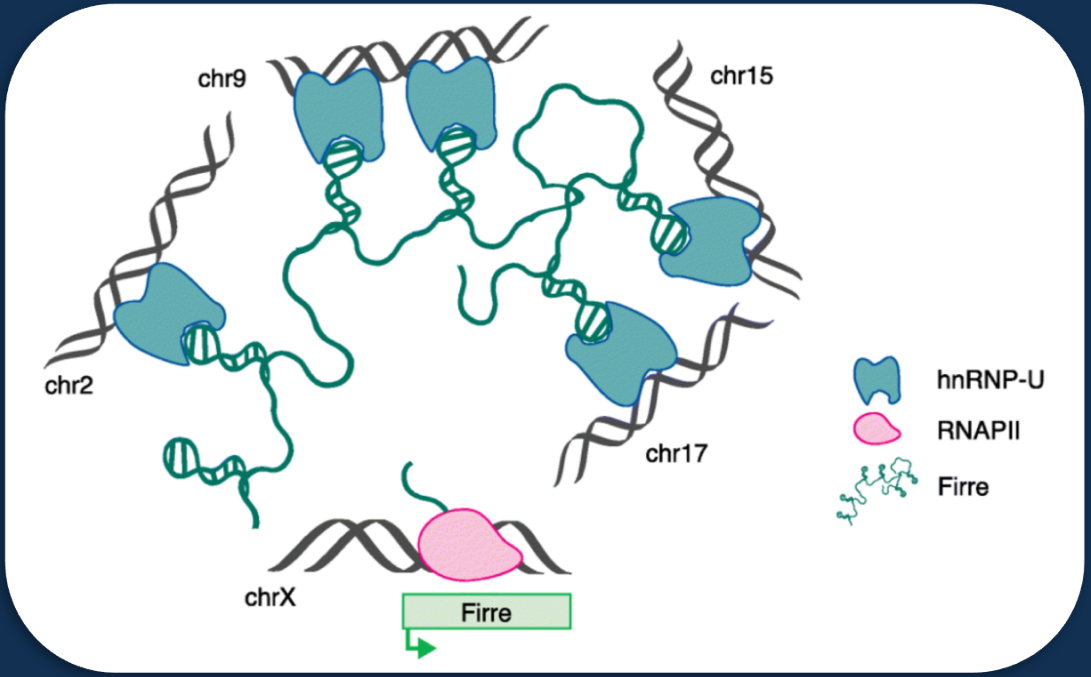
ncRNAs as chromatin modifiers

The newly discovered lincRNA Firre



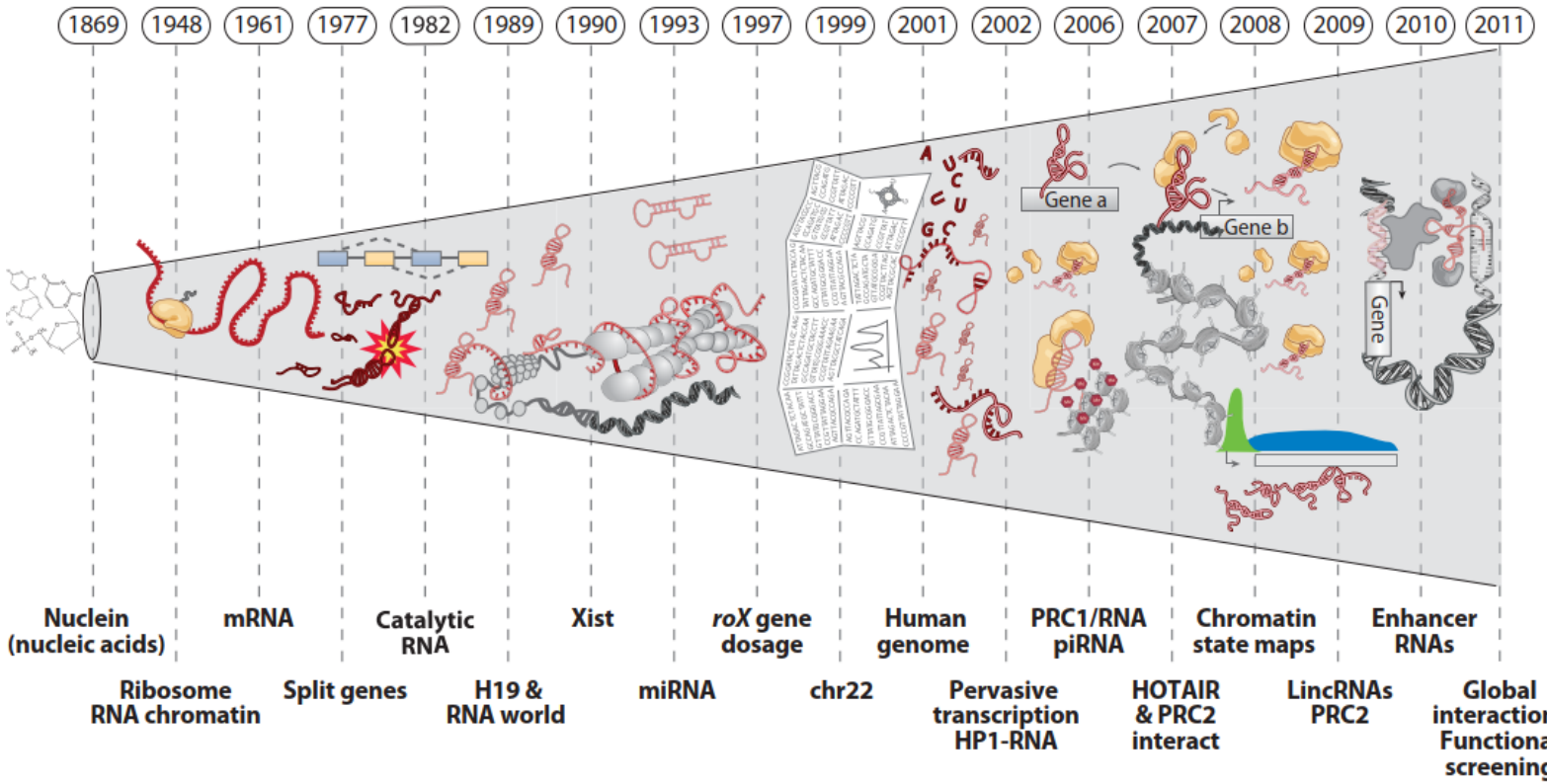
Elisabetta Giordano
Sefora Naomi Agrò
Ilaria Frascolla

Organization

- Introduction on lncRNAs: structure and functions
- Model systems for Firre study

- Introduction to Firre, features and localization
- Firre locus escapes X-chromosome inactivation

- Functional roles of Firre
- Firre interactions with proteins



Introduction

- In 1961 the paper "Genetic Regulatory Mechanisms in the synthesis of proteins" established the concept of messenger RNAs.
- In the 50 years since this landmark paper, numerous regulatory RNA of all shapes and sizes have been discovered.

Introduction

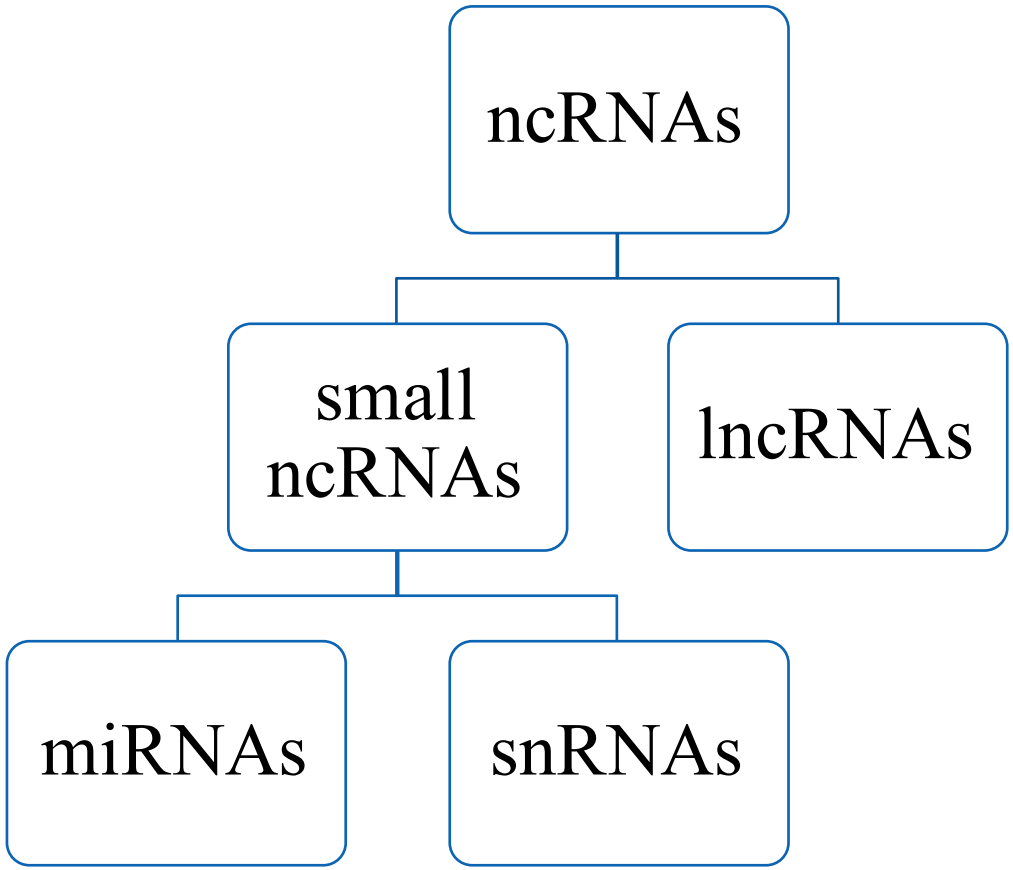
Recent developments in next generation sequencing technologies have led to extensive transcriptomic and bioinformatic analysis of many cell lines and tissues.

