

UNIVERSITÉ PARIS DIDEROT
(PARIS 7)
SORBONNE PARIS CITÉ

ÉCOLE DOCTORALE N° 562 BIO Sorbonne Paris Cité
Biochimie, Biothérapies, Biologie moléculaire, Infectiologie
UMR7216 Epigénétique et Destin Cellulaire

Thesis
in
Genomes, Epigenomes and Cell Fate

Giulia FURLAN

Investigating the contribution of the non-coding gene *Ftx*
to X-chromosome inactivation in mammals

PhD directed by Claire ROUGEULLE
Date of defense: 23rd of September 2016

Jury members

President Reiner VEITIA
Reviewer Joost GRIBNAU
Reviewer Alena SHKUMATAVA
Examiner Elphège NORA
PhD supervisor Claire ROUGEULLE

«Nothing in biology makes sense except in the light of Evolution»

-Theodosius Dobzhansky, 1973-

ACKNOWLEDGEMENTS

I sincerely thank all the members of the jury for having accepted to evaluate my work and for taking the time to read this manuscript: Pr. Reiner Veitia, Pr. Joost Gribnau, Dr. Alena Shkumatava and Dr. Elphège Nora.

I would like to especially thank my supervisor, Claire Rougeulle, who entrusted me with this project and provided constant guidance throughout my PhD, engaging me in stimulating discussions that made me a better scientist.

Special thanks go to Céline Morey, who shared with me happy moments and endless trouble, between failed experiments, black cats and crocodiles. Houston!

I would like to thank Jean-François Ouimette for being there for me, always. (And, especially, for finding six blue-cap cuvettes.)

I thank Melanie Makhoul, a friend inside and outside the lab, for her support, her patience with me, and all what we shared: daily 4 p.m. chips from the vending machine, a fast-growing cactus with five fingers, countless great action movies (and some dubious ones), the best of Homeland, laughter and tears, Exki snacks, Lebanese dinners... I am almost finished- one second, ok?

I thank Olga Rosspopoff for being my friend no matter what, for sharing with me rogue missions and secret whispers, and for forgiving me more than once.

Christophe Huret, for his advices on life and for helping me putting things in perspective.

Celine Vallot, for precious comments and great ideas.

Tharvesh Moideen, for good laughs and special insights on Indian culture.

Miguel Casanova, for reminding me that there are other things in life than science: time to open that Portuguese cheese shop!

Myriam Mohamed, who brightens bad lab days with her smile.

Antonio Romito, who started the project before I joined the team, and all the members of the unit, for their help and advice.

I thank my parents and my sister for supporting me all these years and for believing in me.

Last but not least, Andrea Frapporti. He knows why.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	1
SUMMARY	i
RÉSUMÉ	iii
INTRODUCTION	1
1 Evolutionary origin of sex chromosomes and the need for dosage compensation mechanisms	1
1.1 Evolutionary origin of sex chromosomes.....	1
1.1.1 Evolution of the XY/XX and ZW/ZZ sex determination systems	2
1.1.2 Molecular origin of the modern therian XX/XY sex determination system	4
2 A variety of dosage compensation strategies	8
2.1 Dosage compensation in Mammals	10
2.1.1 Dosage compensation in Prototheria (Monotremes).....	10
2.1.2 Dosage compensation in Theria.....	12
2.1.2.1 Developmental regulation of X chromosome inactivation in Theria	13
2.1.2.2 <i>Ex-vivo</i> cellular systems	16
3 Long non-coding RNA-based dosage compensation mechanisms in Theria	21
3.1 Non-coding genome	21
3.1.1 The discovery of non-coding genome	21
3.1.2 Long non-coding RNAs	22
3.2 Evolution of the long non-coding RNAs <i>Xist</i> and <i>Rsx</i>	25
3.2.1 The eutherian <i>X inactivation center</i>	25
3.2.2 <i>Rsx</i> , a metatherian X-linked long non-coding RNA	33
4 <i>Xist</i> , the master regulator of X chromosome inactivation in Eutheria	35
4.1 <i>Xist</i> discovery and its structure in different eutherian species.....	35
4.2 <i>Xist</i> cellular localization and expression profiles in different eutherian species	36
4.3 <i>Xist</i> requirement in X chromosome inactivation	41
4.3.1 <i>Xist</i> and the initiation of X chromosome inactivation	41
4.3.1.1 <i>Xist</i> upregulation and stabilization.....	43
4.3.1.2 <i>Xist</i> spreading	44
4.3.1.3 <i>Xist</i> functional regions and protein interactors	47

