L2.3

Monoallelic expression

Imprinting

X-chromosome inactivation

Trends in Genetics

Textbook

CellPress
REVIEWS

Special Issue: Gene Expression in Time and Space

Review

Parallels between Mammalian Mechanisms of Monoallelic Gene Expression

Ahmed Amine Khamlichi[1,*] and Robert Feil[2,*]

Different types of monoallelic gene expression are present in mammals, some of which are highly flexible, whereas others are more rigid. These include allelic exclusion at antigen receptor loci, the expression of olfactory receptor genes, genomic imprinting, X-chromosome inactivation, and random monoallelic expression (MAE). Although these processes play diverse biological roles, and arose through different selective pressures, the underlying epigenetic mechanisms show striking resemblances. Regulatory transcriptional events are important in all systems, particularly in the specification of MAE. Combined with comparative studies between species, this suggests that the different MAE systems found in mammals may have evolved from analogous ancestral processes.

Highlights

MAE systems have diverse biological roles, but show similarities in the underlying epigenetic mechanisms.

Transcriptional events and non-coding RNAs are associated with the initiation of MAE.

Differential modification and architectures of chromatin are also important in MAE.

The different MAE systems found in mammals may have evolved from analogous ancestral processes.

 https://doi.org/10.1016/j.tig.2018.08.005

MAE = Mono Allelic Expression RMAE = Random MAE

Table 1. MAE in the Mouse

Genetic loci	Chromosome (Chr.); gene numbers	Nature of monoallelic expression	Mechanism(s) involved	Biological function(s)
Immunoglobulin gene (Ig) loci: *IgH* *IgL: Igκ*, *Igλ*	*IgH*: Chr. 12 *Igκ*: Chr. 6 *Igλ*: Chr. 22	Random, in B cells; $>10^{10}$ possible different immunoglobulins	Genetic recombination, chromatin alterations, ncRNA	Immune system: expression of one heavy chain and one light chain in each B cell
T cell receptor (TCR) loci: *TCRβ*, *TCRα*	*TCRβ*: Chr. 6 *TCRα*: Chr. 14	Random, in T cells; $>10^{8}$ possible different TCRs	Genetic recombination, chromatin alterations, ncRNA transcription	Immune system: expression of one β and one α polypeptide in each αβ T cell
Olfactory receptor *(OR)* genes	>1400 genes Many gene clusters	Random	Epigenetic, predominantly chromatin modifications	Olfaction: expression of one receptor per olfactory neuron
Protocadherin genes	58 genes Three autosomal gene clusters	Random, tissue-specific	Epigenetic, promoter choice and alternative splicing	Neurons: cell-surface diversity, signaling, neuronal survival
RMAE of unique genes	>1000 genes Autosomal and X chromosome	Random, often tissue-specific	Epigenetic, histone methylation	Not clear: may provide diversity in cell identity
Genomic imprinting at autosomal genes	~120 protein- coding genes Hundreds of regulatory ncRNAs	Deterministic, often tissue-specific	Epigenetic, DNA methylation, chromatin modifications, lncRNAs	Not clear: genes involved in development, metabolism, and behavior
X-chromosome inactivation in females	X-linked genes (~2000)	Random, deterministic (imprinted); early embryo and trophoblast	Epigenetic, lncRNAs, chromatin modifications, DNA methylation	Compensates for differential dosage of gene expression between females and males

Khamlichi & Feil, 2018

In mammals, many genes show monoallelic expression (MAE)

- ✓ X Chromosome inactivation (early embryo) (ca. 2,000 genes)
- ✓ Genomic imprinting (ca. 120 protein-coding genes + ncRNAs)
- ✓ Antigen Receptor (Ig in B cells, TCR in T cells)
- ✓ Olfactory receptor (> 1,400 genes)
- ✓ Individual genes showing random MAE (> 1,000 genes)

Technical:

how can we know which allele is expressed ?

Crossing two mouse strains we will have quite variant genomes

Single gene: allele-specific PCR (3’-end of primers on the SNP)

RNA-Seq: map reads on the two variant genomes

Bibliography

Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control

Jeannie T. Lee

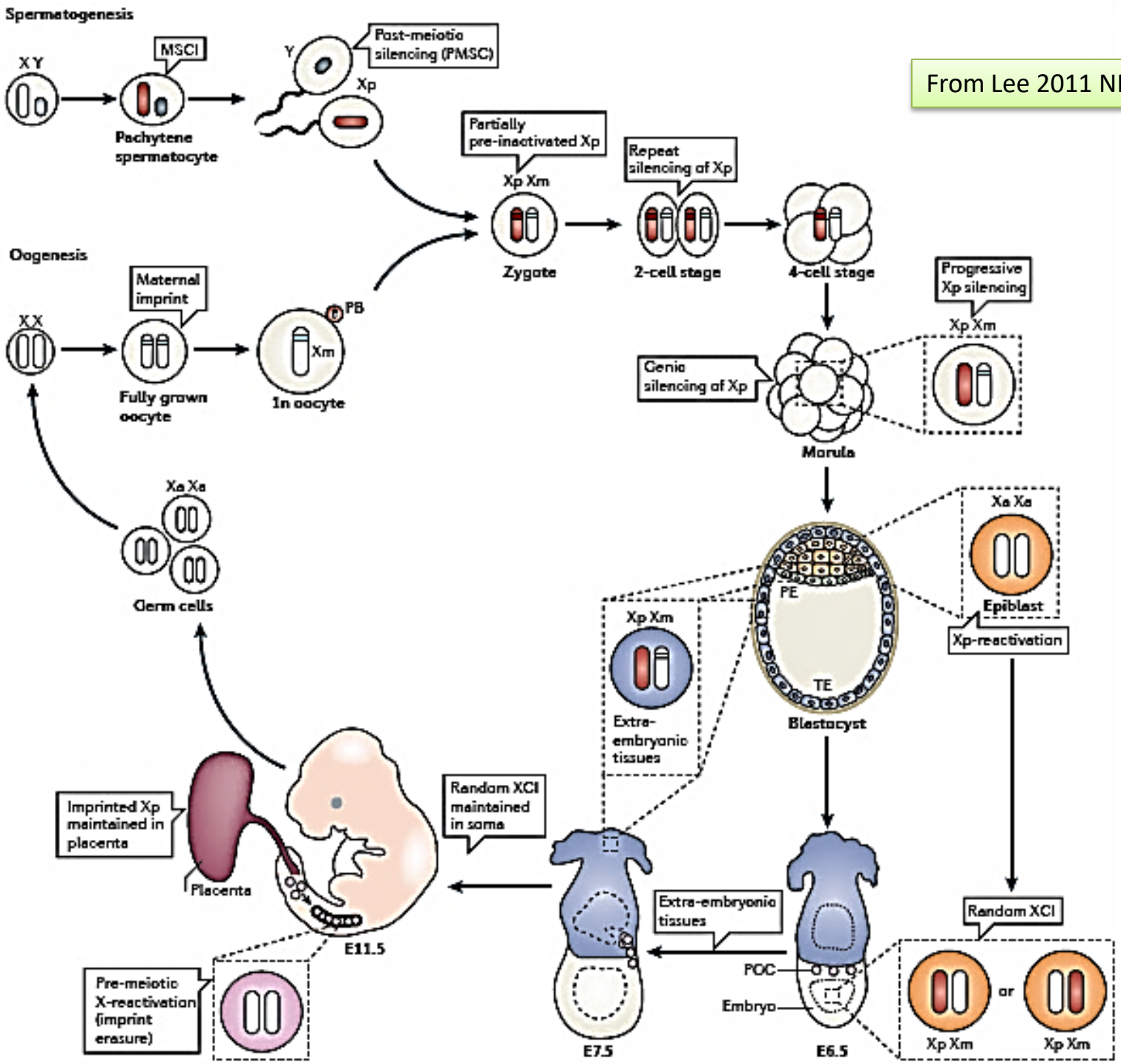
Abstract | The discovery of X-chromosome inactivation (XCI) celebrated its golden anniversary this year. Originally offered as an explanation for the establishment of genetic equality between males and females, 50 years on, XCI presents more than a curious gender-based phenomenon that causes silencing of sex chromosomes. How have the mysteries of XCI unfolded? And what general lessons can be extracted? Several of the cell biological mechanisms that are used to establish the inactive X chromosome, including regulatory networks of non-coding RNAs and unusual nuclear dynamics, are now suspected to hold true for processes occurring on a genome-wide scale.

Dosage compensation

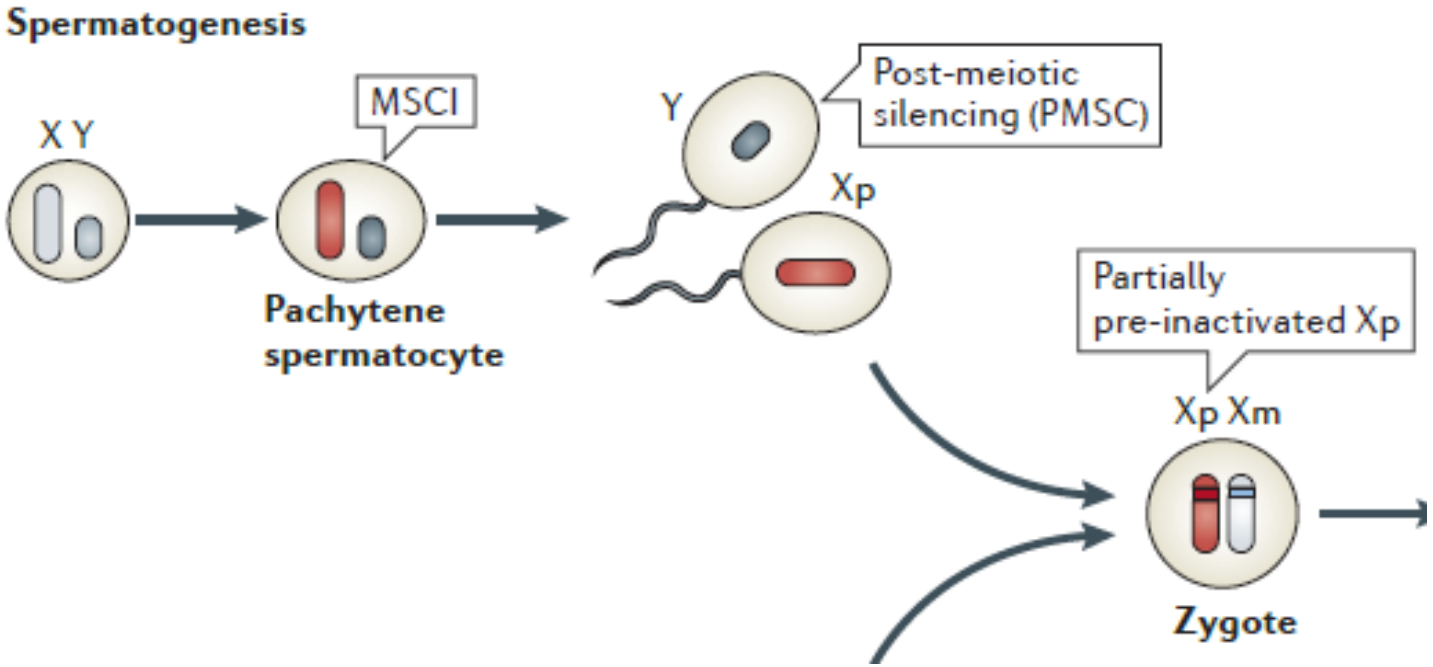
Sex chromosome dosage compensation is a well-established whole chromosome model in which parental-origin effects are studied.
In mammals, females with two X chromosomes achieve parity with males in their X-linked gene dosage through epigenetic inactivation of one X chromosome.

In marsupials, X-chromosome inactivation is imprinted, with the **paternally** inherited X chromosome being **inactive** in somatic cells.

In mice, X-chromosome inactivation is similarly imprinted, but specifically during pre-implantation stages and in extra-embryonic lineages including the placenta.
After that, random inactivation initiates in all embryonic components around the time of implantation.

