Epigenetic regulation by long non-coding RNAs

Methods for lncRNA modulation

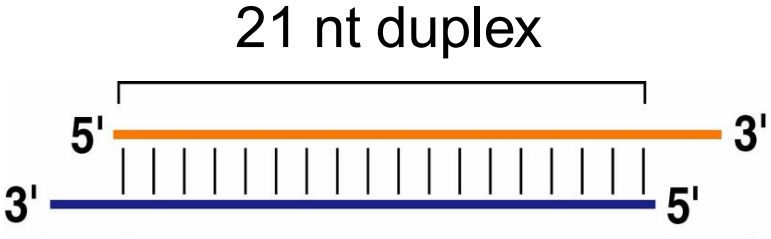
1. **Oligonucleotides:**
   - siRNAs
   - Gapmers

2. **Vector/Lentivirus:**
   - shRNA
   - lncRNA gene

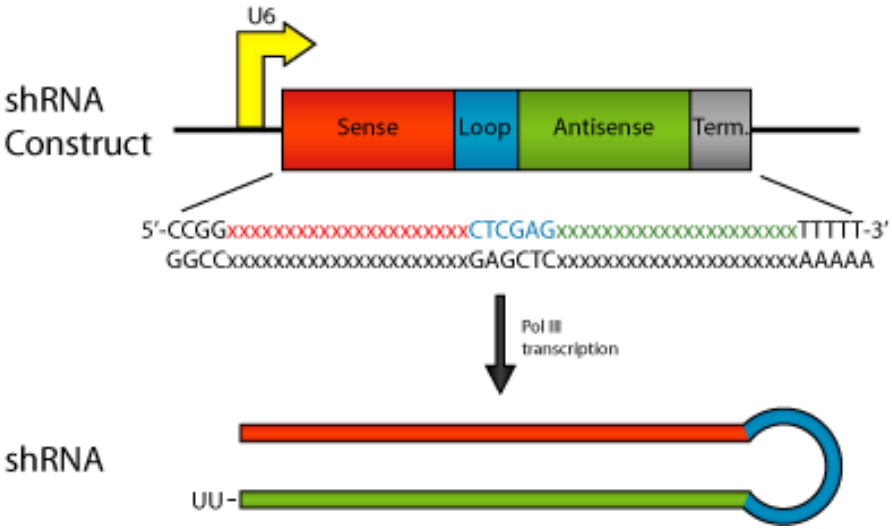
3. **CRISPR-Cas9 or -Cas13:**
   - Knock-out
   - Transcriptional modulation
   - RNA degradation (Cas13)

RNAi methodologies

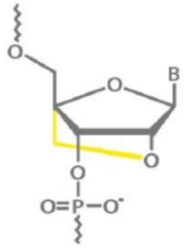
- RNA duplex



- Pol III vector



GAPMERS with LNA™ technology



LNA™

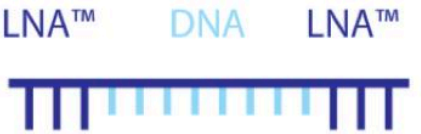
K. Bondensgaard *et al.*, *Chem. Eur. J.* **2000**, *6*, 2687
M. Petersen *et al.*, *J. Am. Chem. Soc.* **2002**, *124*, 5974

LNA™ technology increases binding affinity

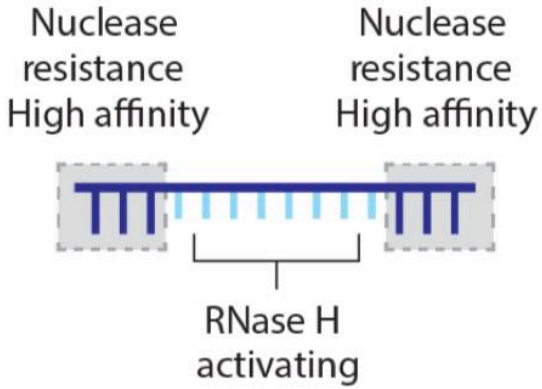
- LNA is a bicyclic high affinity RNA mimic with the sugar ring locked in the 3'-endo conformation
- Obeys Watson-Crick base-pairing rules
- Stable A-helix with good base-stacking

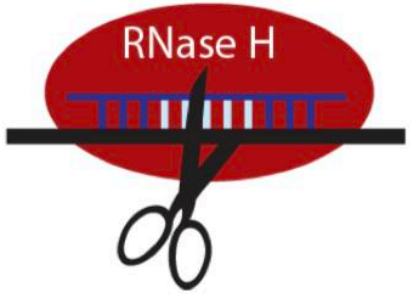
⬇

- Increased Tm (Tm increases by 2 - 8ºC per base)
- Tm normalization (adjust oligos to the same Tm)
- Improved mismatch discrimination
- High sensitivity and specificity in hybridization assays



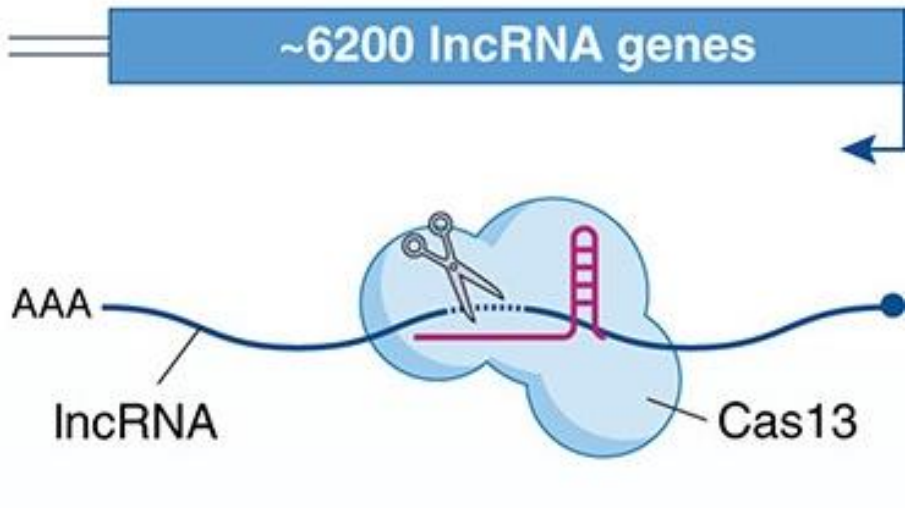
GAPMER





Using the RNA-targeting CRISPR-Cas13 nuclease

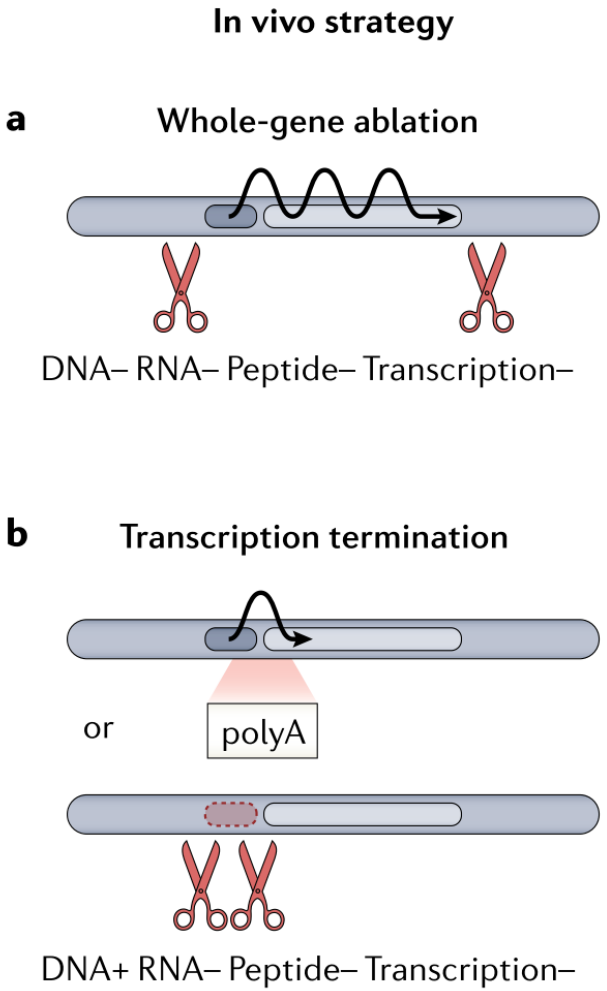
Currently-used methods that target transcripts, typically based on RNAi or antisense oligonucleotides, often have significant non-specific and off-target effects. The newly-reported CRISPR/Cas13 system appears to represent a significant improvement over the existing RNA-targeting methods. First, it has significantly lower off-target effects compared to RNAi. Second, its activity is severely reduced or abrogated by 1–2 mismatches in the center of a guide (g)RNA, allowing for a mismatch control for each targeting gRNA



LOF strategies to understand the lncRNA mechanism of action in *vivo*

a. Whole-gene ablation mouse models are a powerful starting point to investigate whether a functional element is encoded in a lncRNA locus.

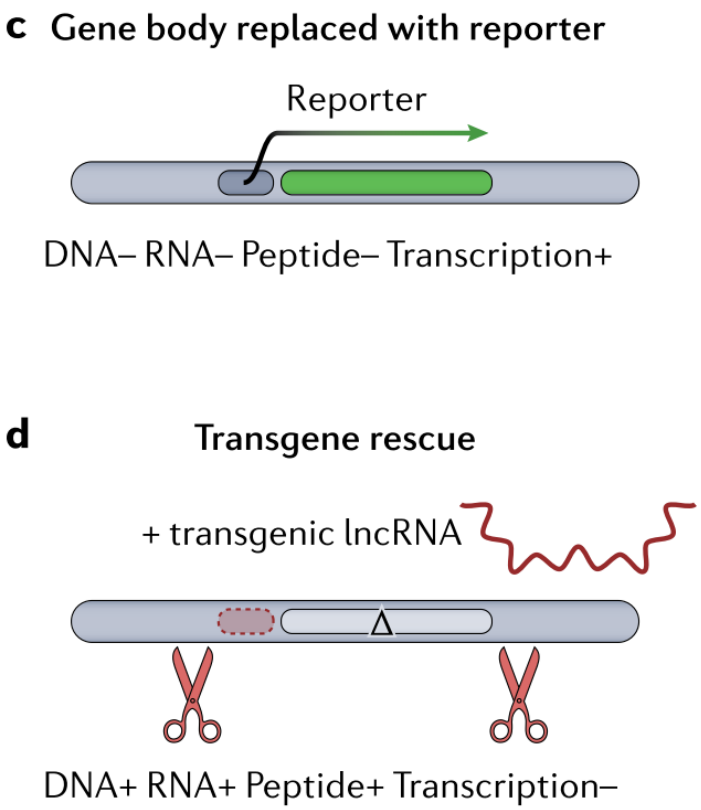
b. Observing a phenotype upon transcriptional termination by polyadenylaton (polyA)-terminator insertion (top) or promoter deletion (bottom) suggests that the functional component of a lncRNA locus is encoded in the lncRNA product, the act of transcription or the lncRNA promoter