4.3.1.4	Features of <i>Xist</i> -mediated transcriptional silencing.....	53
4.3.2	<i>Xist</i> and the maintenance of X chromosome inactivation.....	56
4.3.2.1	XCI maintenance in the embryonic lineages: <i>Xist</i> and other epigenetic pathways..	56
4.3.2.2	<i>Xist</i> -independent XCI maintenance in adult tissues.....	58
5	<i>Xist</i> regulation	59
5.1	<i>Xist</i> repression.....	60
5.1.1	Undifferentiated mESCs: X chromosome inactivation and pluripotency	60
5.1.2	XaXa status and differentiation potential	63
5.1.3	<i>Xist</i> repression and <i>Tsix</i> activation	64
5.1.4	Absence of interaction with <i>Xist</i> binding proteins	67
5.2	<i>Xist</i> activation	67
5.2.1	X-X pairing and the initiation of XCI: <i>Tsix</i> and <i>Xpr</i>	69
5.2.2	<i>Tsix</i> and XCI choice.....	71
5.2.3	Dose-dependent XCI regulators: <i>Rnf12</i> and <i>Jpx</i>	71
5.2.4	Short non-coding transcripts	75
5.2.5	<i>Xce</i> alleles and XCI skewing	75
5.2.6	YY1-dependent <i>Xist</i> transcriptional activation.....	76
5.2.7	<i>Xist</i> repression from the Xa: DNA methylation	77
6	<i>Ftx</i> , an additional non-coding gene in the X inactivation center.....	78
6.1	<i>Ftx</i> structure and conservation across eutherians.....	79
6.2	<i>Ftx</i> cellular localization and expression patterns.....	82
6.3	<i>Ftx</i> expression profiles	84
6.4	<i>Ftx</i> and the hotspot.....	87
	RESULTS	89
1	<i>Ftx</i> deletion strategy.....	89
2	Generation of <i>Ftx</i> mutant female ES cell lines: LF2.....	89
3	<i>Ftx</i> mutant cell lines display female-specific differentiation defects	97
4	<i>Ftx</i> is required for <i>Xist</i> accumulation	106
5	Generation of <i>Ftx</i> mutant female ES cell lines: PGK1	115
6	The lack of <i>Ftx</i> results in cell death and impaired <i>Xist</i> expression also in PGK1 cells	121

7	<i>Ftx</i> is required for <i>Xist</i> accumulation in <i>cis</i>	127
8	Insights into the mechanisms of <i>Ftx</i> -mediated <i>Xist</i> regulation.....	133
8.1	At the level of <i>Xist</i>	133
8.2	At the level of <i>Ftx</i>	135
	CONCLUSIONS	143
	DISCUSSION	145
1	<i>Ftx</i> and cell differentiation	145
2	<i>Ftx</i> and the regulation of <i>Xist</i>	146
3	Mechanisms of <i>Ftx</i> action	148
4	Transcript, act of transcription and/or genomic region?	149
4.1	Exploring the role of the <i>Ftx</i> transcript.....	151
4.2	Exploring the role of ongoing transcription	152
4.3	Exploring the role of the genomic region.....	153
5	<i>Ftx</i> and other XCI regulators	158
6	<i>Ftx</i> and XCI in other eutherian species.....	159
7	From an <i>ex vivo</i> to an <i>in vivo</i> model	160
8	<i>Ftx</i> expression patterns: a role for the RNA transcript?	164
8.1	<i>Ftx</i> as a miRNA precursor	165
8.2	A role for the <i>Ftx</i> transcript in other contexts.....	166
	MATERIALS AND METHODS	169
	REFERENCES	175
	ANNEXES	213

TABLE OF FIGURES

Introduction

Figure 1. Sex determination signals in evolution	2
Figure 2. Evolutionary origin of sex chromosome systems	4
Figure 3. Origin of the eutherian Y chromosome	6
Figure 4. Developmental regulation of XCI in the mouse	14
Figure 5. <i>Ex vivo</i> cellular systems	20
Figure 6. Mechanisms of action of lncRNAs.....	24
Figure 7. Emergence of <i>Xist</i> and <i>Rsx</i>	25
Figure 8. Murine and human X inactivation centers	28
Figure 9. <i>Xic</i> lncRNAs evolution from protein-coding genes	32
Figure 10. <i>Xist/XIST</i> structure in mouse and human	36
Figure 11. Developmental patterns of <i>Xist</i> expression	40
Figure 12. <i>Xist</i> spreading.....	47
Figure 13. Functions, activities and interactors of <i>Xist</i> RNA repeat elements	53
Figure 14. Features of <i>Xist</i> -mediated silencing	55
Figure 15. Absence of X chromosome inactivation in undifferentiated mESCs.....	62
Figure 16. Female-specific and monoallelic <i>Xist</i> upregulation in differentiating mESCs	68
Figure 17. <i>Rnf12</i> activates <i>Xist</i> expression in a dose-dependent manner	73
Figure 18. Delimitation of the <i>X controlling element</i> region	76
Figure 19. <i>Ftx</i> and the murine <i>Xic</i>	78
Figure 20. <i>Ftx</i> locus and transcript isoforms.....	79
Figure 21. Conservation of <i>Ftx</i> across species	81
Figure 22. <i>Ftx</i> subcellular localization in undifferentiated mESCs.....	82
Figure 23. The <i>Ftx</i> locus generates multiple RNA isoforms.....	83
Figure 24. <i>Ftx</i> expression profiles in different tissues.....	85
Figure 25. <i>Ftx</i> expression profiles in differentiating mESCs.....	86
Figure 26. Expression profiles of <i>Ftx</i> and the miR cluster in differentiating female ESCs	86

Figure 27. <i>Ftx</i> impacts on <i>Xist</i> expression levels and chromatin structure at the <i>Xist</i> locus	88
---	----

