Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre

Di Santo Martina
Senci Sergio
Zangrando Luca

Seminar Outline

- *types and roles of ncRNAs*

PART 1: Genome structure, an overview

- *Nuclear spatial organization*
- *Nuclear compartmentalization*
  - *TADs*
  - LADs

PART 2: Xist as a key example of genome-organizing ncRNAs

- Xist repeats
- *XIC overview*
- *Xist localization*
- *hnRNPU structure and functions*

PART 3: Linking ncRNAs to cell-differentiation

- *general features of differentiation*
- *discovery and validation of Firre lincRNA*

PART 4: Characterization of the topological distribution of Firre with functional implications

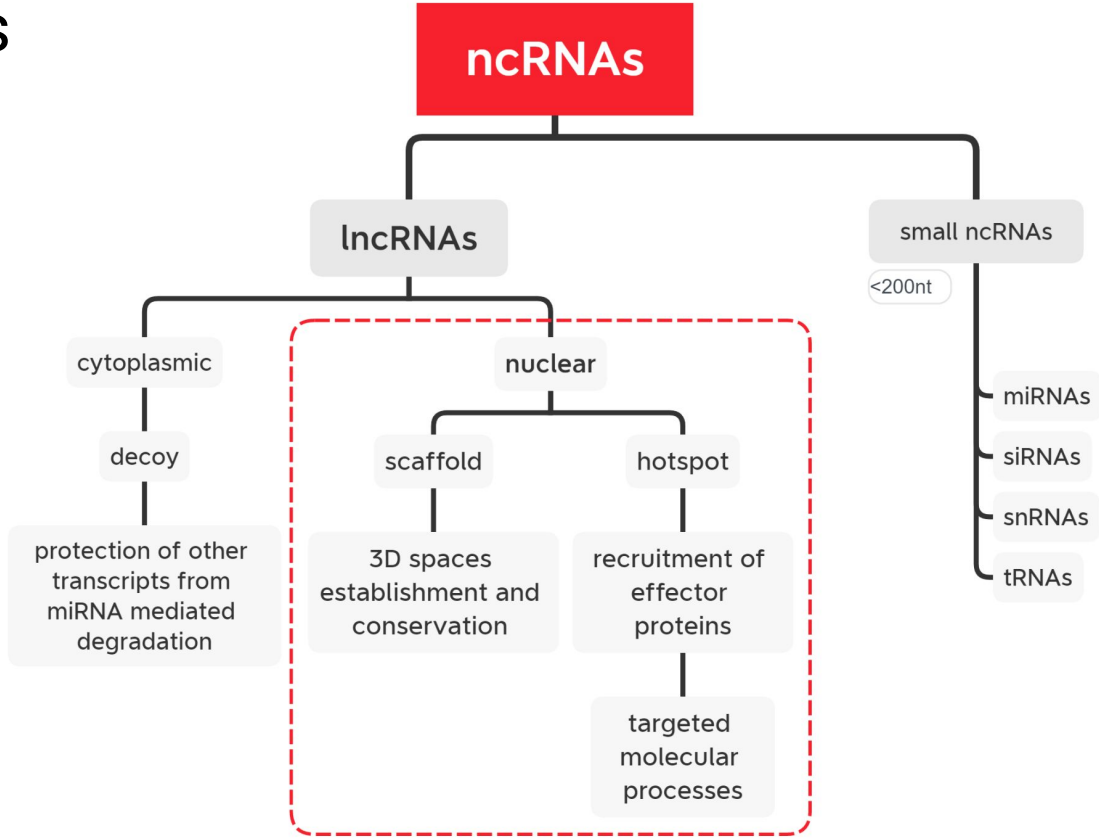
- *Firre structural features*
- Nuclear localization of Firre
- *Cis and Trans-chromosomal interactions*
- Loss-of function models
- Molecular interactors *of Firre: hnRNPU*

Conclusions

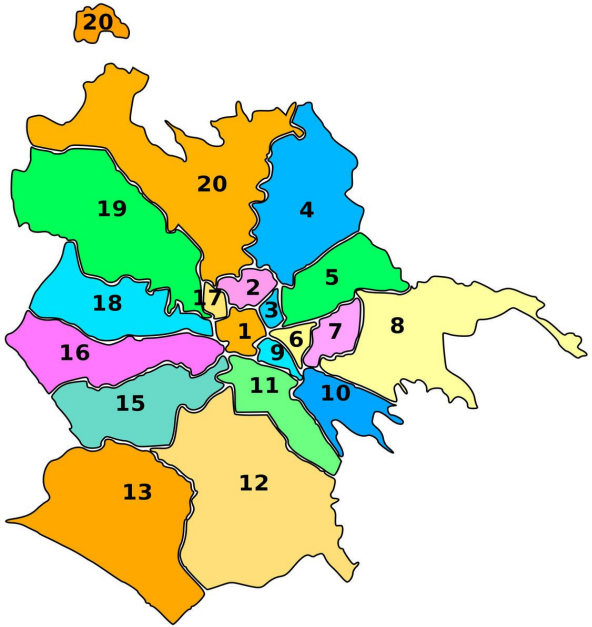
Long non-coding RNAs

Non-coding RNAs (**ncRNAs**) are RNA molecules that are not translated into a protein.

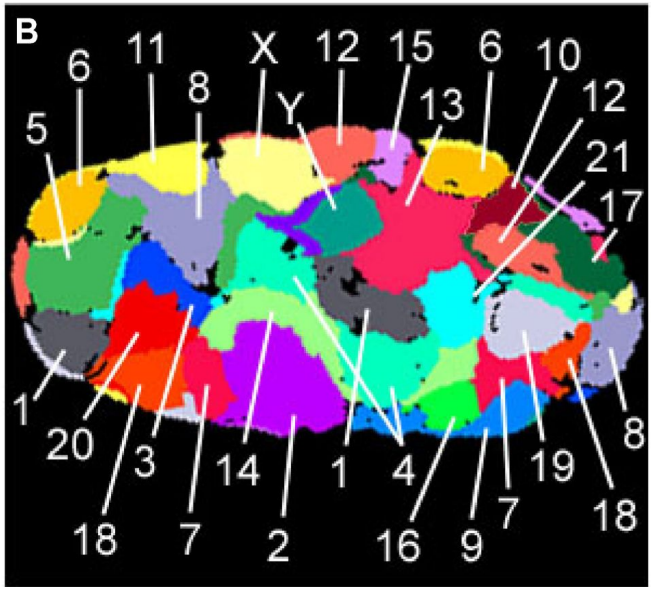
LncRNAs are involved in epigenetic regulation, mainly modulating higher order chromatin structures and epigenetic histone marks.



Nuclear spatial organization



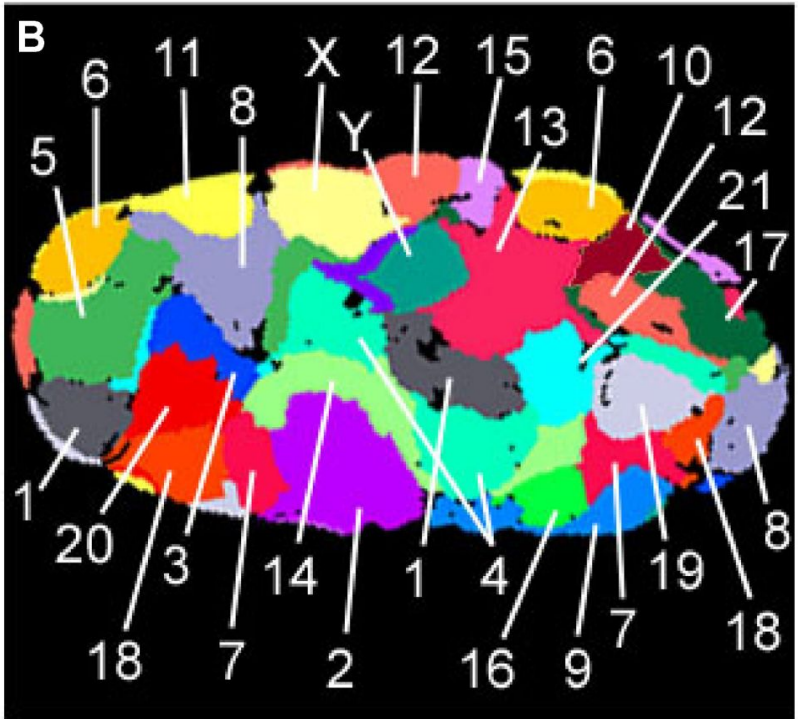
Neighborhoods of Rome



Chromosomal territories

Distinctive internal patterns of organization

Nuclear spatial organization



- Chromosomes occupy cell specific locations in the nucleus during interphase.
- Spatial proximity is fundamental for co-regulation of independent loci.
- Spatial proximity is the result of a number of targeted molecular interactions

Nuclear compartmentalization

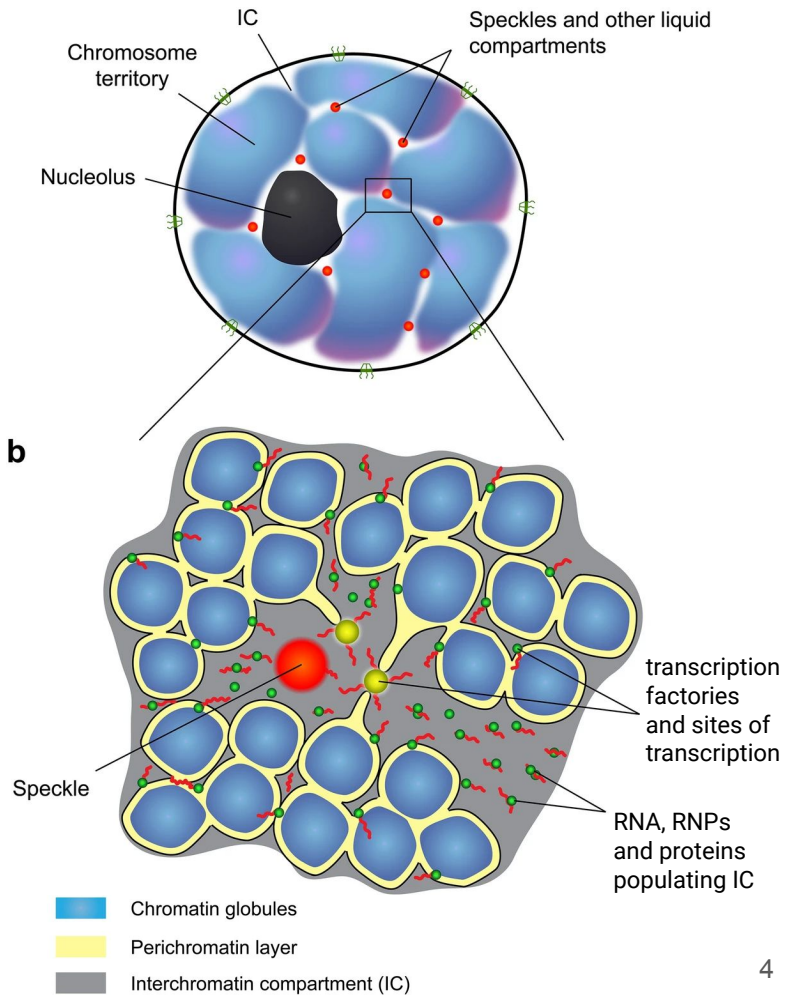
Chromosomal territories

- ❏ are composed by **chromatin globules.**
- ❏ surrounded by **Inter-chromatin space** (IC)

Transcription usually occurs at the **perichromatin layer** (border of chromatin globules and IC).

V

Then, transcripts end up in the **IC** and exploit it to **reach nuclear pore complexes**.