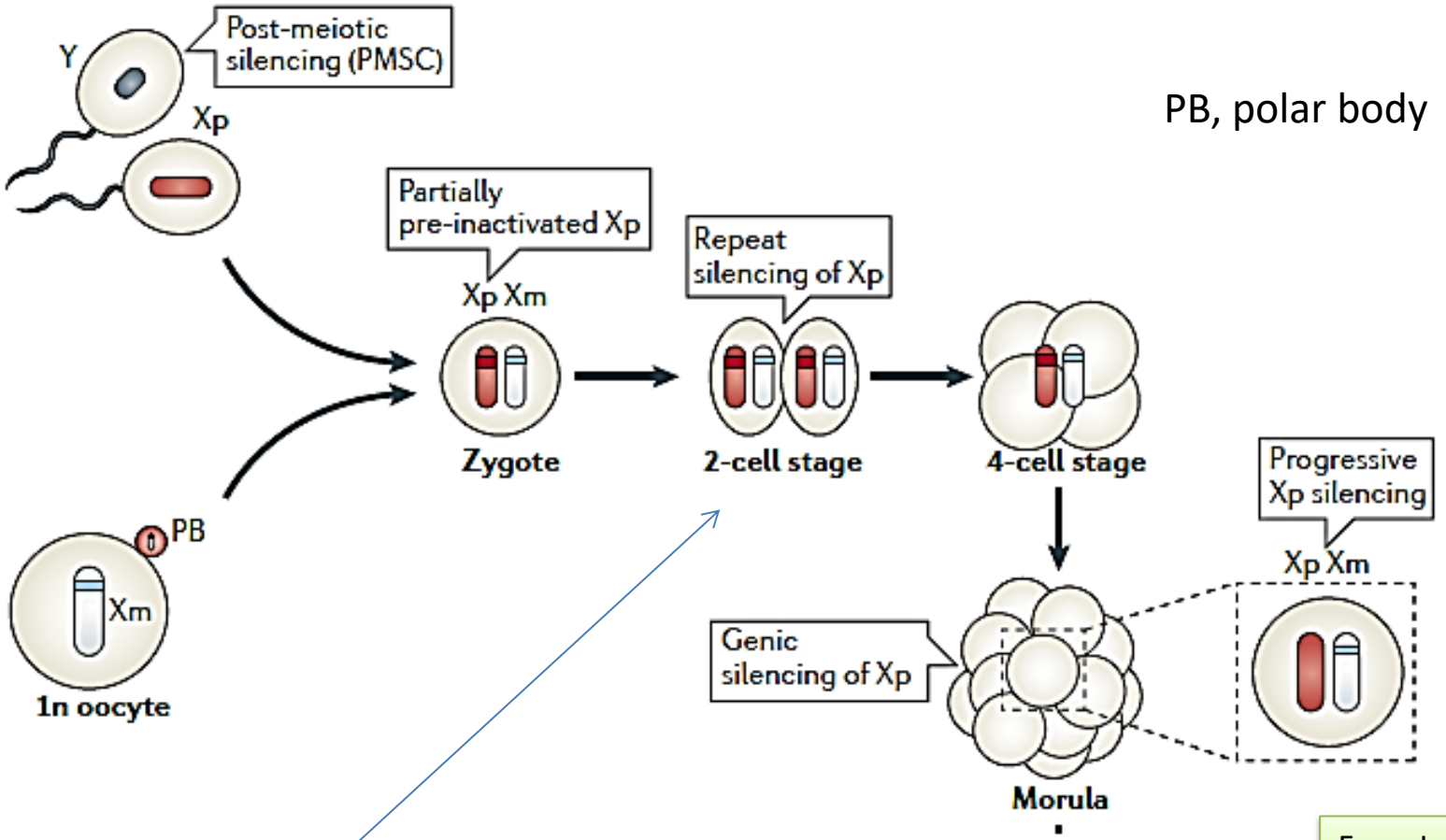


From Lee 2011 NRMCB



The X-chromosome inactivation (XCI) cycle begins during the first meiotic prophase of spermatogenesis, when non-homologous regions of the X and Y chromosomes undergo ‘meiotic sex chromosome inactivation’ (MSCI); the majority of the ‘post-meiotic sex chromatin’ (PMSC) remains transcriptionally suppressed throughout spermiogenesis

From Lee 2011 NRMCB



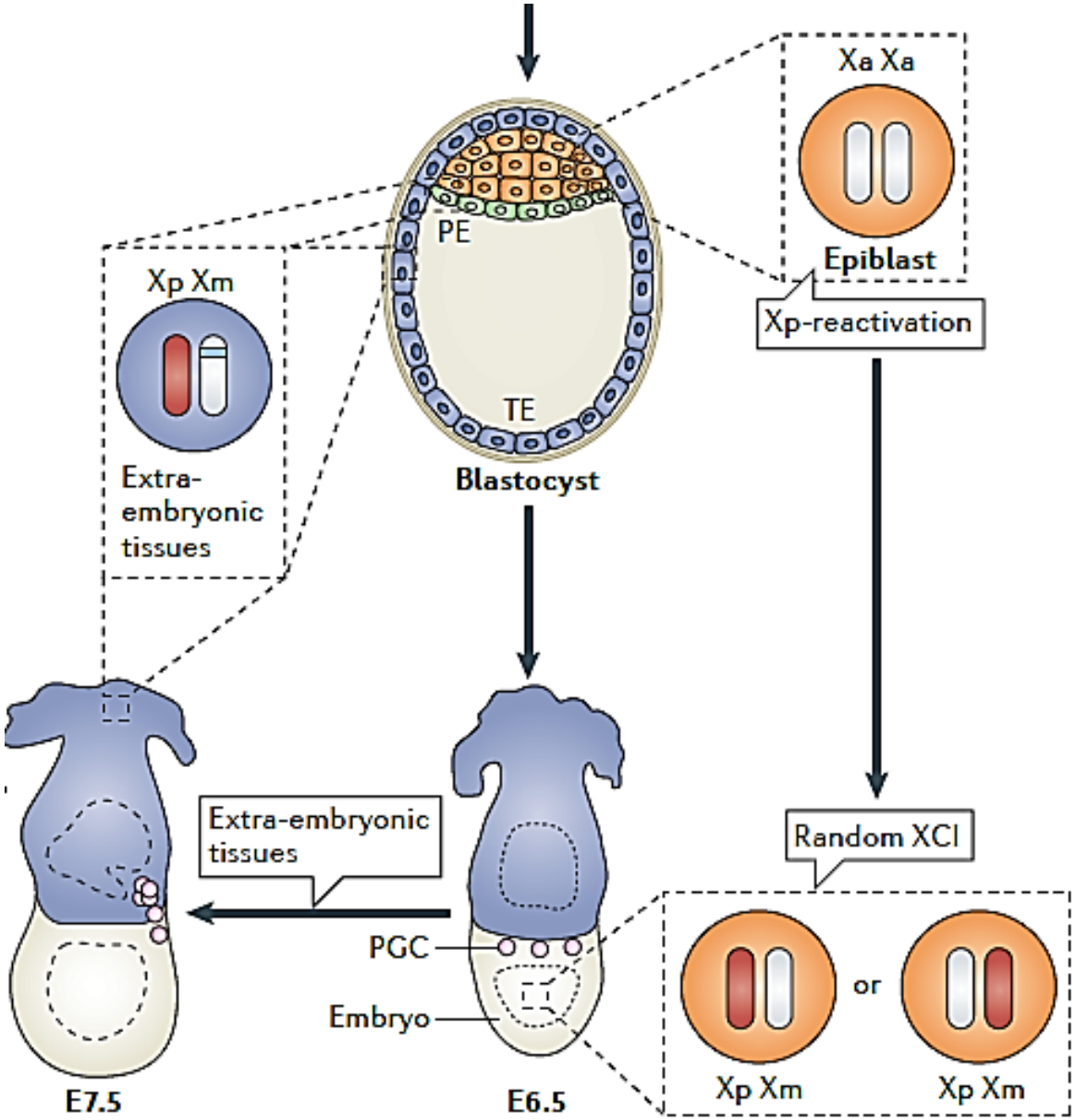
From Lee 2011 NRMCB

When zygotic gene activation takes place at the two-cell stage, repetitive elements of the paternal X chromosome (Xp) (but not of the maternal X chromosome (Xm)) are already inactive, mirroring their inactive state in the male germline. Although X-linked coding genes on the Xp are initially active, they are progressively inactivated during preimplantation development. This completes the 'imprinted' form of XCI, and by the time the early blasto-cyst has formed (around embryonic day 3.5 (E3.5)), every cell of the embryo displays a silent Xp

After E3.5 of peri-implantation development, the epigenetic fates of embryonic and extra-embryonic tissue lineages **diverge**.

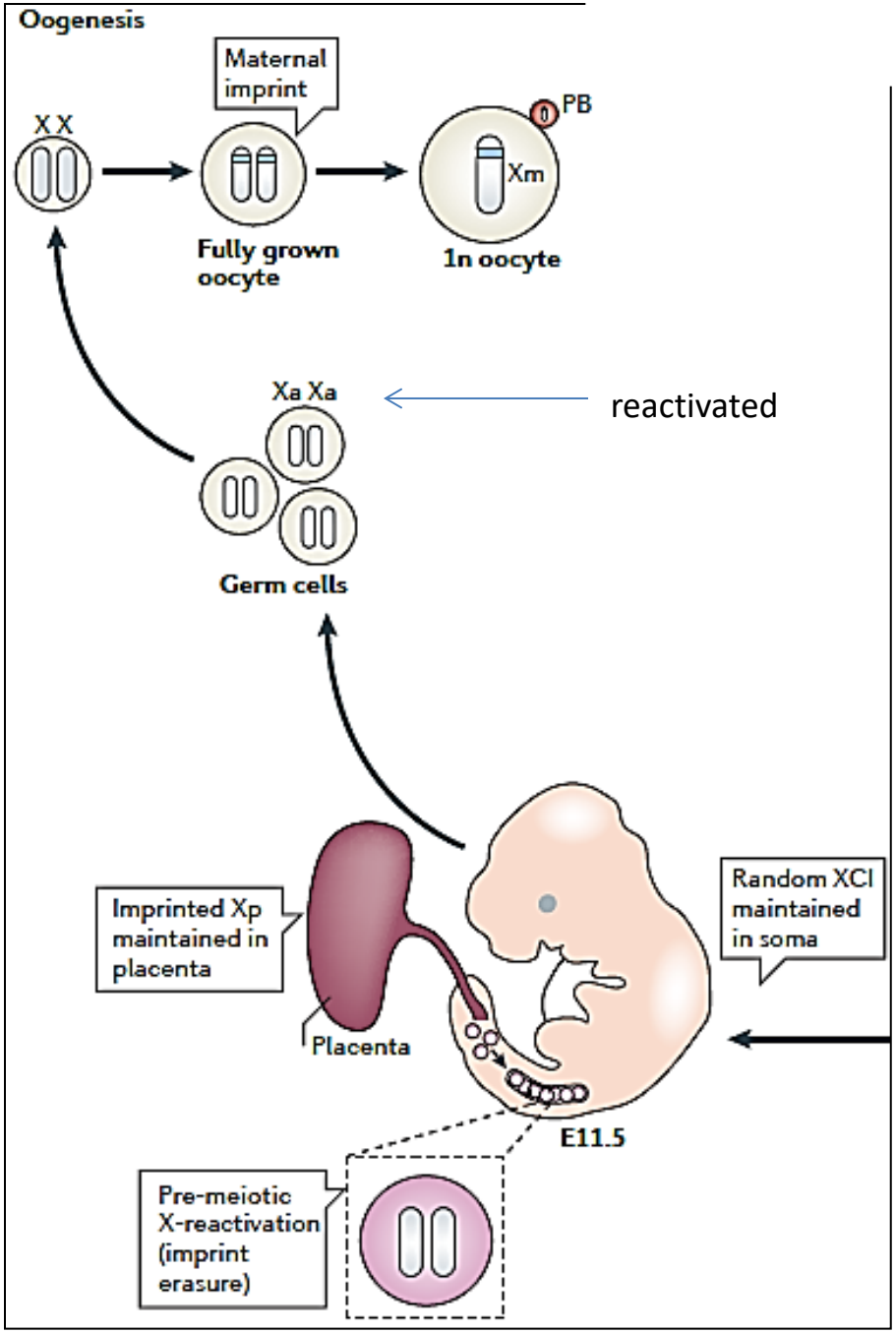
Whereas extra-embryonic tissues, including the primitive endoderm (PE) and the trophectoderm (TE), maintain the imprinted state of Xp inactivation,

the **epiblast lineage** reactivates the Xp and undergoes a new round of 'random' XCI without a parent-of-origin bias. This random XCI in the epiblast is tightly linked to differentiation into the three germ layers and is mediated by the X-inactivation centre (*Xic*).