Results

Figure 1. Generation of LF2 <i>Ftx</i> mutant cell lines.....	94
Figure 2. <i>Ftx</i> expression profiles in female and male WT and <i>Ftx</i> mutant cells	101
Figure 3. Female-specific <i>Ftx</i> -dependent differentiation defect.....	103
Figure 4. <i>Ftx</i> is required for <i>Xist</i> accumulation and proper XCI progression independently of other <i>Xic</i> -linked genes	110
Figure 5. <i>Ftx</i> and <i>Xist</i> are mostly co-expressed	112
Figure 6. Generation of PGK1 <i>Ftx</i> mutant cell lines	118
Figure 7. The lack of <i>Ftx</i> results in cell death and impaired <i>Xist</i> expression in PGK1 cells	124
Figure 8. <i>Ftx</i> is required for <i>Xist</i> accumulation in <i>cis</i>	130
Figure 9. Insights into the mechanisms of <i>Ftx</i> -dependent <i>Xist</i> regulation at the level of <i>Xist</i>	139
Figure 10. Knocking down the <i>Ftx</i> transcript	141
Figure S1. Generation of LF2 <i>Ftx</i> mutant cell lines	96
Figure S2. <i>Ftx</i> expression profiles and female-specific <i>Ftx</i> -dependent differentiation defect.....	105
Figure S3. <i>Ftx</i> is required for <i>Xist</i> accumulation and the effects is independent of the subcloning procedure	114
Figure S4. Generation of PGK1 <i>Ftx</i> mutant cell lines.....	120
Figure S5. The lack of <i>Ftx</i> results in cell death and impaired <i>Xist</i> expression in PGK1 cells	126
Figure S 6. <i>Ftx</i> is required for <i>Xist</i> accumulation in <i>cis</i>	132

Discussion

Figure 1. Working models.....	151
Figure 2. <i>Ftx</i> promoter usage in differentiating female ES cells	152
Figure 3. Binding sites mapping to the X _{Pgk} deletion in PGK1 clone B (ΔFtx_{129}).	156
Figure 4. Protein factors binding sites mapping to the <i>Ftx</i> deleted region.....	158

SUMMARY

X-chromosome inactivation (XCI) is a chromosome-wide regulatory process that, in eutherians, ensures dosage compensation for X-linked genes between sexes. XCI is controlled by a *cis*-acting locus on the X-chromosome, the X-inactivation center (*Xic*), which is particularly enriched in genes producing long non-coding RNAs (lncRNAs) and organized in a bipartite three-dimensional structure. *Xist*, the best-characterized lncRNA of the *Xic*, is the master player of XCI, ultimately responsible for X-chromosome-wide silencing. *Xist* produces a non-coding transcript that accumulates in *cis* on the X-chromosome and recruits the machinery responsible for initiation and propagation of silencing through cell divisions. Other lncRNA genes that map to the *Xic* are involved in the fine-tuning of *Xist* expression. These lncRNAs loci are rapidly evolving, and their plasticity might be at the basis of the different XCI strategies that exist in eutherians and that underlie the determination and maintenance of cell identity across species.

Years ago, the laboratory has identified an additional *Xic*-linked non-coding gene, *Ftx*, localized upstream to *Xist* and transcribed in the same orientation. In the mouse, *Ftx* expression increases during differentiation of embryonic stem cells (ESCs), an extremely useful model to study early development *ex vivo*, thanks to their potential to differentiate in multiple cell types and the ability of female cells to recapitulate XCI upon exit from pluripotency. In female cells, *Ftx* escapes from silencing and its increase in expression is concomitant with *Xist* upregulation at the onset of XCI.

The fact that *Ftx* is upregulated during ESC differentiation, in a profile that correlates with that of *Xist* in females, suggests that *Ftx* could have a role in cell differentiation and/or in XCI. Since these two processes are tightly intertwined in the mouse, we explored the possibility that *Ftx* could play a role in XCI either directly or indirectly, through effects on differentiation. In order to distinguish between the two scenarios, we generated female ESC lines (that differentiate and undergo XCI) (this study) and male ESC lines (that differentiate without undergoing XCI) (Chureau et al., 2011) with impaired *Ftx* expression.

By using a panel of complementary approaches, I could show that, while differentiating mutant male cells show no obvious defects, differentiating mutant

female cells die at high rate. This female-specific phenotype is highly suggestive of defects in XCI that are independent of differentiation.

In female *Ftx*^{-/-} lines, XCI is strongly impaired, with a significant decrease in the levels of *Xist* expression and in the percentage of cells showing normal *Xist* accumulation patterns. Importantly, a high proportion of the cells that still retain *Xist* expression show abnormal X-chromosome coating, with diffused *Xist* signals rather than a compact *Xist* “cloud”, and a decreased ability to silence X-linked genes. These data reveal that *Ftx* is a positive *Xist* regulator and it is required for proper XCI establishment. In female *Ftx*^{+/-} lines, the levels of *Xist* expression and the percentage of cells showing normal *Xist* accumulation patterns are also decreased, albeit to a lower extent compared to *Ftx*^{-/-} lines, suggesting that *Ftx* works in a copy-dependent manner. In addition, a high proportion of *Ftx*^{+/-} cells display skewed X-inactivation, with preferential inactivation of the wild-type X chromosome. This suggests that *Ftx* role on *Xist* accumulation is mostly restricted in *cis*. Taken together, these results demonstrate that *Ftx* is required for XCI establishment, where it functions as a strong *Xist* activator.