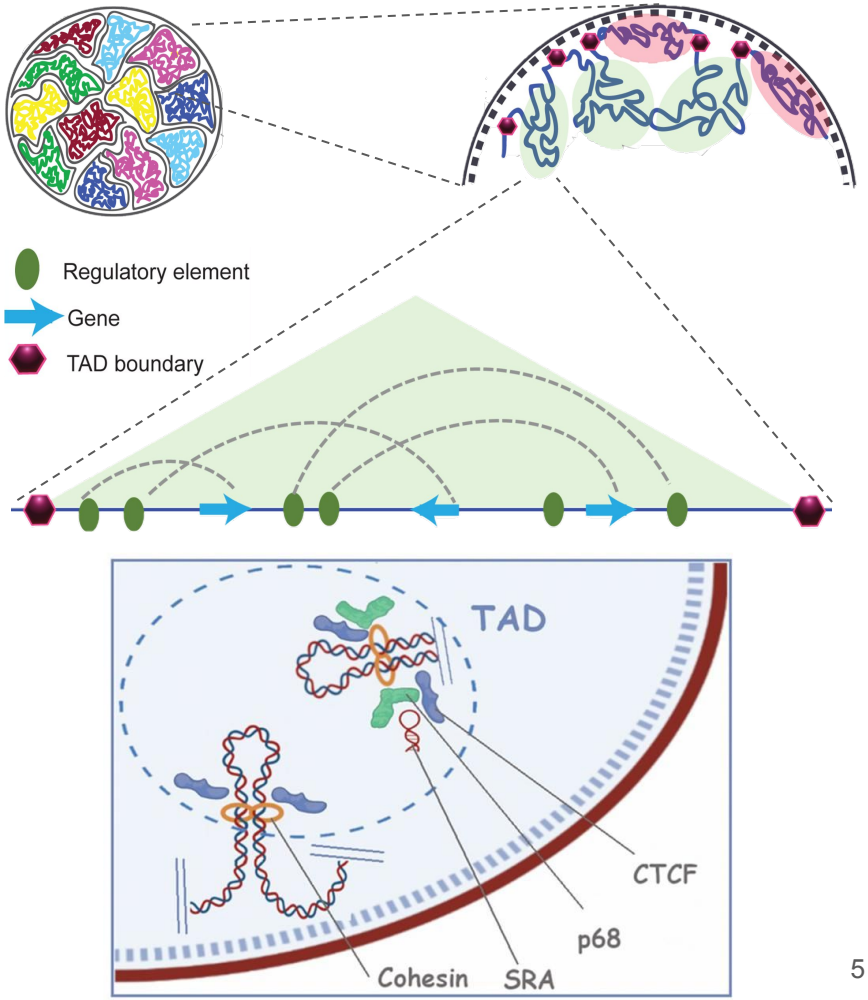


TADs

TADs exhibit more **intradomain interactions** than interdomain ones.

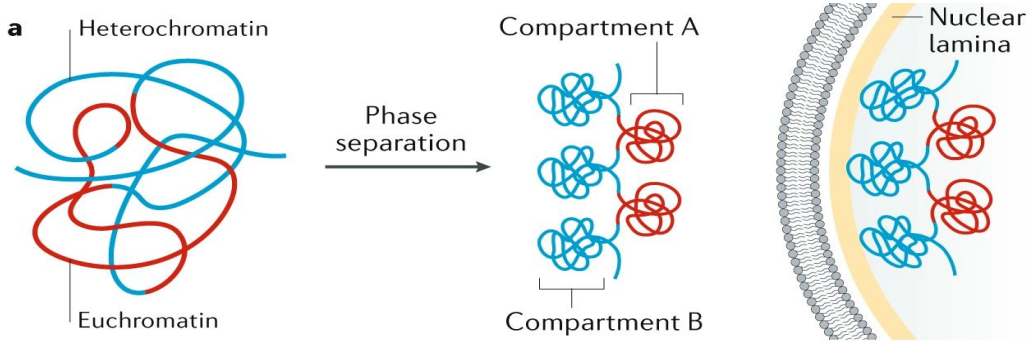
TADs are enclosed by **TAD boundaries**, which are enriched for the insulator binding protein **CTCF**, housekeeping genes, transfer RNAs, and SINEs.

TADs formation happens by **loop extrusion**, mediated by **cohesin assembly** only at insulator sequences bound by **CTCF**. Cohesin-CTCF interaction is tightly regulated and can be modulated by a number of other actors including **p68** (RBP) and its associated ncRNA, called **Steroid receptor RNA Activator** (SRA).



CTCF: CCCTCF binding factor.

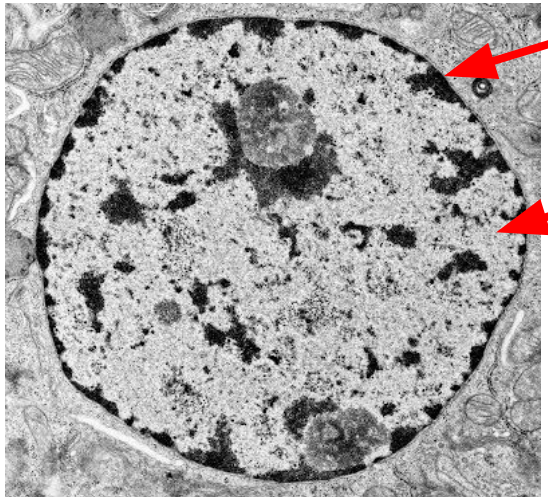
Higher order of TADs: type A and type B compartments

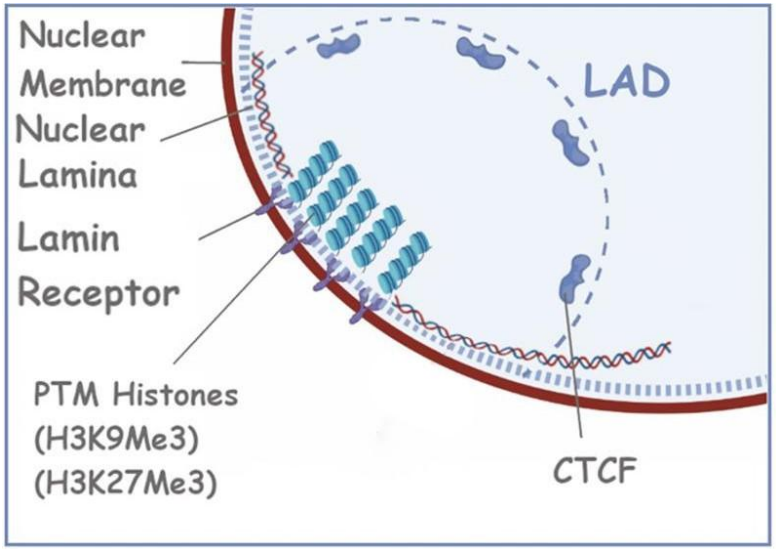


- The organization of the genome into TADs is critical for coordinated transcriptional regulation, chromatin states, and DNA replication.
- These structures dynamically change during differentiation and are perturbed in disease. Over the years, multiple studies have described the role of different types of ncRNAs in TADs formation and maintenance.
- Generally, the mammalian genome is arranged into compartments of active and inactive chromatin.
- Moreover, linearly non-contiguous TADs can contact each other, defining long-range interactions that can vary between cell types and during differentiation

Lamina-associated domains (LADs)

The **nuclear lamina** is a mesh of **lamin** proteins. It's the key organizer of the radial arrangement of chromatin, by creating a nuclear compartment where inactive chromatin clusters in **lamina-associated domains** (**LADs**).



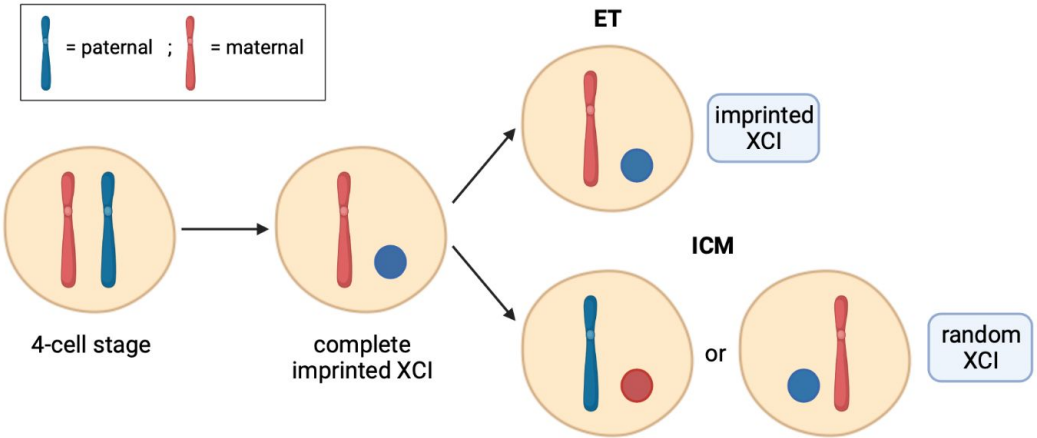


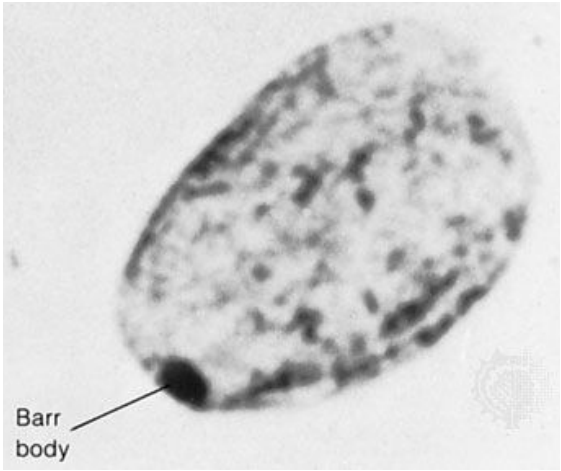
LADs (0.1–10 Mb) are usually **gene-poor**, **heterochromatic**, and show low gene activity. LADs aids functional organization of the genome and enables a spatio-temporal regulation of replication and transcription.

X Chromosome Inactivation (XCI)

As females have two X chromosomes, **X-inactivation** prevents an abnormal protein dosage and other lethal consequences. In mouse the event mainly proceeds <u>in 2 steps</u>:

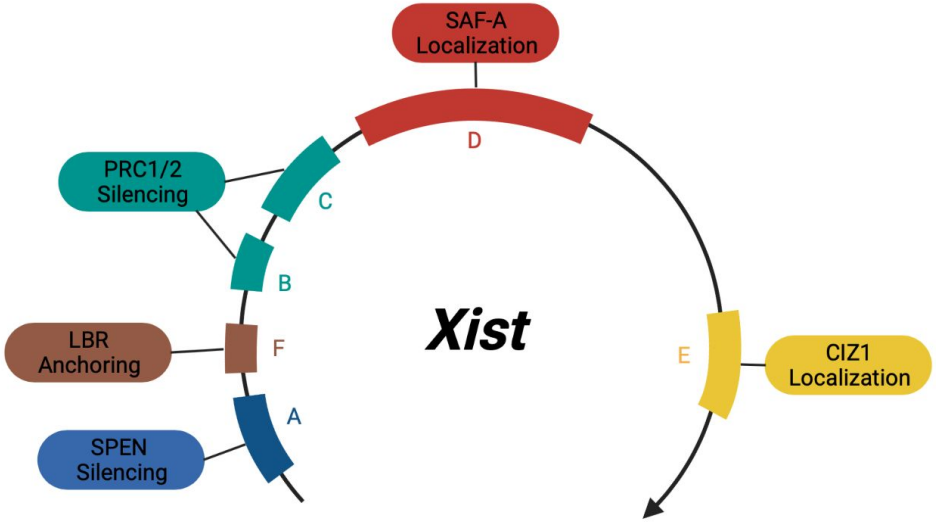
1. At 4-cell stage → paternal X chromosome is inactivated and then reactivated;
2. At implantation stage → random XCI only in cells of the inner cell mass.





Key example: *Xist*

Xist *(X inactive specific transcript)* is a lncRNA required for **X chromosome inactivation (XCI)**. *Xist* is mainly expressed by the inactivated X, then it spreads in *cis* and **acts as hotspot** for the recruitment of several factors.

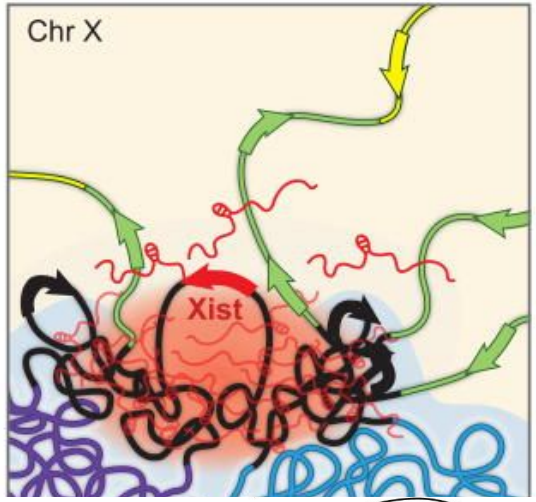


Repeat A binds SPEN (SHARP), recruiting HDAC3
Repeat B and C bind hnRNPK, recruiting PRC1/2
Repeat D binds hnRNPU (SAF-A)
Repeat E binds CIZ1
Repeat F (together with A/E) binds lamin B receptor, anchoring the heterochromatic Xi to the nuclear periphery (Barr body).