Xa, active X chromosome

From Lee 2011 NRMCB



The epiblast gives rise to somatic cells of the embryo, and thus this randomly inactivated X chromosome is maintained throughout embryogenesis. The epiblast is also the lineage that gives rise to the primordial germ cells (PGCs).

To ensure equal meiotic segregation of the Xm and the Xp, a second round of X-chromosome reactivation takes place in **PGCs**. During the oocyte growth phase, a maternal germline mark is then imprinted upon both X chromosomes, which enables this X chromosome (the future Xm) to resist imprinted XCI in the next generation as the XCI cycle begins anew.

From Lee 2011 NRMCB

We observe two phenomena:

1. Imprinted inactivation of Xp

2. Random epigenetic inactivation of one X

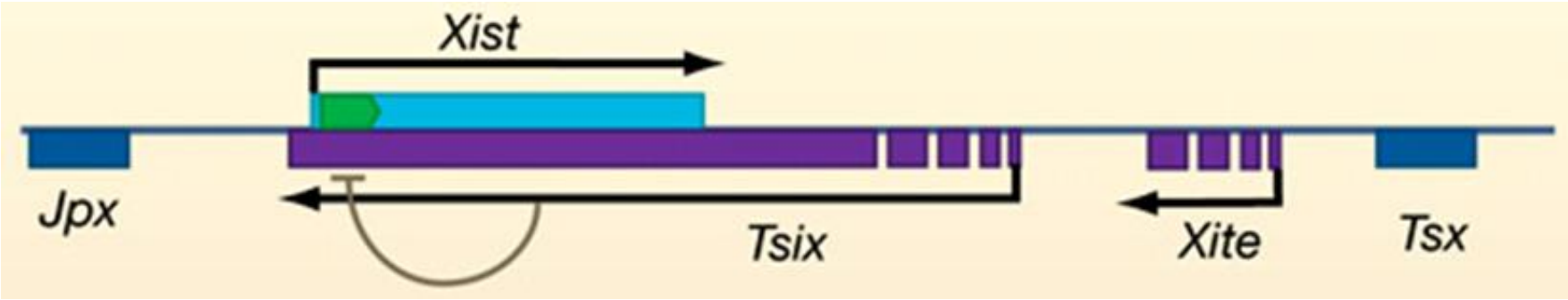
This follows the normal rules of imprinting, i.e. widespread CpG methylation throughout the Xp

This depends on random choice of the Xic activity on one X-chromosome

Xic= X-inactivation center

The best-characterized long noncoding transcripts at the Xist locus (mouse)

Xist RNA is transcribed from the inactive X chromosome



(XIST)

http://genome-euro.ucsc.edu/cgi-bin/hgTracks?db=hg19&position=chrX%3A73008383-73104691&hgsid=207585419_UIxMxKvgGrVLMst0Vb2nziEpz7Pj

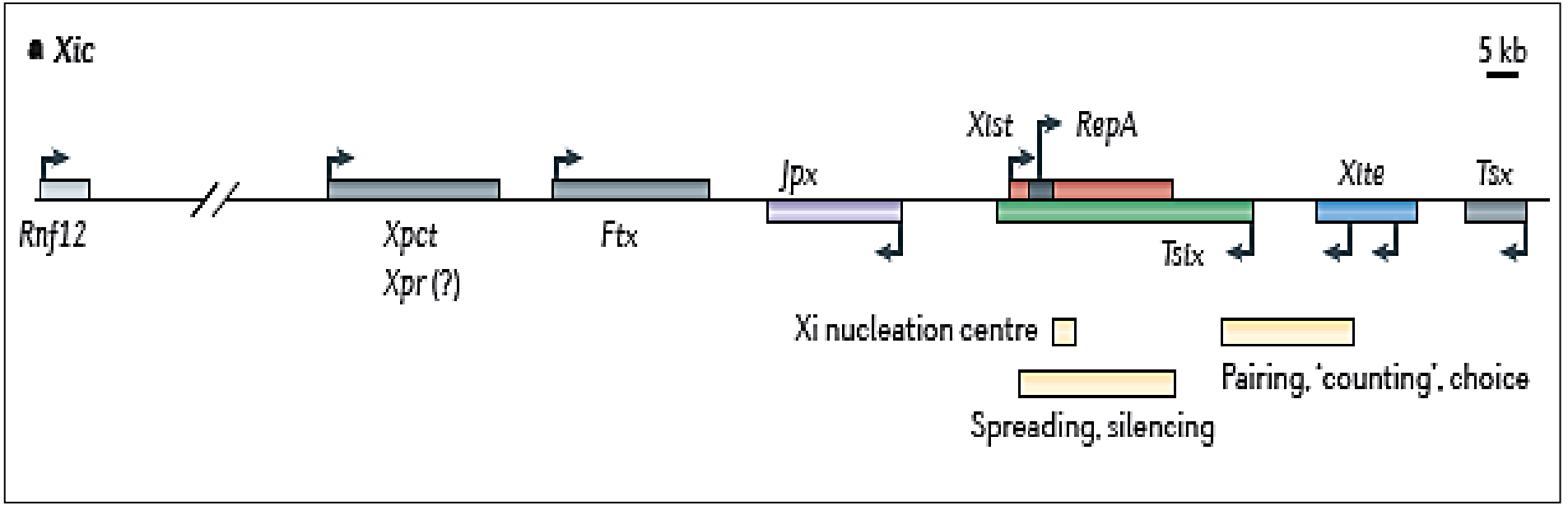


Figure 2 | **Non-coding genes of the *Xic*.**
a | The X-inactivation centre (*Xic*) and surrounding regions. Non-coding genes with established functions in regulating X-chromosome inactivation (XCI) are shown in colour, whereas those in grey have been proposed as candidate regulators. The role of the X-pairing region (*Xpr*) in pairing is under debate. RING finger 12 (*Rnf12*) is a coding gene located ~500 kb upstream of X-inactivation specific transcript (*Xist*). Yellow bars indicate regions involved in the different steps of XCI.

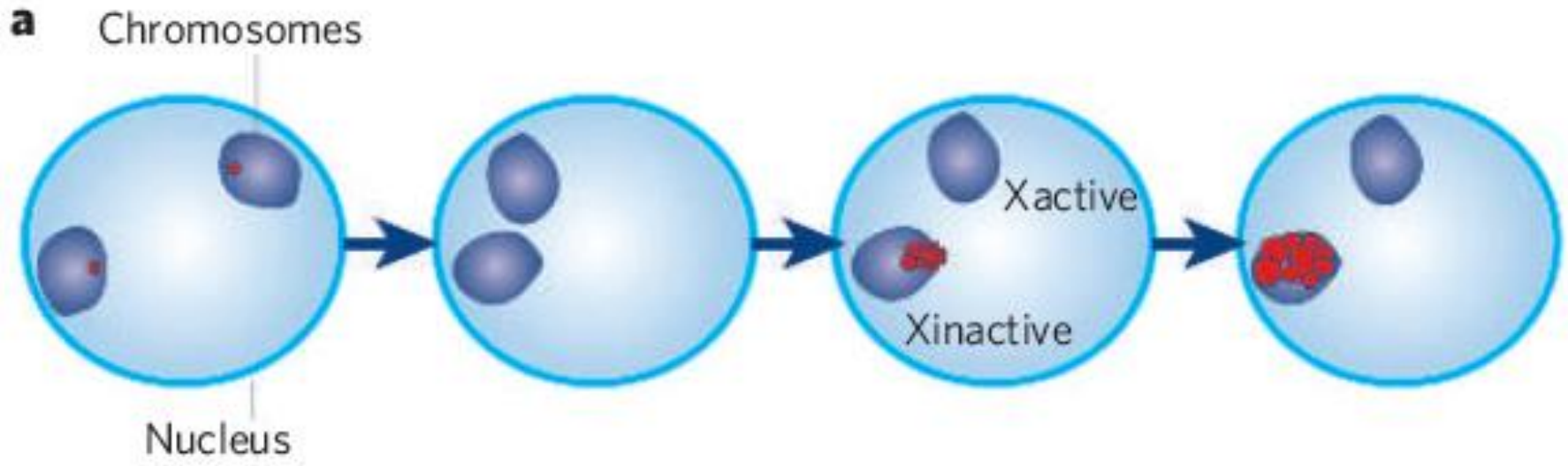


Figure 1 | Events of nuclear reorganization during X-chromosome inactivation. a, Soon after female embryonic stem cells start to differentiate, the two X chromosomes (purple) come together in the nucleus, and the X-inactivation centres, which initiate X-chromosome inactivation, interact.
These events occur concomitantly with the process of X-chromosome counting and choice and lead to upregulation of *Xist* transcription (red) from the future inactive X chromosome (Xinactive).

