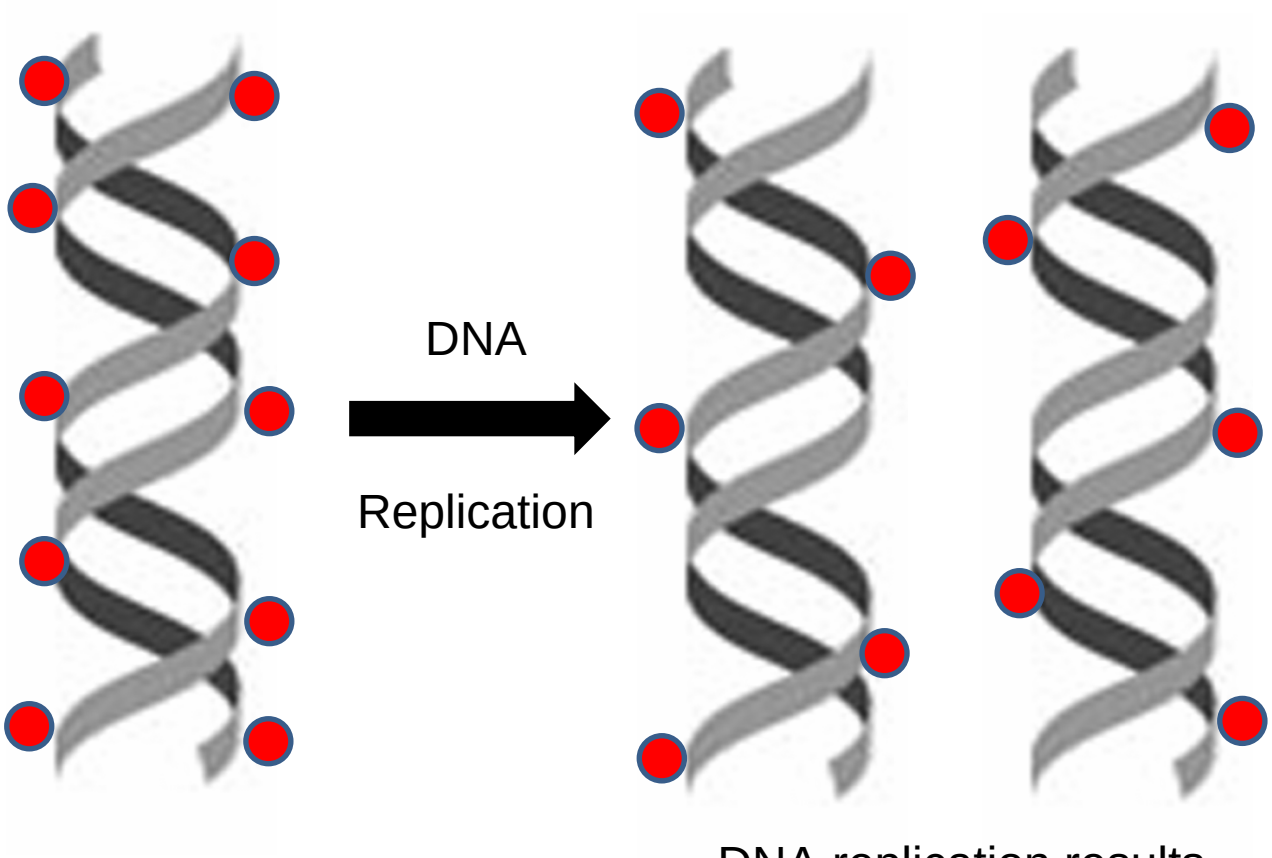
These methyl marks are then faithfully maintained during subsequent cell divisions by the action of the maintenance methyltransferase, DNMT1, which has a preference for hemi-methylated DNA.