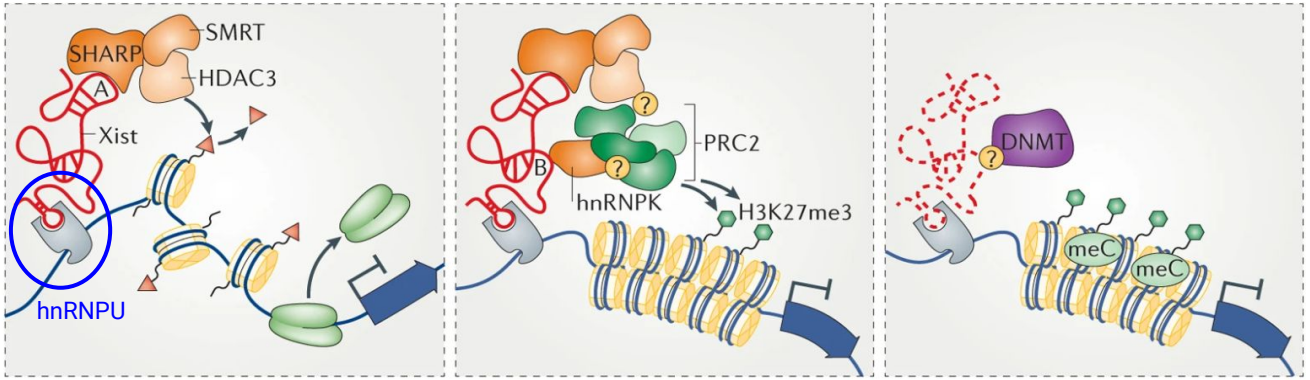
SPEN: SPlit ENDs; SHARP: SMRT/HDAC1 Associated Repressor Protein ; SAF-A: Scaffold Attachment Factor A ; CIZ1: CDKN1A Interacting Zinc Finger Protein 1; PRC: Polycomb Repressive Complex.

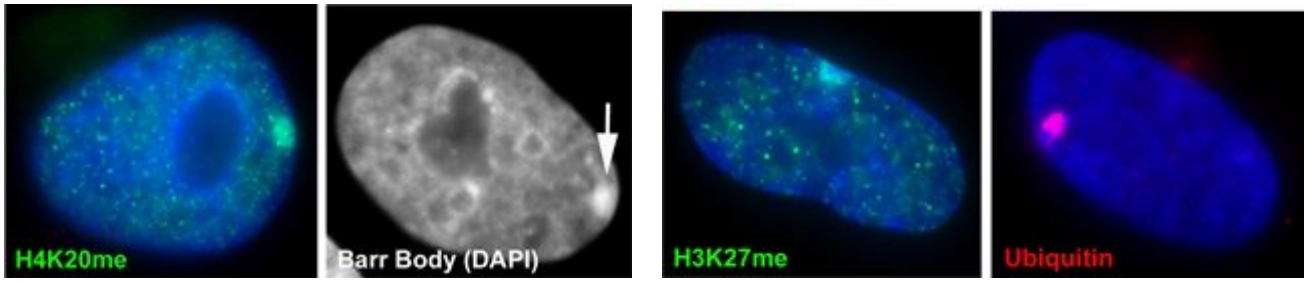
XCI: molecular mechanism

1. *Xist* expression is repressed by the binding between **CTCF** at *Xist* promoter.
2. ***Jpx*** (lncRNA) activates *Xist* expression by evicting CTCF protein.
3. *Xist* is transcribed.
4. *Xist* spreads in *cis* → entire X chr covered. Xist localization depends on **hnRNP-U** (Heterogeneous Nuclear Ribo-Nucleo-Protein U, previously called Scaffold Attachment Factor - A, SAF-A)
5. Establishment of the inactive state interacting with:
   a. **hnRNPK** → recruits **PRC1** → H2K119ubi → recruitment of **PRC2** → H3K27me3;
   b. **SHARP** (SPEN) → recruits **HDAC3** → silencing of transcription



Xist localization depends on hnRNP-U (SAF-A)





- Xist is transcribed and it diffuses to the **interchromatin space.**
- It is bound and stabilized by hnRNPU.
- There, it recruits the effectors responsible for heterochromatin formation.

hnRNP-U (SAF-A)

Heterogeneous Nuclear Ribo-Nucleo-Protein U

hnRNP-U can bind AT-rich dsDNA sequences, known as **SARs** (*Scaffold Attachment Regions*).

<u>hnRNP-U:</u>

- ❏ Is required for the maintenance of nuclear compartmentalization.
- ❏ Consistently binds to CTCF and RAD21 at active chromatin sites.
- ❏ Interacts with lncRNAs;
- ❏ Is involved in higher order chromatin organization.

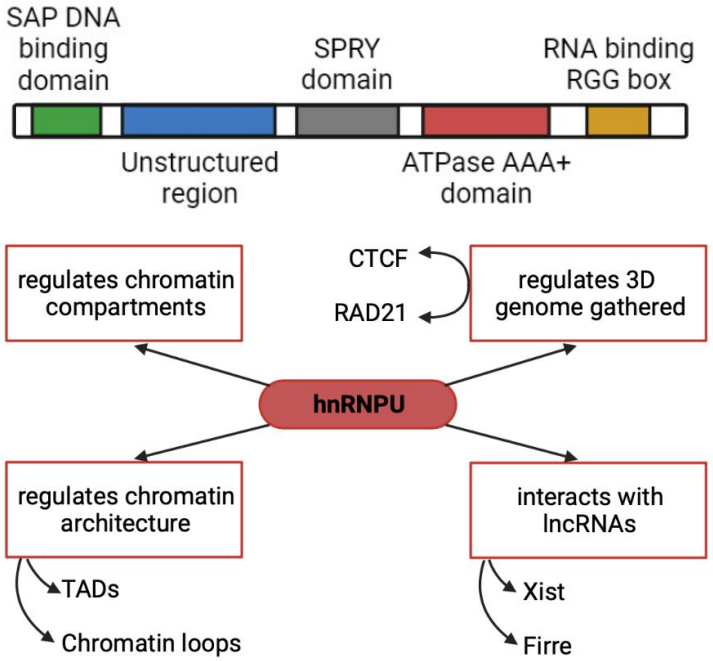
Review > Semin Cell Dev Biol. 2019 Jun;90:161-167. doi: 10.1016/j.semcdb.2018.07.006. Epub 2018 Jul 20.

The role of nuclear matrix protein HNRNPU in maintaining the architecture of 3D genome

Linlin Zhang [1], Dongli Song [1], Bijun Zhu [1], Xiangdong Wang [2]

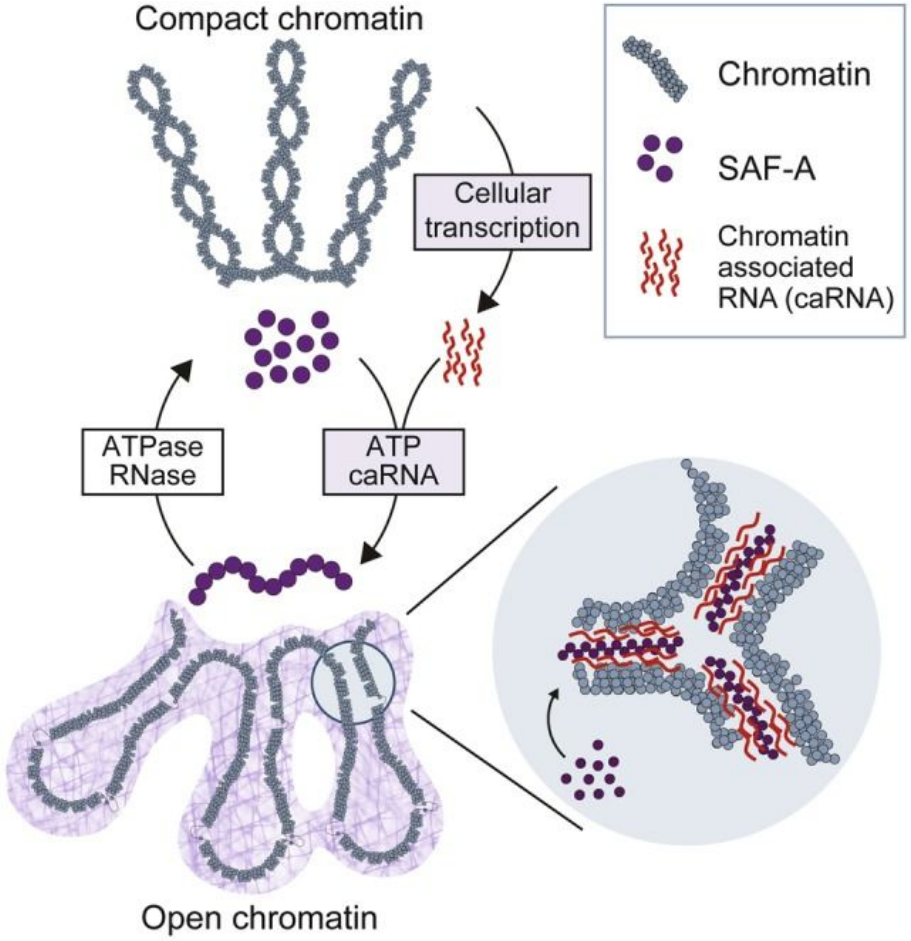
Affiliations + expand

PMID: 29981443 DOI: 10.1016/j.semcdb.2018.07.006

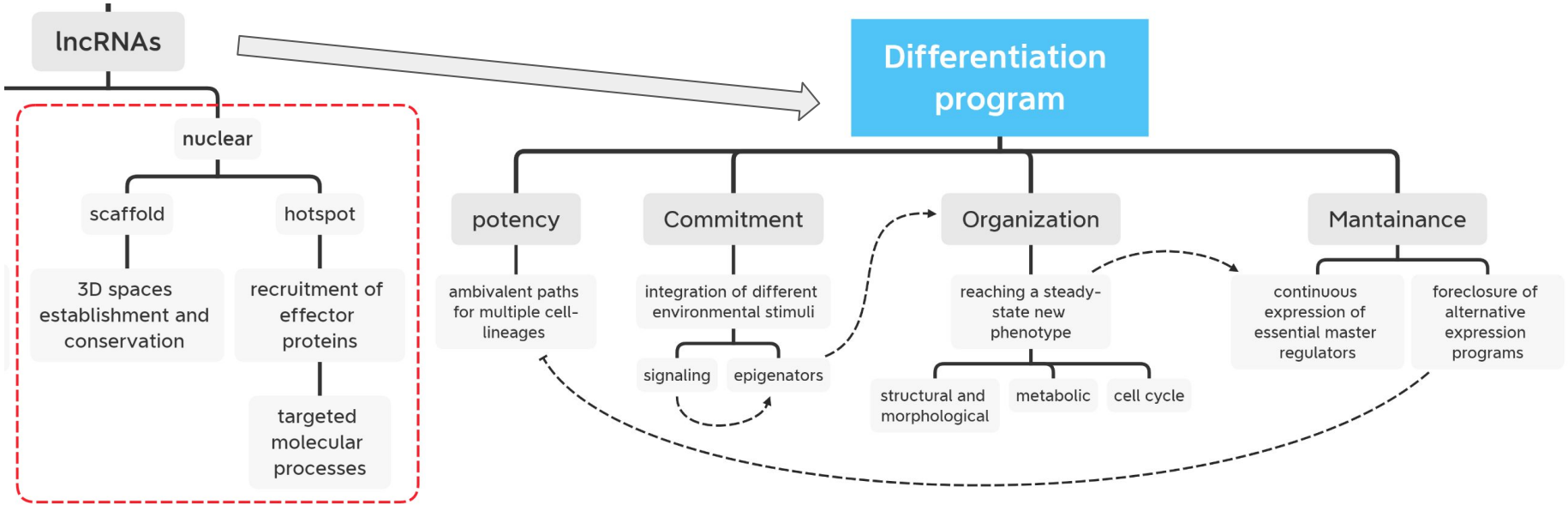


hnRNP-U RNA binding activity

- Is mediated by RGG domain.
- Is exploited to target chromatin associated lncRNA localization.
- hnRNP-U oligomerizes via ATP binding in vivo and through assembly and disassembly regulates large scale chromatin structure in response to RNA binding.