Fraser & Bickmore 2007

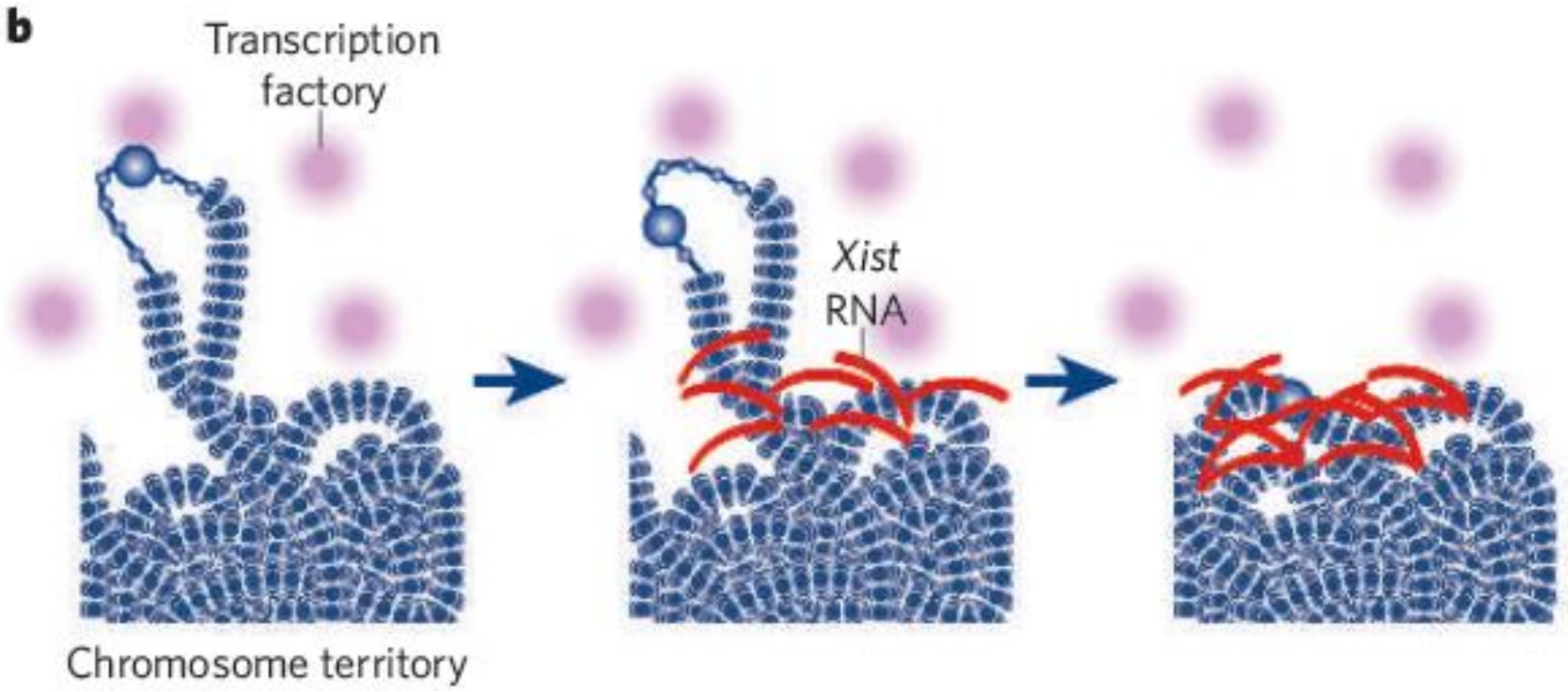
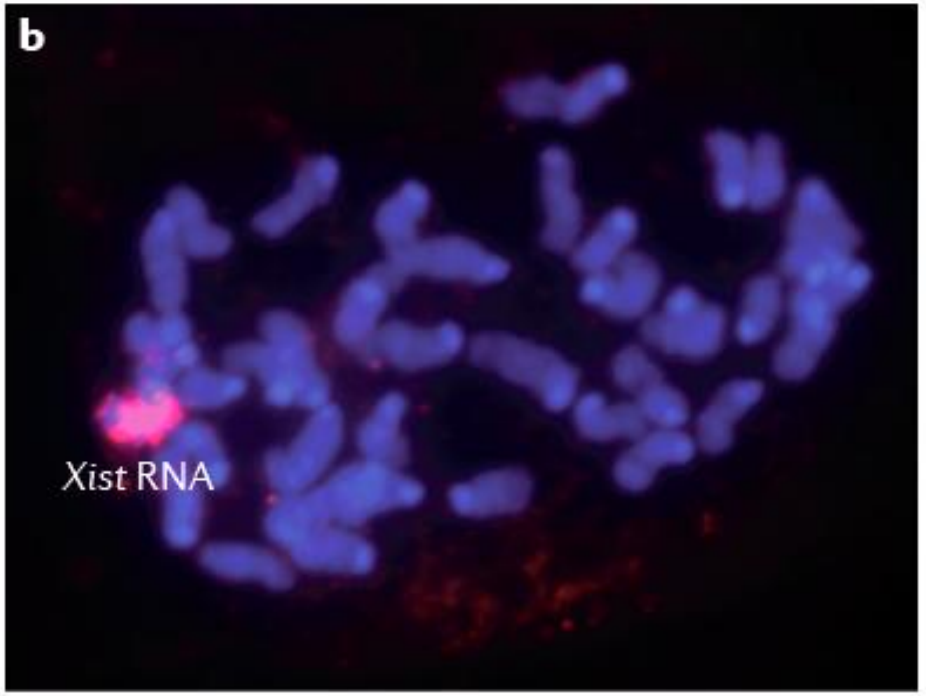
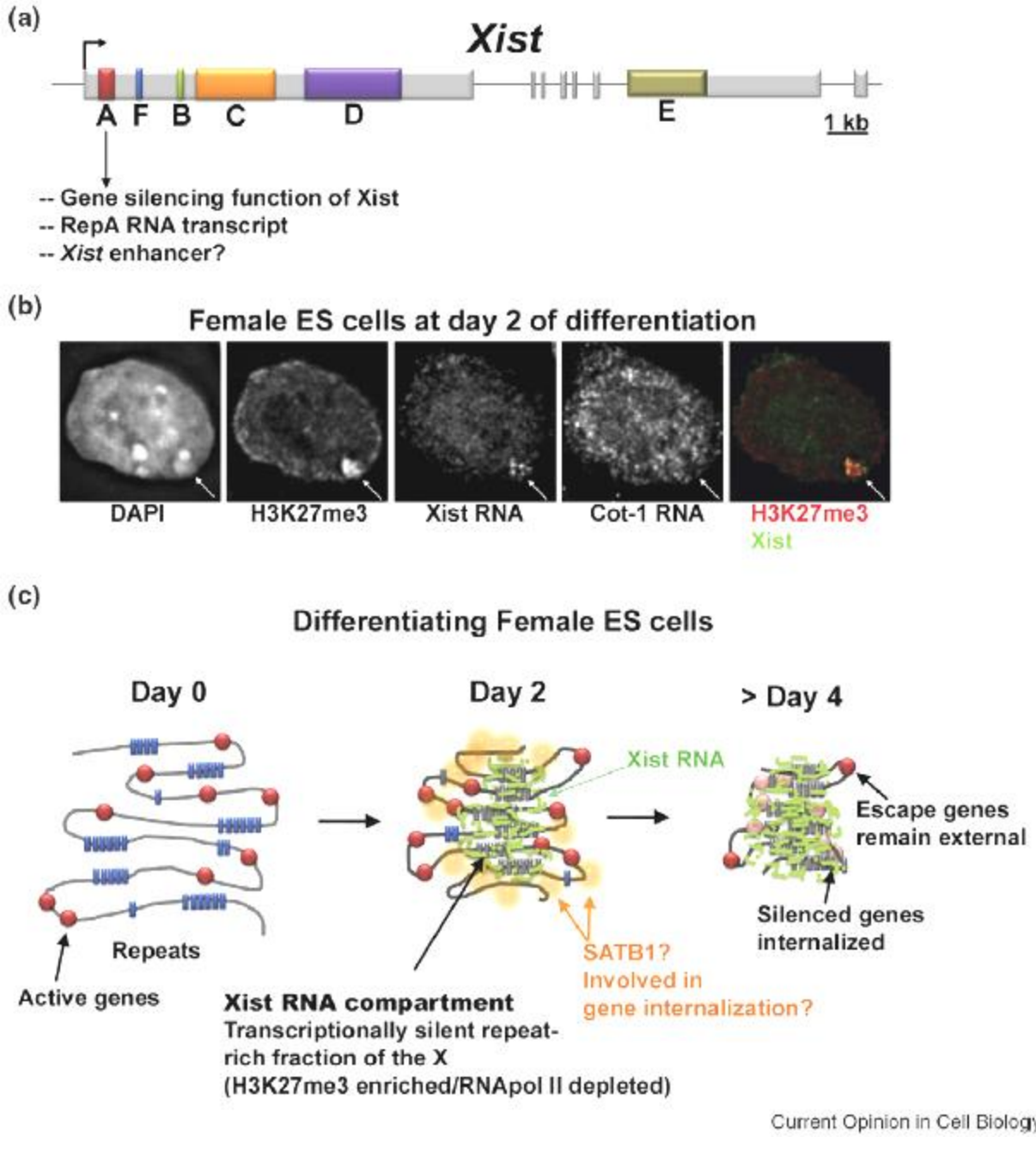


Figure 1 | Events of nuclear reorganization during X-chromosome inactivation.
b, The coating of the inactive X chromosome by *Xist* RNA molecules excludes Pol II and the transcriptional machinery (pink) from the inactive X-chromosome territory. Genes initially located outside the domain (purple circles) coated by *Xist* RNA are retracted back inside the *Xist* compartment as they become silenced through a mechanism dependent on the A repeats of *Xist* RNA.

Fraser & Bickmore 2007

b | *Xist* RNA coats the inactive X chromosome (Xi) on a metaphase spread of a female mouse embryonic fibroblast, as shown by RNA fluorescence *in situ* hybridization (*Xist* RNA, red; mouse chromosomes, blue).

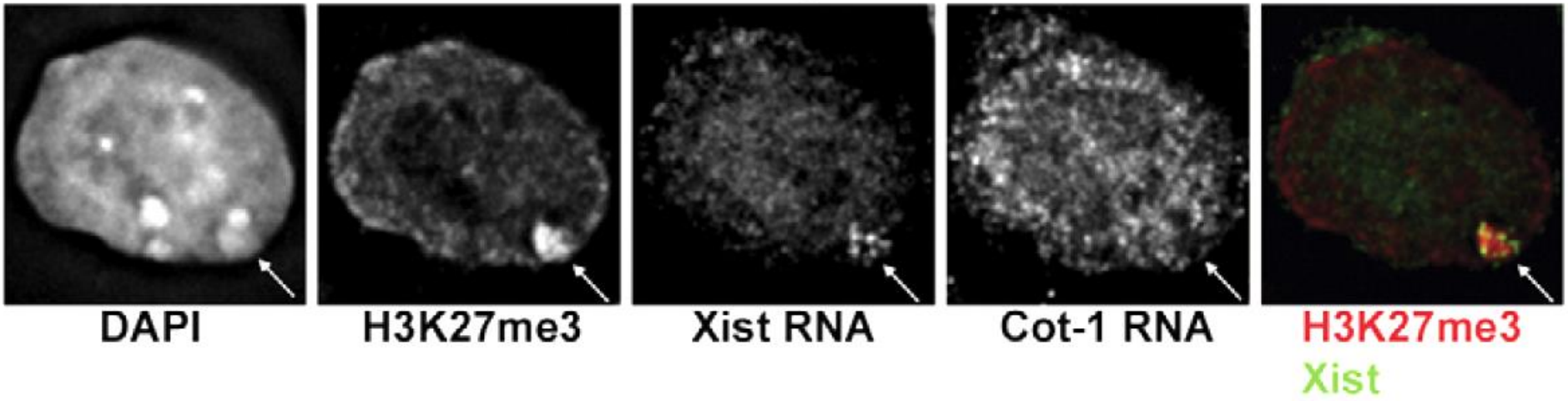




(a) Structure of the Xist gene with the conserved repeat regions labeled A–F. The A region (red) denotes the conserved A-repeat region essential for gene silencing.
(b) Combined RNA-immuno-fluorescence on day 2 differentiated female ES cells, showing the Xist-coated transcriptionally silent compartment which is enriched for H3K27me3.
(c) Model. Xist coating induces the formation of a transcriptionally silent repetitive compartment. As genes are silenced they are recruited into this compartment. A possible mediator for this internalization may be the matrix-associated protein SATB1/2.

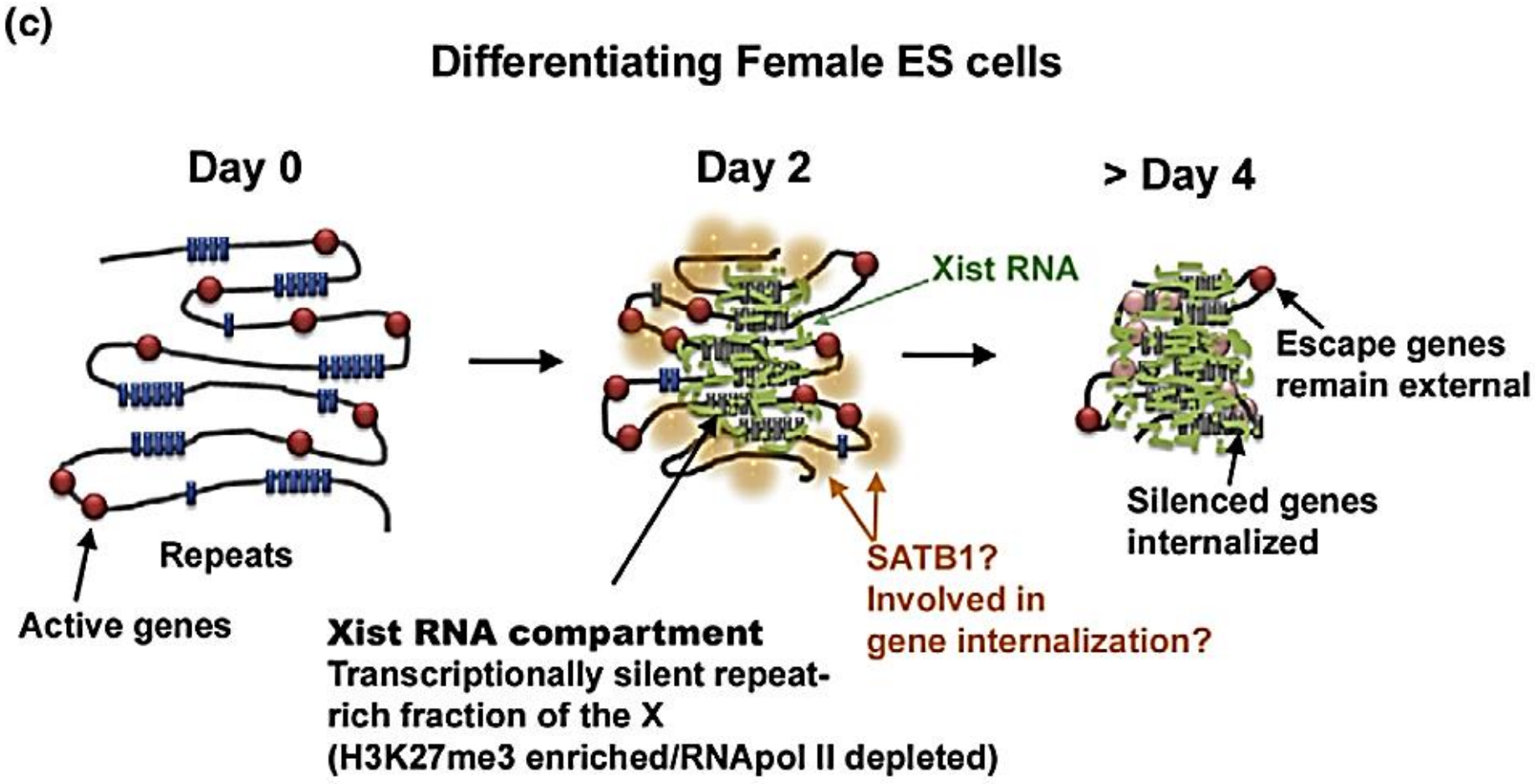
from Chow, Current Opinion in Cell Biology 2009, 21:359–366