RÉSUMÉ

L'inactivation du chromosome X (XCI) est un mécanisme qui permet l'extinction transcriptionnelle d'un des deux chromosomes X présent chez la femelle. Bien que ce processus soit essentiel chez les mammifères, il est mis en place selon des stratégies propres à chaque espèce. L'inactivation du chromosome X, qui permet d'équilibrer la dose de transcrits provenant du chromosome X entre les mâles et les femelles, est régulé par une région spécifique nommée centre de l'inactivation du chromosome (*Xic*). L'étude de cette région a révélé deux caractéristiques essentielles à son fonctionnement: 1) son organisation tri-dimensionnelle en deux sous-domaines et 2) une abondance élevée de gènes produisant de longs ARNs non codants (lncRNAs). Parmi ces lncRNAs, le transcrit *Xist* est l'effecteur principal pour l'inactivation du X. En effet, lorsqu'il est produit en quantité suffisante, ce transcrit peut s'accumuler en *cis* sur le chromosome et recruter la machinerie qui permettra l'initiation et la propagation de l'extinction transcriptionnelle à l'échelle du chromosome. Plusieurs autres loci du *Xic* produisant des lncRNAs ont été décrits comme des régulateurs «fins» de l'expression de *Xist*; leur délétion produit des effets mineurs sur l'expression de *Xist*. Ces loci, dont la séquence est divergente entre les espèces, pourrait contribuer à la diversité des mécanismes de mise en place de l'inactivation observée au sein des mammifères.

Le laboratoire d'accueil a identifié un nouveau gène du *Xic* qui produit le lncRNA *Ftx*, qui est situé en amont de *Xist* et transcrit dans la même orientation. Dans la souris, l'expression de *Ftx* augmente au cours de la différenciation des cellules souches embryonnaires (ESC), un modèle *ex vivo* classique pour l'étude des étapes précoces du développement embryonnaire. Les ESC femelles déclenchent notamment l'inactivation du chromosome X lorsqu'elles sont mises dans un milieu de différenciation. Dans ces conditions, l'expression de *Ftx* est augmentée en parallèle de celle de *Xist*, avant le déclenchement de l'XCI. Ceci nous a conduit à explorer le rôle que *Ftx* pourrait avoir dans la régulation de l'inactivation du chromosome X et la différenciation, deux processus qui sont connus pour être interconnectés chez la souris. Afin de discriminer la contribution de *Ftx* dans ces deux processus, nous avons délété *Ftx* dans des lignées ESC femelles (qui se différencient et déclenchent

l'XCI) et mâles (que se différencient sans XCI).

En utilisant une combinaison d'approches expérimentales complètes, j'ai pu montré que la délétion de *Ftx* n'entraîne pas de défauts évidents au cours de la différenciation des ESC mâles. Au contraire, chez les femelles, la délétion de *Ftx* entraîne une mort cellulaire massive au cours de la différenciation, ce qui suggère un problème spécifique dans l'XCI plutôt qu'un défaut de différenciation cellulaire. En effet, l'inactivation du chromosome X est fortement perturbée dans les cellules *Ftx*^{-/-} et s'accompagne par une forte baisse du niveau d'expression et d'accumulation de *Xist*. Dans ce contexte, certaines cellules parviennent à maintenir l'expression de *Xist* mais le profil de couverture du chromosome X par *Xist* est anormal, présentant un profil diffus; ceci est associé à une extinction transcriptionnelle déficiente des gènes liés à l'X. Il apparaît donc que *Ftx* est un activateur de *Xist* et qu'il est requis pour une mise en place efficace de l'ICX. Dans les lignées hétérozygotes *Ftx*^{+/-}, l'expression et l'accumulation de *Xist* est aussi affectée mais dans une moindre mesure, si bien qu'il apparaît que le nombre de copies de *Ftx* soit important pour sa fonction. Par ailleurs, l'inactivation du chromosome X dans les cellules *Ftx*^{+/-} est biaisée de telle sorte que le chromosome X portant une copie fonctionnelle de *Ftx* est préférentiellement inactivé, suggérant un rôle en *cis* de *Ftx*. Ces résultats montrent que *Ftx* est un activateur de *Xist* et qu'il est essentiel pour la mise en place de l'inactivation.

INTRODUCTION

1 Evolutionary origin of sex chromosomes and the need for dosage compensation mechanisms

Something as important as sex determination might seem static and unchanging, but in reality sex determination systems are constantly evolving. Sex can be determined by chromosomes, where one sex is homogametic for the sex chromosome and the other is heterogametic. As a consequence of these differences, the heterogametic sex is functionally monosomic for that chromosome, which generally encodes genes involved in many processes required for life, and not just sex-specific processes. To cope with this disparity, dosage compensation equalizes sex chromosome expression between males and females. Proper chromosome dosage is essential for the viability and fitness of an organism and, despite different dosage compensation strategies exist, they all serve the same purpose and share common molecular mechanisms.