Linking lncRNAs to cell differentiation



Firre: the discovery

Firre was first identified as a regulator of adipocyte precursors differentiation and so named **lnc-RAP-1** (Regulated in Adipogenesis-1).

It was shown to **positively regulate** the expression of **four key adipocyte markers**: *Pparγ*, *Cebpα*, *Fabp4*, and *AdipoQ*.

- ❏ How *Firre* was discovered?
- ❏ Is it really relevant in adipogenesis?

› Proc Natl Acad Sci U S A. 2013 Feb 26;110(9):3387-92. doi: 10.1073/pnas.1222643110. Epub 2013 Feb 11.

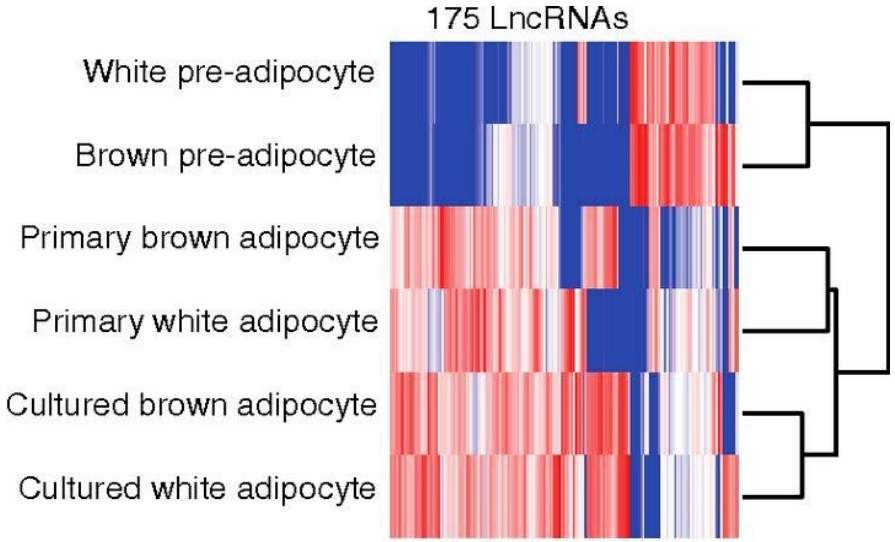
Long noncoding RNAs regulate adipogenesis

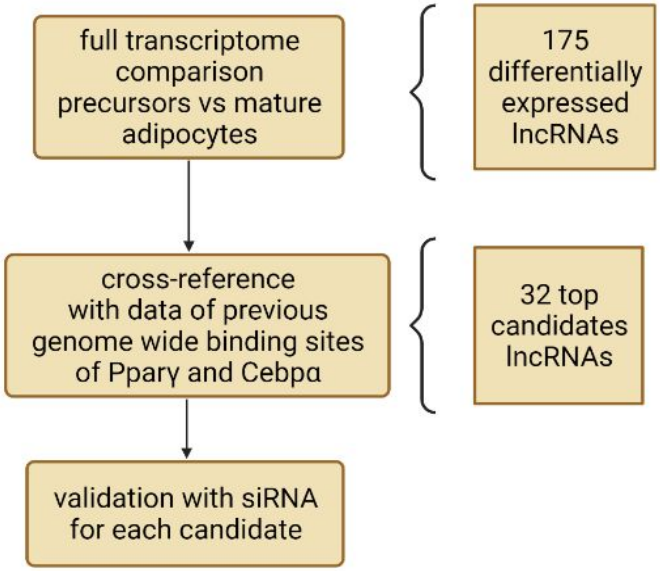
Lei Sun [1], Loyal A Goff, Cole Trapnell, Ryan Alexander, Kinyui Alice Lo, Ezgi Hacisuleyman, Martin Sauvageau, Barbara Tazon-Vega, David R Kelley, David G Hendrickson, Bingbing Yuan, Manolis Kellis, Harvey F Lodish, John L Rinn

Pparγ: Peroxisome proliferator-activated receptor gamma; *Cebpα:* CCAAT/enhancer-binding protein alpha; *Fabp4:* fatty-acid-binding protein; *AdipoQ:* adiponectin.

How *Firre* was discovered

Top-down approach: precursors vs mature adipocytes full transcriptome comparison by RNA-seq



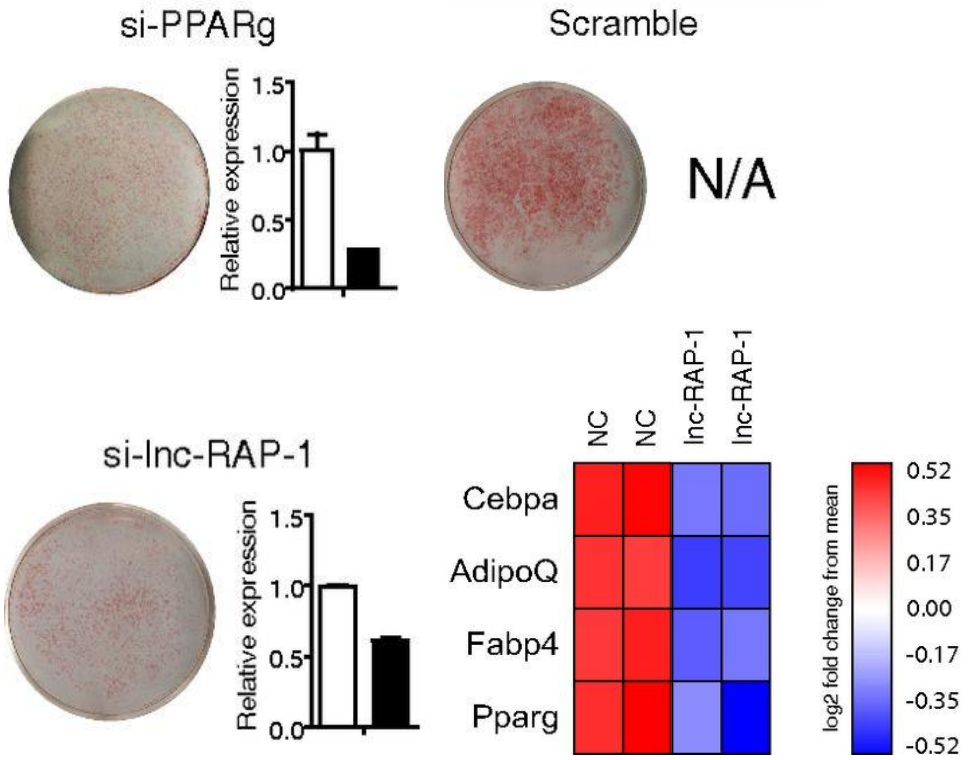


Key point: lncRNA genes are **bound** and **activated** by **transcription factors** responsible for **coordinating adipogenesis**.

How *Firre* was discovered: validation

Primary white preadipocytes were isolated and cultured. One day before differentiation, cells were transfected with siRNAs for each lncRNA.
After 4 days of differentiation, lipid accumulation was detect by ORO staining and RT-PCR was performed to examine the expression of each lncRNA and several adipogenesis marker genes.

Conclusion: together with other lncRNAs, **lnc-RAP-1** (*Firre*) plays a key role in **induction** of multiple **adipocyte-specific genes**.



How *Firre* was studied?

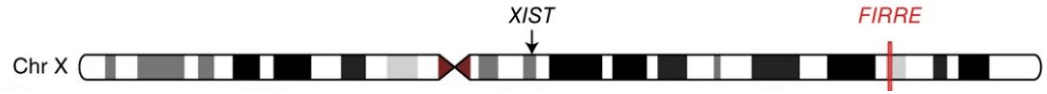
nature structural & molecular biology

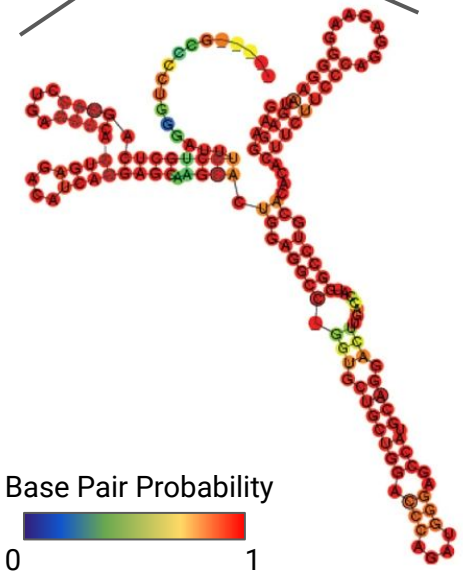
› Nat Struct Mol Biol. 2014 Feb;21(2):198-206. doi: 10.1038/nsmb.2764. Epub 2014 Jan 26.

Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre

Ezgi Hacisuleyman [1], Loyal A Goff [2], Cole Trapnell [3], Adam Williams [4], Jorge Henao-Mejia [4], Lei Sun [5], Patrick McClanahan [6], David G Hendrickson [3], Martin Sauvageau [3], David R Kelley [3], Michael Morse [7], Jesse Engreitz [7], Eric S Lander [7], Mitch Guttman [8], Harvey F Lodish [9], Richard Flavell [10], Arjun Raj [6], John L Rinn [11]

Firre

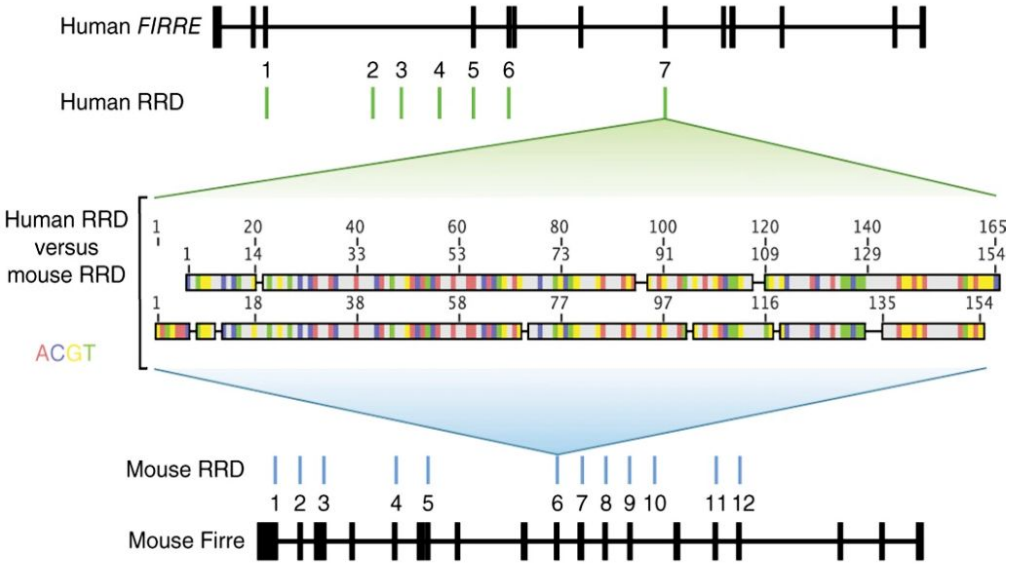




- *Firre* (Functional Intergenic Repeating RNA Element) is a lncRNA whose locus is contained in the Xq.
- Firre contains a peculiar 156-bp **Repeating RNA Domain (RRD)**.
- The RRD is unique to *Firre* and is not found in any other genomic region, and a secondary structure-prediction tool (AliFold) revealed that the RRD is highly structured.

Human and murine *Firre*

Firre presents **numerous alternatively spliced isoforms** with differential inclusion or exclusion of RRD sequences.



RRD occurs **16** times in *Mus musculus* and **8** times in *Homo sapiens*.