DNMT1	maintenance methyltransferases
DNMT3A *DNMT3B*	*de novo* methyltransferases – highly expressed at embryo implantation when waves of de novo methylation are occurring in the genome

De novo DNA methylation

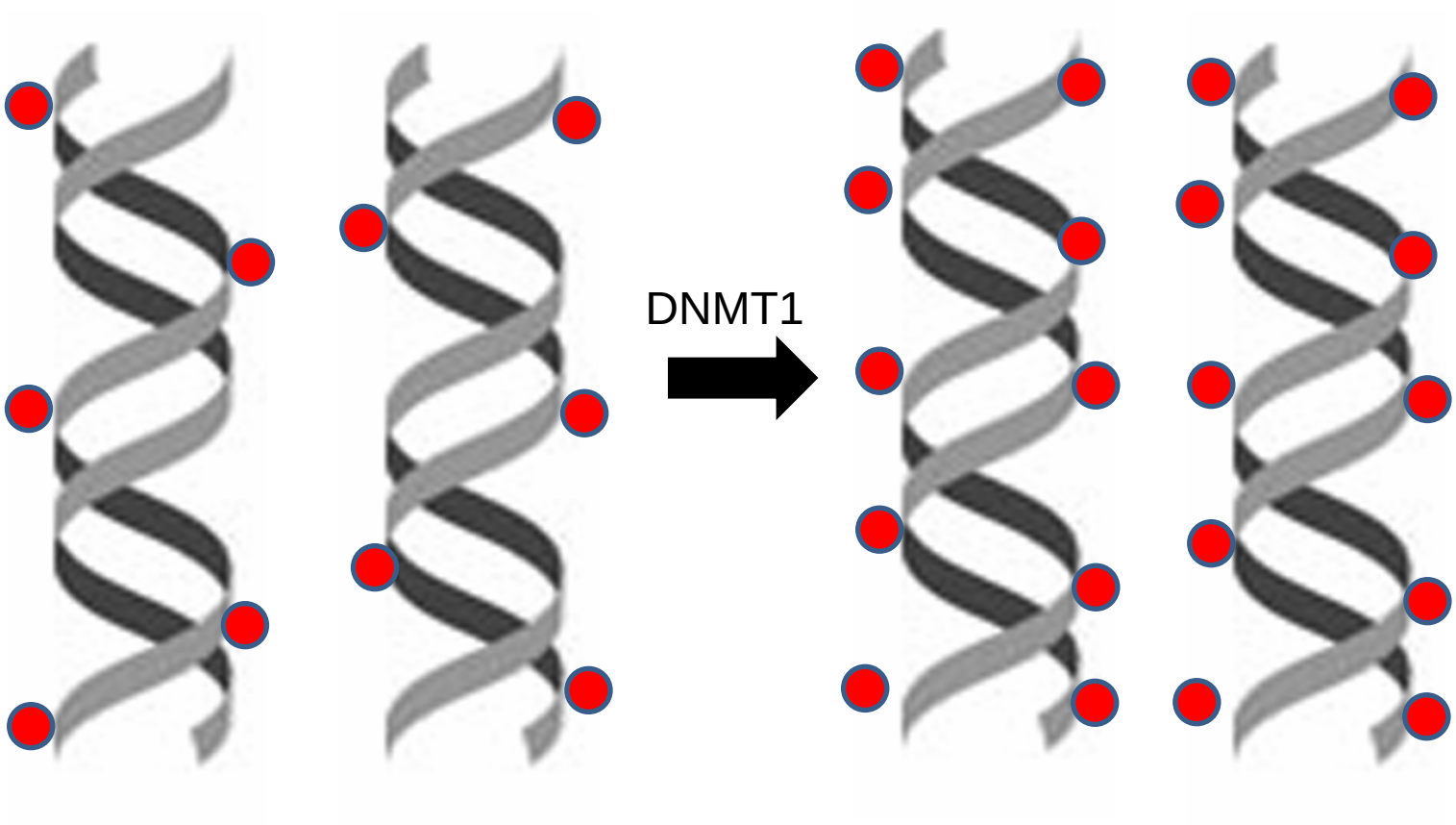


During early development, methylation patterns are initially established by the *de novo* DNA methyltransferases, DNMT3A and DNMT3B.



DNA replication results in the generation of hemi-methylated DNA

Maintainence of DNA methylation



DNMT1 acts on the hemimethylated DNA and methylates the newly synthesized DNA and thus maintains the methylation status of DNA

DNA METHYLATION PLAYS A VERYIMPORTANT ROLE

IN DIFFERENTIAL GENE EXPRESSION DURING

EMBRYONIC DEVELOPMENT

GENOME IMPRINTING

X CHROMOSOME INACTIVATION

DNA METHYLATION ,
IMPRINTING
AND
TRANSCRIPTIONAL SILENCING

WHA T IS IMPRINTING?

Selective inactivation of the maternal or paternal allele such that only the paternal or maternal allele is expressed.