1.1 Evolutionary origin of sex chromosomes

In many vertebrate species, a bipotential gonad develops into either testes or ovaries based on the action of an initial sex determination signal. Sex determination signals vary widely among species and can be environmental or genetic (Figure 1) (Bachtrog et al., 2014). Environmental sex determination is widespread in lower vertebrates, and genetic sex determination has probably evolved from it several times. An example of environmental sex determination is found in reptiles (Gilbert, 2000), where external factors, such as incubation temperature during embryonic development after egg fertilization, specify the sex of the offspring. A major disadvantage of environmental sex determination is that the existence of a species can be threatened by sudden changes in the environment. According to one theory (Miller et al., 2004), climate changes caused the extinction of the dinosaurs, because their temperature-dependent sex-determination system might have forced them to produce predominantly male or female offspring. Sex is instead genetically determined in species relying on internal reproduction, the appearance of which

challenged temperature-dependent sex determination mechanisms and favored the evolution of chromosome-based sex determination systems.

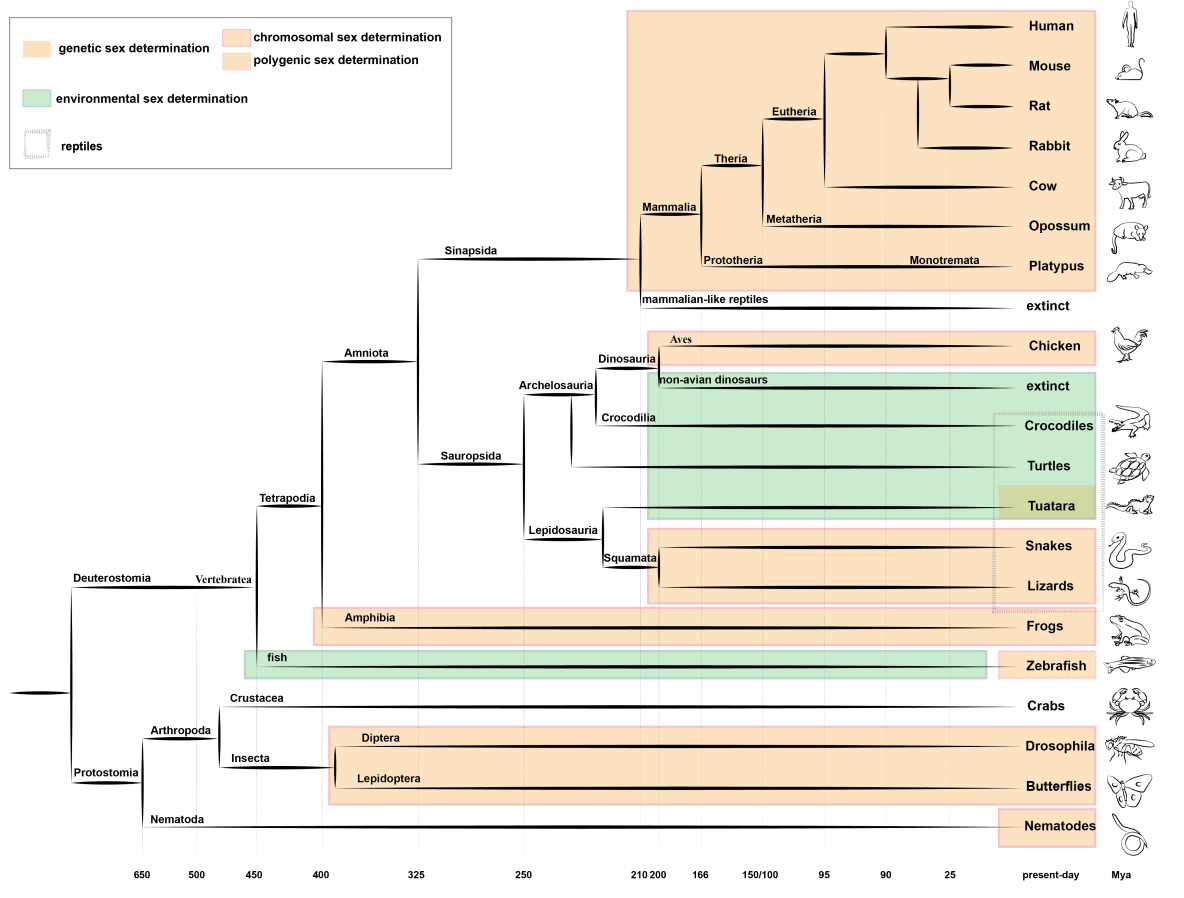


Figure 1. Sex determination signals in evolution

Phylogenetic tree illustrating the different types of sex determination signals.

1.1.1 Evolution of the XY/XX and ZW/ZZ sex determination systems

Sex chromosome systems observed in present-day species vary considerably but a common feature is the presence of morphologically distinct sex chromosomes. In the XY sex chromosome system (such as in mammals and many insects), females have two X chromosomes, and males are XY or XO (male heterogametic systems). In the ZW system (such as in birds), males are ZZ, and females are ZW (female heterogametic systems).

Both male heterogamety and female heterogamety have independently evolved numerous times (Charlesworth, 1991). For example, while monotremes and therians both have XY/XX systems, their X and Y chromosomes are thought to have evolved independently (Figure 2). The initial observation that the monotreme X chromosomes have significant sequence similarities to the avian Z chromosome (Veyrunes et al.,

2008), led to the hypothesis that basal amniotes displayed early prototypes of the avian W and Z chromosomes that were passed along both sauropsid and synapsid lines, and subsequently kept in the bird lineage but modified in the monotreme lineage. After the prototherian/therian divergence the monotremes began evolving from a ZW/ZZ system to a XY/XX system. Along the line to marsupials and placental mammals (Eutheria) our ancestors lost their dependence upon the ZW/ZZ system and evolved the modern therian X and Y chromosomes.