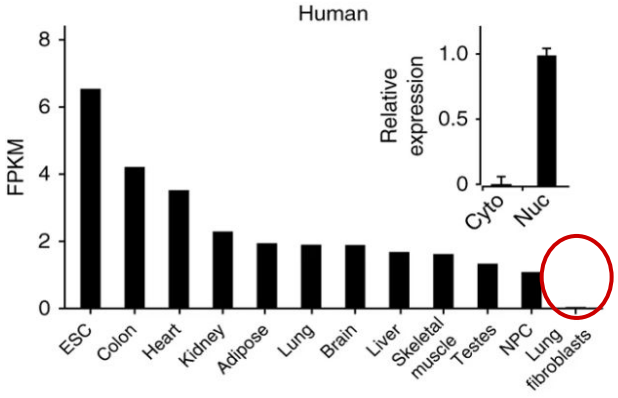
RRDs present:

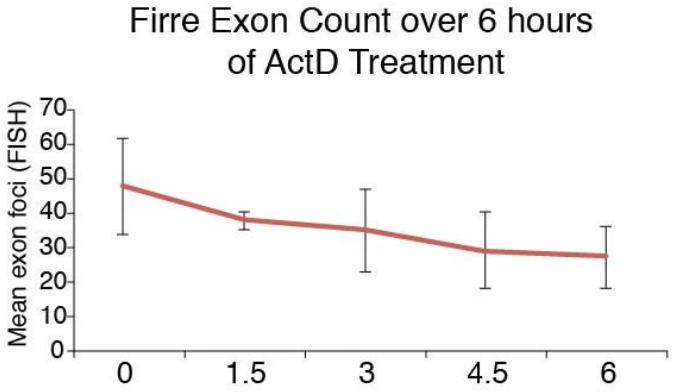
- ❏ a **96%** sequence identity **within** species;
- ❏ **68% across species**.

Firre features

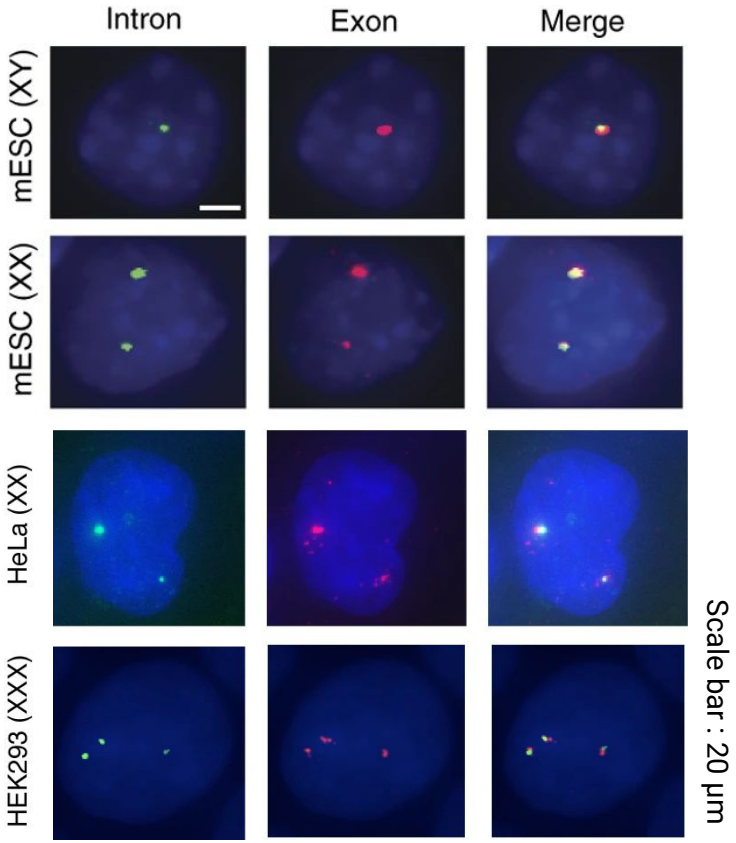
Firre has also:

- A **different expression pattern *in vivo*** (RT-qPCR data)**;**
- A strong **stability** → *Firre* transcripts remains stable even after 6 hours of Actinomycin-D (ActD) treatment.





Nuclear localization of *Firre*

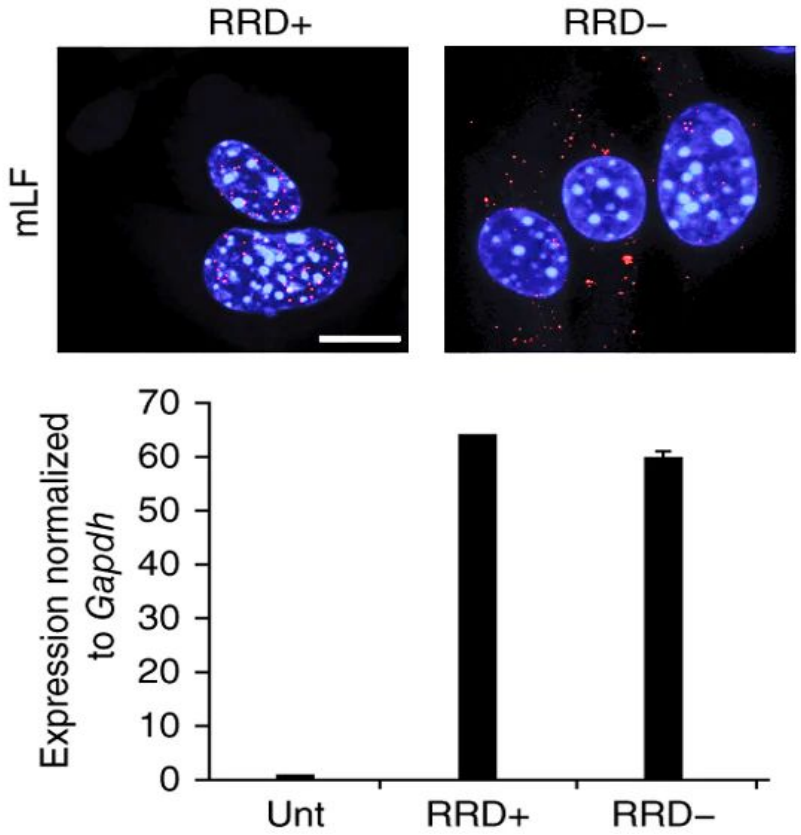


Dual labelling strategy RNA FISH:
- intronic probes to mark the transcription site of *Firre*
- exonic probes to target the mature *Firre* transcript localization.

Colocalization observed, with a slightly extension of the mature *Firre* transcript over the transcription site. Same results obtained in hESC.

Firre is **nuclear localized** and forms **expression foci** on both X chromosomes **before and after X-chromosome inactivation** (XIC).

Is RRD involved in the nuclear localization of *Firre*?



Viral overexpression of two isoforms of *Firre*, one containing and one not containing the RRD.

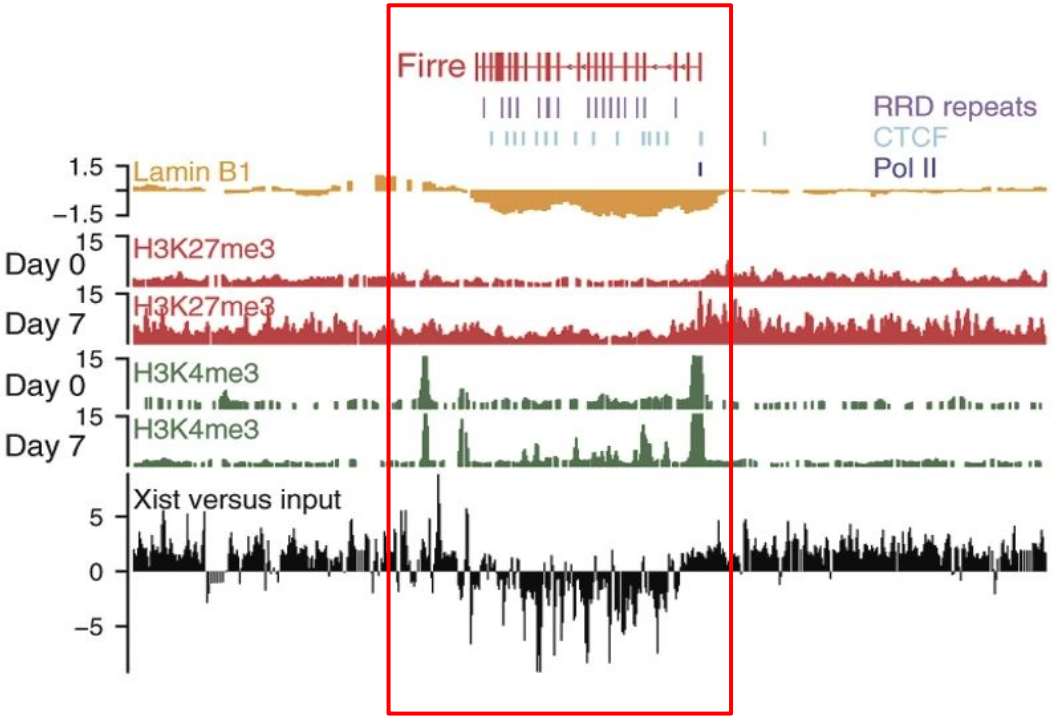
Lung fibroblast cells were used because they do not express *Firre*. Same results obtained with hLFs.

RNA FISH of the RRD- construct provide evidence for nuclear exportation.

The RRD domain is **required** for the **focal nuclear localization** of *Firre*.

Does the *Firre* locus escape XCI?

ChIP-seq data for the Firre locus during female mESC development



colocalization with CTCF (insulator molecule)

depletion of lamin B1 (heterochromatin marker)

depletion of epigenetic repressive markers and enrichment in activating markers

strong and focal depletion of Xist

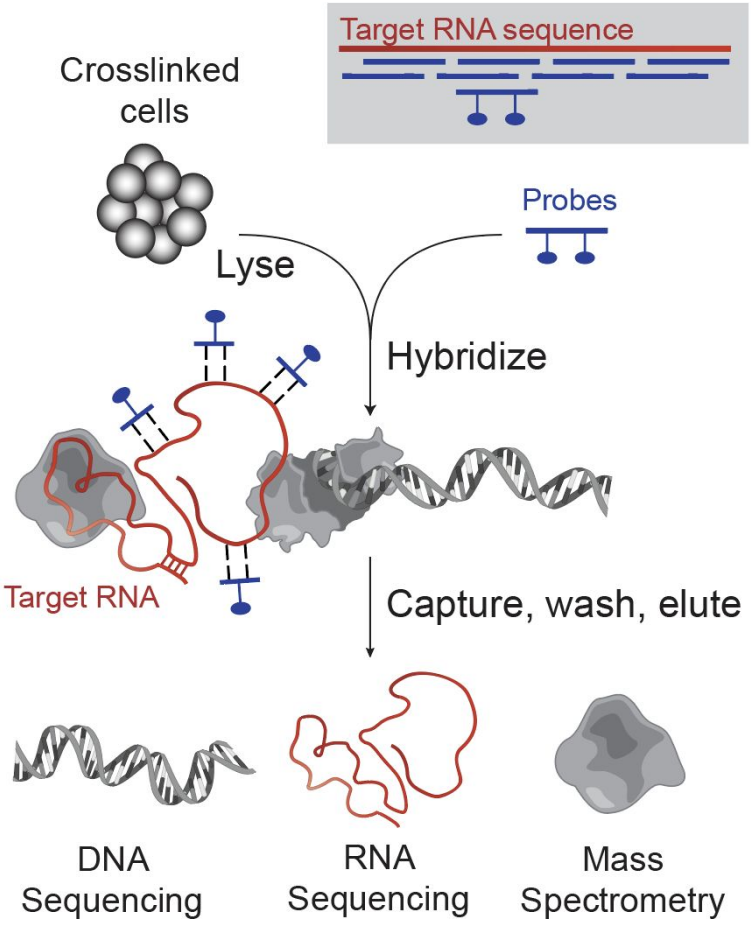
Firre locus **escapes XCI**.

RAP: RNA Antisense Purification

Technique used to analyze genome-wide DNA binding sites of a RNA and eventually the proteins bound by that RNA.