These analyses show that **more than 75%** of the human genome is **transcribed,** only a small portion of the transcripts are translated into final protein products.

The rest of the transcripts that do not possess any protein-coding capacity are annotated as **noncoding RNAs (ncRNAs)**.

Classification

- Small ncRNAs are shorter than 200 nucleotides .
- Long ncRNAs are longer than 200 nucleotides.
- 15.000 lincRNAs are encoded by the human genome.



LincRNAs

- By definition, these RNAs must not have open reading frames that encode proteins, but recent evidence indicates that short polypeptides may be encoded by these transcripts.
- Many of these transcripts are encoded by RNA polymerase II, are spliced and are polyadenylated.
- LincRNAs were found to participate in a vast variety of biological processes.

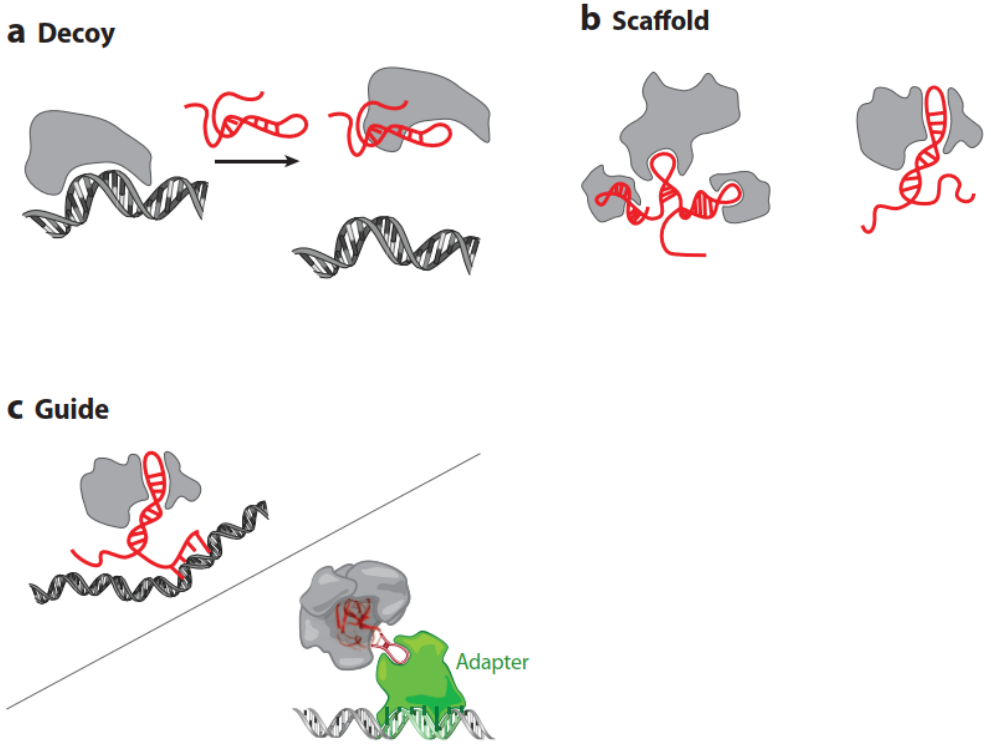
LincRNAs can be found in the nucleus or in the cytoplasm

Nuclear lincRNAs may function as:

- Decoys
- Scaffolds
- Guides

LincRNAs have many functions in the cytoplasm:

- mRNA stability
- sequester miRNA(miRNA sponges)
- translation of target mRNAs



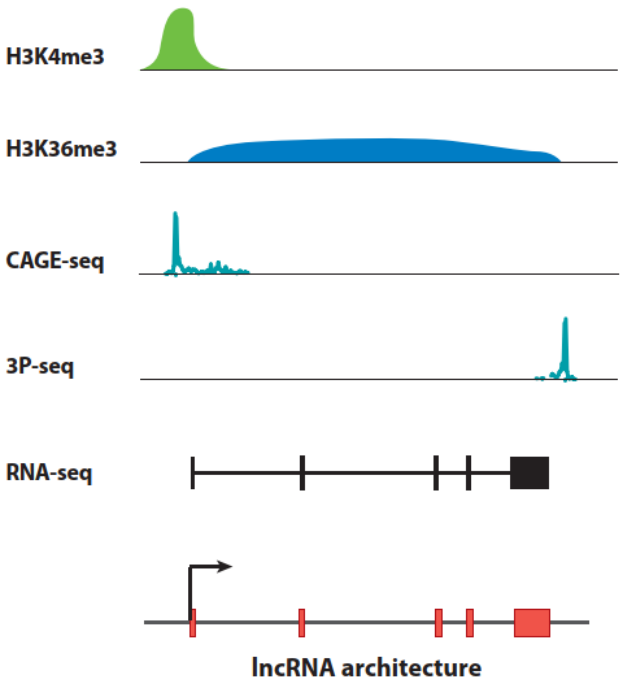
Genomic discovery of lncRNAs

Can the complexity of different organisms be explained by the number of classic protein-coding genes, their splicing diversity or new types of regulation?

Tiling microarrays: new technologies were emerging to understand the regulation of gene expression and de novo identification of new genes. The combined power of microarrays and draft genome sequences reported initial estimates that there may be as many lncRNA genes as protein-coding genes.

Conclusion: a vast majority of the genome was transcribed. LncRNAs also have a specific regulation.

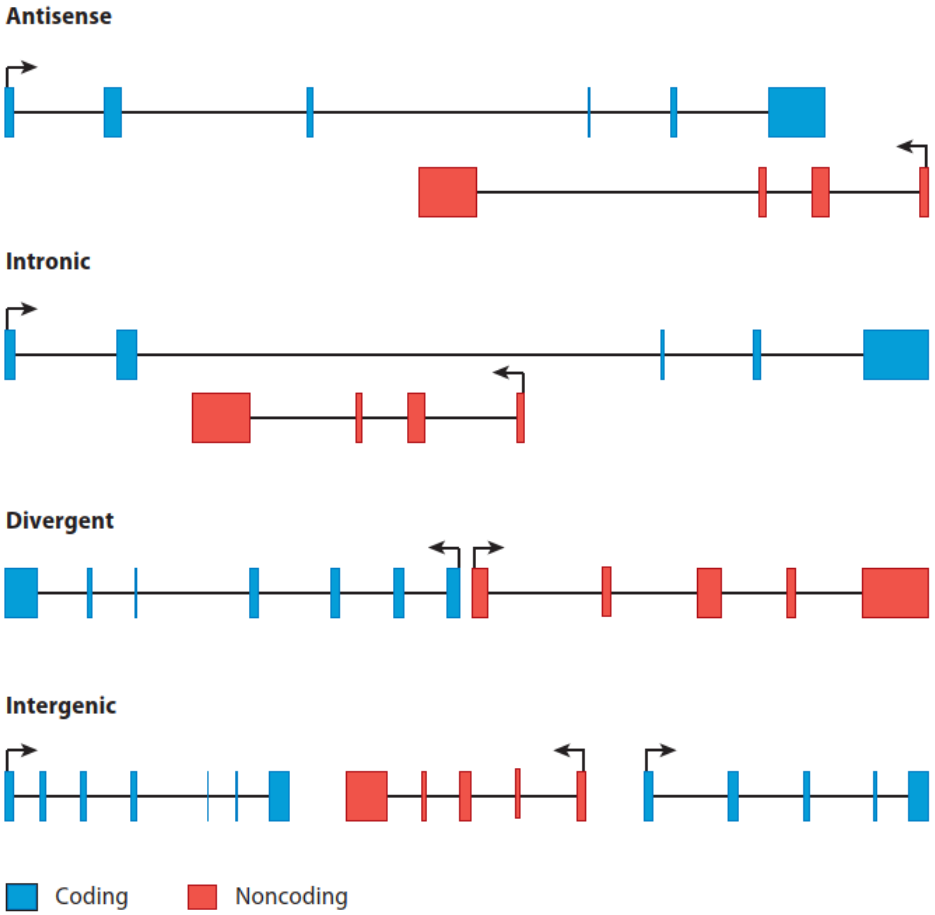
Genomic discovery of lncRNAs



- Chromatin marks: a clear signature of polymerase II-transcribed genes occupied by **H3K4me3** at the promoters of genes, followed by **H3K36me3** along the transcribed unit. Chromatin marks in several cell types revealed that about 5000 K4-K36 domains represented lncRNAs.
- RNA sequencing: it can directly define the primary structure of lncRNA. Several applications have emerged to perform RNA-seq. For example, CAGE seq and polyadenylation ends (3P-seq) define the precise beginnings and ends of transcripts.