Female ES cells at day 2 of differentiation



Combined RNA-immuno-fluorescence on day 2 differentiated female ES cells, showing the Xist-coated transcriptionally silent compartment which is enriched for H3K27me3.

from Chow, Current Opinion in Cell Biology 2009, 21:359–366



Model. Xist coating induces the formation of a transcriptionally silent repetitive compartment. As genes are silenced they are recruited into this compartment. A possible mediator for this internalization may be the matrix-associated protein SATB1/2

from Chow, Current Opinion in Cell Biology 2009, 21:359–366

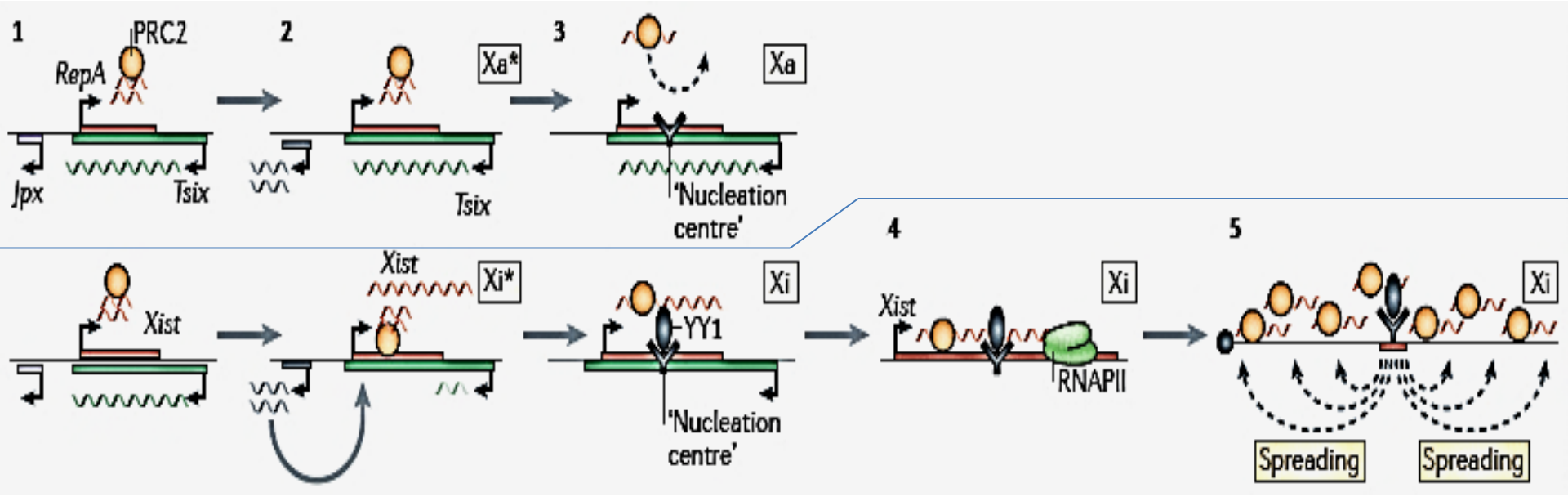
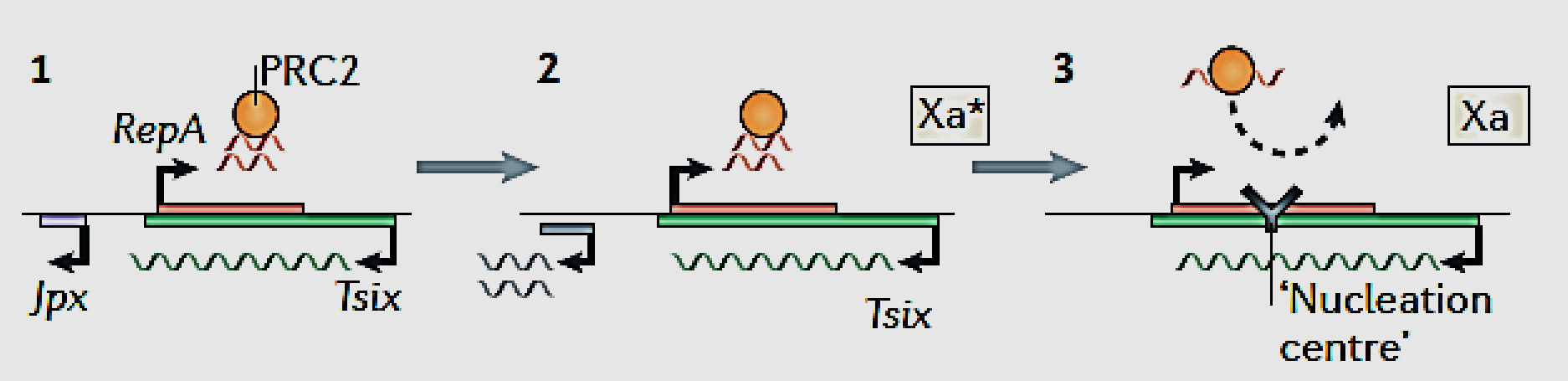


Figure 3 | **Molecular events at the initiation of XCI.** Co-transcriptional recruitment of **Polycomb repressive complex 2 (PRC2)** and tethering to YY1 explain the *cis*-acting nature of X-inactivation specific transcript (*Xist*) RNA. First, the biallelic *Xist* antisense gene (*Tsix*) prevents loading of repeat A RNA (*RepA*)–PRC2 and initiation of X-chromosome inactivation (XCI) (step 1). *Xist* expression is then permitted during cell differentiation by: first, induction of the *Jpx* activator; and second, monoallelic loss of *Tsix* on the future (denoted with an asterisk) inactive X chromosome (Xi) following X–X pairing, which allows *RepA*–PRC2 to load (step 2). *Xist* then co-transcriptionally recruits PRC2, and YY1 binds the Xi 'nucleation centre', but is blocked from binding the active X chromosome (Xa) (step 3). The *Xist*–PRC2 complex co-transcriptionally loads onto the YY1-based nucleation centre (step 4) and from here, *Xist*–PRC2 spreads in a *cis*-limited fashion to inactivate the X chromosome (step 5). RNAPII, RNA polymerase II.

At the X-active:

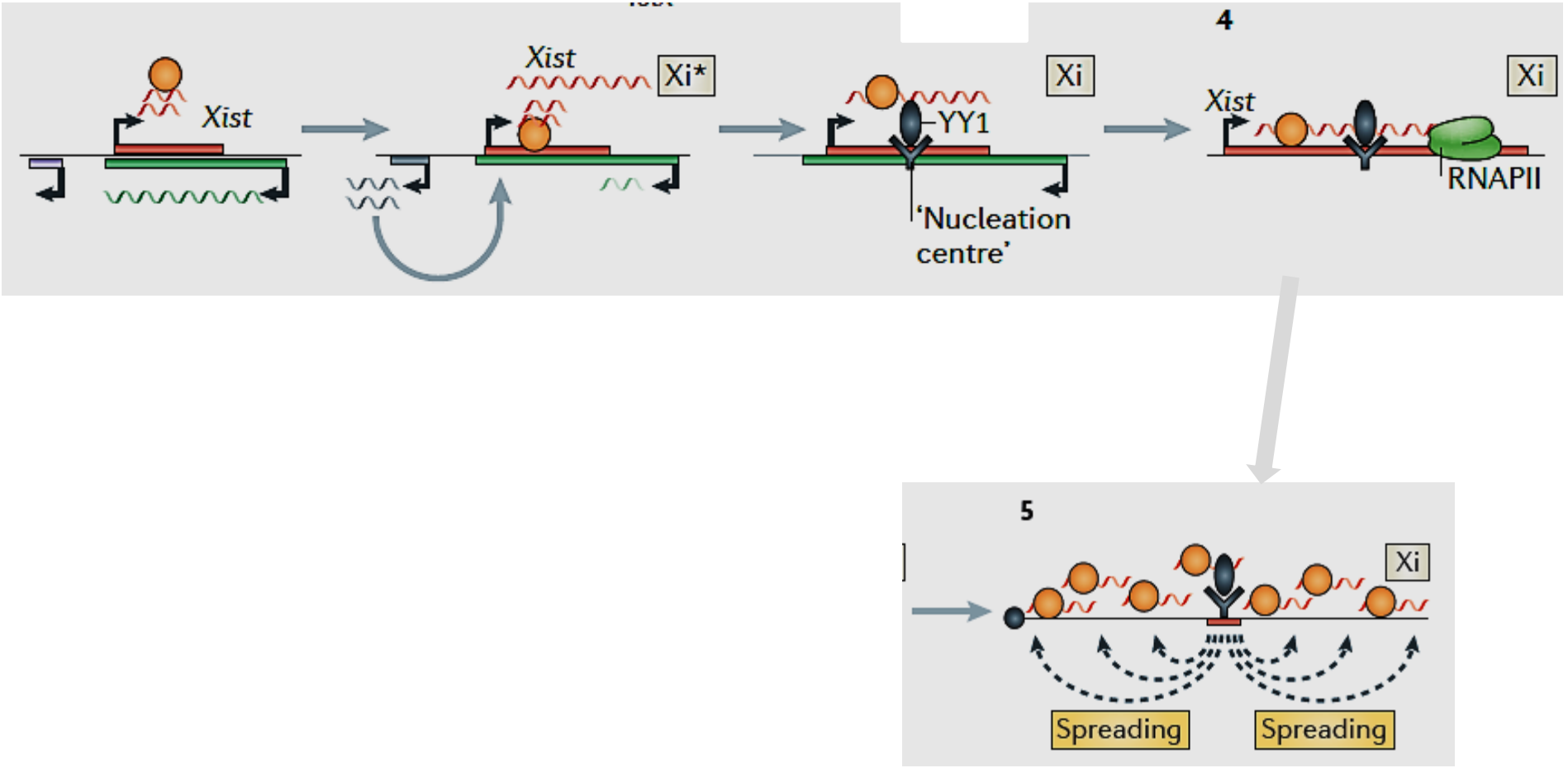


Antisense Tsix RNA is transcribed → only the small RepA part of Xist is transcribed

RepA RNA binds PRC2 (Polycomb Repressive Complex 2) and diffuse away.

Tsix antisense transcription prevents YY1 factor from binding the nucleation center. binding RepA does not contain YY1 site.