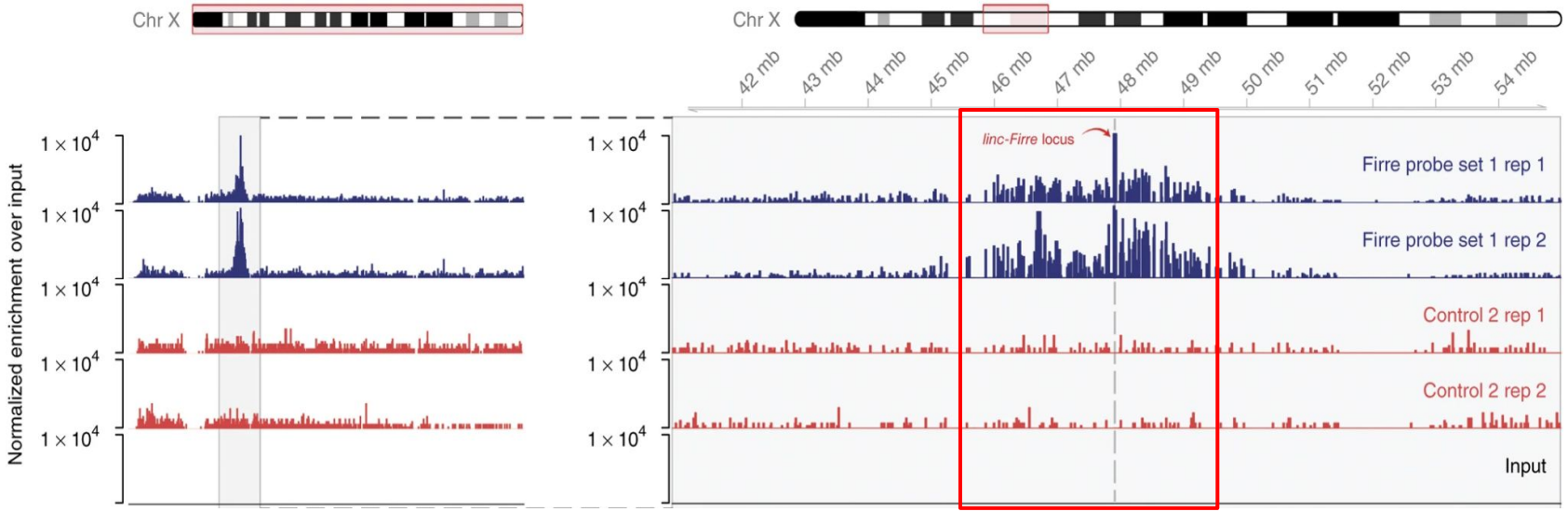
Steps:

1. Biotinylated antisense oligos for the RNA of interest,
2. Cross-link cells and lysis,
3. Hybridization,
4. Pulldown using streptavidin beads,
5. DNA seq or mass spectrometry.



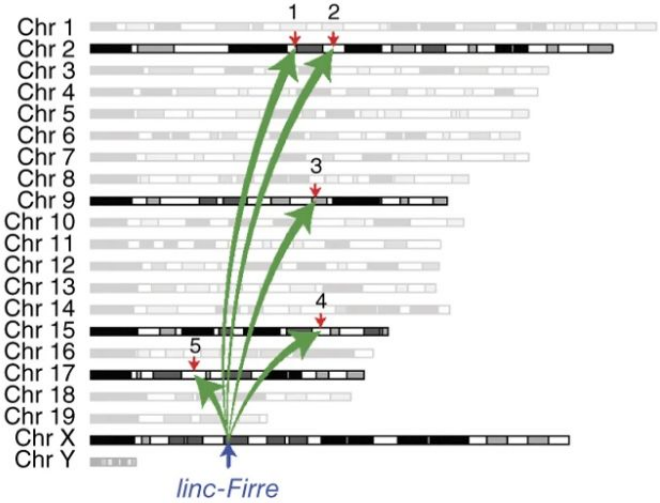
Where does *Firre* localize in the chromatin *in cis*?

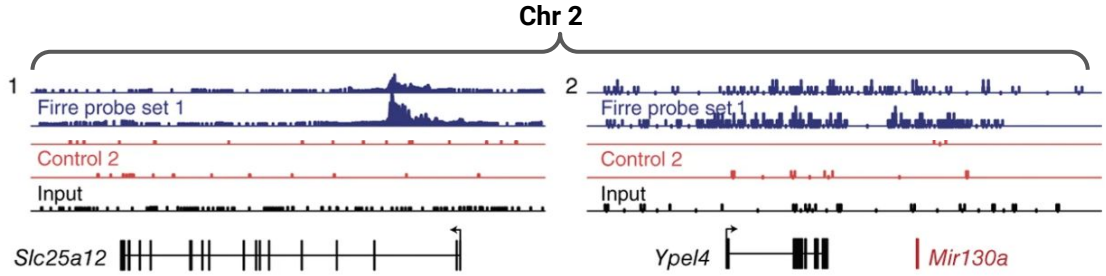
RAP-DNA seq data by using antisense probes for *Firre* RNA



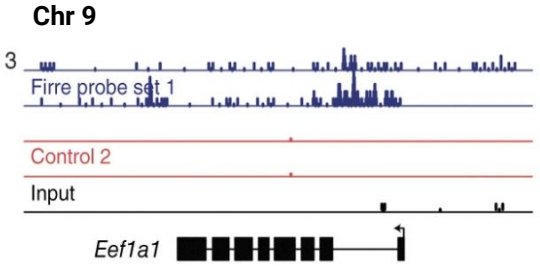
Firre enrichment in the X chromosome **extends** on the **~5-Mb region flanking** the *Firre* locus.

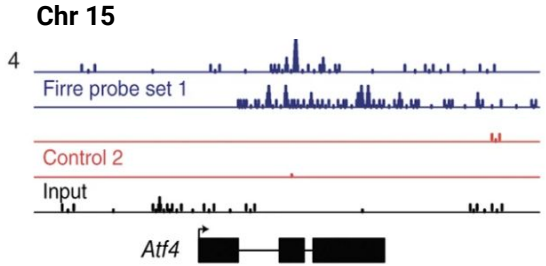
Where does *Firre* localize in the chromatin *in trans*?

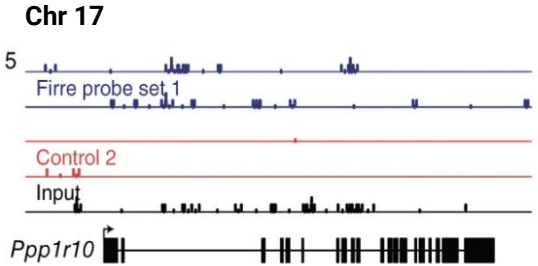




Interestingly, *Slc25a12*, *Atf4*, *Mir130a* and *Ppp1r10* have regulatory roles during **adipogenesis**.







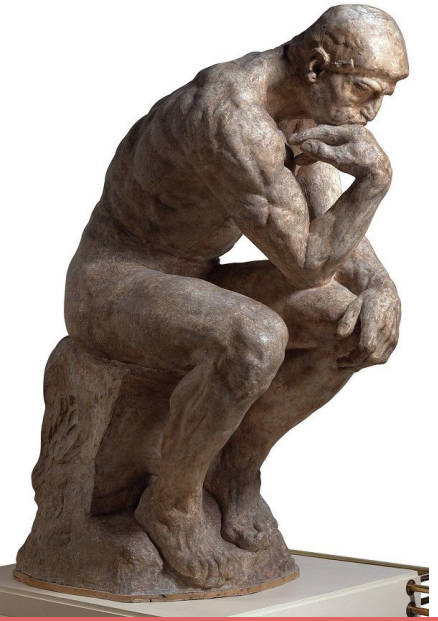
Where does *Firre* localize in the chromatin?

Conclusions so far:

- ❏ *Firre* is predominantly present in the **nucleus** thanks to its **RNA Repeat Domain**
- ❏ *Firre* locus **escapes X-chromosome inactivation**
- ❏ *Firre* localizes **around its transcription site**, forming thus 1 focus on male cells and 2 foci on female cells
- ❏ *Firre* can localize on **multiple chromosomes**.

Key point: *Firre* localizes to chromatin both ***in cis*** and ***in trans***.

How does *Firre* localize in the chromatin?



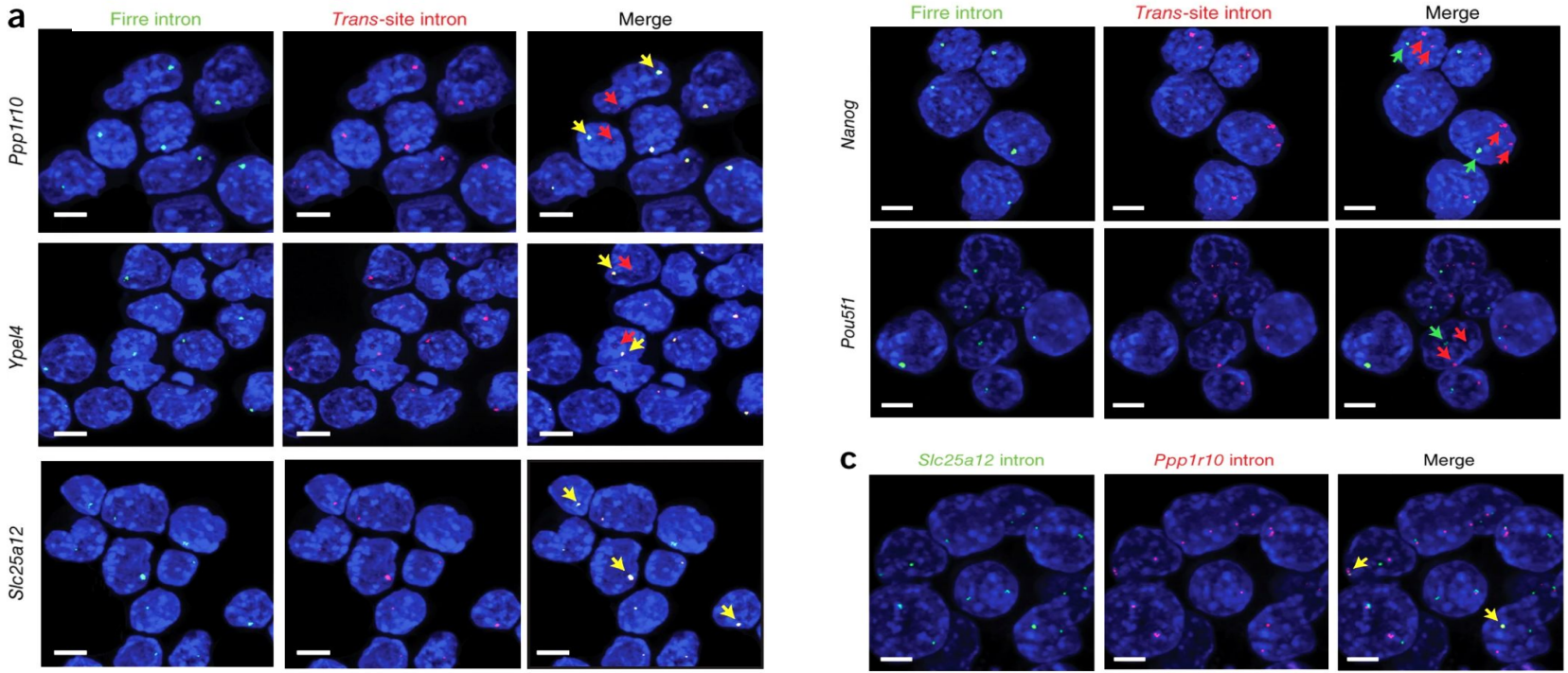
Results so far suggest **two possible models**:

1. **Shuttling hypothesis**: *Firre* could be shuttled from its transcription site to the other chromosomes;
2. **Focal localization hypothesis**: The focal localization of *Firre* serves to bring the trans-interacting sites in three-dimensional proximity to the *Firre* locus on the X chromosome.

Which model is correct?

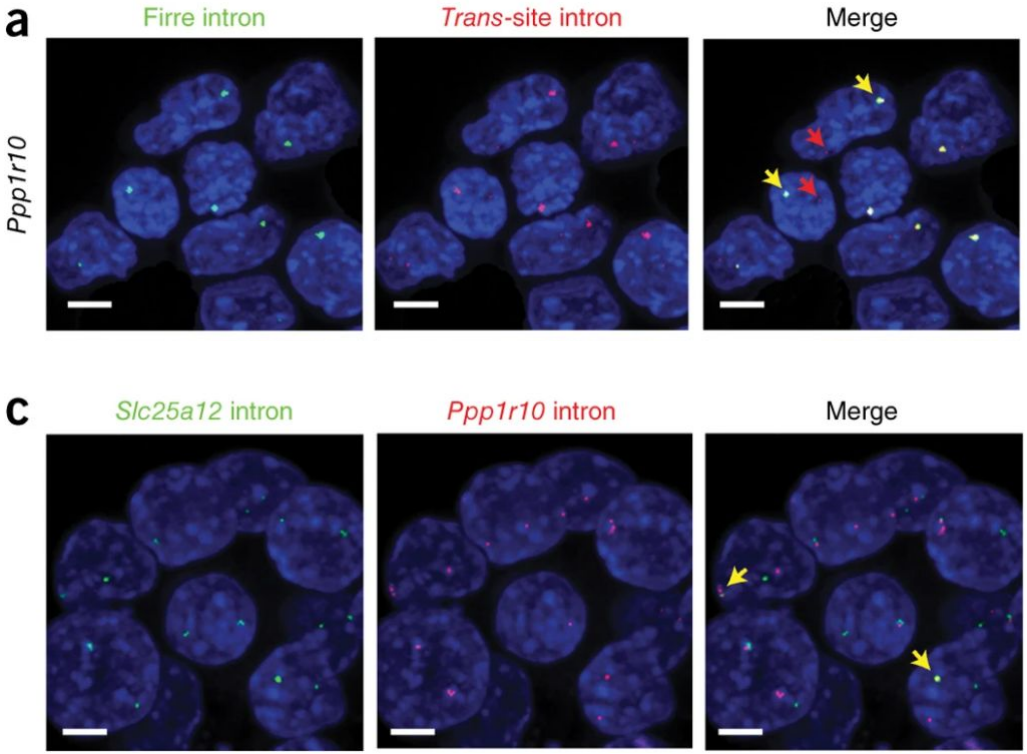
How does *Firre* localize in the chromatin?