Conclusion: there are highly conserved promoter regions that recruit the binding and direct regulation of key transcription factors.

Anatomy of long noncoding RNA (lncRNA) loci



- Through the addition of information (such as conservation, coding potential patterns, and anatomical properties), lncRNA gene families have been identified.
- LncRNAs are often defined by their location relative to nearby protein-coding genes.

Conclusion: An understanding of the sequence and structural elements that relate to lncRNA function will allow classification of lncRNA families as has occurred for protein families with similar structural domains.

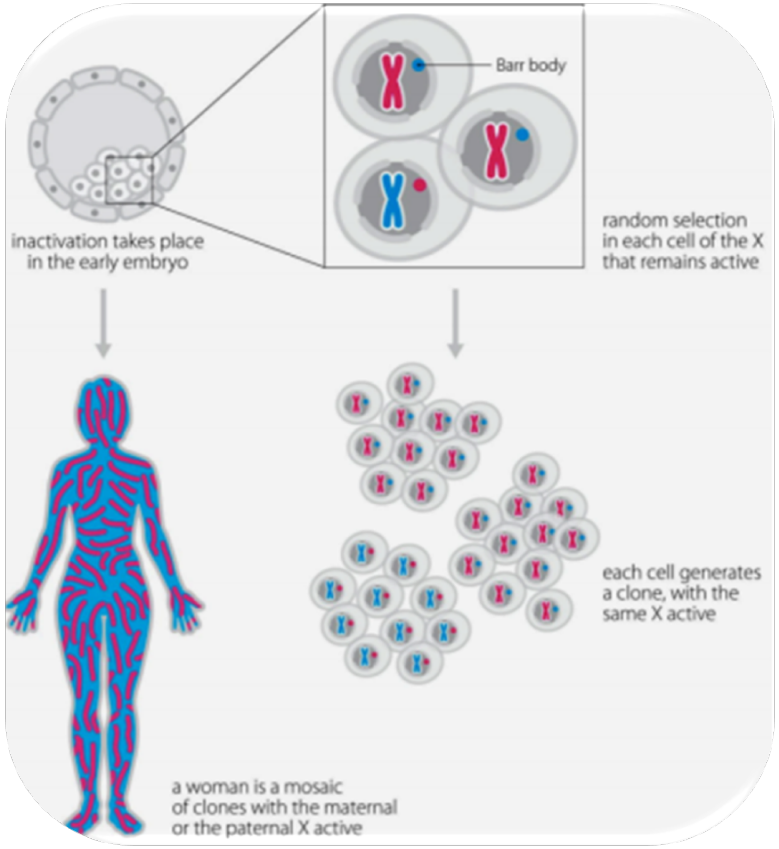
Model system for lncRNA FIRRE

- Scientists have used mESCs and hESCs as model systems.
- Many ES cell cultures protocols are based on the use of monolayer of primary mouse embryonic fibroblasts (MEF) feeder cells .
- MEF cells perform two important roles in stem cell culture:

1. Growth factors
2. Cellular matrix

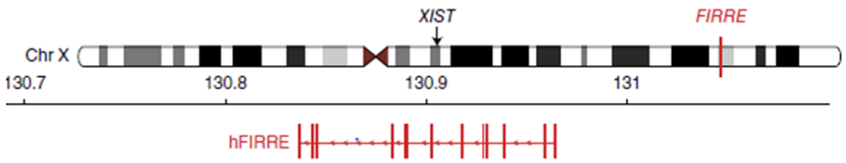
X chromosome inactivation

- One of the X chromosome in female (XX) is inactivated so as to have a chromosomal balance with male (XY).
- The X chromosome inactivation is random.
- It occurs in the blastocyst.
- The region responsable of the inactivation codifies for lncRNA XIST.

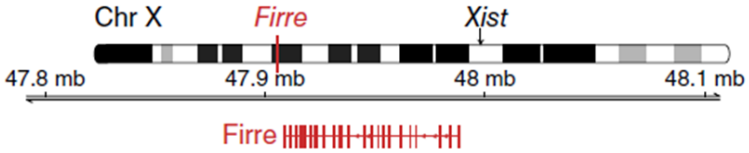


Escape from X inactivation

- Some genes can escape X inactivation
- Some genes must escape from inactivation (pseudoautosomic regions)



✓ Human *FIRRE* locus located on the X chromosome



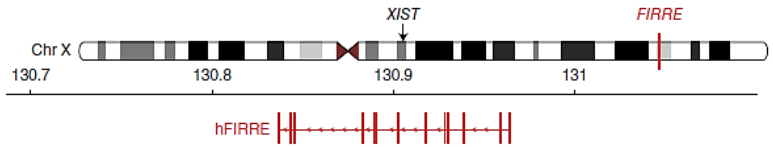
✓ Mouse *FIRRE* locus located on the X chromosome

- lncRNA Firre escapes from X inactivation.

lincRNA Firre

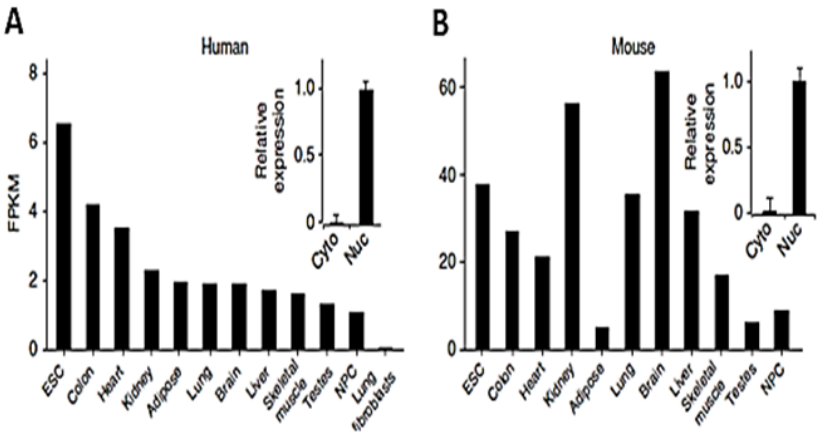
KEY POINTS

- **Firre** is a lincRNA, termed functional intergenic repeating RNA element.
- An ~5-Mb domain localizes across around Firre site of transcription on the X chromosome. This domain of Firre localization is also in spatial proximity to at least five other trans-chromosomal loci within the nucleus. This cross-chromosomal colocalization requires Firre, indeed the **genetic deletion** of Firre results in a loss of spatial proximity between its trans-chromosomal binding sites.
- A unique 156-bp repeating RNA domain **(RRD)** in Firre sequence is required for both interaction with the nuclear-matrix factor **hnRNPU** and localization of Firre transcripts in a punctate manner in the nucleus. Strikingly, **knockdown of hnRNPU**, similarly to deletion of the Firre locus, results in a loss of spatial proximity between the Firre locus and its trans-chromosomal binding sites.
- lncRNAs such as Firre can interface with and modulate higher-order nuclear architecture across chromosomes.