Interestingly, genes on the human X chromosome have orthologs on platypus chromosome 6, an autosome (Veyrunes et al., 2008). The ancestor of this autosome probably evolved into the human X chromosome, but remained autosomal in monotremes, indicating that, while the monotreme and therian lineages both inherited the ancestral proto-W and proto-Z chromosomes, both lineages then evolved their own X and Y chromosomes to replace these (Veyrunes et al., 2008).

Since the marsupial and placental mammal sex chromosomes are homologous, it is safe to assume that the therian XY/XX system evolved before the marsupials and placental mammals diverged. However, these chromosomes have undergone subsequent changes in both lineages (Veyrunes et al., 2008).

On a side note, although they are not homologous, the therian XY and the avian ZW display similar features. The bird Z and the mammalian X are highly conserved among species and possess a sex-biased gene content, while the bird W and the mammalian Y are both degrading to a certain extent during evolution.

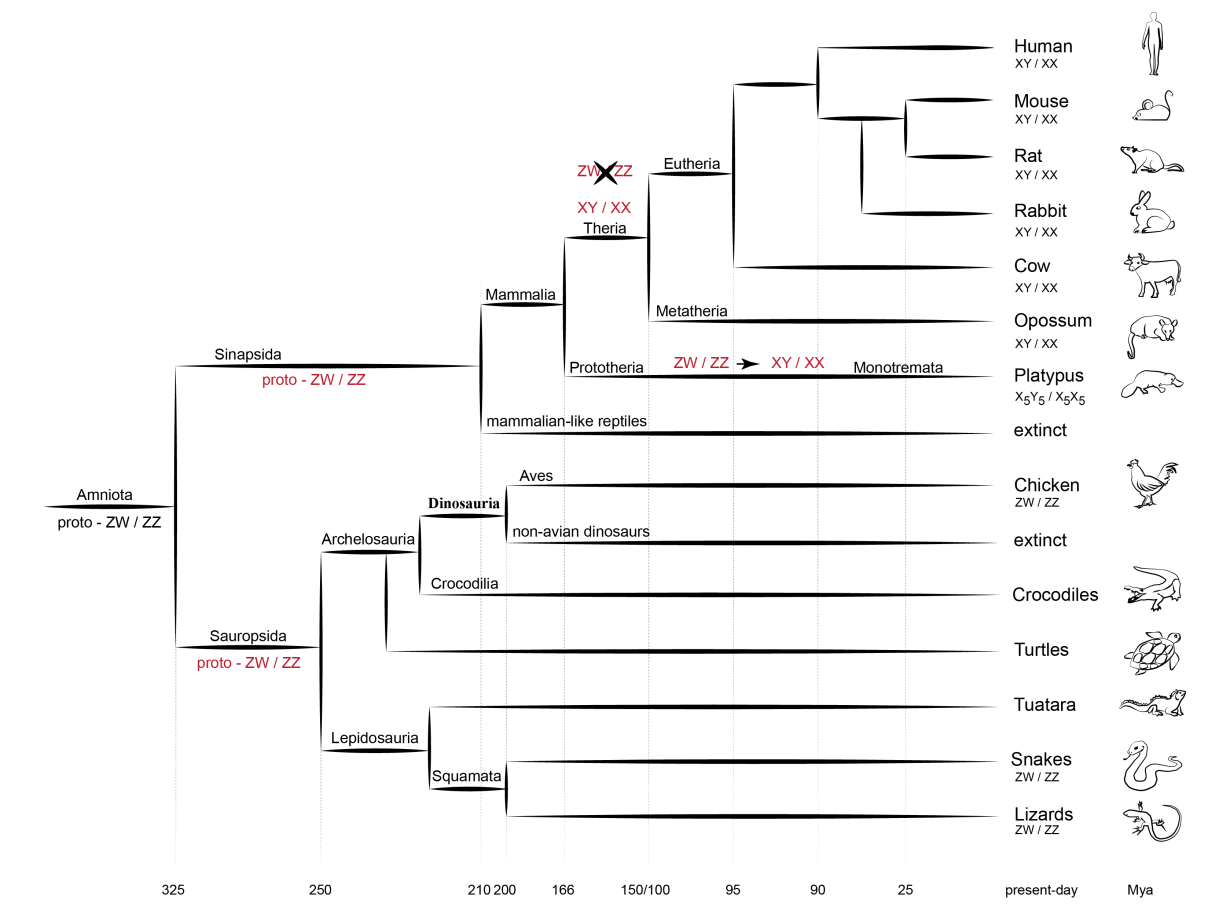


Figure 2. Evolutionary origin of sex chromosome systems

Phylogenetic tree showing the origin of present-day XY/XX and ZW/ZZ sex chromosome systems.