In diploid organisms, somatic cells possess two copies of the genome.

Each autosomal gene is therefore represented by two copies, or alleles, with one copy inherited from each parent at fertilisation.

For the vast majority of autosomal genes, expression occurs from both alleles simultaneously.

However, in mammals, a small proportion (<1%) of genes, only one of the alleles (either paternal or maternal) is expressed such that gene expression occurs from only one allele.

Selective inactivation of one of the alleles is known as IMPRINTING

The term "imprinting" was first used to describe events in the insect *Pseudococcus nipae* (mealbugs).

In these insects, both the male and female develop from a fertilized egg.

In females, all chromosomes remain euchromatic and functional.

In embryos destined to become males, one haploid set of chromosomes becomes heterochromatinized after the sixth cleavage division and remains so in most tissues. Thus, males are functionally haploid.

In insects, imprinting describes the silencing of the paternal genome in males, and thus is involved in sex determination.

The fact that the maternal and paternal genomes are not exactly same became evident from nuclear transplantation experiments in mouse zygotes in the early 1980s.

The vast majority of mouse parthenogenones/gynogenones (with two maternal or egg genomes) and androgenones (with two paternal or sperm genomes) die at, or before, the blastocyst/implantation stage.

These studies indicated that normal development requires the contribution of both the maternal and paternal genomes.

Parthenogenetic/gynogenetic embryos have twice the normal expression level of maternally derived genes, and lack expression of paternally expressed genes, while the reverse is true for androgenetic embryos.

In mammals, genomic imprinting describes the processes involved in introducing functional inequality between two parental alleles of a gene.

Two well-studied examples if imprinting in humans are:

H19: only the maternal allele is expressed

insulin-like growth factor 2 (*Igf2):* only the paternal allele is expressed

Mammalian genome has > 100 imprinted genes in clusters

These genes are Imprinted by selective methylation of one of the alleles

H19 is a long noncoding RNA that is transcribed only from the maternally inherited allele; the paternal H19 allele is not expressed.

A major function of H19 is to regulate the expression of another gene known as the Igf2.

The regulatory elements which control the imprinting are known as imprinting control regions (ICRs) and differentially methylated regions (DMRs).

H19 contains a DMR that also acts as an ICR.

ICR is differentially methylated at its CpGs according to parental inheritance.

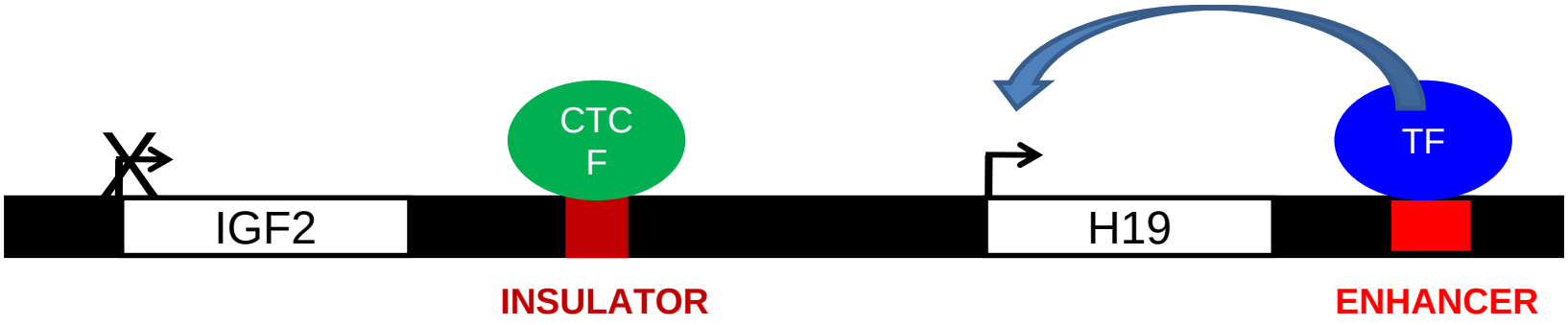
Usually, the paternal copy of H19 is methylated and silent while the maternal copy is hypomethylated or unmethylated and expressed in the offspring cell.

When H19 promoter DNA is methylated, its expression is shut off. However, at the same time , the expression of IGF2, a neighboring gene on chromosome 11, increases.