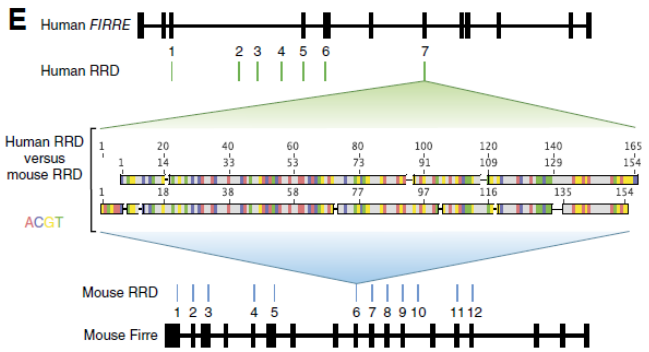


✓ Human *FIRRE* locus located on the X chromosome

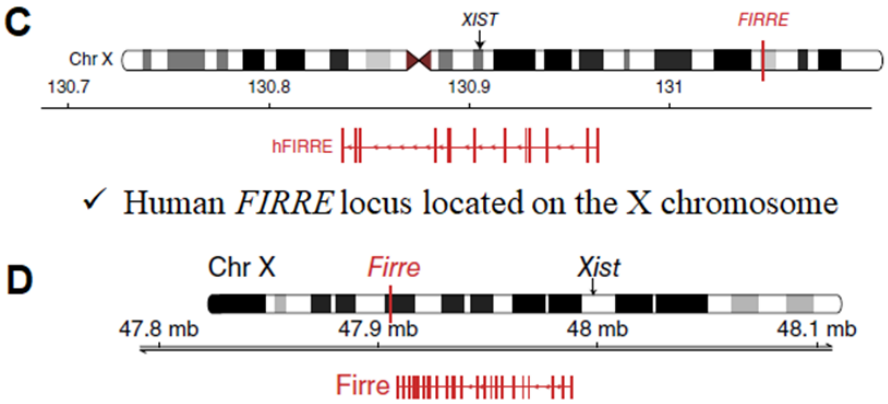
Features of Firre



✓ Expression pattern of Firre in Vivo



✓ **RRD** = unique 156-bp **R**epeating **R**NA **D**omain

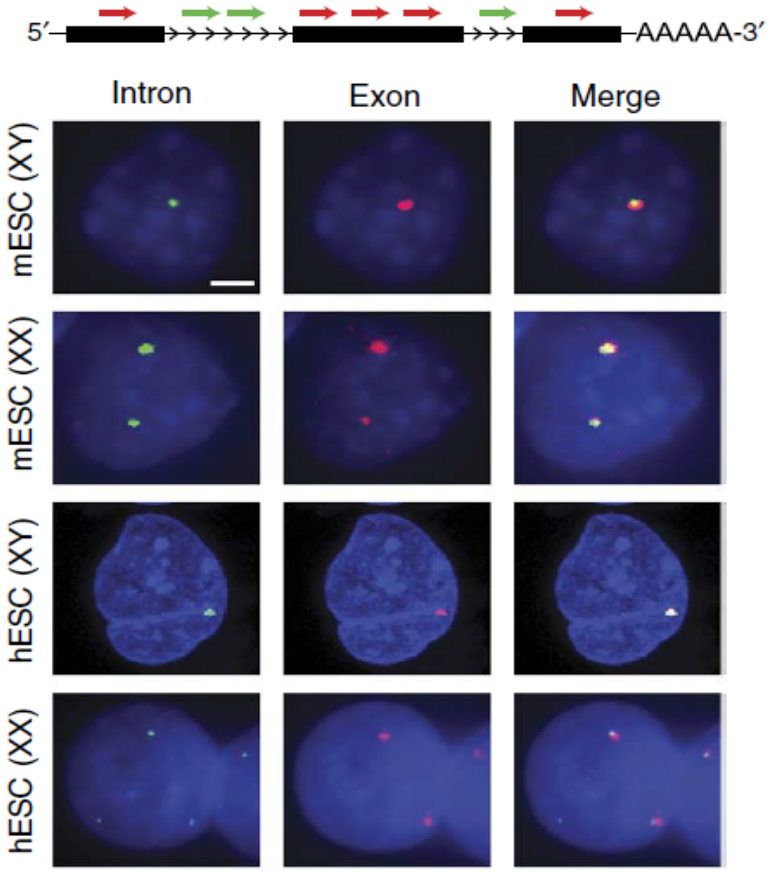


✓ Human *FIRRE* locus located on the X chromosome

✓ Mouse *FIRRE* locus located on the X chromosome

- Firre is a nuclear-retained and chromatin-associated lncRNA
- Firre is being required for proper adipogenesis in a loss-of-function screen.

RNA FISH for determining Firre localization and Firre expression foci



PROTOCOL

Template**:** male (XY) and female (XX) mouse and human embryonic stem cells (mESCs and hESCs, respectively).

Analysis: single-molecule RNA FISH.

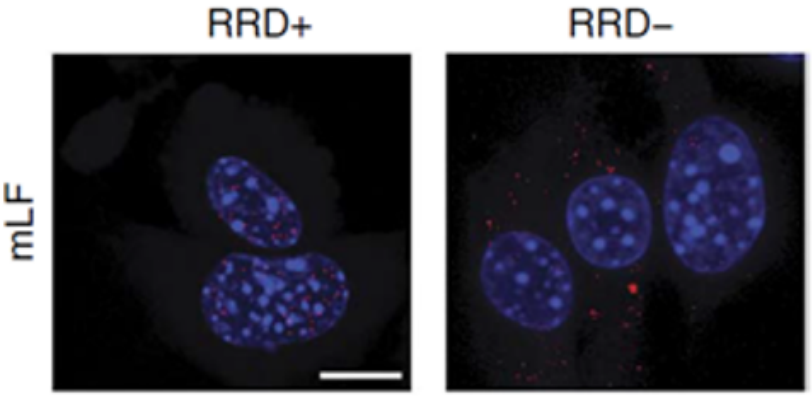
Strategy: dual labeling strategy to independently target the introns and exons of Firre, thus marking:

- the site of transcription on the X chromosome (intronic probes)
- and the location of the mature transcripts (exonic probes).

CONCLUSION:

- It is revealed an exclusively nuclear and focal distribution for Firre in all cells tested.
- Firre exhibited strong expression foci near its site of transcription in both male and female mESCs and hESCs.
- Subcellular localization and expression of Firre in other cell lines with and without inactive X chromosomes was similar to those observed in ESCs.

What does the nuclear location of Firre depend on?



PROTOCOL:

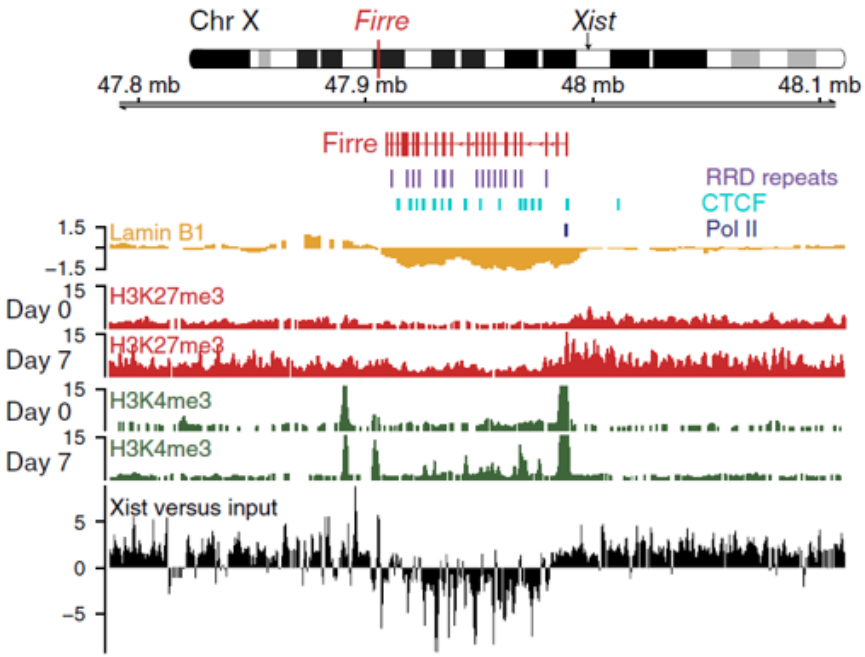
Template**:** murine lung fibroblasts (mLF).

Analysis: single-molecule RNA FISH and retroviral overexpression of Firre.

Strategy: Firre is express via retrovirus-mediated integration in human and mouse lung fibroblasts, which do not express Firre. In this case, it is use isoforms of Firre either with or without the RRD.

CONCLUSION: In the absence of the RRD, the nuclear localization of Firre is disrupted, and it is detected Firre RNA in the cytoplasm → **RRD is required for the focal nuclear localization of Firre.**

The Firre locus escapes X-chromosome inactivation



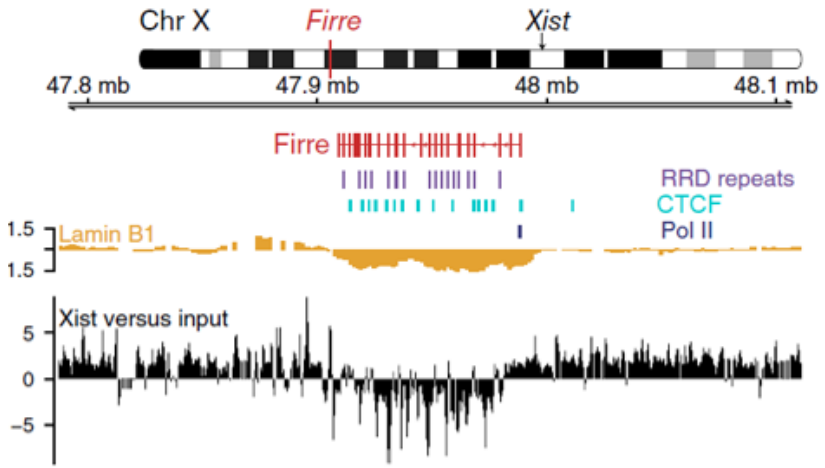
PROTOCOL:

Analysis: chromatin immunoprecipitation sequencing (ChIP-seq).

Consideration:

1. There is an considerable depletion of lamin B1 across the mouse Firre locus and across the human FIRRE locus in various cell lines. The domain of lamin B1 depletion extends precisely across the body of the *Firre* gene.
2. The *Firre* locus was specifically and significantly depleted of tri-methylated histone H3 K27 (H3K27me3) in differentiated mESCs and in human cells before and after X-chromosome inactivation.
3. The Firre locus was enriched for tri-methylated histone H3 K4 (H3K4me3) with and without Firre transcription.
4. There is a striking localization pattern of CCCTC-binding factor (CTCF) adjacent to almost every exon of Firre.

The Firre locus escapes X-chromosome inactivation



PROTOCOL:

Template**:** mouse lung fibroblasts (mLFs).