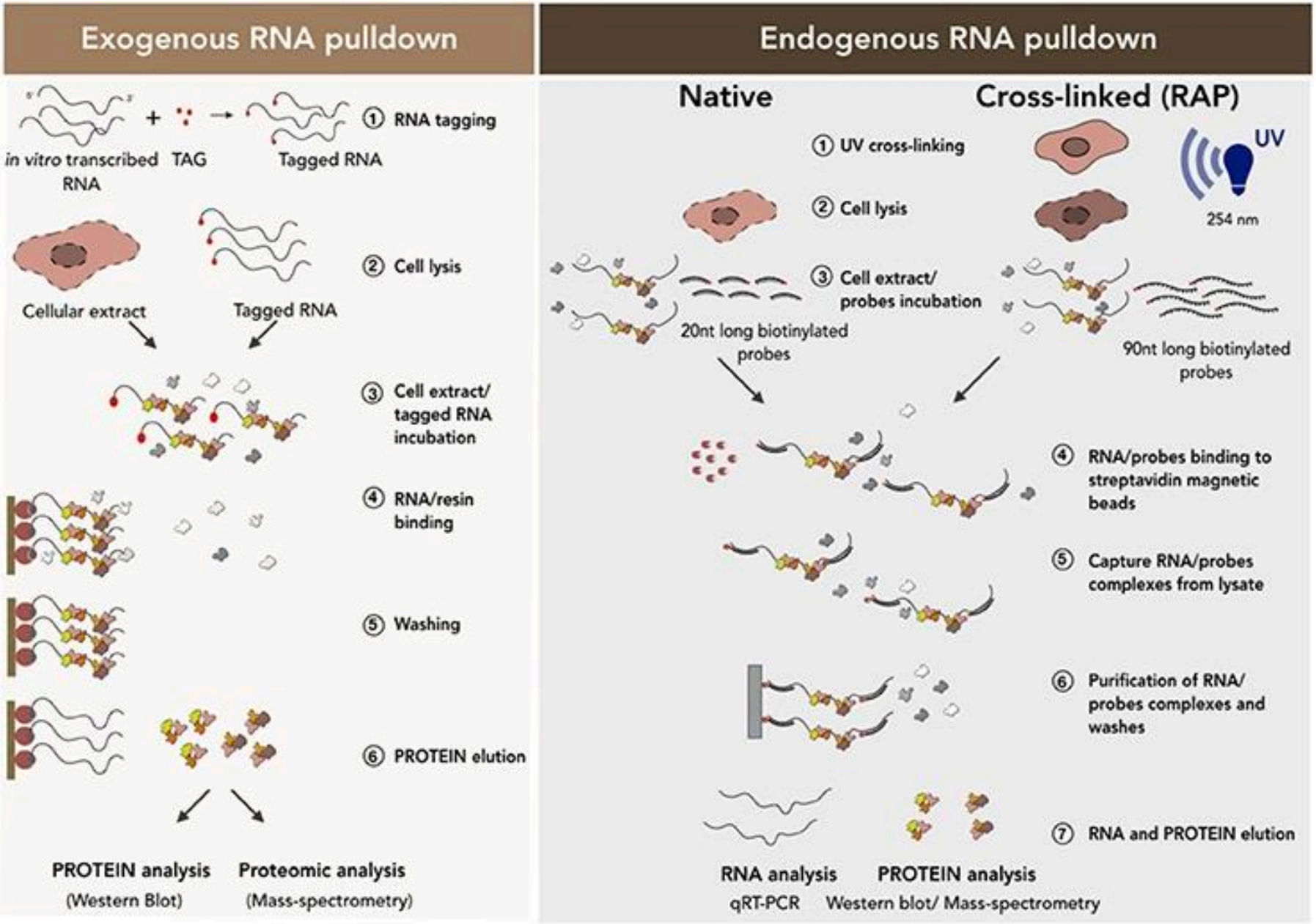
LOF strategies to understand the lncRNA mechanism of action in *vivo*

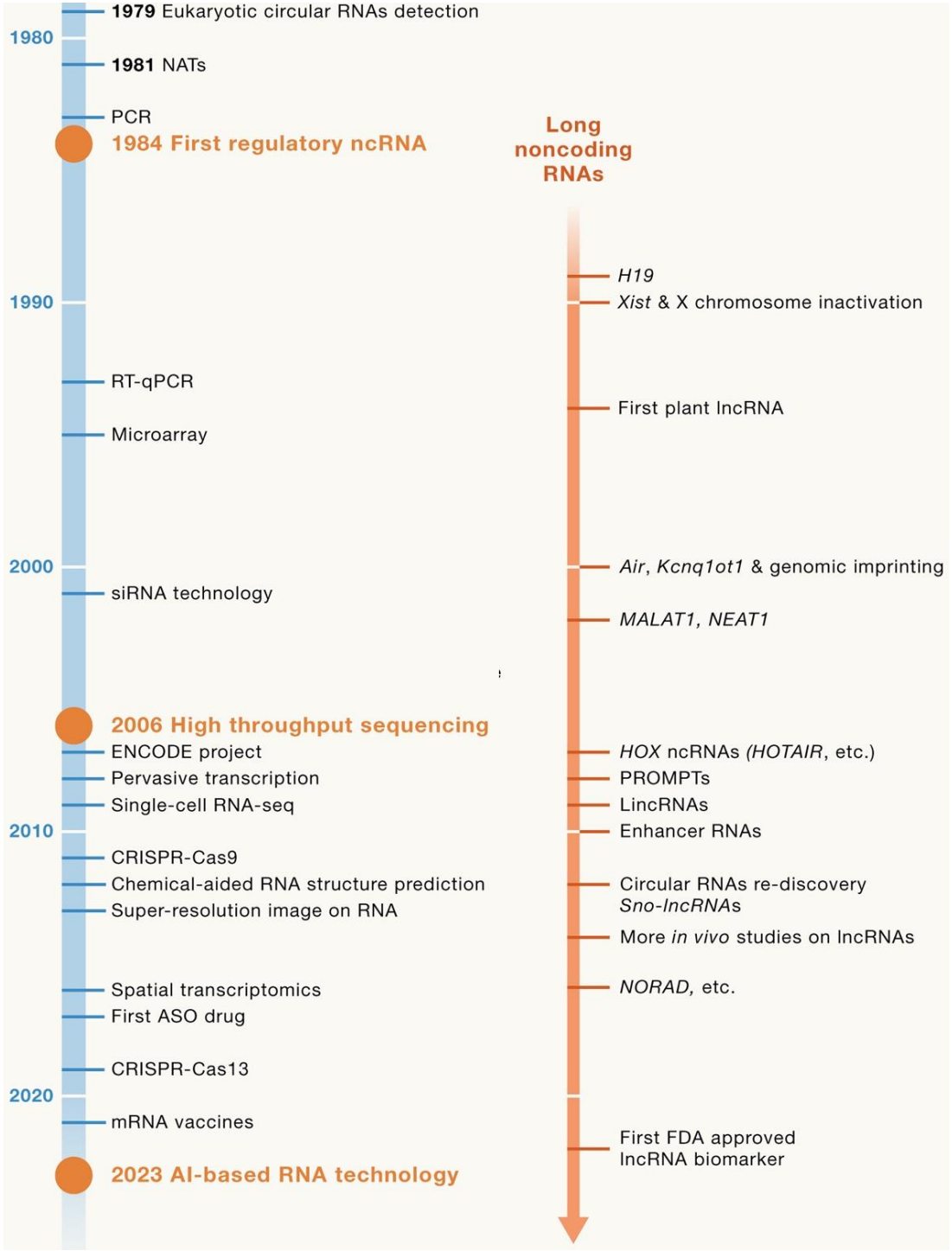
c. Replacement of the lncRNA gene body with a reporter (driven by the endogenous promoter) controls for the act of transcription.

d. Transgene rescue or gain of function is considered the gold standard to prove trans activity of a lncRNA molecule. The processed lncRNA of interest or transgene is inserted on a different chromosome in the mutant background causing the phenotype. Rescues of the phenotype demonstrate that the lncRNA mediates its function *in trans* via the lncRNA molecule.



Methods for the identification of lncRNAs interactors



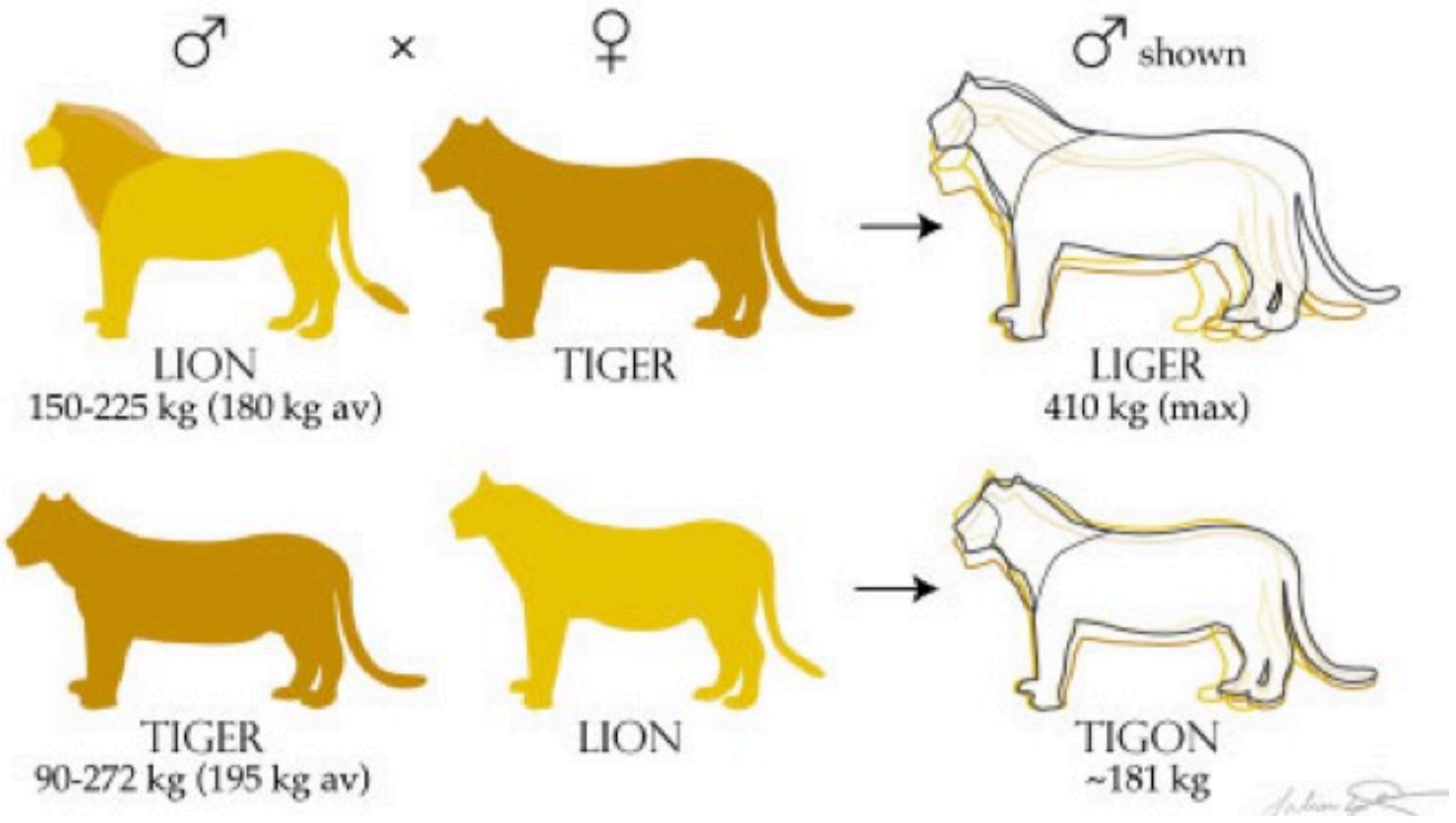


Imprinting and X inactivation

Imprinting

Imprinting and X dosage compensation

Genomic Imprinting



Genomic Imprinting

- Nonequivalence in expression of alleles at certain gene loci dependent on the parent of origin.

- Only one parental allele is expressed.

- The current number of imprinted genes in the mouse is approximately 150 (http://www.mousebook.org/catalog.php?catalog=imprinting).

- The imprinted genes are typically located in clusters of 3–12 (or more) genes that play essential roles in the growth and development of the fetus, as well as in post-natal behavior and metabolism.

Genomic Imprinting

Theory

- Conflict exists between the interests of the paternal and maternal genes
- Imprinting is a way to ensure the offspring's survival both in situations of scarce resources and in situations where resources are abundant.
- For optimal fitness of the father, paternal genes maximize acquisition of maternal resources to ensure larger sized offspring
- Maternal genes are sparing in the demands of maternal resources, so that the mother has a better chance to bear further offspring
- Paternally-expressed genes generally stimulate growth
- Maternally-expressed genes generally repress growth

Genomic Imprinting

- Most imprinted clusters contain **protein coding genes** and **noncoding RNAs** (ncRNAs). The ncRNAs are of different varieties (microRNAs, snoRNAs, and lncRNAs), some of which are essential to the mechanism that imprints these genes in *cis*.