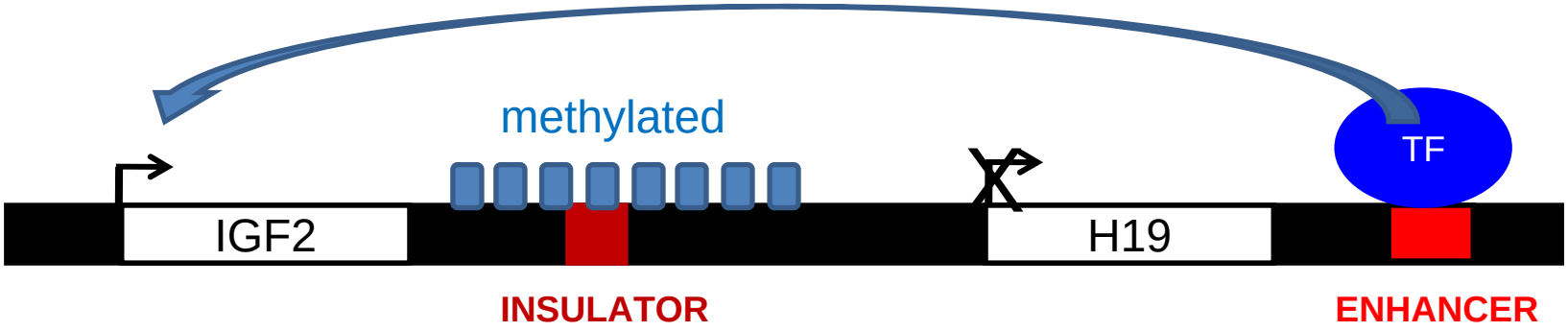
Specific cis-acting elements called

insulators

play a key role in the differential regulation of H19 and IGF2



HUMAN MATERNAL CHROMOSOME 11



HUMAN PATERNAL CHROMOSOME 11

H19: only the maternal allele is expressed

insulin-like growth factor 2 (*Igf2):* only the paternal allele is expressed

Application:

A gene inserted at random into the mammalian genome is often “silenced”, and placing insulators upstream and downstream of that gene can protect the gene from silencing.

The expression of non-coding RNAs, such as *Air* on mouse chromosome 17 and *KCNQ1OT1* on human chromosome 11p15.5, also play a key role in gemome imprinting

The *Air* Noncoding RNA: An Imprinted *cis*-silencing Transcript
BRAIDOTTI, et al., doi:10.1101/sqb.2004.69.55
Cold Spring Harb Symp Quant Biol 2004. 69: 55-66

Abnormal imprinting during development results in genetic diseases:

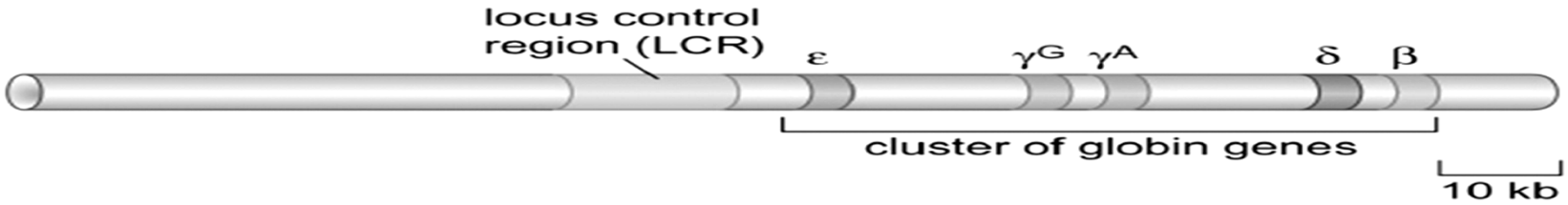
Beckwith-Wiedemann syndrome,
Silver-Russell Syndrome,
Angelman Syndrome
Prader-Willi Syndrome.

Regulation of certain genes requires locus control region (LCR)

Human and mouse globin genes are clustered in genome and differently expressed at different stages of development

A group of regulatory elements collectively called the locus control region (LCR), is found 30-50 kb upstream of the cluster of globin genes.

It binds regulatory proteins that cause the chromatin structure to “open up”, allowing access to the array of regulators that control expression of the individual genes in a defined order.