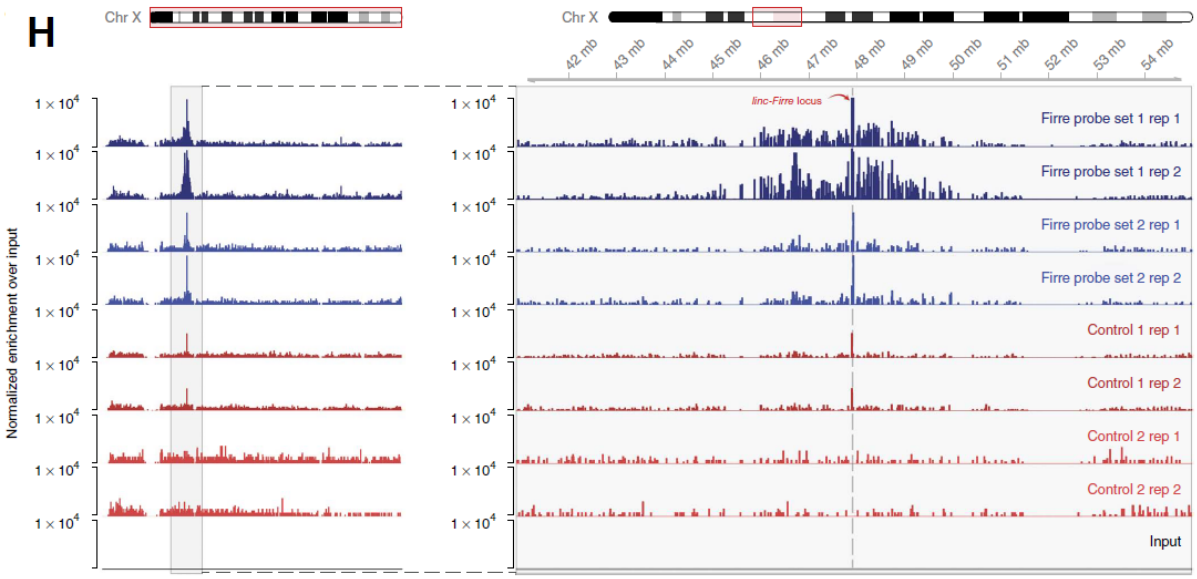
Analysis: Xist RNA antisense purification (RAP).

Consideration:

1. In contrast to the enrichment of Xist across most of the X chromosome, a strong and focal depletion was present in Xist binding at the *Firre* locus; this was similar to what was observed at genes known to escape X-chromosome inactivation.
2. Interestingly, the Xist-depleted boundaries are consistent with the previously identified boundaries for the lamin-B1 depleted regions.

CONCLUSION: Collectively, these data indicate that the *Firre* locus escapes X-chromosome inactivation and has a notable enrichment for CTCF and H3K4me3 and depletion for H3K27me3 and lamin B1.

Firre localizes to chromatin in cis



PROTOCOL:

Template**:** male mESCs.

Analysis: RNA antisense purification (RAP) shown along the X chromosome.

Strategy: RAP technique use two sets of 120-bp antisense probes targeting Firre and two sets of sense probes (blue) as negative controls (red). This was followed by sequencing to identify genomic regions directly bound by Firre.

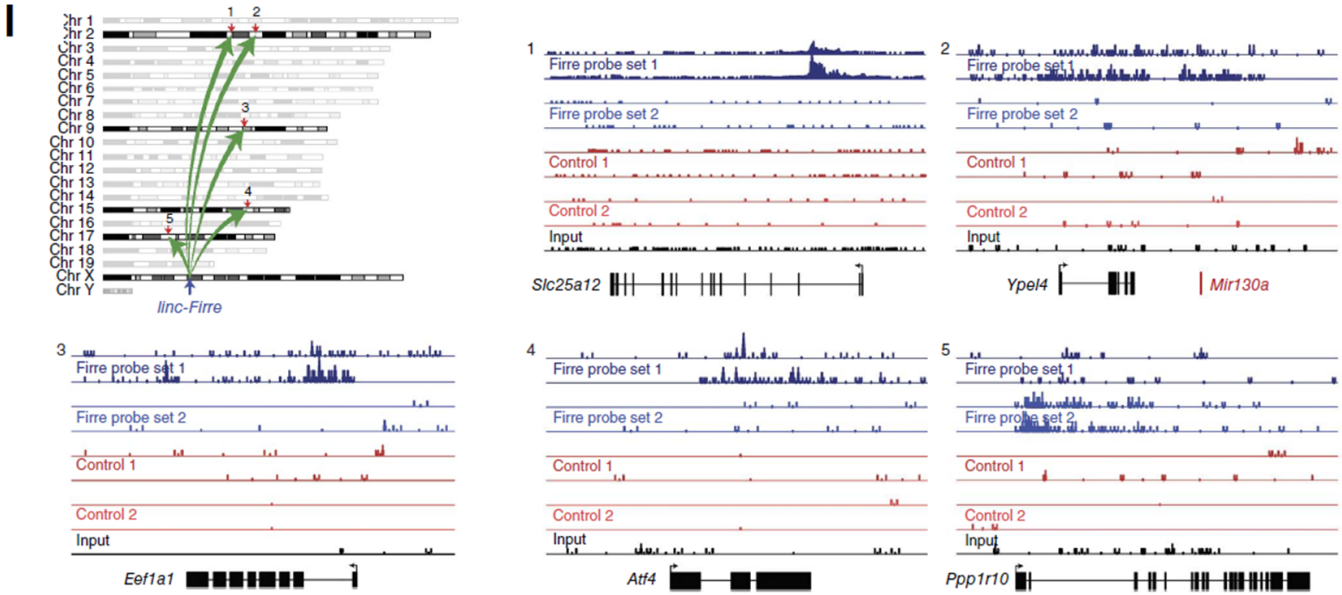
CONCLUSION: The peaks highlighted in figure H indicate an ~5-Mb domain of Firre localization around the Firre locus.

Firre localizes to chromatin in trans

Template**:** male mESCs.

Analysis: RAP by Firre shown for five distinct interchromosomal genomic loci.

Strategy: RAP technique use two sets of 120-bp antisense probes targeting Firre and two sets of sense probes (blue) as negative controls (red). This was followed by sequencing to identify genomic regions directly bound by Firre.



CONCLUSION:

- There are also five significantly enriched peaks of Firre located on chromosomes 2, 9, 15 and 17 that overlap known genes including Slc25a12, Ypel4, Eef1a1, Atf4 and Ppp1r10.
- Notably, Slc25a12, Eef1a1, Ppp1r10 and Atf4 have regulatory roles during **adipogenesis.**

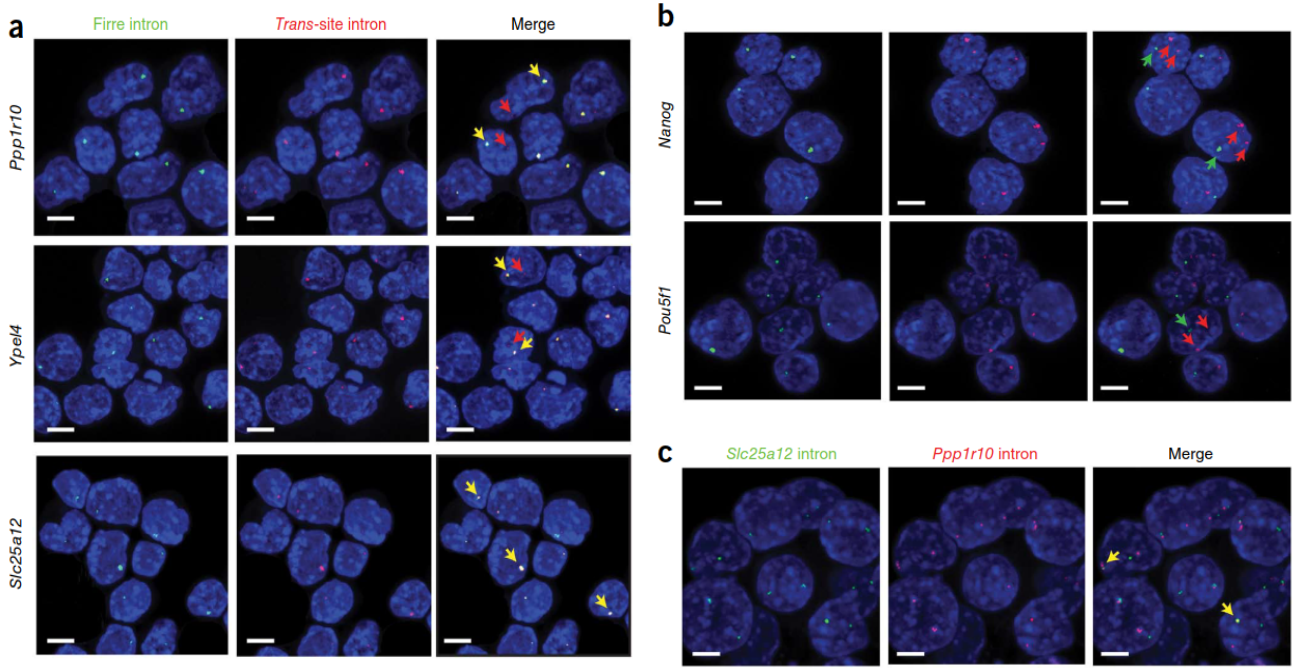
Conclusion

Collectively these data suggest that Firre is localized on multiple chromosomes yet has only one predominant nuclear localization site in male and two in female cells around its site of transcription. These observations suggest two possible models:

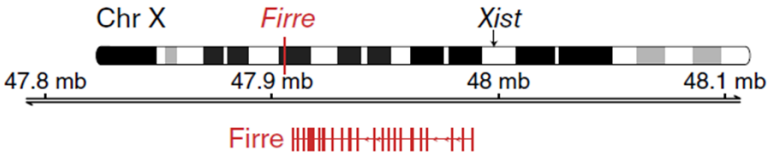
- One possibility is that Firre could be shuttled from its site of transcription to these sites on other chromosomes.
- Alternatively, the focal localization of Firre to its own genomic locus could mean that it serves as a regional organizing factor to bring the trans-interacting sites into the three-dimensional proximity of the Firre locus on the X chromosome.