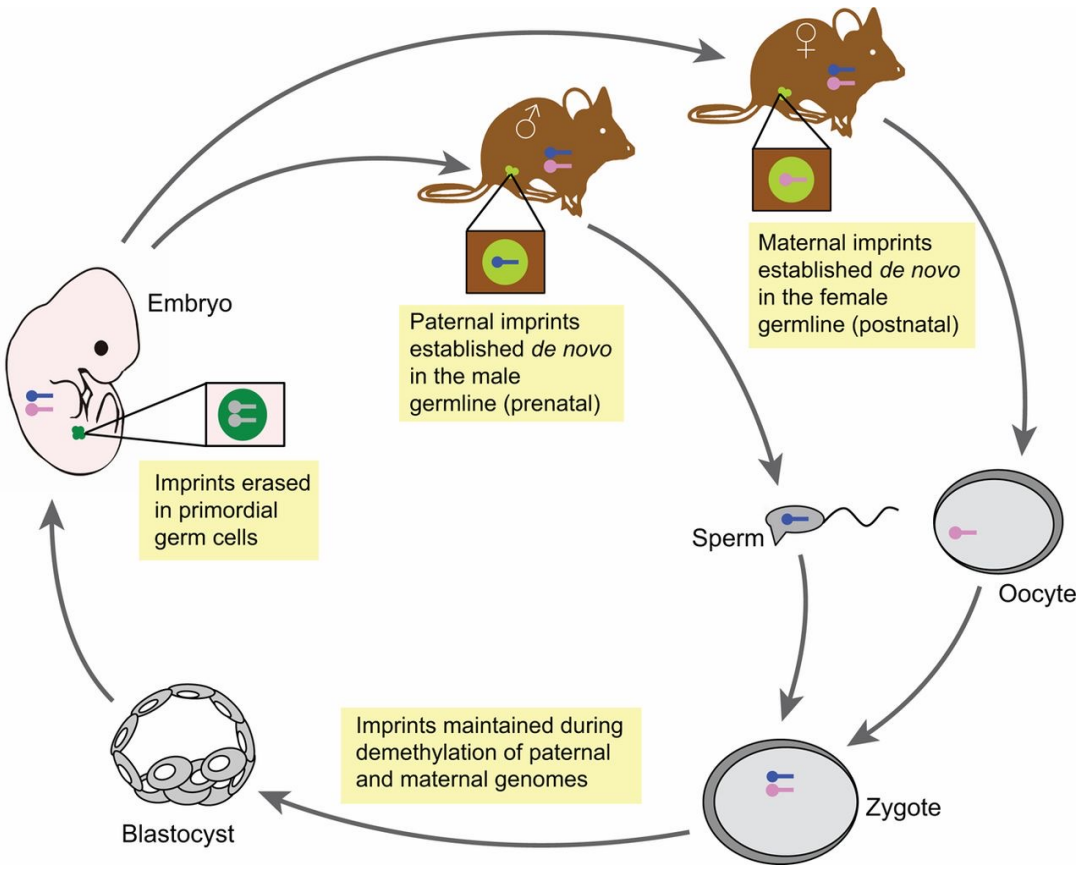
- These clusters typically contain genes that are expressed exclusively from the maternally or paternally inherited chromosomes

- While many imprinted genes are ubiquitously imprinted, some exhibit tissue-specific or temporal-specific imprinting patterns.

- Each well-studied cluster has a discrete **imprinting control region (ICR)** that exhibits parent-of-origin-specific epigenetic modifications (DNA methylation and post-translational histone modifications) and governs the imprinting of the locus. Although the mechanism(s) that confer the allele-specific epigenetic modifications is poorly understood.

Timeline of epigenetic reprogramming

The first wave of genome-wide DNA demethylation takes place shortly after fertilization. Imprinted ICRs are maintained despite this demethylation event. *De novo* DNA methylation occurs around implantation. In both male and female primordial germ cells another wave of DNA demethylation initiates as the cells migrate toward the genital ridge. All ICRs are also erased at this time. In male germ cells methylation imprints are acquired around E15.5–17.5. In the female germline methylation imprints are not acquired until after birth, in growing oocytes. Activities in male and female germ cells are represented in blue and red, respectively.



Imprinting is required for normal development

An individual normally has one active copy of an imprinted gene. Improper imprinting can result in an individual having two active copies or two inactive copies. This can lead to severe developmental abnormalities, cancer, and other problems.

Prader-Willi and Angelman syndrome are two very different disorders, but they are both linked to the same imprinted region of chromosome 15.

Prader-Willi syndrome • Symptoms include learning difficulties, short stature, and compulsive eating. • Individuals are missing gene activity that normally comes from dad. • Happens when dad's copy is missing, or when there are two maternal copies.

Angelman syndrome • Symptoms include learning difficulties, speech problems, seizures, jerky movements, and an unusually happy disposition. • Individuals are missing gene activity that normally comes from mom. • Happens when mom's copy is defective or missing, or when there are two paternal copies.

Genomic Imprinting

Two mechanisms have been described for mediating imprinting in clusters:

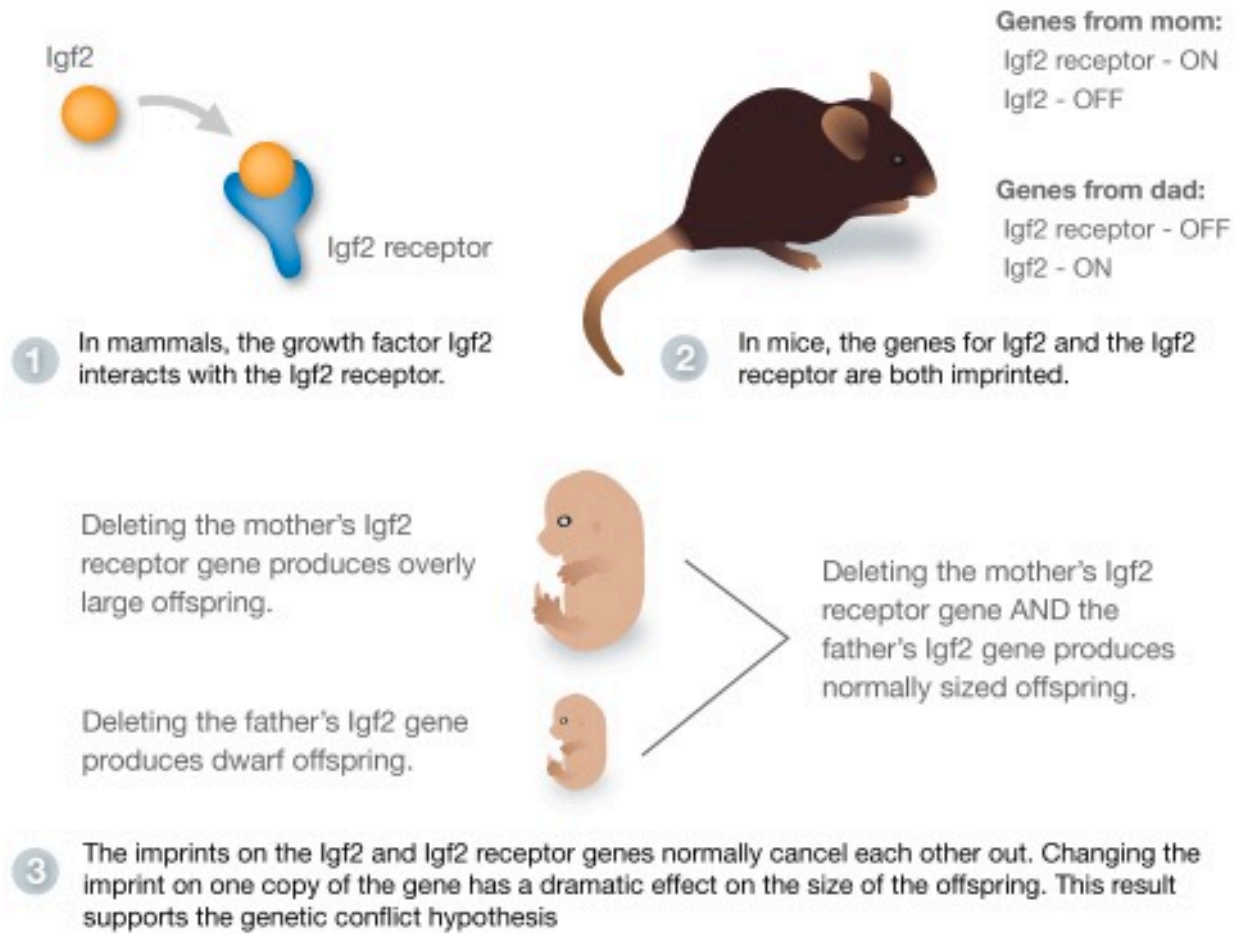
1. Silencing through the **insulator**

2. Silencing through ***in cis*** **ncRNA (macro-ncRNA) transcription**

It is possible that both mechanisms are used, but in a tissue-specific manner.

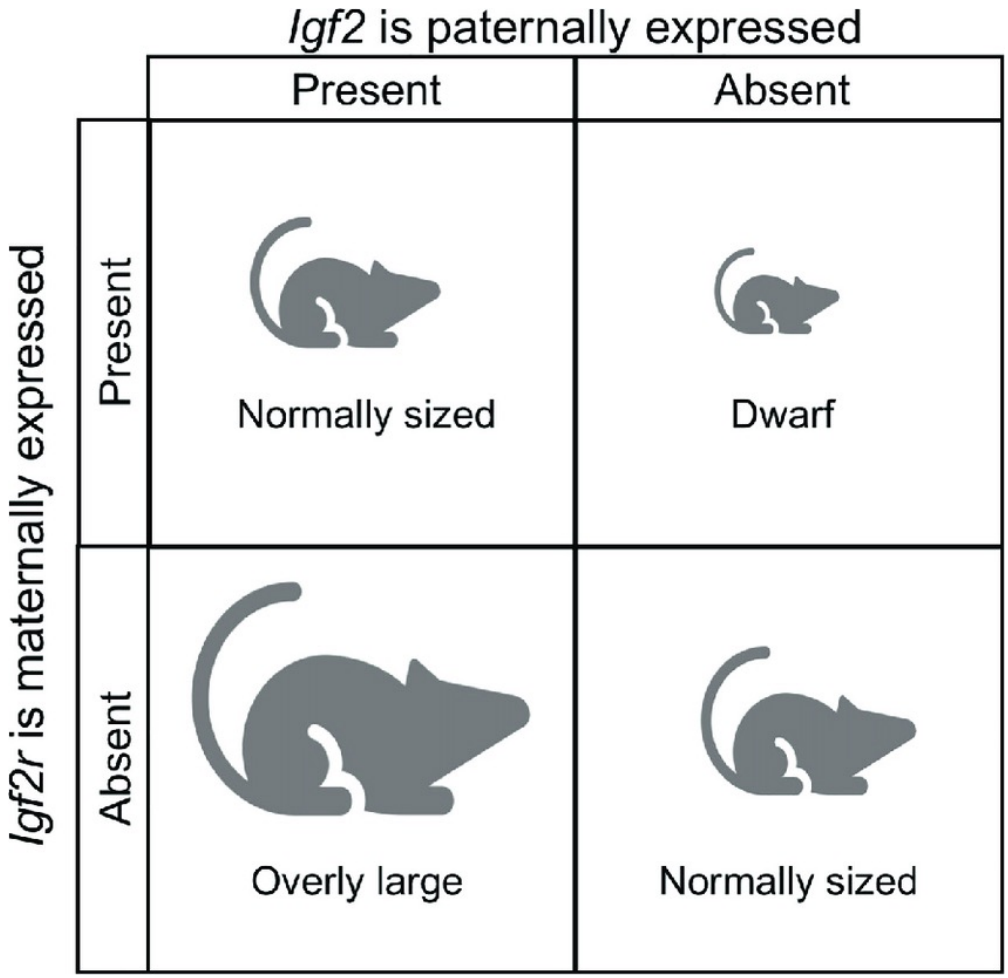
Imprinting at the Igf2/Igf2r loci

The Igf2 gene codes for a hormone that stimulates growth during embryonic and fetal development. It exerts its effects by binding to the IGF-1 receptor. IGF2 may also bind to the IGF-2 receptor, which acts as a signalling antagonist; that is, to prevent IGF2 responses.



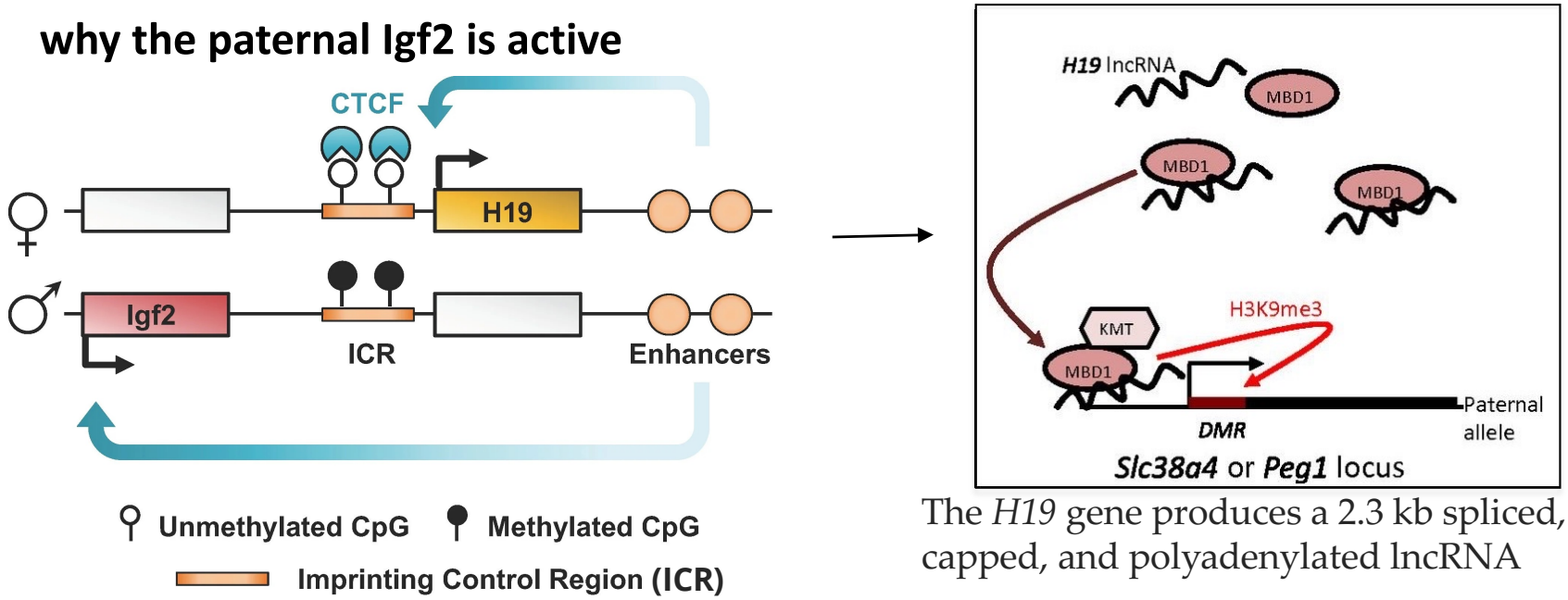
The Igf2 gene (but not the Igf2 receptor gene) is also imprinted in humans. Activation of the maternal Igf2 gene causes Beckwith-Wiedemann Syndrome (BWS). Babies with BWS are born larger than 95% of their peers. They also have an increased risk of cancer, especially during childhood.