Dual labelling strategy RNA-FISH on mESCs to detect *Firre* on *trans* interacting genes



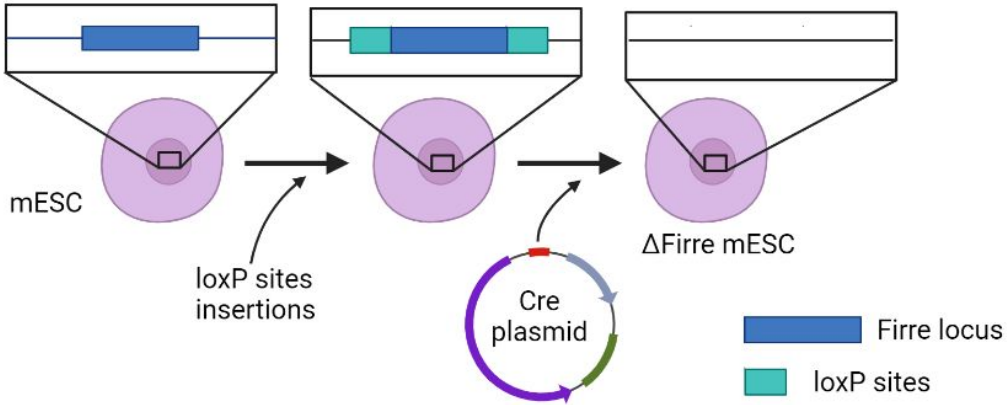
scale bar 40µm

Firre localization in the chromatin

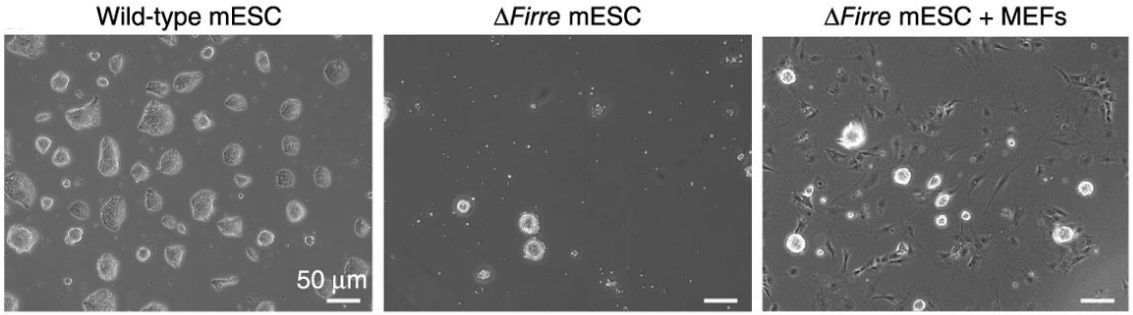


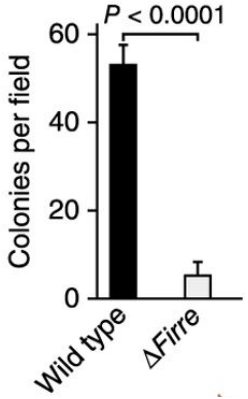
Conclusions: the dual labeling RNA FISH supports the focal hypothesis since the signals merge

What is the function of *Firre*?



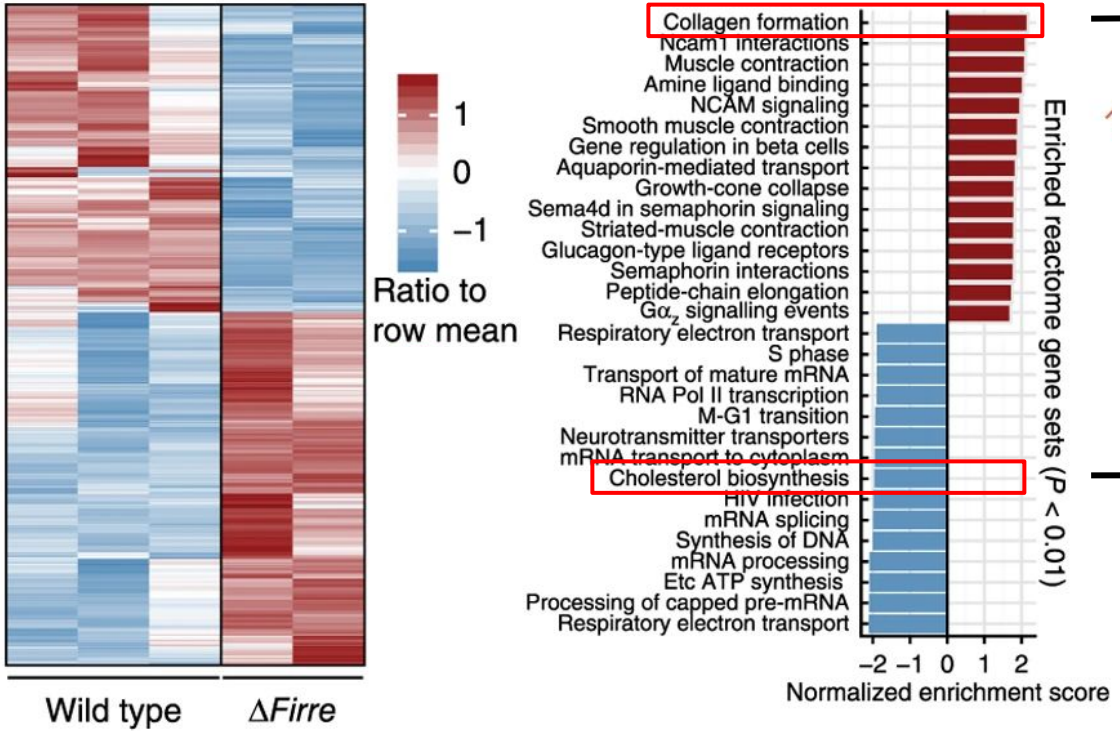
Conclusion: Δ*Firre* mESCs show a marked retardation in growth rate and colony formation in comparison with wt mESCs.





What is the function of *Firre*?

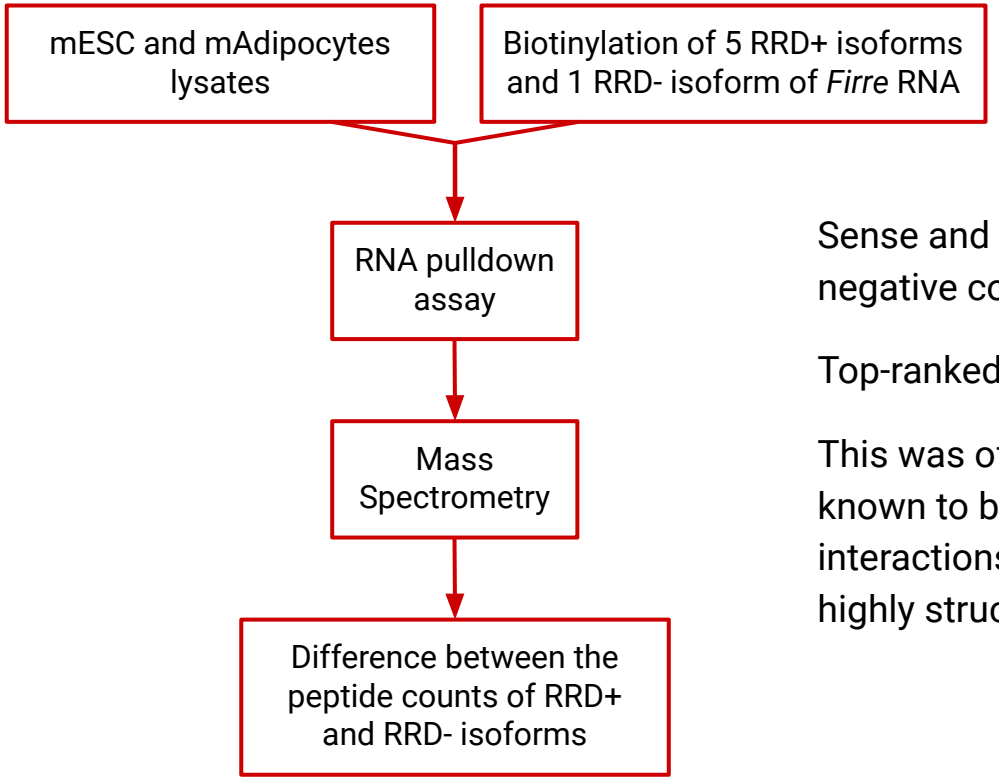
Differential gene expression analysis on wt and Δ*Firre* male mESC with RNA-seq (paired-end Illumina), followed by GSEA (Gene Set Enrichment Analysis)



TGFβ signalling is increased in Δ*Firre* mESCs, which is known to be an inhibitor of adipogenesis

Adipogenesis pathways are downregulated in Δ*Firre* mESCs

Are there any proteins interacting with *Firre* RRD?



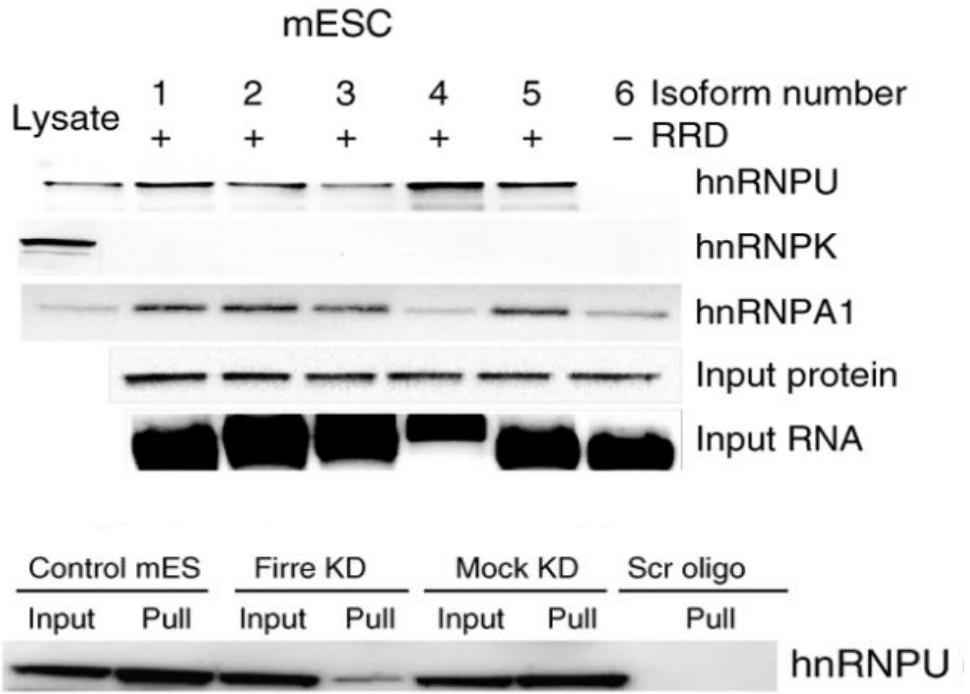
Sense and antisense TERT RNA were used as negative control in all the process.

Top-ranked candidate: **hnRNPU.**

This was of particular interest because it was already known to be involved in Xist localization, DNA interactions with scaffold attachment regions and for highly structured chromatin territories.