1.1.2 Molecular origin of the modern therian XX/XY sex determination system

Albeit independent in many different taxa, the parallel evolution of modern sex chromosome determination systems seems to have repeatedly started with the modification of a pair of identical (non sex-determining) autosomes, that diverged from each other over the course of time after the initial acquisition of a sex-determining role (Charlesworth, 1991).

The best characterized example is the therian system, whose Y chromosome appeared around 160 million years ago, when one of the autosomes mutated, acquiring the *Sry* (Sex-determining region Y) gene, that limited the Y chromosome to the male sex (Bachtrog, 2013). The *Sry* gene is believed to have arisen by mutational change of the *Sox3* gene, whose protein product, SOX3, is still capable of mimicking the action of SRY if expressed at high doses (Sutton et al., 2011).

The therian *Sry* interestingly maps on platypus chromosome 6, but platypuses

actually lack SRY, suggesting that the appearance of this functional gene in therians was the trigger for the evolution of the modern therian XY/XX system. As this gene began to take a role in male sexual differentiation, selection favored the limitation of recombination between the proto-Y chromosome and the proto-X chromosome and the segregation of male-specific genes on the new Y chromosome. The resulting loss of homology between regions of the new Y and X chromosomes containing and flanking *Sry* led, over time, to a decrease in the rate of meiotic recombination between the two chromosomes. Other chromosomal rearrangements such as fusions and translocations and the following gradual loss of gene functions on the Y chromosome further limited genetic exchange, accelerating the morphological differentiation of the Y from the X.

Although they share a common evolutionary origin, the eutherian (Figure 3) and metatherian sex chromosomes further evolved along the two lineages, modifying their gene content to a different degree.