Igf2/Igf2r


Igf2 is paternally expressed
Present
Absent
Igf2r is maternally expressed
Present
Absent
Normally sized
Dwarf
Overly large
Normally sized

Insulator model of Imprinting

The insulator model is exemplified by the **H19/Igf2 locus**. Here, the intergenic ICR is paternally methylated. On the unmethylated maternal allele, **CTCF** binding prevents enhancers from interacting with the Igf2 promoter. Instead, the enhancers activate H19 lncRNA expression, which inhibits paternal allele expression by recruiting MBD1 and KMT (histone lysine methyltransferase). On the paternal allele, methylation of the ICR spreads to the H19 promoter, silencing its expression, and prevents CTCF from binding the ICR, allowing the enhancers to activate Igf2 expression.



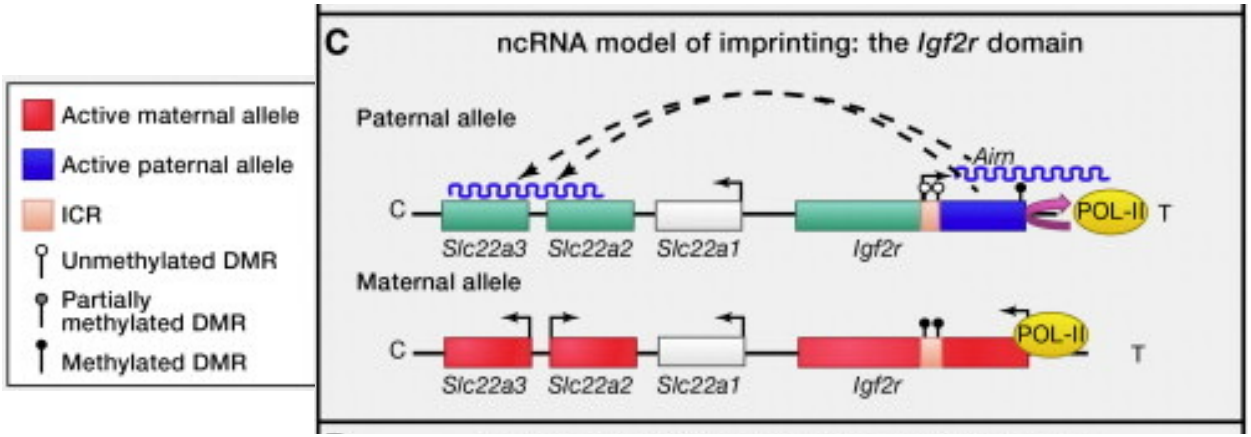
The *H19* gene produces a 2.3 kb spliced, capped, and polyadenylated lncRNA

It remains to be determined how DNA methylation is recruited to specific CpGs. Recently, a link has been found between histone methylation and DNA methylation suggesting that patterns of histone methylation could dictate patterns of DNA methylation, which could then be stably inherited.

ncRNA model of Imprinting (Igfr2)

For the **Igf2r locus**, transcription of the **Airn lncRNA** is governed by a promoter within the ICR and is expressed from the unmethylated paternal allele. In somatic cells, transcription of Airn over the Igf2r promoter precludes Igf2r expression, in part by kicking RNA polymerase II (POL-II) off of the promoter. In extraembryonic lineages (placenta), Airn lncRNA recruits enzymes that confer repressive histone modifications to silence genes in cis (Scl22a1/a2).

why the paternal Igf2r is inactive

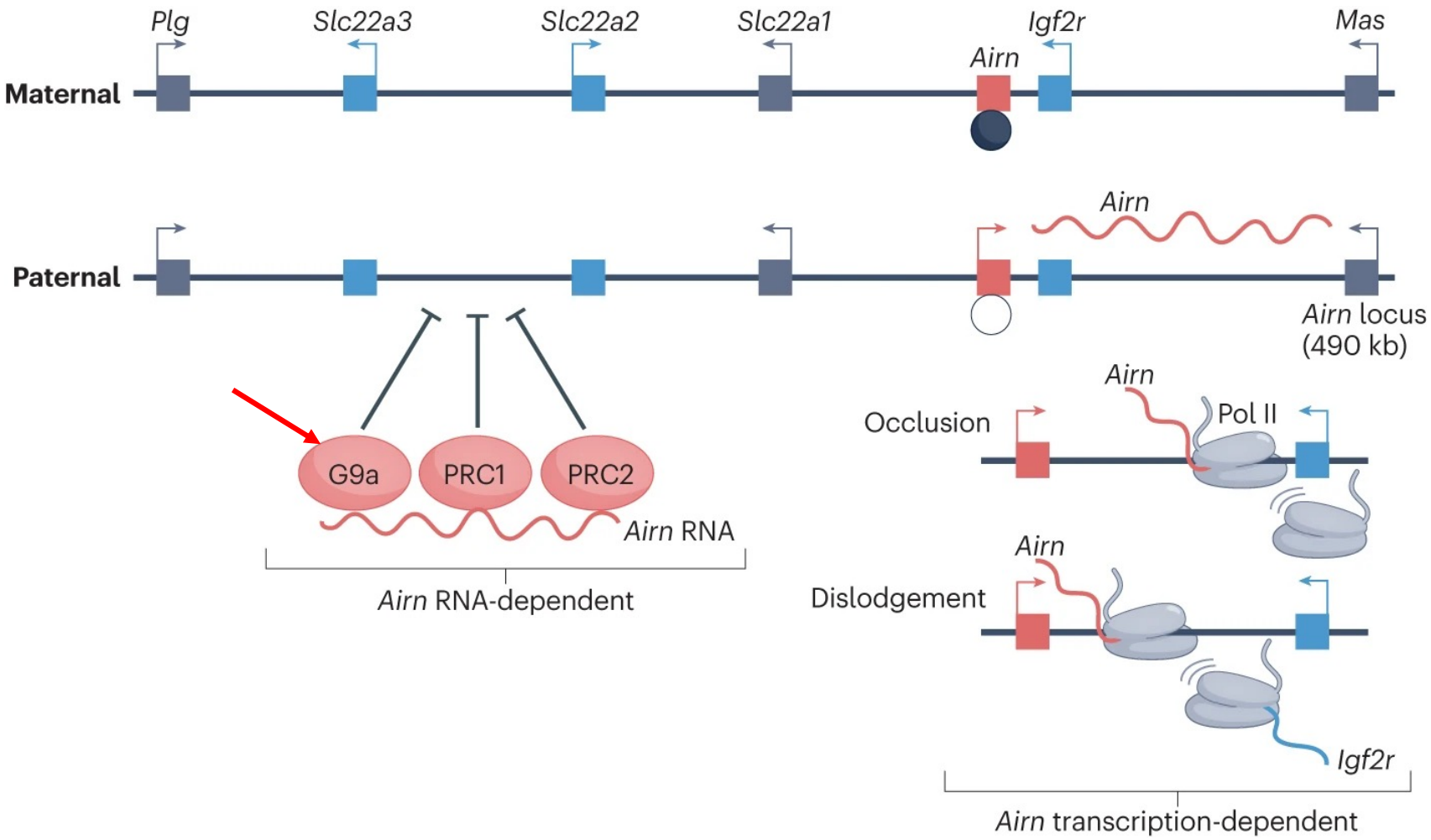


unmethylated

Interestingly, mechanisms that control imprinting in the placenta—a short-lived organ—may differ from those mechanisms that regulate imprinting in the much longer-lived somatic lineages.

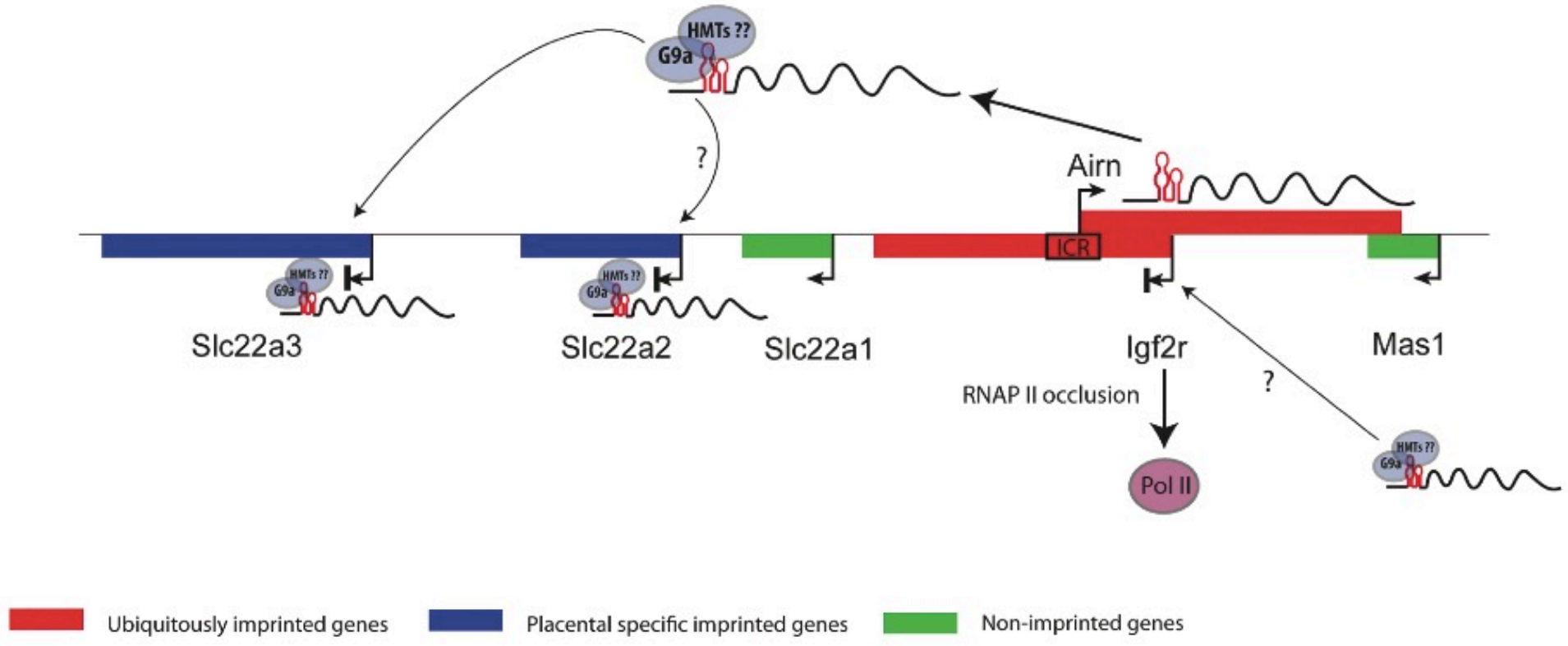
Silencing through *in cis* ncRNA transcription

In the Igf2r locus, the silencing of the overlapping Igf2r gene is mediated mainly by the Airn transcription itself, whereas Airn RNA mediates the silencing of the non-overlapping placentally imprinted genes through the recruitment of **G9a** -> H3K9me.