Locus control regions
Qiliang Li, Kenneth R. Peterson, Xiangdong Fang, and George Stamatoyannopoulos
Blood, 1 November 2002, Vol. 100, No. 9, pp. 3077-3086

X Chromosome Inactivation
during embryonic development

X chromosome inactivation

Chromosomal dosage and compensation

- Women are XX, men are XY
- How are levels of all essential X-encoded gene products similar between men and women if women have twice the number of alleles?

Mary Lyon – 1961

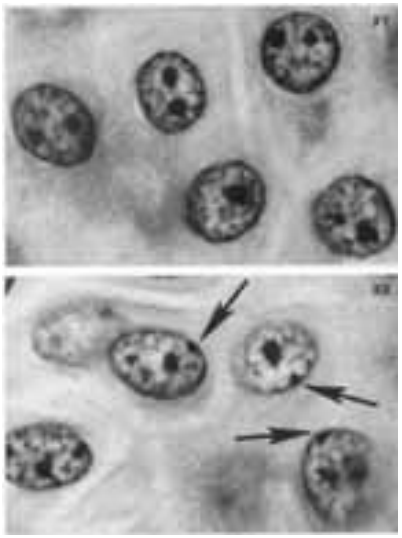
In cells with multiple X chromosomes, all but one is inactivated during mammalian embryogenesis – the "*Lyon effect*"

X-inactivation; which X?
Usually random
but always paternal in marsupials
and variable in calico cats representing regional expression of differing pigmentation genes on alternate X chromosomes

Lyon, M. (1961). "Gene action in the X-chromosome of the mouse (Mus musculus L.)". *Nature* 190: 372-3.

X-inactivation

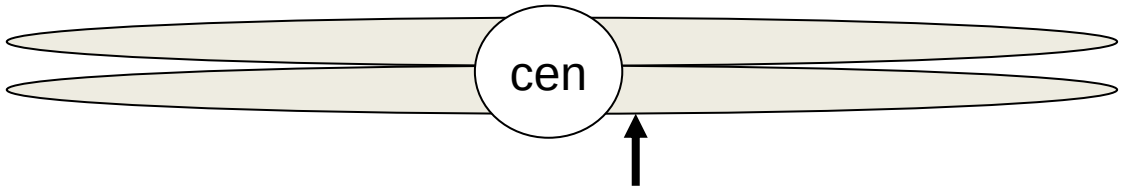
The repressed X-chromosome condenses to form a Barr body



The Lyon Hypothesis (Lyonization)

- Random selection of X chromosome
  - Inactive throughout cell's lifetime
  - Xa = active X chromosome
  - Xi = inactive X chromosome

- X inactivation center (XIC)
  - Near centromere
  - Contains 12 genes
    - 7 genes code for proteins
    - 5 genes code for untranslated RNA



XIC – the X inactivation centre

required for X-inactivation

introduction of XIC to ANY chromosome leads to silencing

XIC encodes two genes: *XIST*, and *TSIX*

- Xist and Tsix
  - Two genes actively involved in inactivating an X chromosome
  - Antagonistic roles
    - Xi → ↑ Xist expression
      ↓ Tsix expression
    - Xa → ↓ Xist expression
      ↑ Tsix expression

Human Xist is a 17 kb long RNA that is expressed on the inactive chromosome and not on the active one.

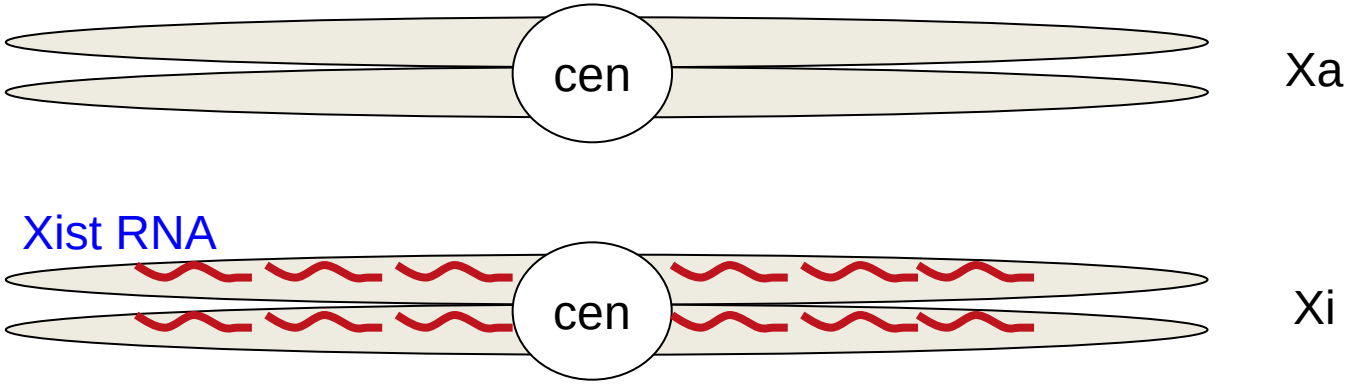
It is processed similarly to mRNAs, through splicing and polyadenylation, however, it remains untranslated.