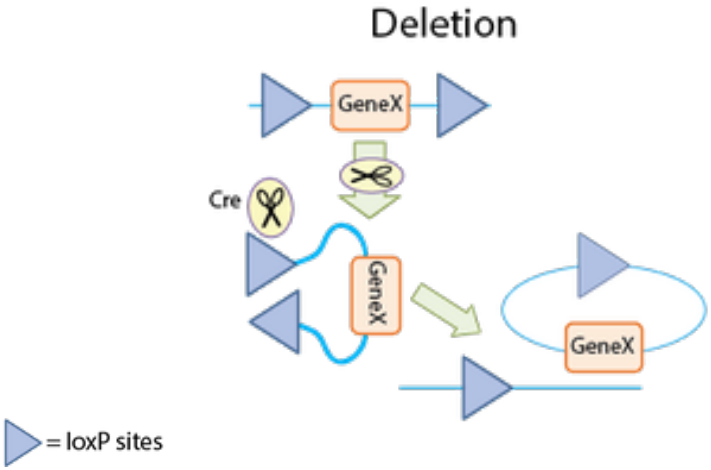
Firre localization in the nucleus

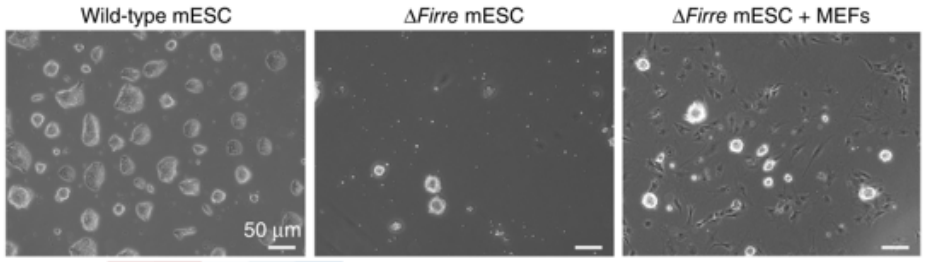
- Single-molecule RNA FISH in mESCs
- Firre colocalizes with the three-tested trans-interacting genes (Slc25a12, Ypel4, Ppp1r10)
- Firre does not colocalize with the control genes (Nanog, Pou5f1)
- The trans-genes colocalize in the nucleus



Conclusion: Firre locus resides in three-dimensional proximity to these trans-chromosomal binding sites







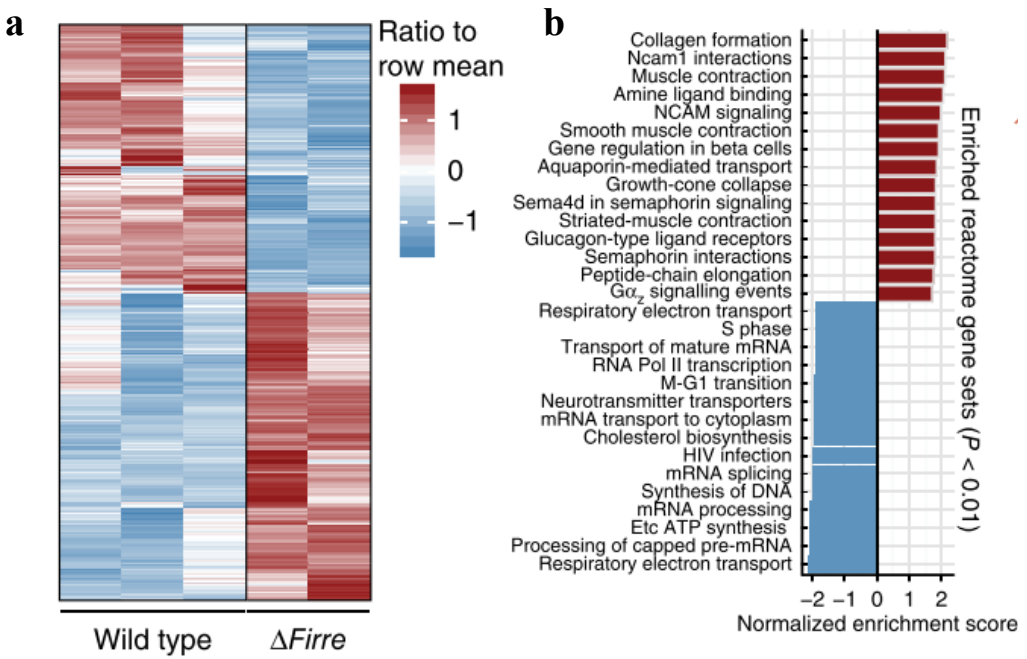
Functional roles of Firre

- Firre KO with site specific recombination (Cre- LoxP system) in male mESCs
- ΔFirre cells show retarded growth rate and lower number of colonies
- ΔFirre cells grown on a fibroblast feeder layer can form more colonies than ΔFirre not grown on feeders in the same amount of time

Conclusion: Firre has a role in regulating cell growth

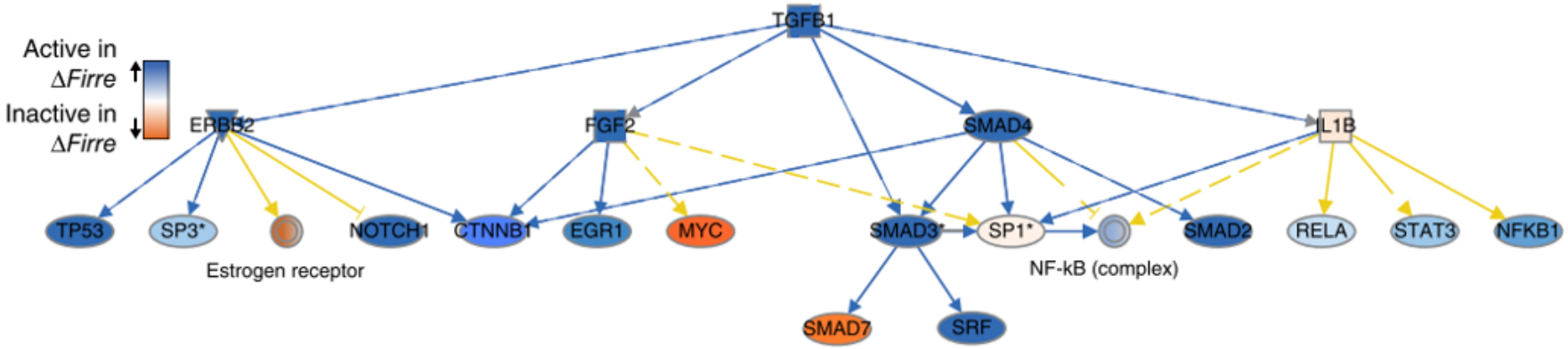
Does Firre deletion alter any pathway?