Silencing through *in cis* ncRNA transcription

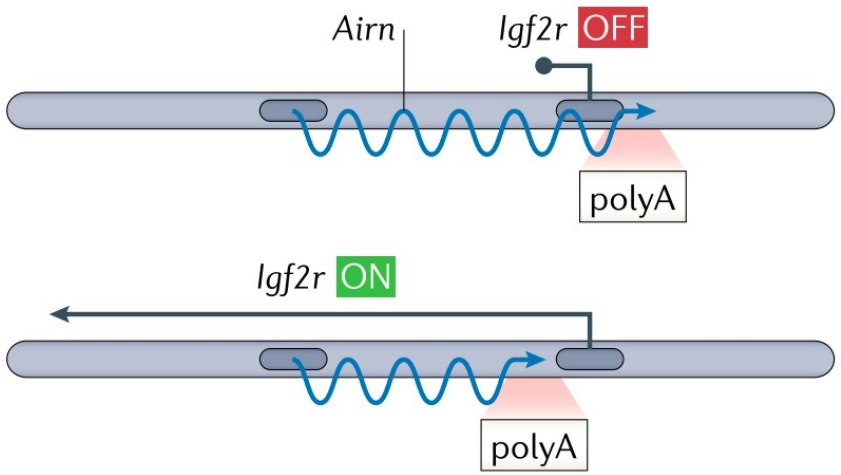
In the Igf2r locus, the silencing of the overlapping Igf2r gene is mediated mainly by the Airn transcription itself, whereas Airn RNA mediates the silencing of the non-overlapping placentally imprinted genes through the recruitment of **G9a** -> H3K9me.



Airn can works by transcriptional overlap

Igf2r silencing does not require any part of the *Airn* lncRNA, but only transcription through the *Igf2r* promoter. To test the role of Airn transcription versus product in Igf2r silencing, they used homologous recombination in embryonic stem (ES) cells to insert polyadenylation (polyA) cassettes on the paternal chromosome that truncate Airn to different lengths.

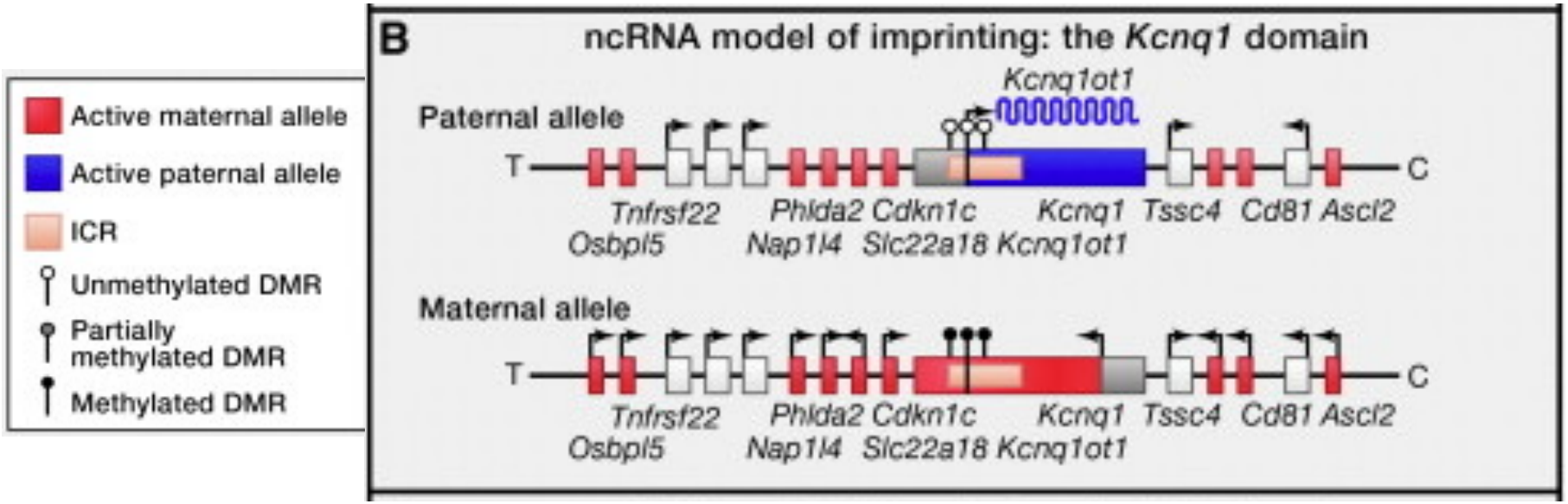
Truncation of *Airn* from the 3′ end after but not before the *Igf2r* promoter maintains *Airn*-mediated silencing



→ Conclusion: *Airn* transcriptional *Igf2r* promoter overlap is required for silencing

ncRNA model of Imprinting (Kcnq1)

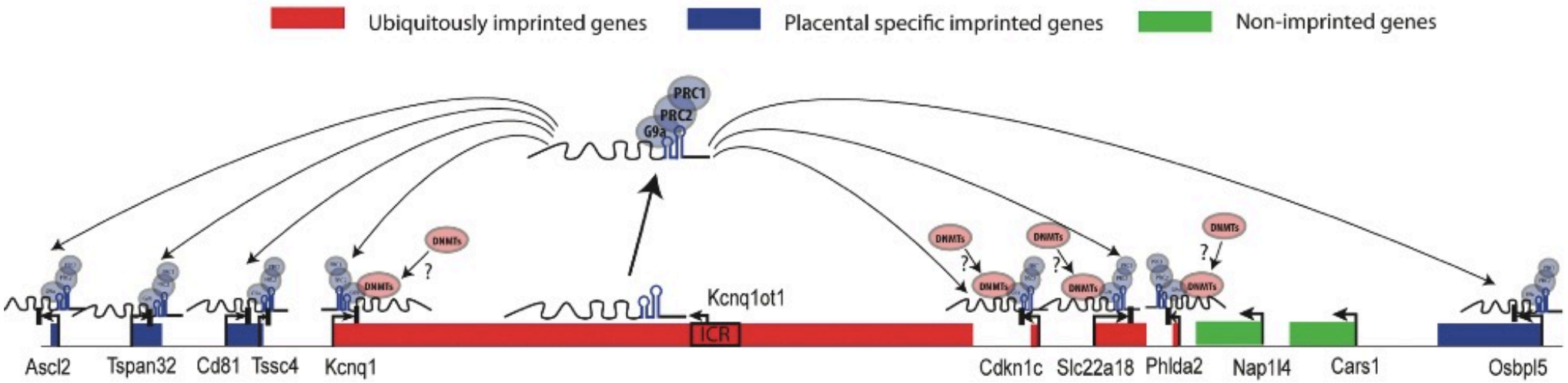
For Kcnq1 domain, the ICR contains the promoter of the **Kcnq1ot1 lncRNA**. On the paternal allele, the ICR is unmethylated, allowing Kcnq1ot1 expression. Kcnq1ot1 expression silences the paternal allele of the linked genes in cis. On the maternal allele, Kcnq1ot1 is not expressed due to methylation of the ICR, and the adjacent imprinted genes are expressed.



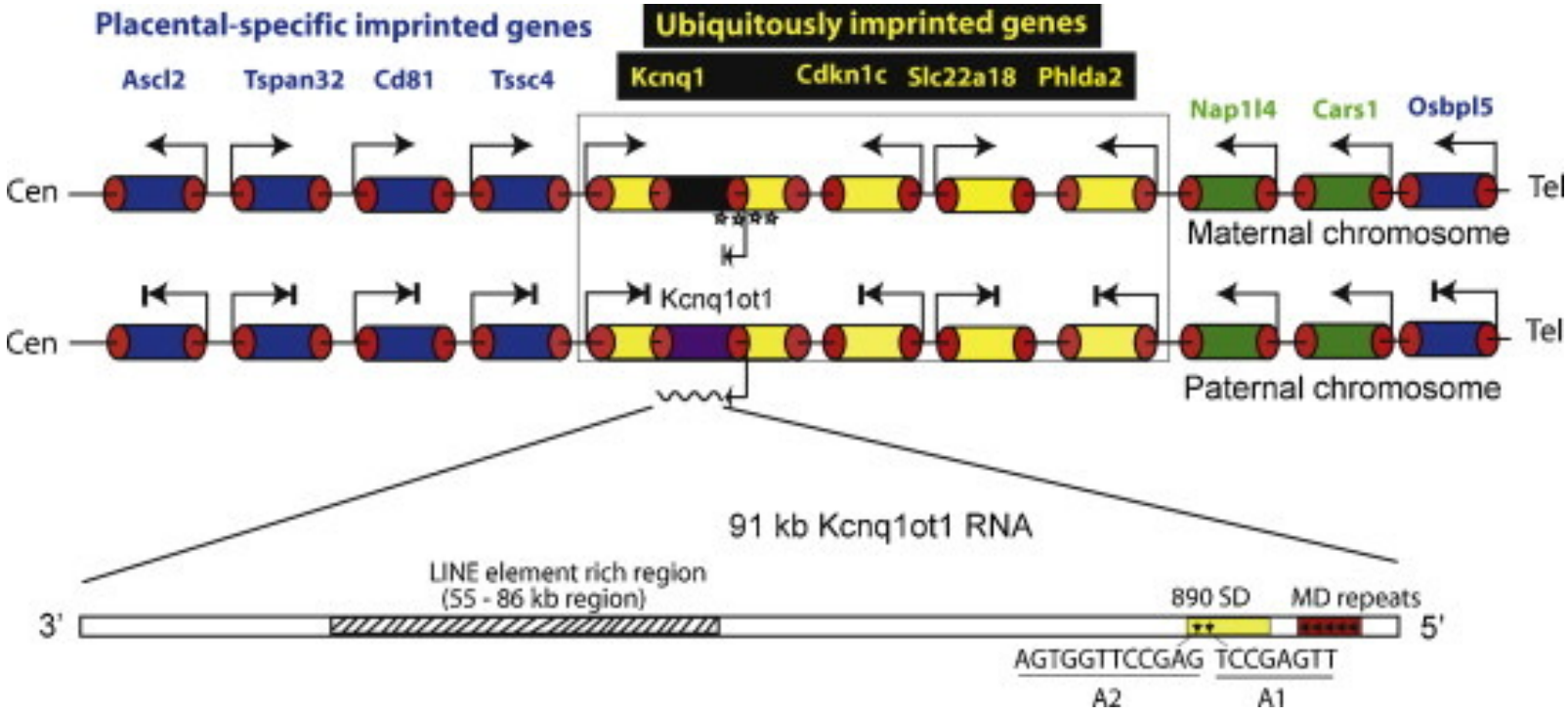
mechanisms that control imprinting in the placenta differ from those mechanisms that regulate imprinting in somatic lineages.

Mechanisms for lncRNA-mediated imprinting

Paternally expressed ***Kcnq1ot1*** **RNA** interacts with chromatin in cis and recruits **histone modifying complexes**, such as **PRC2**, **PRC1** and **G9a**, to the promoters of imprinted genes. This promotes the establishment of repressive heterochromatin and silencing of placental-specific imprinted genes. Kcnq1ot1 RNA establishes the repressive chromatin of imprinted genes in the Kcnq1 locus via a multilayered silencing pathway involving both repressive histone modifications and DNA methylation.