The inactive X is coated with this transcript, which is essential for the inactivation.

X lacking Xist will not be inactivated, while duplication of the Xist gene on another chromosome causes inactivation of that chromosome

XIST is the only gene expressed from the inactive X (X_i) and not expressed from the active X (X_a)

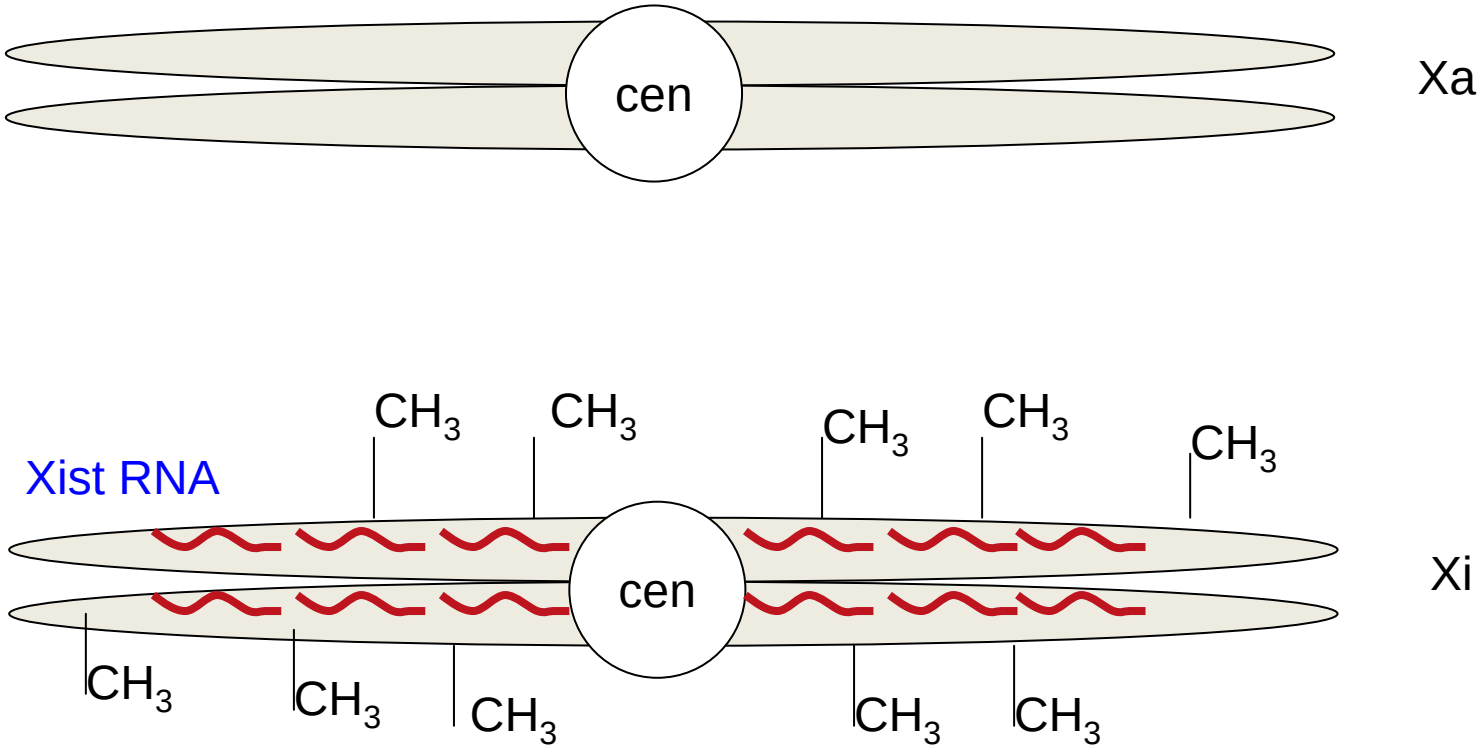
Initially, both X weakly express *XIST*, then on the future X_i, XIST expression is increased while on future X_a, XIST expression is repressed