At X-inactive



Genomic Imprinting

Monoallelic expression of some genes in Mammals is not random, but depends upon **genomic imprinting** determined during oocyte or spermatocyte production

(*germinal epigenetic inheritance*, different from somatic epigenetic inheritance)

We say these genes are **imprinted**

and the phenomenon is known **as genomic imprinting**

Bibliography

TIMELINE

Genomic imprinting: the emergence of an epigenetic paradigm

Anne C. Ferguson-Smith

Abstract | The emerging awareness of the contribution of epigenetic processes to genome function in health and disease is underpinned by decades of research in model systems. In particular, many principles of the epigenetic control of genome function have been uncovered by studies of genomic imprinting. The phenomenon of genomic imprinting, which results in some genes being expressed in a parental-origin-specific manner, is essential for normal mammalian growth and development and exemplifies the regulatory influences of DNA methylation, chromatin structure and non-coding RNA. Setting seminal discoveries in this field alongside recent progress and remaining questions shows how the study of imprinting continues to enhance our understanding of the epigenetic control of genome function in other contexts.

Evidence of genomic imprinting

Early observations, particularly in insects and plants, indicated that the appearance of a particular visible trait in offspring could differ depending on whether it was transmitted from the mother or the father. In some of the early studies, imprinting effects were observed cytogenetically and, as such, were seen to affect whole chromosomes. However, genetic experiments suggested that parental-origin effects could also act at the level of the gene.

Whole chromosome effects. Historically, there have been several examples of parental-origin-specific 'marking' of whole chromosomes documented in the literature. Indeed, the term 'imprinting' was first coined by the cytogeneticist Helen Crouse[1] in 1960 to describe the programmed elimination of one or two paternally derived X chromosomes in sciarid flies. Sciarid zygotes inherit two paternally derived and

Note

Epigenetic modifications are established during gametogenesis, and are stable (not erased) during the zygotic clearance.

Parental imprinting: character that is passed form father or mother to the pups.

Trans-generational imprinting: this character was already present before germinal cells determination.

(character is passed on through more than one generation)

At the gene level, genomic imprinting was confirmed twenty-five years ago, as an outcome of nuclear transplantation experiments in the mouse.

About 150 genes are known to be imprinted in mice and humans, and imprinting is conserved in ruminant species as well.

Many imprinted genes are involved in fetal development and growth, and some influence behavior.

Recent methylome analysis in mice by Me-DIP or bisulfite-Seq has evidenced some more imprinted loci.

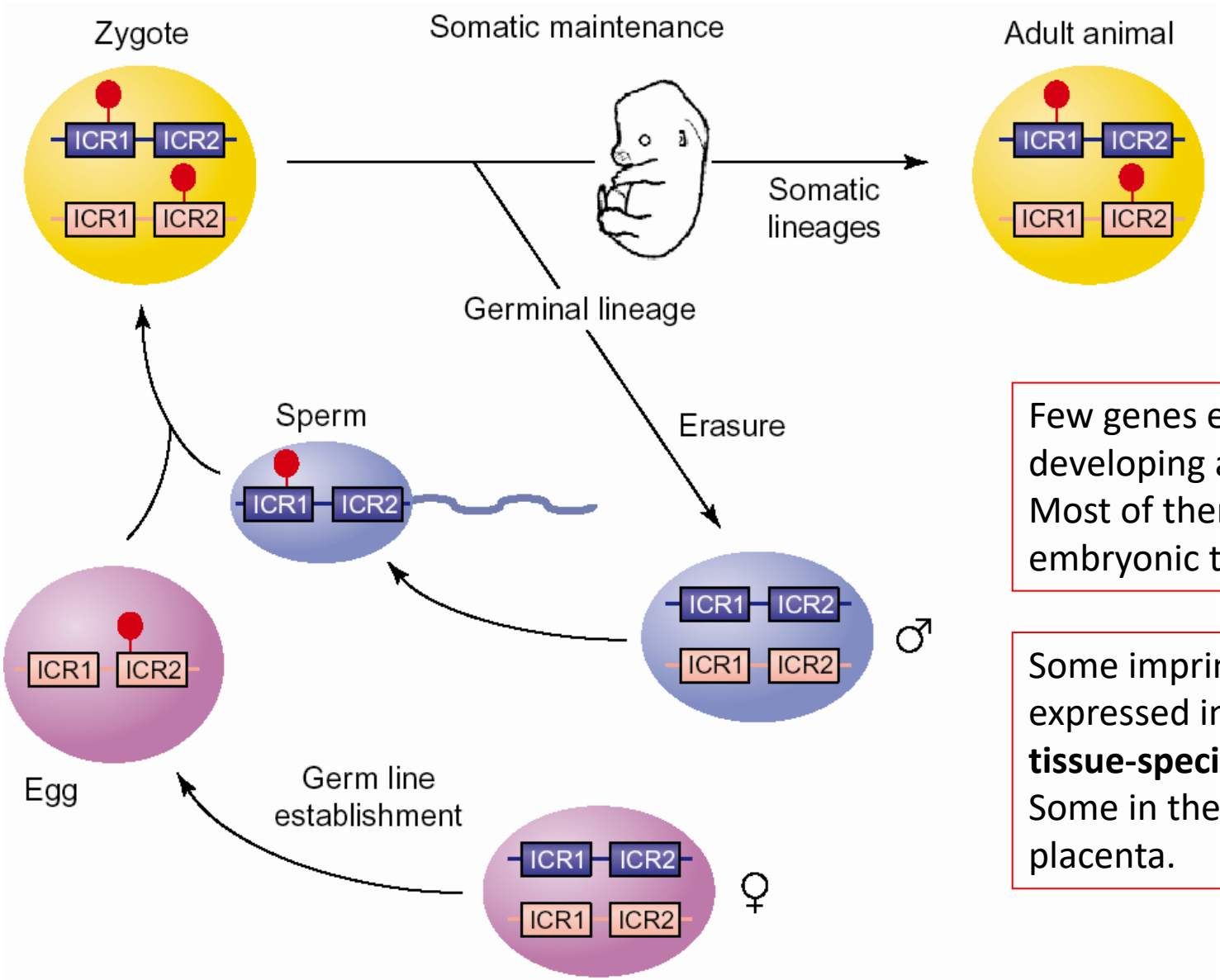
Imprinted genes

- usually clustered (grouped and co-regulated) Large Chr. Domains (LCD)
- They include a CG-rich region (island) called ICR - imprinting control region
- ICR is CpG hypermethylated in one allele (DMR=differentially methylated region).
- Methylation is done in germinal cells
- Each imprinted locus has at least one lncRNA

At most ICRs, the allelic methylation originates from the egg. At only a few, it is established during spermatogenesis.

Following fertilization, allelic methylation marks are maintained throughout development and they mediate imprinted expression.

Recently, an oocyte driven imprinting was reported, not depending on CpG methylation but on H3K27me3 on a specific region of the locus. This is inactivating, as in the case of X-chromosome.



Few genes expressed in developing and adult tissues
Most of them in extra-embryonic tissues

Some imprinted genes are expressed in a **cell- or tissue-specific** fashion.
Some in the brain, many in placenta.

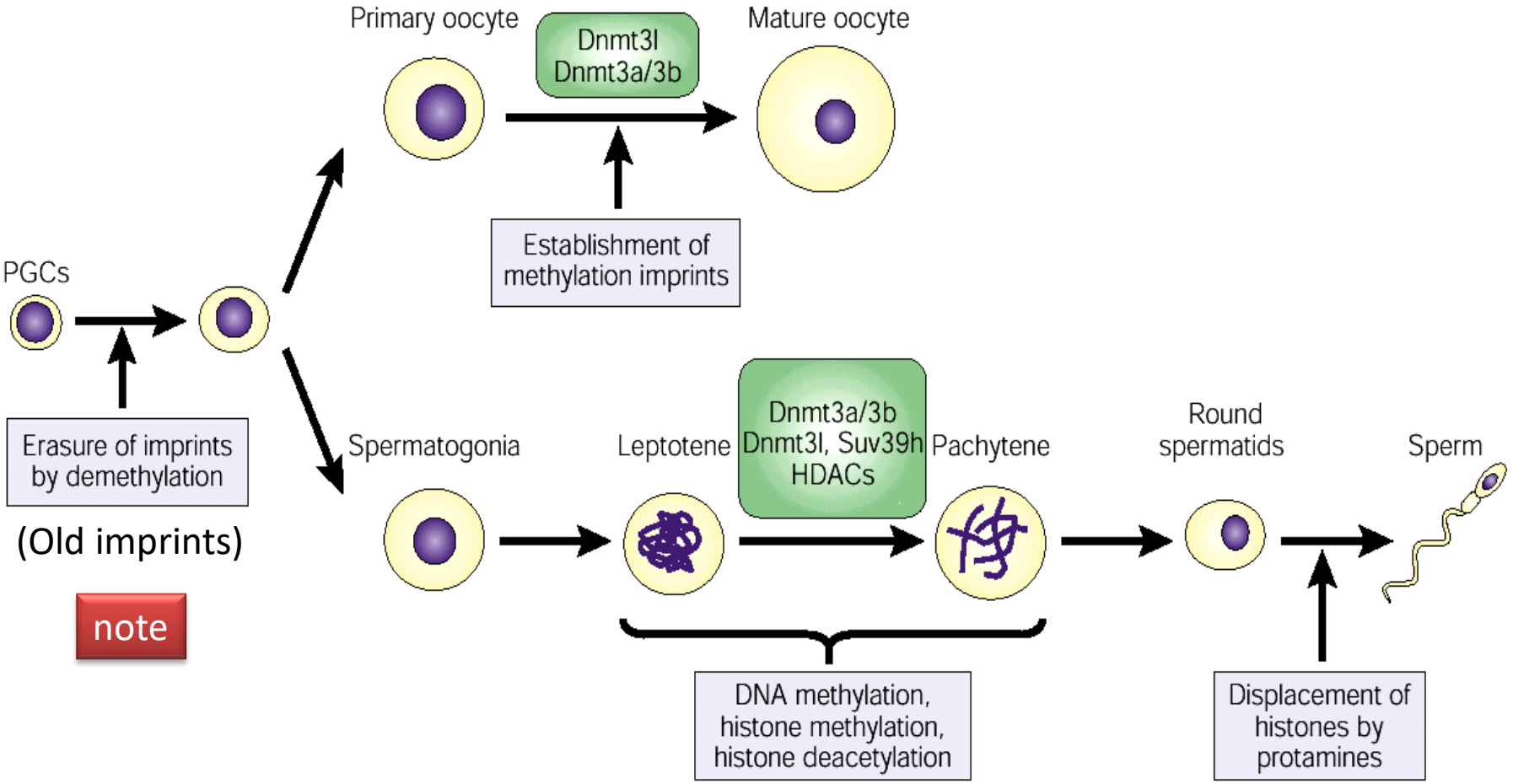
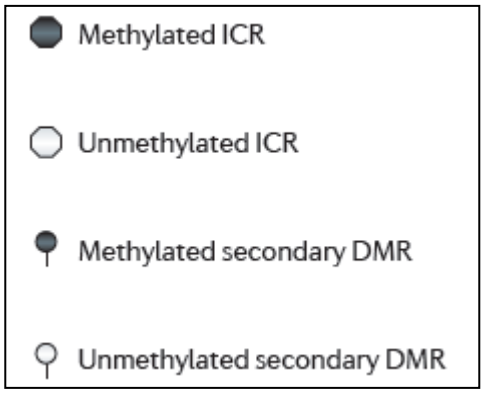
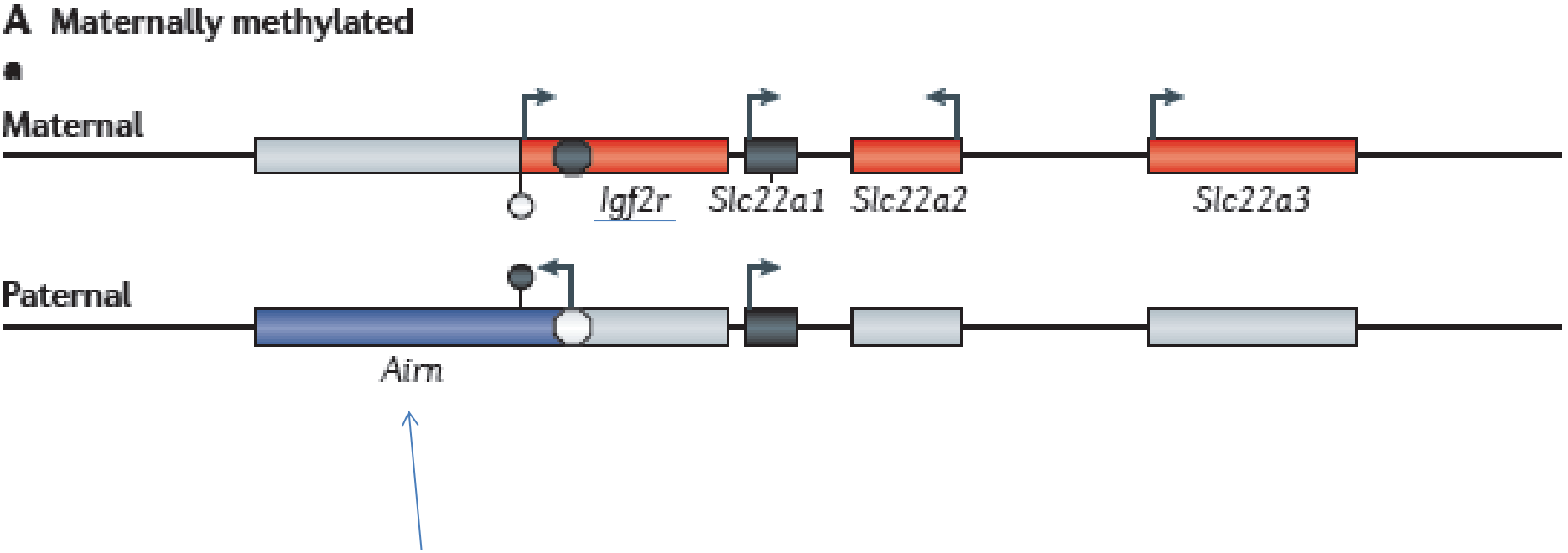