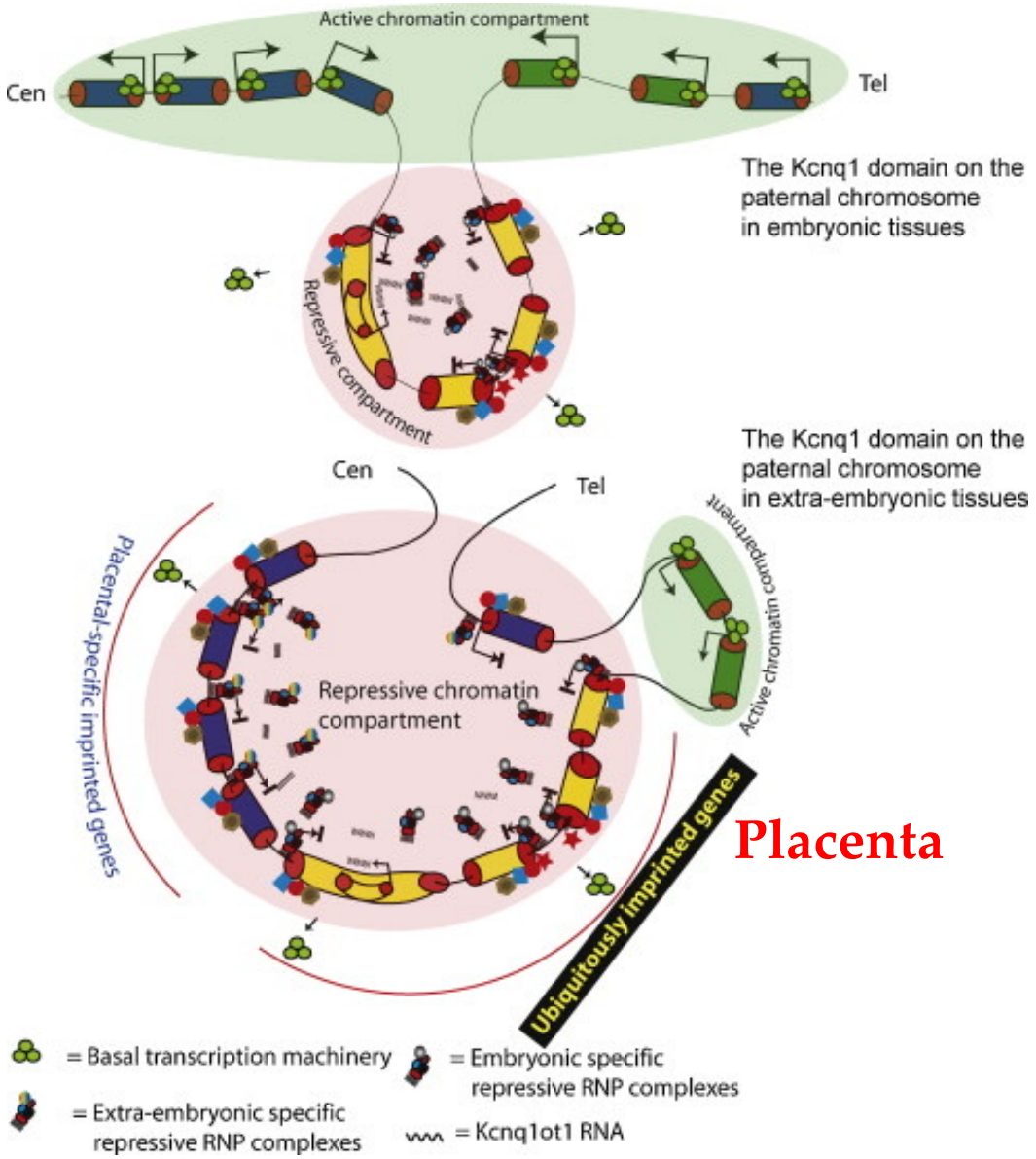


Physical map of the Kcnq1 locus, located on the mouse chromosome 7, showing placental-specific and ubiquitously imprinted genes.



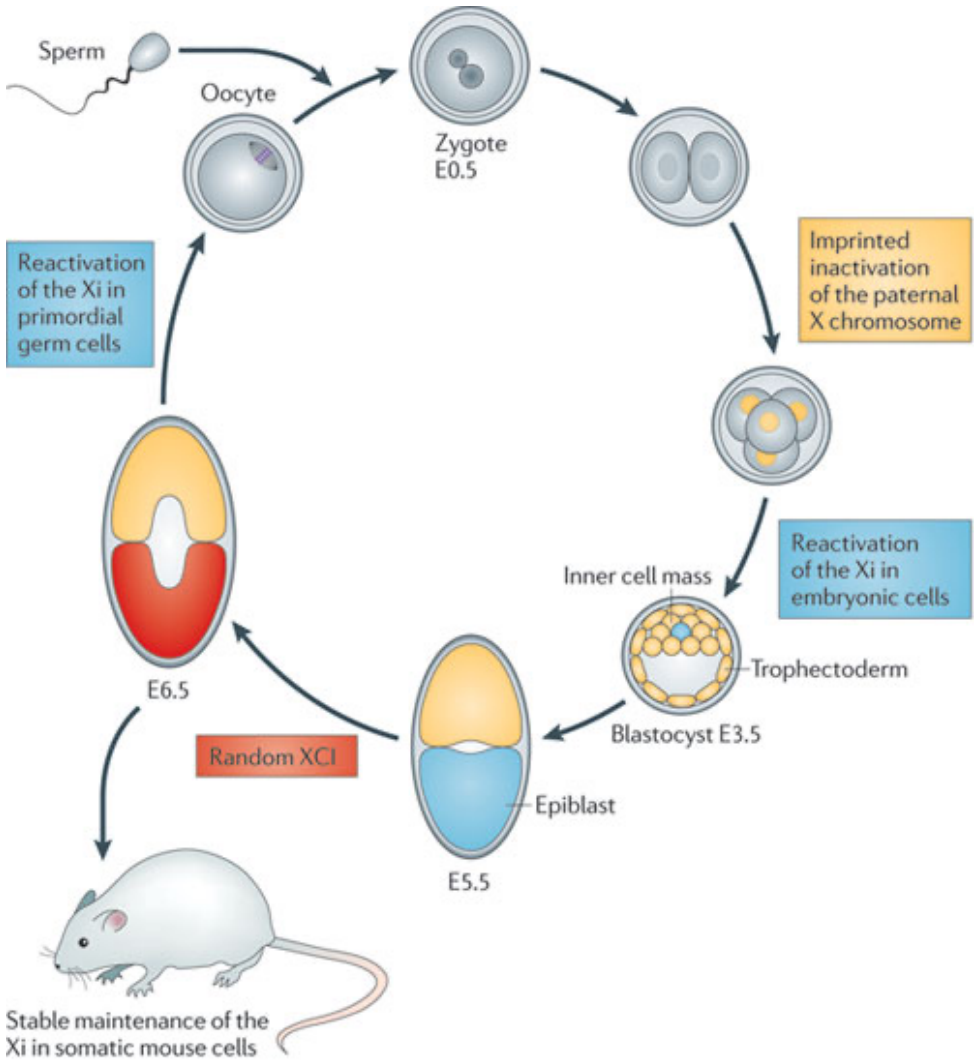
Mechanisms for lncRNA-mediated imprinting

Hypothetical models explaining the Kcnq1ot1 mediated lineage-specific transcriptional silencing in the Kcnq1 domain by organizing higher-order repressive chromatin compartments devoid of RNAPII machinery. In **the embryonic tissues** Kcnq1ot1 silences the ubiquitously imprinted genes on the paternal chromosome by recruiting both DNA and chromatin modifying enzymes, whereas in **placenta Kcnq1ot1** silences the ubiquitously as well as placental-specific imprinted genes by exploiting both the embryo and placental-specific mechanisms.



X Chromosome Inactivation (XCI)

Inactivation of the paternally inherited **X** is initiated in cleavage-stage embryos. **Imprinted inactivation** (yellow) of the paternally inherited **X** is maintained in the developing extra-embryonic lineages. In the **inner cell mass** of blastocysts, reactivation of the inactive **X** (**Xi**) occurs specifically in cells that will form the embryo (blue). Subsequently, both X chromosomes are active in the cells of the developing epiblast between embryonic day 3.5 (E3.5) and E5.5 (blue). Random XCI (red) of either the paternally or maternally inherited X chromosome is initiated around E5.5. Once established, the XCI pattern is maintained in the somatic lineages of female mice. In the developing germ line, the Xi is reactivated such that both X chromosomes are active in oogenesis.



X Chromosome Inactivation Center (Xic)

- XCI is regulated by cis-acting master control regions the **X Chromosome Inativation Center (Xic)** that has been mapped to a 100–500 kb region. Genetic analyses based on knockouts, gain-of-function mutations, and transgenic overexpression have shown that the Xic is necessary and sufficient to regulate XCI.

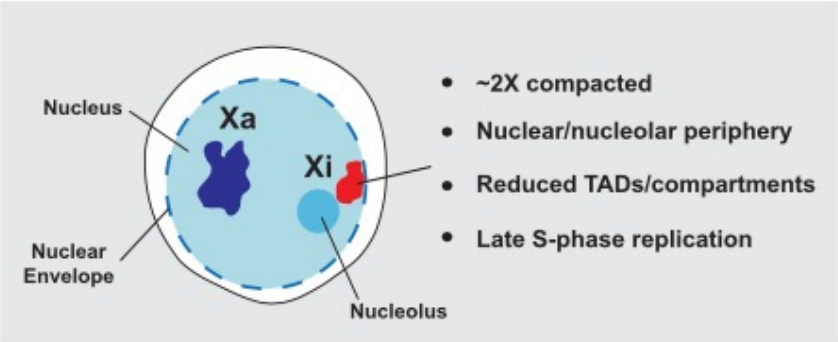
- XIC controls expression of the ***XIST*** gene (X-inactive-specific transcript) that produces a 17-20 kb ncRNA molecule. Xist ncRNA "coats" the entire *local* X-chromosome – *cis*-acting.

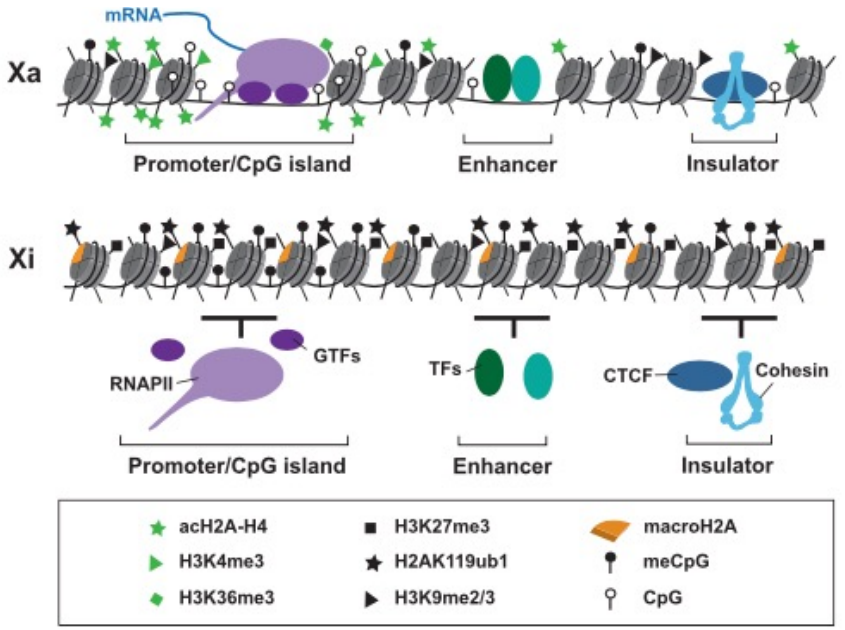
- Deleting the noncoding locus Xist results in loss of silencing capability (1996) and female-specific lethality.

Xist localization to the X chromosome

X Chromosome Inactivation (Xi)

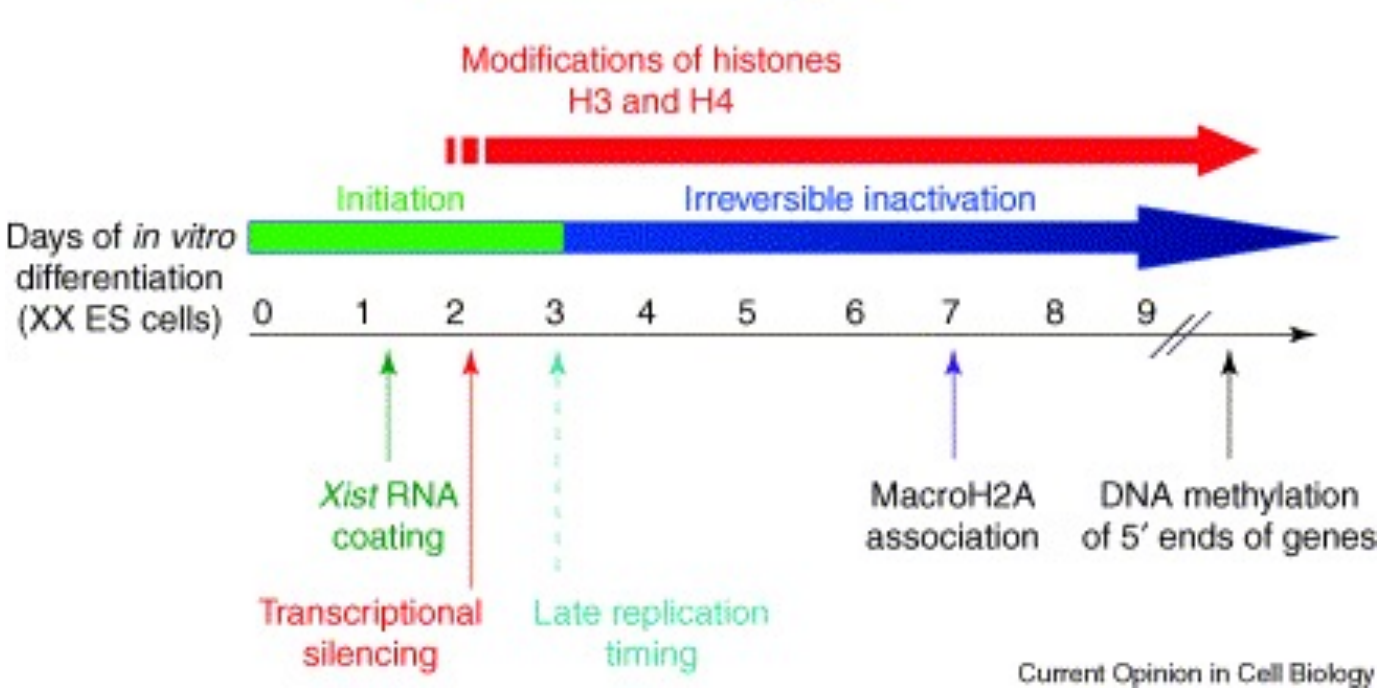
- Initial *XIST* RNA expression and coating
- Association of chromatin modifying proteins (HDACs, PRCs and other proteins)
- DNA methylation of X-chromosome genes
- Modification of histones by HMTs (H3K27me3 and H3K9me3) and hypoacetylation
- Recruitment of the histone variant macroH2A
- Formation of condensed chromatin