Firre interacts with hnRNPU: confirmation

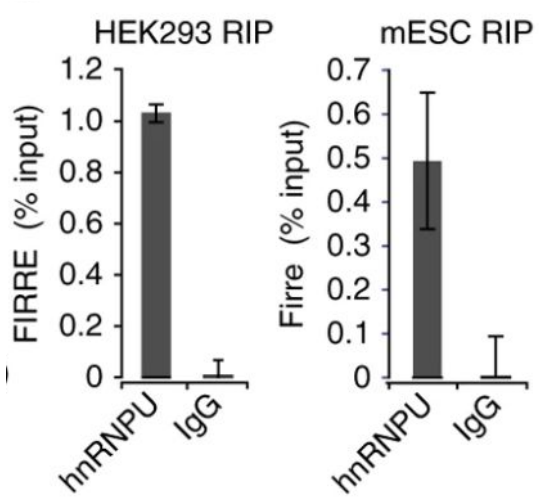
- To confirm this interaction, RNA pulldown assays + WB were performed, using both RRD+ and RRD- isoforms. Also, additional members of the hnRNP family were tested, showing either no association or independent association with RRD.
- To determine if this interaction has a biological relevance at endogenous levels, endogenous *Firre* was captured using desthiobiotin complementary DNA oligos to KD *Firre*.
  Mock KD cells were treated with scrambled oligos as positive control.



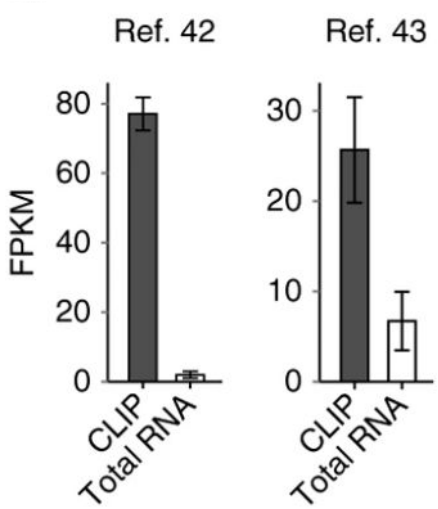
Scr: scrambled (same nucleotide composition, but different sequence; Mock: use of scramble oligos

Firre interacts with hnRNPU: confirmation

RIP for hnRNPU on mESC and HEK293



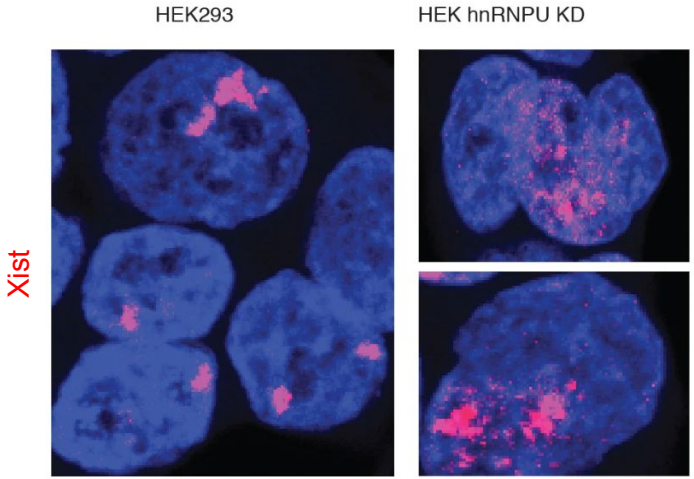
CLIP data for hnRNPU

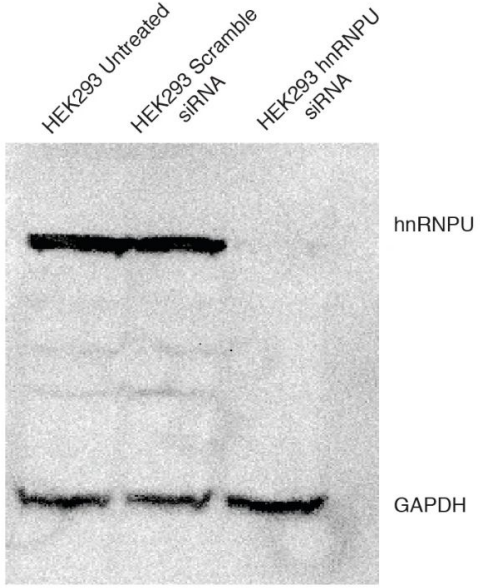


Conclusions: *Firre* interacts with hnRNPU. The **RRD** is both r**equired and sufficient** for this interaction.

Is *Firre* localization regulated by hnRNPU?

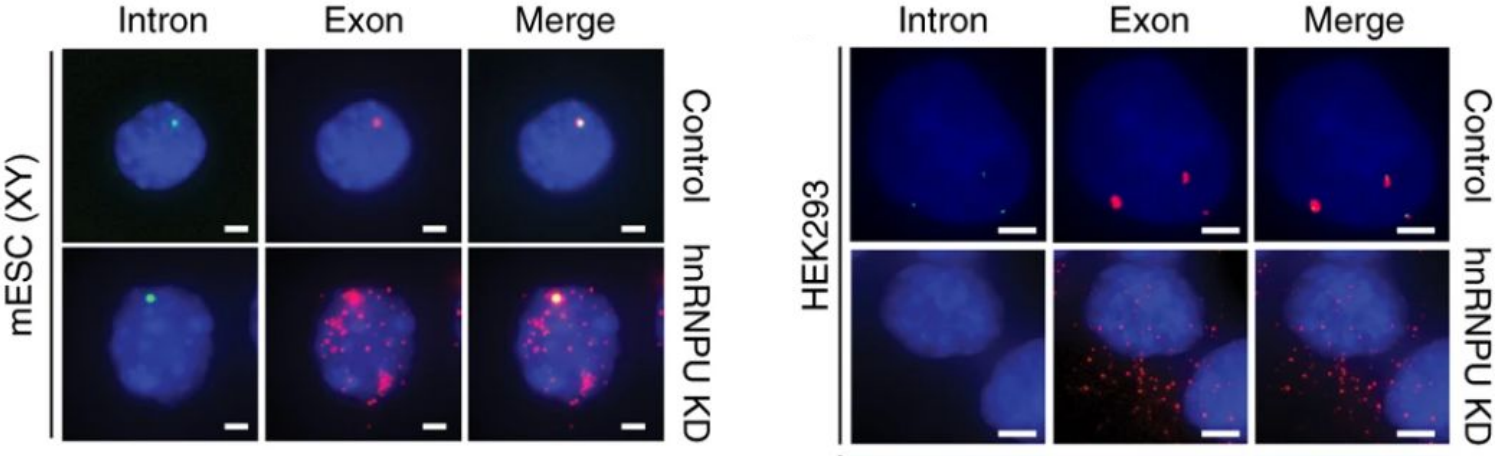
To further investigate if hnRNPU regulates *Firre* localization, a siRNA-mediated depletion model for hnRNPU was used on mESCs and HEK293.
RNA FISH of *Xist* in HEK293 hnRNPU KD confirms previously described role of hnRNPU in *Xist* localization.





These results confirm previous results, validating this experimental model.

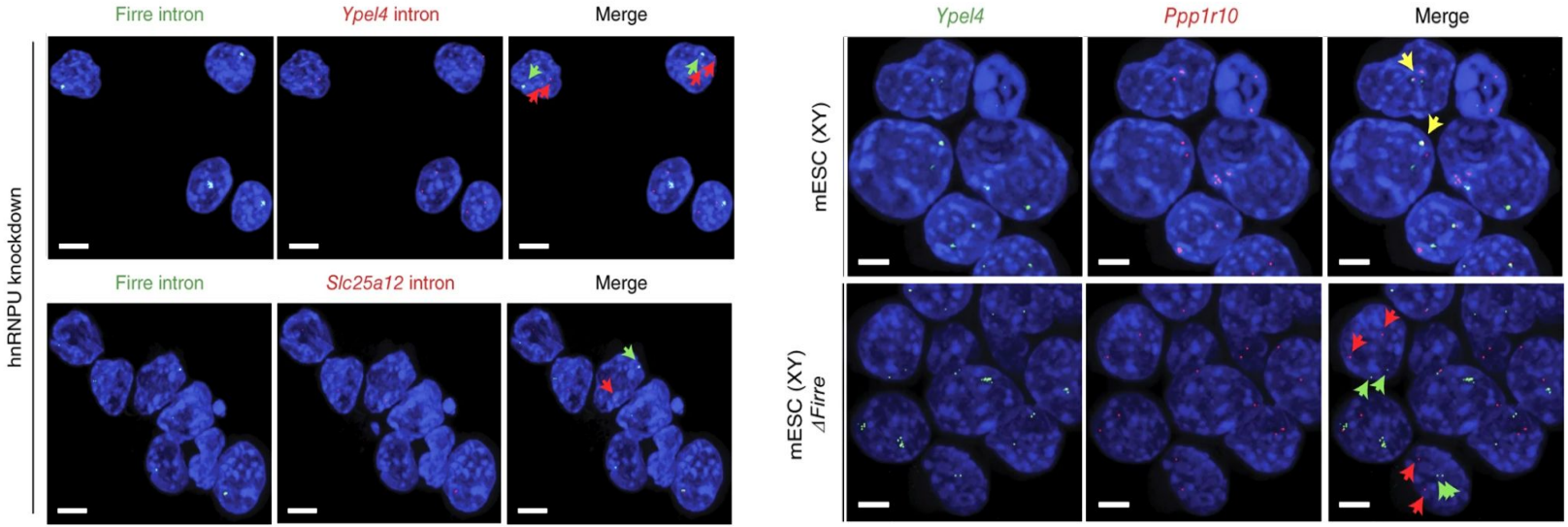
Is *Firre* localization regulated by hnRNPU?



Conclusions:

- **loss of hnRNPU** causes **delocalization** of *Firre*, with even translocation in the cytoplasm.
- **hnRNPU** and **RRD** are **essential for proper localization** of *Firre* in focal nuclear foci.

Is hnRNPU involved in proximal *trans* localization of *Firre*?

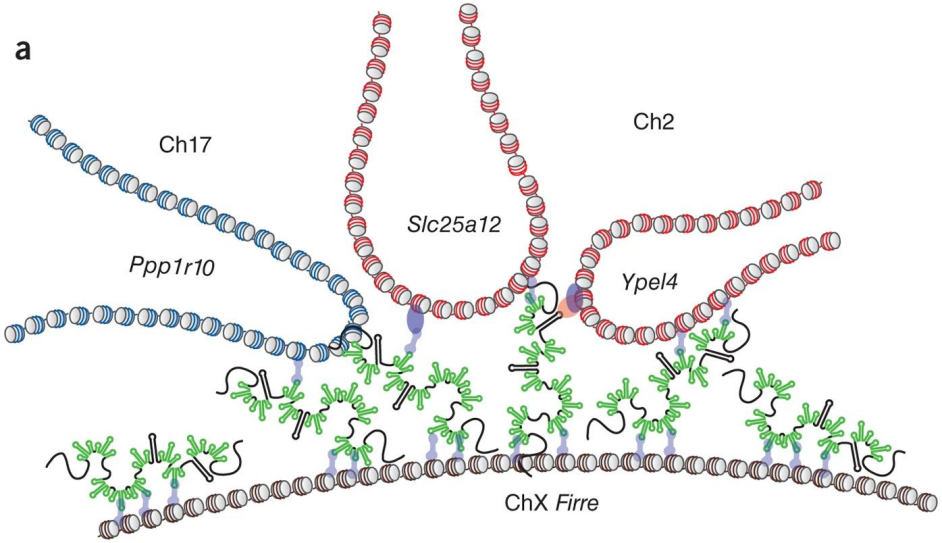


Both *Firre* deletion and hnRNPU siRNA-mediated depletion and results in a decreased colocalization of each trans site with *Firre*.

Firre, along with **hnRNPU**, has a role in **maintaining** and **enstablishing higher-order nuclear architecture**.

A model of *Firre* as a regional organization factor

- Firre provides a platform for the trans-chromosomal interactions of specific genomic loci.
- RRDs are bound by hnRNPU, thus facilitating interactions with trans-chromosomal regions in many possible ways:
  1. interactions with nuclear-matrix components;
  2. binding of hnRNPU to matrix attachment regions in trans;
  3. interactions with other protein complexes to mediate an indirect binding to DNA.




Thanks for your attention