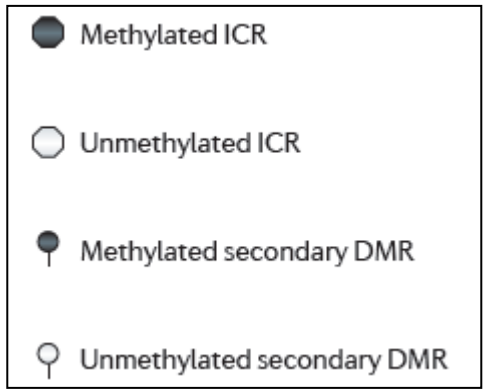
Figure 2 | **Epigenetic reprogramming during gametogenesis.** Primordial germ cells (PGCs) undergo demethylation at imprinted loci, which erases parental imprinting marks at around embryonic day 11.5–12.5 (REFS 50,51). The female PGCs develop to form primary oocytes. During oocyte growth and maturation, the maternal-specific genomic imprints are re-established through the *de novo* methylation activities of the methyltransferases Dnmt3a and Dnmt3b, and an associated protein Dnmt3l[52,53]. During spermatogenesis, several factors seem to function during the differentiation of the spermatocytes from the leptotene to pachytene stages of meiosis. During this period, histones are hypoacetylated, and the functions of Suv39h, Dnmt3a and Dnmt3l are essential[36,52,53,55]. The crucial stage when these factors function is not defined.



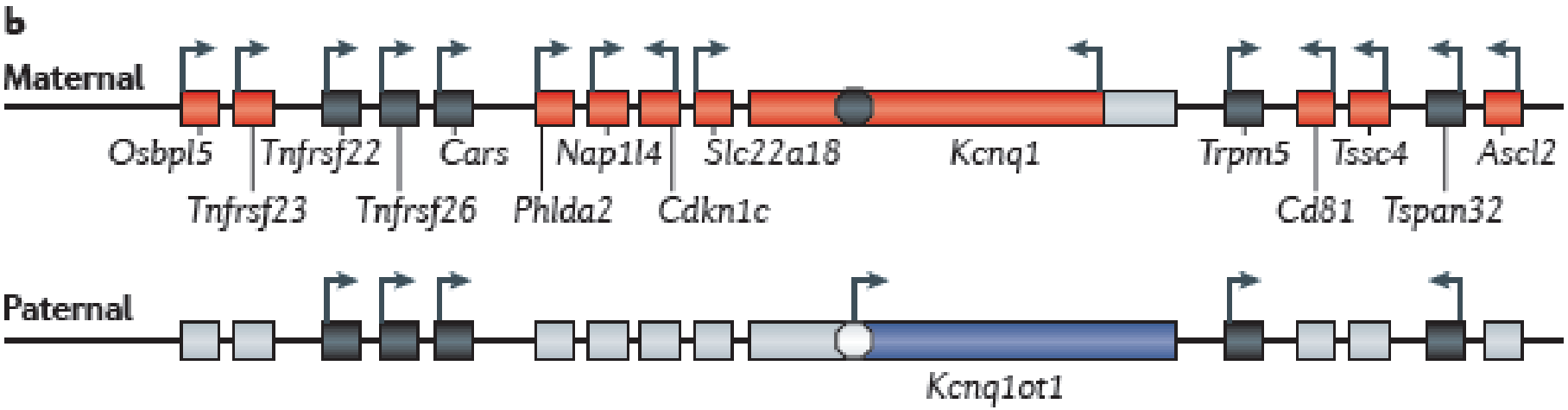


Airn is a noncoding RNA

DMR=differentially methylated region



A Maternally methylated



ICR

lncRNA

Methylated ICR

Unmethylated ICR

Methylated secondary DMR

Unmethylated secondary DMR

Maternally expressed gene

Paternally expressed gene

Bi-allelically expressed gene

Repressed gene

B Paternally methylated

lncRNA

DMR0 DMR1 DMR2

Maternal

CTCF

H19-DMD

H19

Enhancers

Paternal

Ins2

Igf2

ICR

Methylated ICR

Unmethylated ICR

Methylated secondary DMR

Unmethylated secondary DMR

Maternally expressed gene

Paternally expressed gene

Bi-allelically expressed gene

Repressed gene

B Paternally methylated

c

Maternal

miRNAs

miRNAs

IG-DMR

Gtl2

AntiRtl1

C/D SnoRNAs

Mirg

Paternal

Begain

Dlk1

Rtl1

Dio3

ICR

Molecular mechanisms linking ICR CpG methylation to silencing of associated protein-coding genes in genomic imprinting:

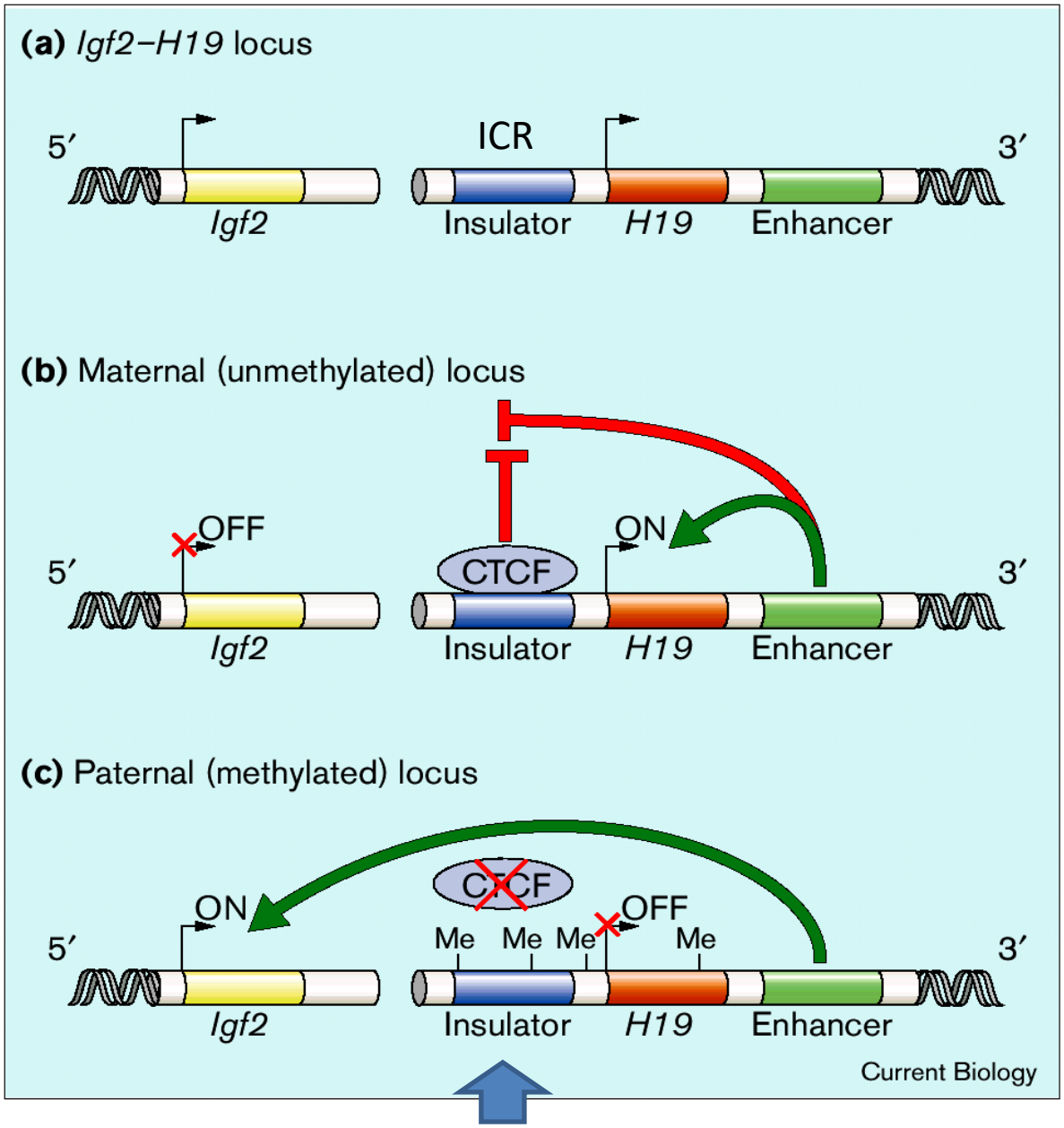
-The Insulators model

-The noncoding RNA model

1st mechanism: insulators

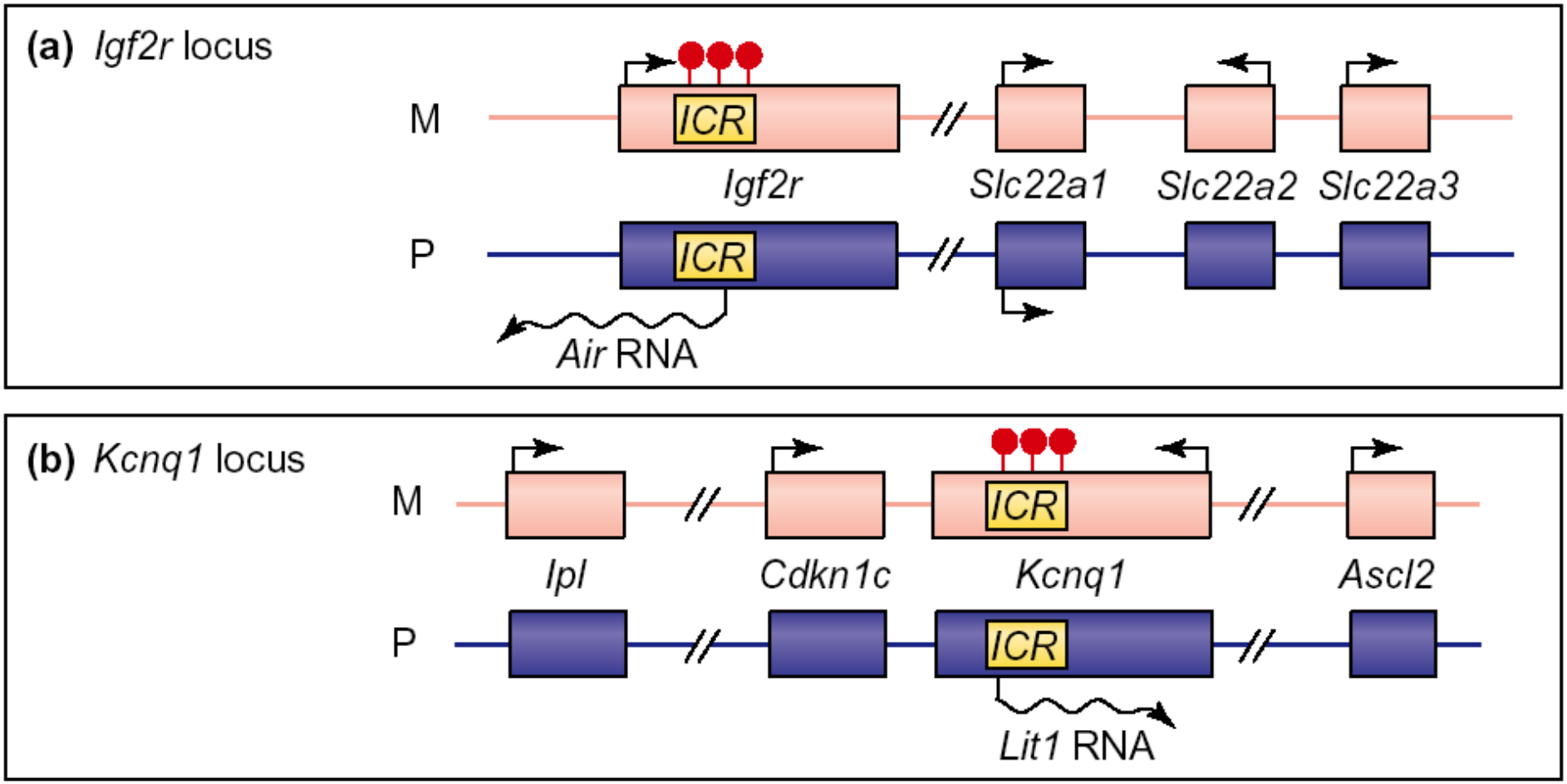
This mechanism was described first, two decades ago