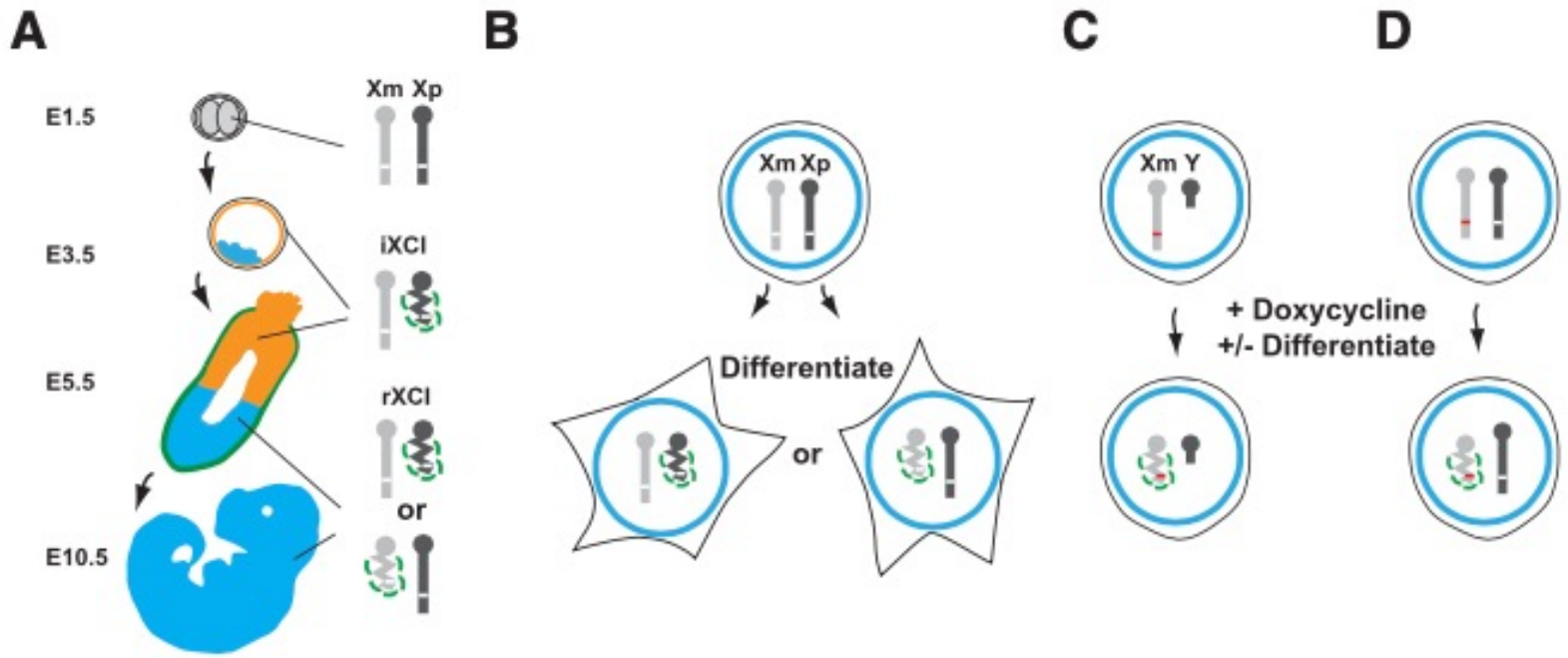
XIST acts Early During Development

The initiation of gene silencing by *Xist* has been studied in mouse ESCs, which are derived from female blastocysts and undergo XCI when they enter differentiation. Xist is required for ESC differentiation. XCI is tightly linked to cell differentiation in the epiblast and the possession of two Xa is a hallmark of pluripotent cells of both mouse ESC and iPSC.



Models for X inactivation studies

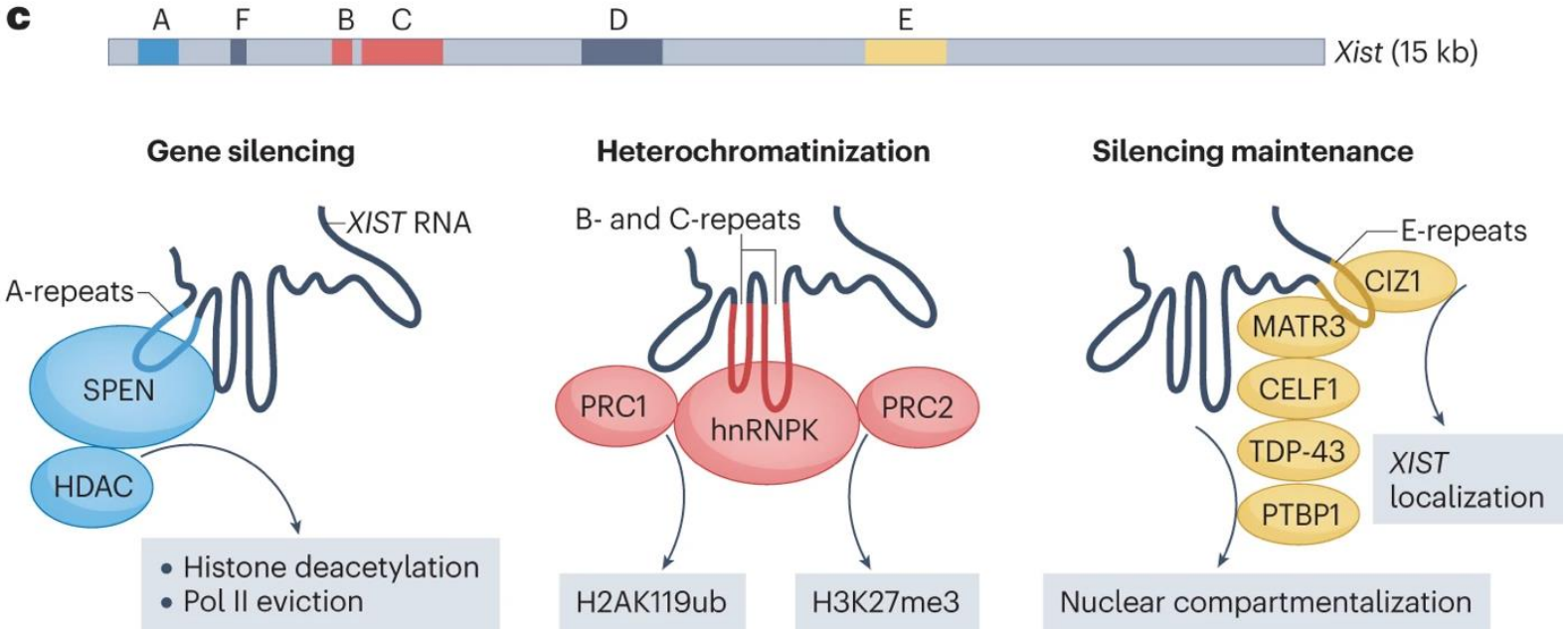
(A) **Early mouse embryos**. X inactivation (indicated as condensed bar with Xist RNA represented as green dashed line) proceeds at the two- to four-cell stage (E1.5–E2.5), initially on the paternally inherited X chromosome (Xp), referred to as imprinted X chromosome inactivation (iXCI). iXCI is maintained in trophectoderm (orange) and extraembryonic endoderm (green) derived tissues through development (shown here E3.5–E10.5), whereas in cells of the embryo proper (blue), initial reactivation of Xp at E3.5 is followed by random X inactivation (rXCI) of Xp or the maternal X chromosome (Xm) at around E5.5. (B) In XX **mouse embryonic stem cells (mESCs)** both X chromosomes are active. Random X inactivation proceeds following the onset of cellular differentiation *in vitro*. (C,D) X inactivation in mESCs can be uncoupled from differentiation by using **an inducible promoter** (red line) to drive Xist expression, illustrated on the single X chromosome in XY mESCs (C) or one of two X chromosome in XX mESCs (D). Silencing can be induced with or without cell differentiation in these models.



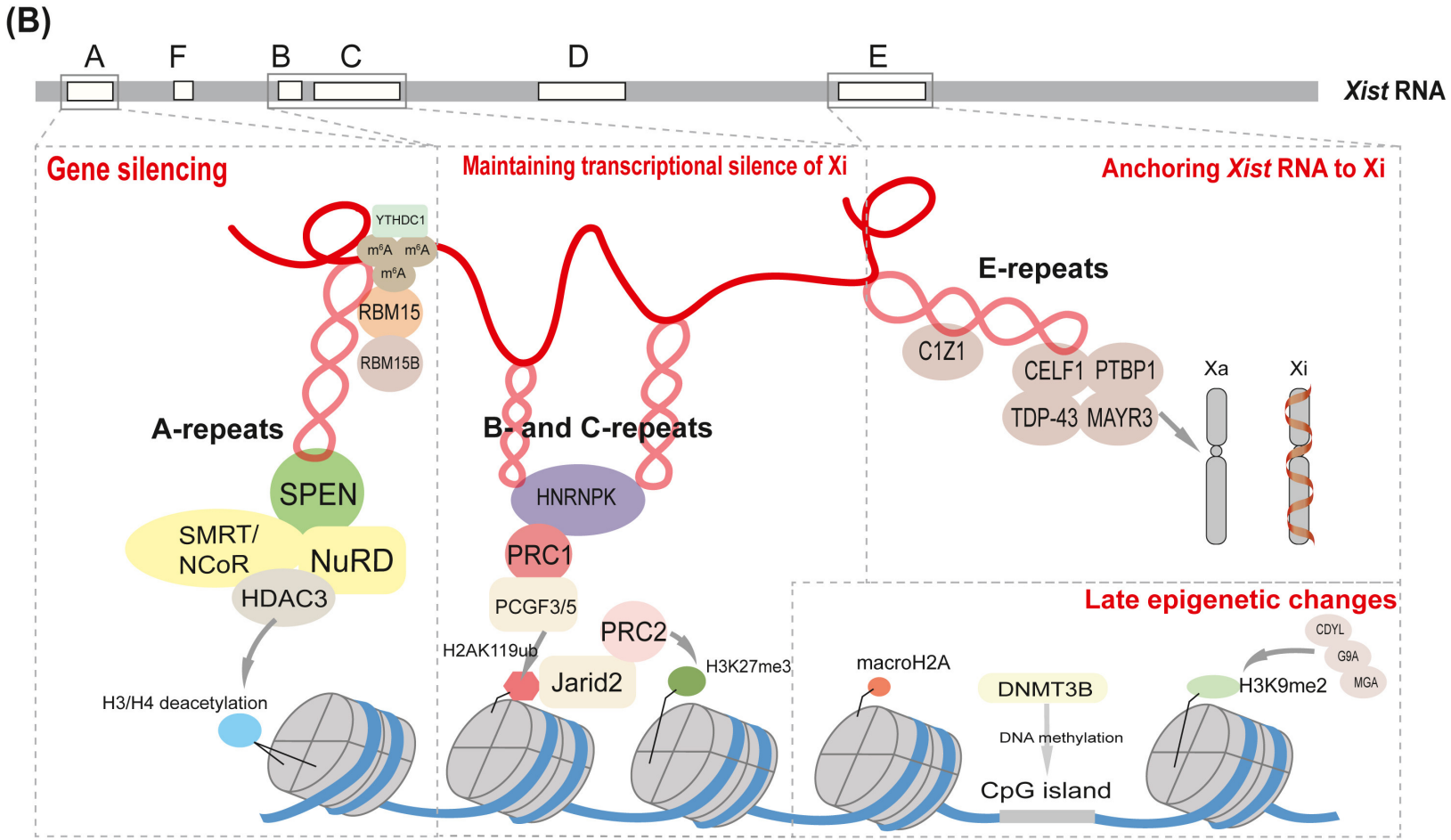
Pathways for Xist-mediated silencing

XIST RNA contains functional repeat regions, A–F.