- 1077 genes showed differential expression in wt cells and ΔFirre cells
- Gene analysis from available data from Gene Set Enrichment Analysis (GSEA)
- Upregulation: extracellular matrix organization, cell-surface receptor-ligand interactions
- Downregulation: mRNA processing, nuclear export, electron transport chain-mediated energy metabolism
- Decreased expression of Ppp1r10



Conclusion: Firre deletion leads to a depletion of genes involved in growth and energy metabolism

Firre involvement in adipogenesis

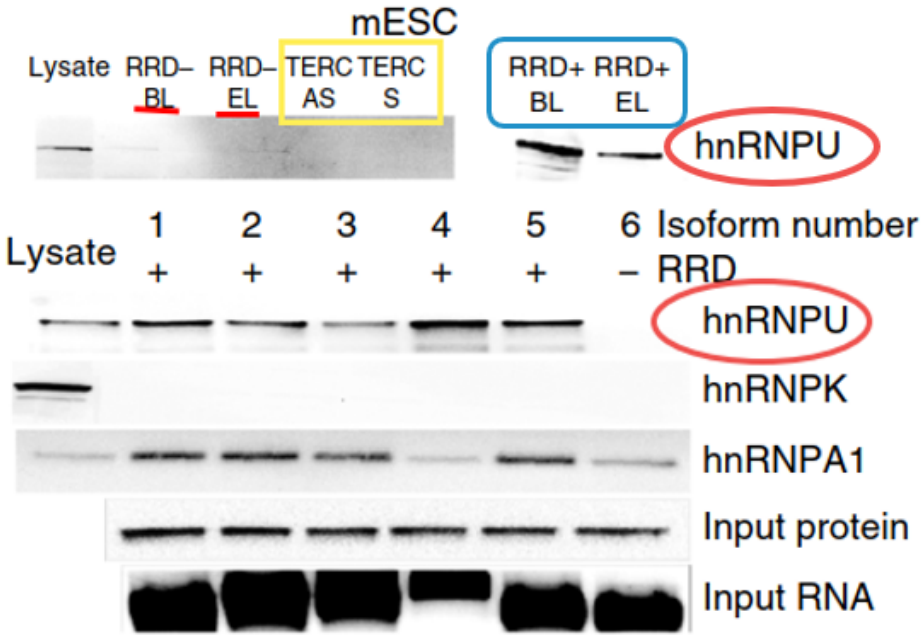


- ΔFirre cells show an increase in TGF-β signaling pathway
- TGF-β is known to be an inhibitor of adipogenesis

Conclusion: Firre is involved in the regulation of adipogenesis together with some of the genes in its trans-sites

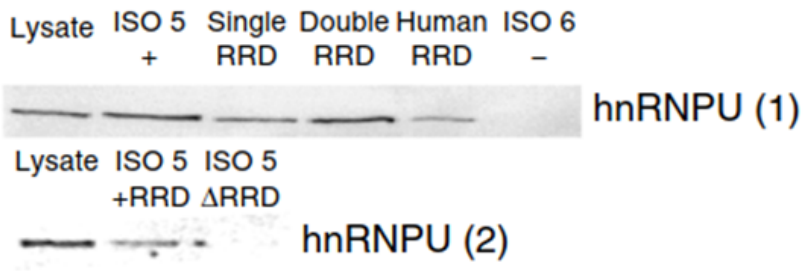
Firre-proteins interactions

- RNA pulldown assay and MS from mESC and mouse adipocytes lysate (5 Firre RRD+ isoform and one RRD– isoform)
- 8 candidate proteins associated with Firre in an RRD-dependent manner. hnRNPU was the most represented one
- WB to validate
- hnRNPU involved in the localization of Xist, in the interaction with scaffold-attachment regions on DNA and in the formation of highly structured chromatin territories
- Specificity of Firre for hnRNPU: no precipitation of hnRNPK and hnRNPA1



Conclusion: lincRNA Firre in the nucleus interacts with hnRNPU through RRD

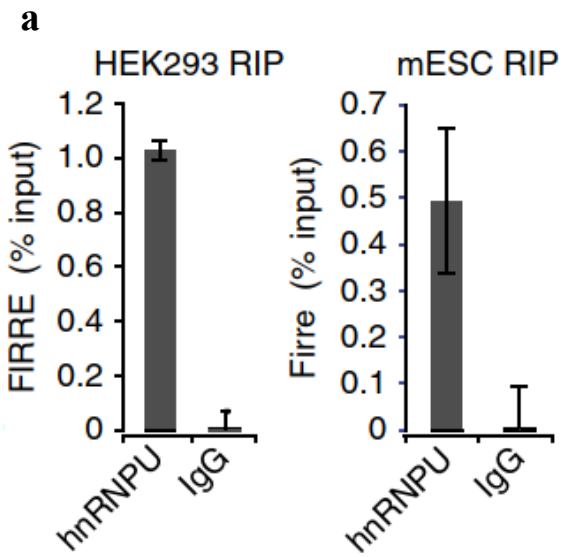
Is RRD sufficient to establish the interaction?

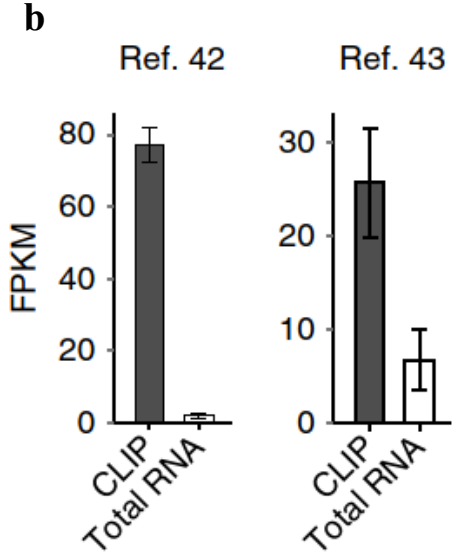


- Western blots for RNA pulldowns
- (1) pulldown performed with in vitro-biotinylated single and double copies of the mouse synthetic RRD constructs and human RRD synthetic sequence in mESC lysates
- (2) pulldown performed with an endogenous isoform with RRD or the same isoform with RRD deleted in mESCs.

Conclusion: hnRNPU does not coprecipitate with the ΔRRD isoforms; it interacts with RRD.

hnRNPU- Firre interaction validation experiments



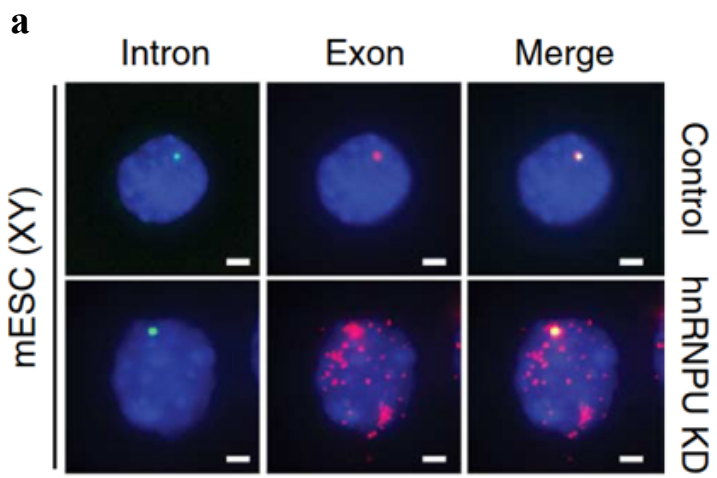


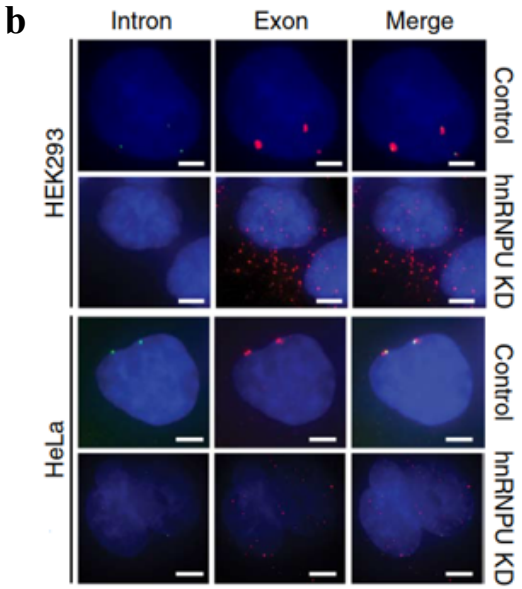
- RNA immunoprecipitation in HEK293 and mESC
- Strong enrichment of Firre after RIP targeting hnRNPU

- Analysis of publicly available data from UV-crosslinked RIP targeting hnRNPU (data showed as gene-level FPKM values)
- Specific binding of Firre to hnRNPU