We re-define this model today, saying that the methylated ICR "separates" two domains



This part, called «insulator» corresponds to the ICR for this locus

2° mechanism: long noncoding RNAs



Why should the transcription of a long noncoding RNA **repress** transcription of neighboring protein-coding genes ?

Most studied model is **X-chromosome inactivation,** i.e. Xist RNA interaction with PRC2.

This mechanism neatly demonstrated in the case of Kcnq1 and Air (less clearly in my opinion)

other mechanisms may operate, as e.g. transcriptional interference

More difficult to understand why some intermixed genes escape

Textbook

REVIEW

Epigenetic Regulation by Long Noncoding RNAs

Jeannie T. Lee

Recent studies show that transcription of the mammalian genome is not only pervasive but also enormously complex. It is estimated that an average of 10 transcription units, the vast majority of which make long noncoding RNAs (lncRNAs), may overlap each traditional coding gene. These lncRNAs include not only antisense, intronic, and intergenic transcripts but also pseudogenes and retrotransposons. Do they universally have function, or are they merely transcriptional by-products of conventional coding genes? A glimpse into the molecular biology of multiple emerging lncRNA systems reveals the "Wild West" landscape of their functions and mechanisms and the key problems to solve in the years ahead toward understanding these intriguing macromolecules.

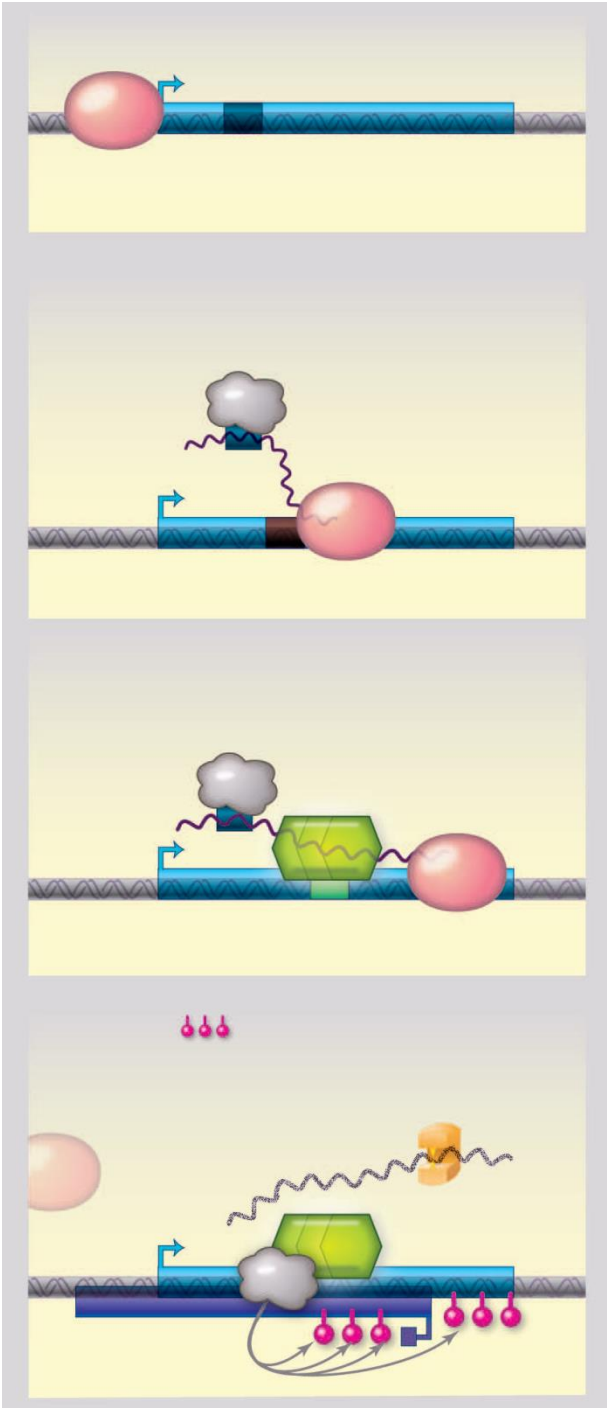


Fig. 2. LncRNAs tether epigenetic complexes to chromatin, enabling allele- and locus-specific regulation. LncRNA that is synthesized binds to an epigenetic complex (such as PRC2) and, together, are loaded onto chromatin cotranscriptionally through DNA-bound factors (such as YY1 for Xist RNA). Epigenetic modifications then silence the gene, and rapid lncRNA turnover prevents its diffusion to other loci.

From: Lee 2012

A number of cases where RNA plays a role in silencing

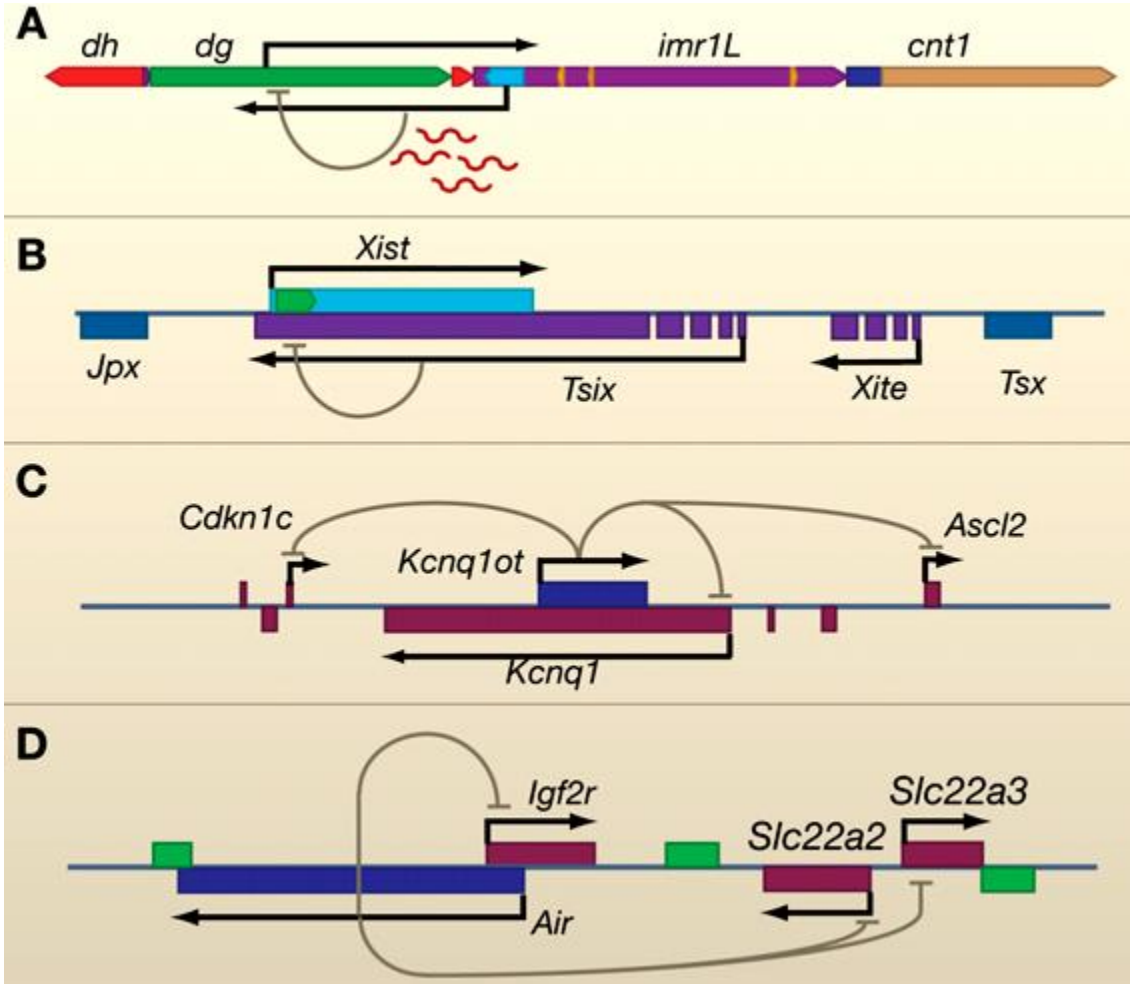


Figure 3. Silencing Transcripts at the S. pombe Centromere and Different Transcript-Mediated Silencing Systems

(A) S. pombe dg-imr transcript in centromere IL. The forward promoter is silenced by constitutive transcription and processing into siRNA of the reverse strand.

(B) The X inactivation center has several noncoding transcripts, and transcription of Tsix silences the Xist promoter.

(C and D) Paternal locus transcription of noncoding transcripts Kcnq1ot and Air influences the expression of overlapping and nonoverlapping genes in the imprinted gene cluster at the telomeric end of mouse chromosome 7 and the Igf2r locus, respectively. Paternally expressed genes are colored blue, maternally expressed genes are colored red, and ubiquitously expressed genes are colored green.

Transgenerational epigenetic inheritance

much attention today since it represents the suggestive possibility that modifications induced from environment and other stimuli can affect imprinting of some genes in the offspring

like:
foods
drugs
pollutants
diseases
lifestyle
drinking
stresses
etc...

phenomena that have been indeed described and filed in medical genetics

Caution !

There are several examples of apparent transgenerational epigenetic inheritance that do not require epigenetic modifications in the **gametes**.

Example:
Rats that are nurtured by stressed mothers are more likely to be stressed. This phenotype is perpetuated across generations and involves the setting of a 'stressed' state by the hypothalamic– pituitary–adrenal axis in the pup.
There is an epigenetic change (DNA methylation) at the glucocorticoid receptor gene promoter in the hippocampus of the pups.

Thus, the establishment is in the pup, not in gametes.
This is **not** genomic imprinting

Antigen Receptors (Ig + TCR) are possibly discussed in the Immunology courses

Sporadic monoallelic genes have been discovered recentlythaks to RNA-seq technology, but little is known on the mechanism.
Hypothesis: limiting TFs or feed-backs.

Olfactory receptors

- 1,400 genes in several clusters on several chromosomes
- only one expressed in each single sensory neuron
- selection is at random
- Repressed OR genes aggregate centrally in the nucleoplasm, forming distinct foci that feature interchromosomal interactions
- the single transcribed allele is localized outside these foci
- expression of the random choosen one at the surface gives a feed-back signla that established repressive chromatin on the foci containing all other alleles.

next lesson: technical

1. discussion of Research Paper 2

2. CpG methylation genome-wide analysis