- **A-repeats** promote the initial steps of gene silencing through SPEN- mediated and histone deacetylase (HDAC)-mediated histone deacetylation and RNA Pol II eviction.
- **B-repeats** and C-repeats induce heterochromatinization through recruitment of PRC1 and PRC2 downstream of hnRNPK.
- **E-repeats** are required for CIZ1-dependent maintenance of XIST localization at Xi and in recruiting RNA-binding proteins to mediate the nuclear compartmentalization of Xi



Pathways for Xist-mediated silencing



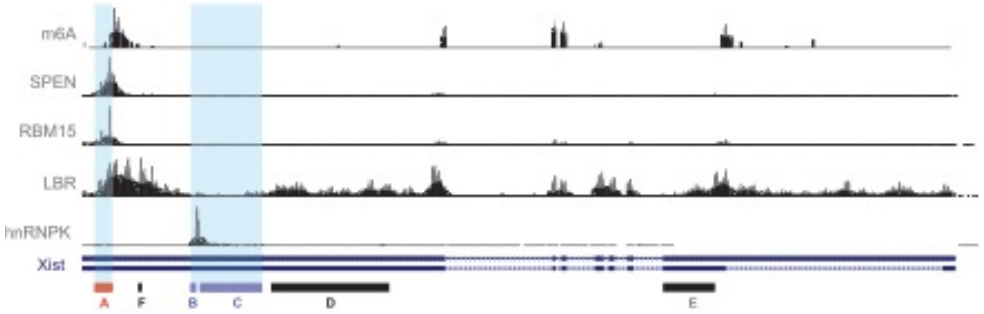
Xist RNA-binding proteins (RBPs) with a role in chromosome silencing

SPEN(SHARP)

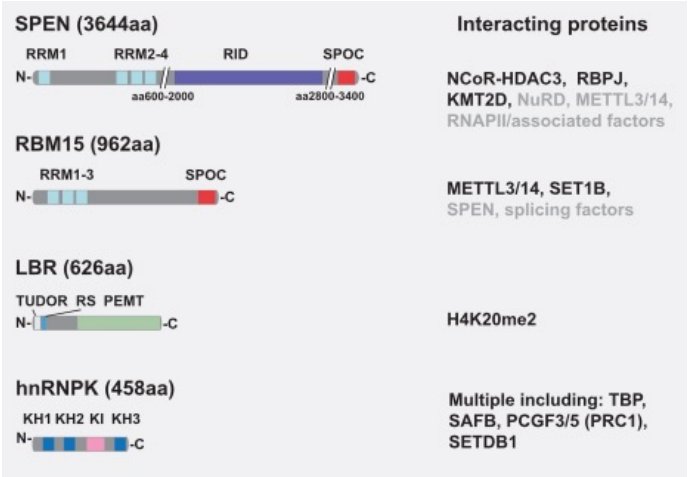
- The RNA binding protein (RBP) SPEN(SHARP) was identified as a key factor for establishment of Xist-mediated silencing in several independent studies.

- SPEN interacts with the SMRT/NCoR–HDAC3 histone deacetylase complex, which represses gene transcription and sequester the H3K4 methyltransferase coactivator complex KMT2D.

CLIP data



A-repeat and B/C-repeat regions are highlighted with pale-blue shading

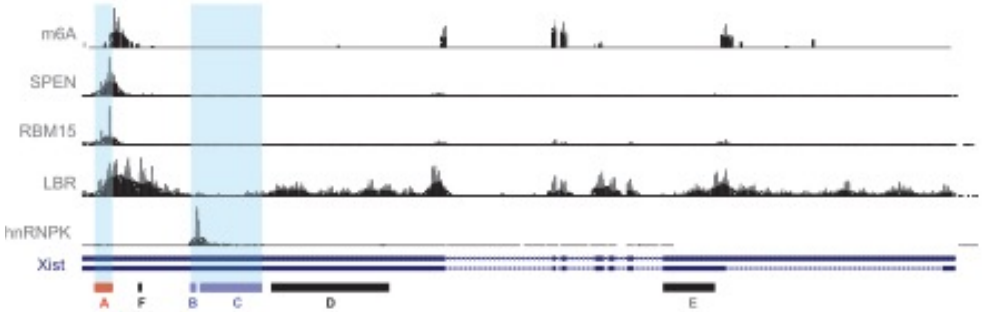


Xist RNA-binding proteins (RBPs) with a role in chromosome silencing

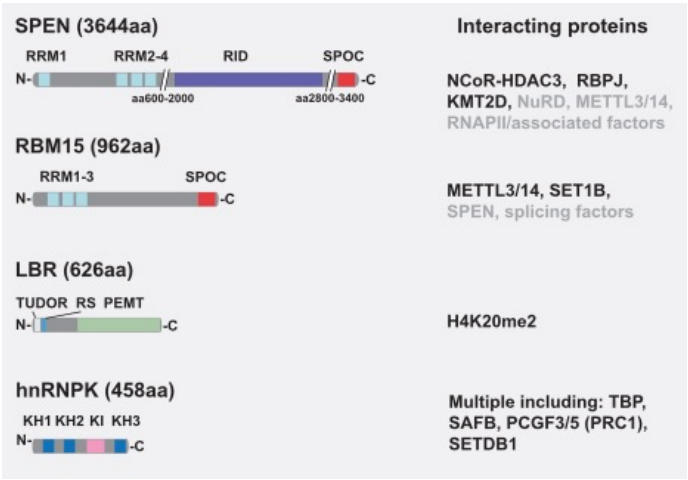
hnRNPK

- the RBP hnRNPK bound to the Xist RNA B/C repeat initiates PRC1 recruitment and H2AK119ub1 deposition followed by PRC2-mediated H3K27me3.
- However, the relevance of PRC activity on Xi is still under debate.
- Similar to SPEN, at least some of the observed silencing deficiency in the absence of the Polycomb system may result from perturbation of Xist–RNA localization rather than Xi chromatin modification.

CLIP data



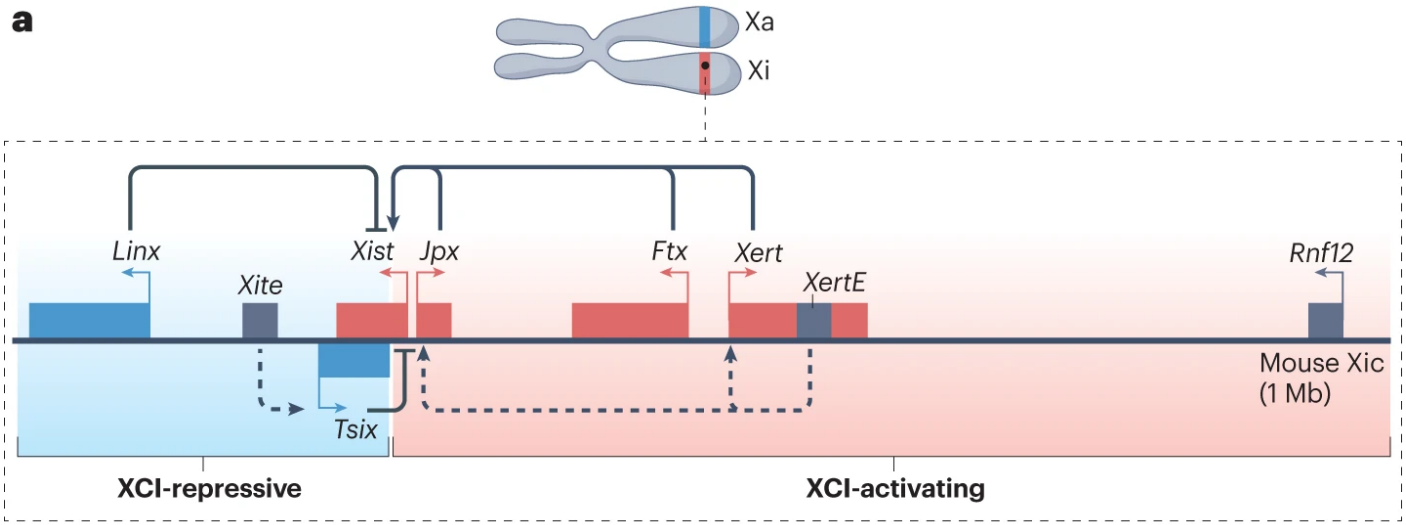
A-repeat and B/C-repeat regions are highlighted with pale-blue shading



Xic lncRNAs

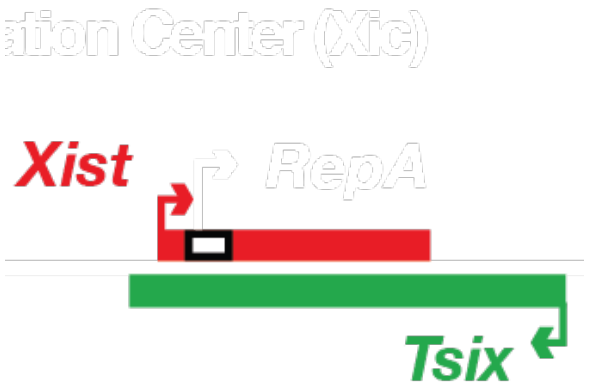
The X-inactivation center contains multiple genes encoding lncRNAs, including **Xist**, **Tsix**, **Xite**, **Jpx/Enox**, **Ftx**, and **Tsx**. Many of these loci regulated Xist epression, some acting negatively (e.g., Tsix), others positively (e.g., Jpx).

However, whether these loci act through their transcription, RNA product or DNA elements, and at which of these levels Xist is regulated, remain open questions.



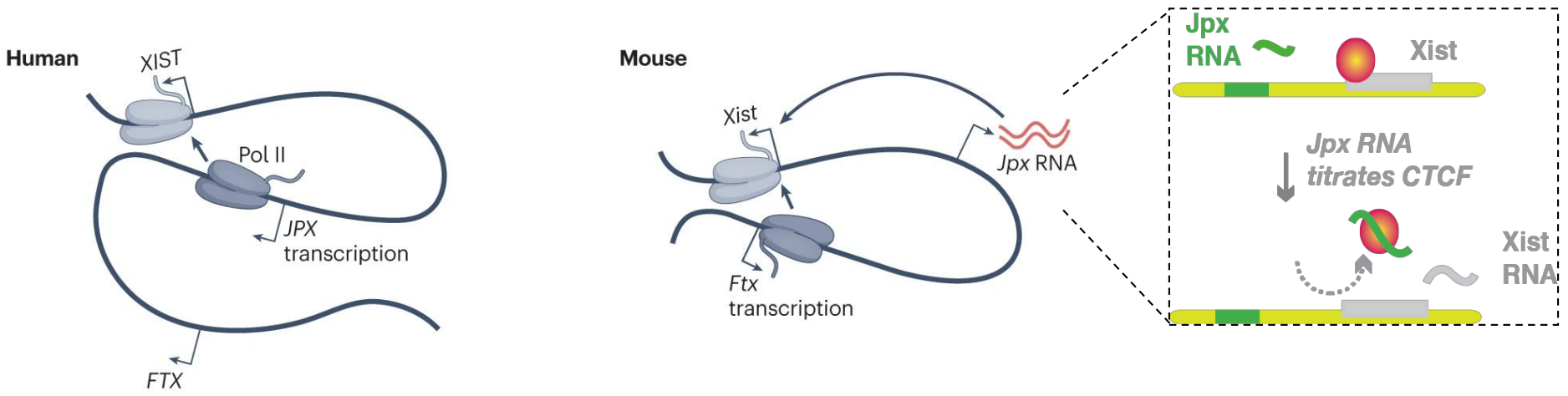
Tsix RNA repressed Xist expression

Tsix is transcribed antisense across Xist, and represses *Xist* by preventing its upregulation *in cis*. Current evidence points to the act of transcription, rather than Tsix RNA, acting repressively on Xist expression.



Similarities and differences between XIST regulation in human and mouse

In human, whereas FTX is not essential for XIST regulation, JPX transcription, but not the mature RNA, contributes to RNA Pol II loading and XIST transcription and accumulation. In mouse, Ftx transcription promotes Xist transcription, whereas mature Jpx transcript is responsible for Xist transcriptional activation and accumulation. CTCF (CCCTC-binding factor) represses Xist transcription. At the onset of XCI, Jpx RNA is upregulated, binds CTCF, and extricates CTCF from one Xist allele.



The regulatory network of Xist

Pluripotency factors repress Xist expression either by acting on the locus directly or indirectly through Tsix stimulation. **Rnf12** and upstream cis-regulators of Xist promote its upregulation in the absence of the pluripotency-mediated repression. More recently, factors such as **YY1** and **DNMT1** have been implicated as trans-acting activators of Xist expression, while **MOF-associated complexes** have a negative effect. Pluripotency regulates XCI through Xist and initiation of XCI feeds back on the pluripotency network, so that cell differentiation does not proceed without dosage compensation being achieved.