Conclusion: Firre directly and specifically binds to hnRNPU

hnRNPU regulates the spatial expression of Firre





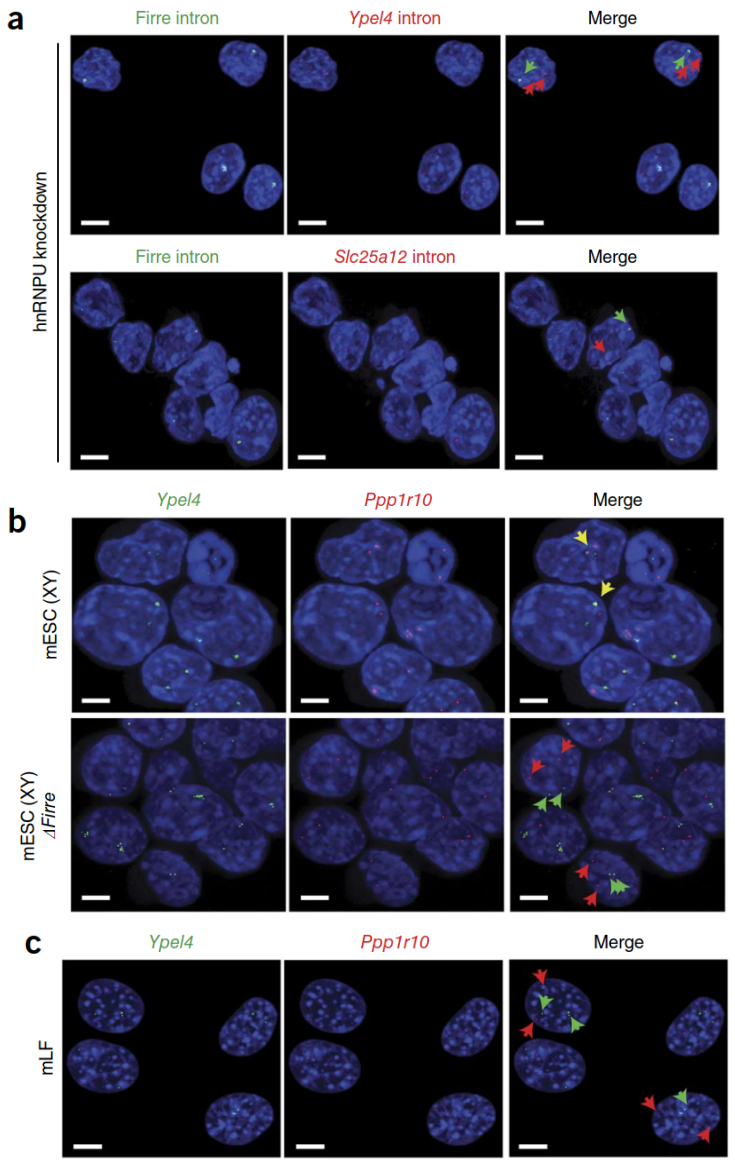
- Knockdown of hnRNPU with a siRNA in male mESCs and human cell lines (HEK293, HeLa)
- FISH targeting Firre
- hnRNPU KD leads to the delocalization of Firre in the cell (mESC, HEK293, Hela)

Conclusion: Firre nuclear localization is depending on hnRNPU

Decreased colocalization of each trans site with Firre in absence of hnRNPU

- Knockdown of hnRNPU with a siRNA
- RNA FISH in hnRNPU KD male mESCs and in mLFs not expressing Firre

Conclusion: hnRNPU is responsible for the colocalization of Firre and its trans-sites



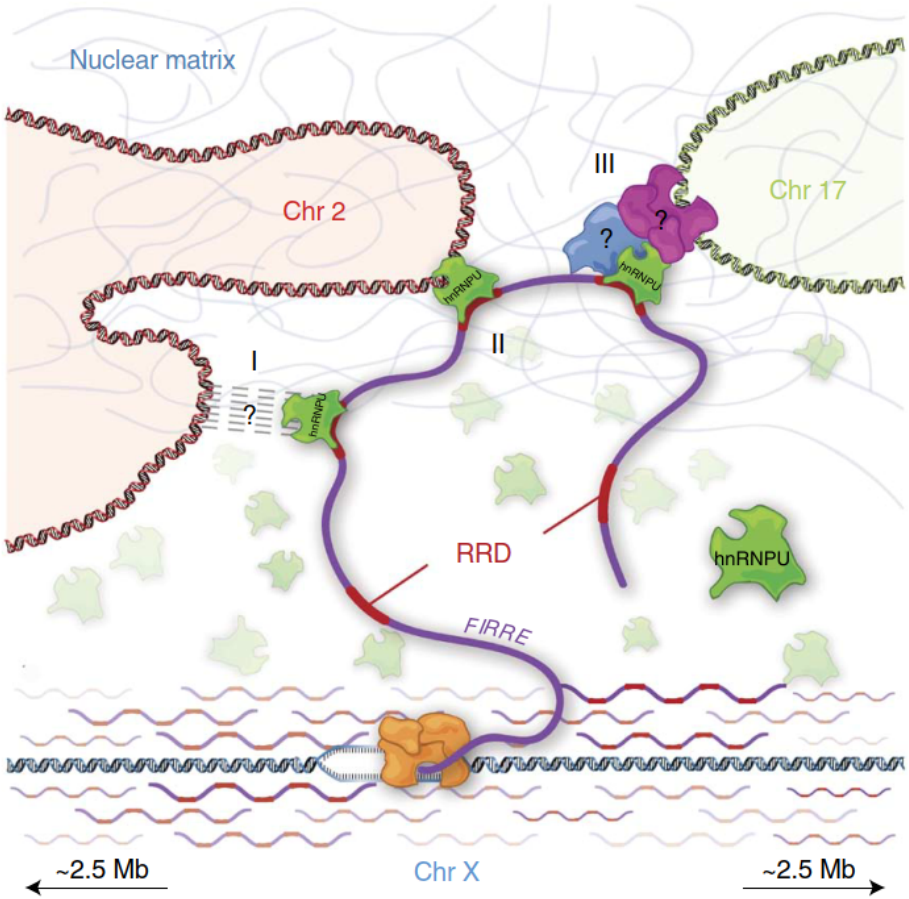
hnRNPU facilitates the trans-chromosomal interactions

Model for hnRNPU interactions with trans-chromosomal regions:

I - Tertiary interactions with nuclear-matrix components

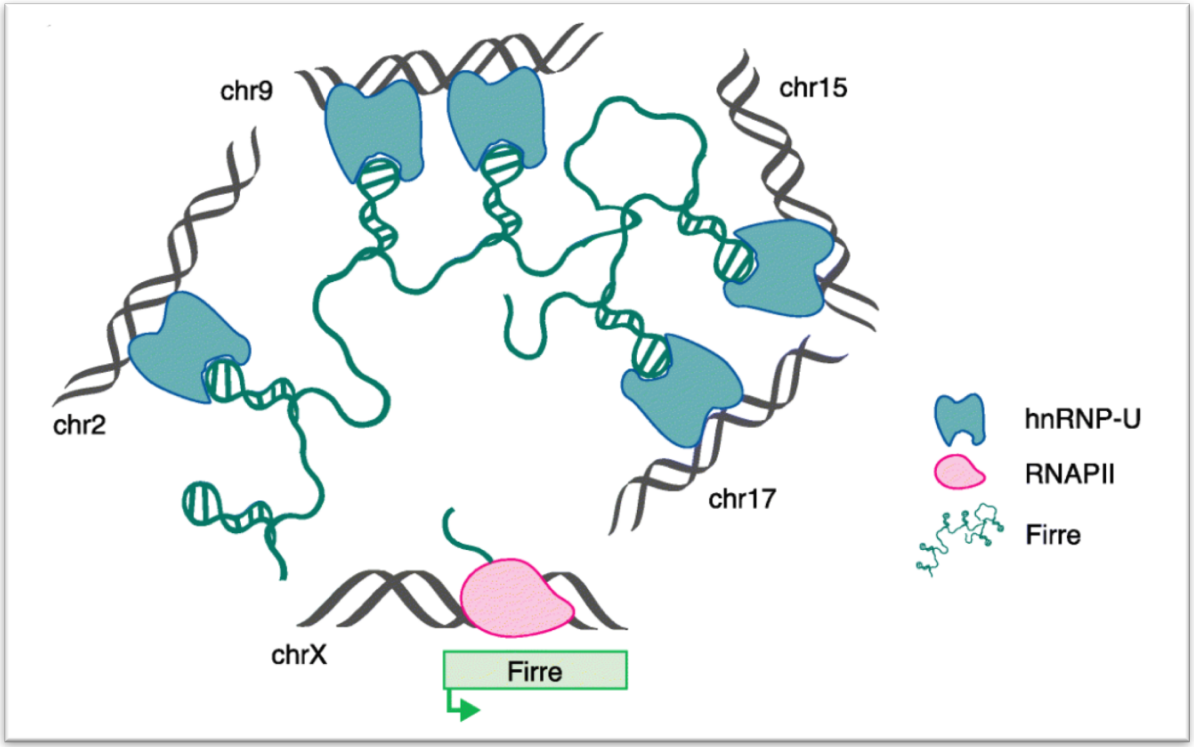
II - Direct binding of hnRNPU to matrix attachment regions in trans

III - Interactions with other protein complexes to facilitate the indirect binding to the DNA



Summary

- Identification of the new ncRNA-encoding gene, Firre
- Firre localizes in the nucleus, around its site of transcription on X chr
- It escapes X chr inactivation
- Firre interacts with trans-sites on other chromosomes through hnRNPU
- Firre plays a role in pluripotency and adipogenesis



Conclusion

- Important roles in cell physiology and nuclear architecture of lincRNA Firre
- Future genetic studies in mouse models will be required to further illuminate the role of Firre in mammalian development and disease.

Thanks for your attention!

References

- Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. Ezgi Hacisuleyman, Loyal A Goff, Cole Trapnell, Adam Williams, Jorge Henao-Mejia, Lei Sun, Patrick McClanahan, David, G Hendrickson, Martin Sauvageau, David R Kelley, Michael Morse, Jesse Engreitz, Eric S Lander, Mitch Guttman, Harvey F Lodish, Richard Flavell, Arjun Raj, John L Rinn Nat Struct Mol Biol 2014 Feb;21(2):198-206.
- Genome Regulation by Long Noncoding RNAs. John L. Rinn1 and Howard Y. Chang2 Annu. Rev. Biochem. 2012. 81:145–66
- Long noncoding RNA (lincRNA), a new paradigm in gene expression control. Emre Deniz1 & Batu Erman2, Funct. Integr. Genomics (2017) 17:135–143