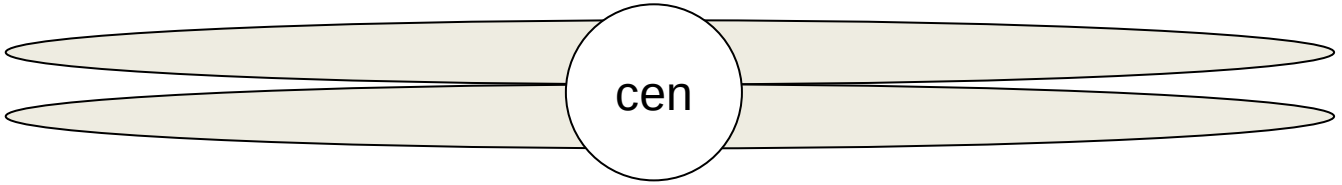


XIST coats the X_i chromosome, moving out from the XIC

The activity of *TSIX* is reciprocal to that of *XIST*, expressed by X_a suppresses *XIST*

Alleles bearing a deletion of TSIX are much more likely to be inactivated.


cen
Xa
CH3
CH3
CH3
CH3
CH3
Xist RNA
cen
Xi
CH3
CH3
CH3
CH3
CH3

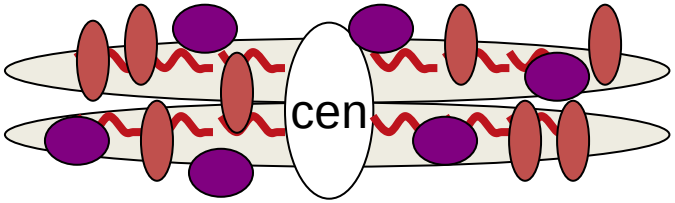


Xa

High levels of histone H2A
High levels of histone H3 methylation

H2A

H3



Xi

Hall,L.L., Lawrence,J.B., (2003), The Cell Biology of a Novel Chromosomal RNA: Chromosome Painting By XIST/Xist RNA Initiates a remodeling cascade. j.semcdb 14, 369-378

Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C. D. (2001), Methylation of Histone H3 at Lys-9 Is an Early Mark on the X Chromosome during X Inactivation. Cell 107, 727-738.

Lee, J. T., Davidow, L. S., Warshawsky, D., (1999), Tisx, a gene antisense to Xist at the X-inactivation centre. Nat. Genet. 21, 400-404.

Understanding epigenetic modifications during early embryonic development has several implications in:

Nuclear cloning

Differentiation of embryonic stem cells to specific cell types

Inducible pluripotent stem cells (iPS) cells

Epigenetics, development and nuclear cloning

Nuclear cloning

Enucleated mature oocyte & introduce a somatic or embryonic stem cell nucleus

Activate development with electric current

Implant embryo into surrogate mother

Relatively low success rate; many births with developmental abnormalities

Donor chromatin must be remodeled in the oocyte

Developmental abnormalities look like parent-of-origin imprinting errors suggesting the imprints, which should remain in place, are improperly altered / remodeled in the mature oocyte

Reik et al. 2001. Science 293: 1089
Rideout et al. 2001. Science 293:1093)

Transient/short term gene silencing

Developmental genes that are needed during the later stages of development are transiently held in a repressed state during early development. This is achieved through short-term epigenetic marks such as histone modifications, which can be removed before or within a few cell divisions.

Permanent/ long term gene silencing

On the contrary, certain other regions of the genome are marked with epigenetic information that is stably maintained and heritable after many cell divisions. For example, imprinted genes, transposons and the inactive X chromosome require long-term silencing that is sustained throughout the development and lifespan of an organism. This is generally achieved by DNA methylation, an epigenetic mark that refers to the addition of a methyl group to the fifth carbon of base C. Because DNA methylation provides heritable, long-term silencing that is crucial for an organism,

Active DNA demethylation: many roads lead to Rome

Susan C. Wu and Yi Zhang

Nature Reviews Molecular Cell Biology |
AOP, published online 4 August 2010