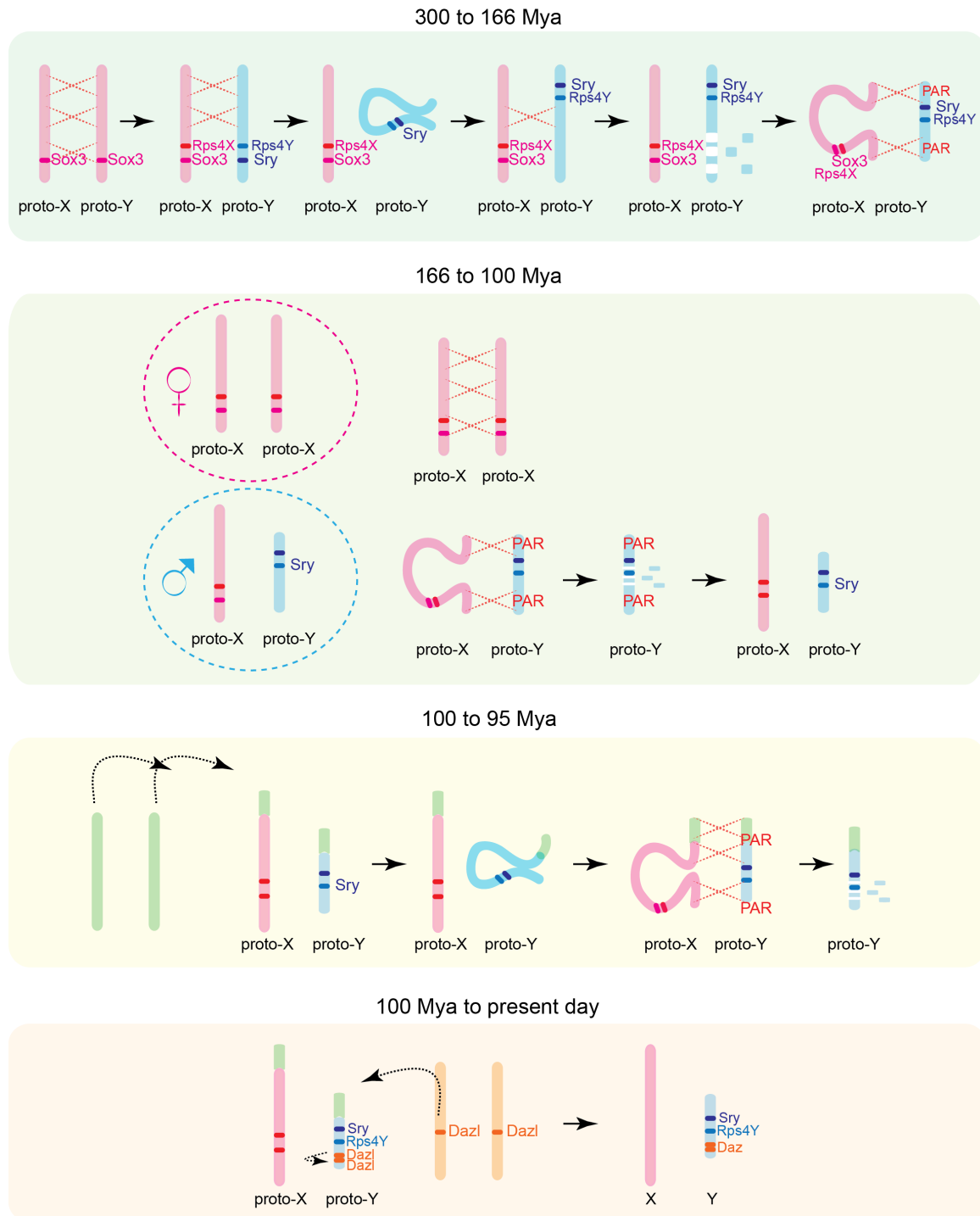


Figure 3. Origin of the eutherian Y chromosome

Around 300 mya, a mutation in the *Sox3* gene generated *Sox3*, a critical determinant of maleness, on the proto-Y. While the functions of *Sox3* and *Sry* rapidly diverged over time, a flanking gene, *Rpsx*, retained a similar function on both the X and Y chromosomes. Inversions (internal recombination events) facilitated the rearrangement of genes on the proto-Y, limiting recombination with the proto-X. The non-recombining portions of the proto-Y became susceptible to deletions, and the chromosome shrank. With the exception of two small pseudoautosomal regions (PARs), the proto-Y could not recombine with its counterpart. After the divergence of Theria from Metatheria, another inversion occurred on

the proto-Y. While in females recombination continued across the full length of the two identical proto-Xs, in males only the extremities of the chromosome were left able to recombine with homologous genes on the proto-X. Around 100 mya, a translocation occurred, whereby a block of autosomal genes extended the length of both the proto-X and proto-Y chromosomes. Additional inversions, absence of recombination and further gene loss caused the proto-Y to reduce its size once again. More recently, an autosome contributed a copy of the *Daz* gene (implicated in spermatogenesis) on the proto-Y chromosome, which duplicated a couple of times (probably to ensure that at least one copy is maintained in the case of further deletion events) and finally produce the present-day Y chromosome. The present-day Y chromosome is about 1/3 of the X chromosome in length and contains, in addition to male-specific genes, a series of dosage-sensitive genes with important ubiquitous regulatory roles.

2 A variety of dosage compensation strategies

Regardless the system (ZW/ZZ or XY/XX), since the heterogametic sex has (mostly) only a single copy of each sex chromosome, two major consequences arise: the need to equalize the dosage of the expression of genes on the common chromosome (Z or X) between the sexes, and between the sex chromosomes and the autosomes (Charlesworth, 1996).

In most model organisms studied so far, dosage compensation seems to be an essential requirement for successful development, and failure in dosage compensation leads to embryonic lethality (Belote and Lucchesi, 1980; Takagi and Abe, 1990). However, the extent to which the dosage is compensated can vary among and within species, from complete to partial (Table 1).

Species or clade	Karyotype (male/female)	Sex chromosome dosage compensation
Eutherian mammals	XY/XX	complete for dosage-sensitive genes
Metatherian mammals	XY/XX	incomplete
Aves	ZZ/ZW	incomplete
<i>Caenorhabditis elegans</i>	XO/XX	complete
<i>Drosophila melanogaster</i>	XY/XX	complete
<i>Ornithorinchus anatinus</i>	XY/XX	incomplete

Table 1. Extent of dosage compensation across species

Extent of dosage compensation in selected species or clades, with either XY/XX or ZW/ZZ sex determination systems.

Globally, female heterogametic ZW/ZZ systems seem to lack a global dosage compensation mechanism (Ellegren H 2007, Itoh Y 2007): even if with variations across developmental stages and tissues (Itoh et al., 2010; Mank and Ellegren, 2009), the averages of male to female ratios of Z-linked gene expression deviate from 1 (full compensation), but rarely reach 2 (complete absence of compensation). This suggests a low efficiency of global dosage compensation and argues in favor of a region-specific dosage compensation pattern.

Dosage compensation is more complete in the different XY/XX male heterogametic systems, as reported by extensive studies in the fly *Drosophila melanogaster*, the nematode *Caenorhabditis elegans* and several mammalian species. In flies and worms XX embryos become females (or hermaphrodites), while XY (or XO) embryos become males. In mammals, XX embryos become females and XY embryos become males. The dosage compensation strategies adopted by these species include two-fold X upregulation (flies) (Conrad and Akhtar, 2011), two-fold X downregulation

(worms) (Strome et al., 2014), and complete X transcriptional silencing (mammals) (Figure 4).

The genes encoding the dosage compensation machinery differ in these species, suggesting that they were put in place after each lineage divergence occurred. However, several of the dosage compensation genes have orthologs in other species and similarities have been reported regarding the mechanism used to neutralize the differences in X-linked gene dose: in all cases, specialized complexes are targeted to the X chromosome of only one sex to regulate transcription, reflecting the common need to achieve regulation of a consistent number of genes that are, in most cases, functionally unrelated but physically linked.

For example, dosage compensation in *Drosophila* shares interesting parallels with dosage compensation in Theria: both the upregulation of the single male X in *Drosophila* and the inactivation of one of the two Xs in female therians involve not only histone modifications but also the presence of non-coding RNAs (*roX* and *Xist*) (Lakhotia, 2015). *RoX* and *Xist* are not homologous and constitute very different systems, but in both cases the sex-specific RNA is transcribed from, and accumulates across the targeted sex chromosome, and is required for the targeting of the silencing complex to chromatin.

Like in mammals, worm X chromosome-wide regulation represents a repressed chromatin state that can spread in *cis* (Csankovszki et al., 2004). In all these cases, the selective recruitment of dosage compensation complexes establishes a transcriptional regulation of the X chromosome that is epigenetically maintained in the somatic cell lineages throughout the life of the individual.

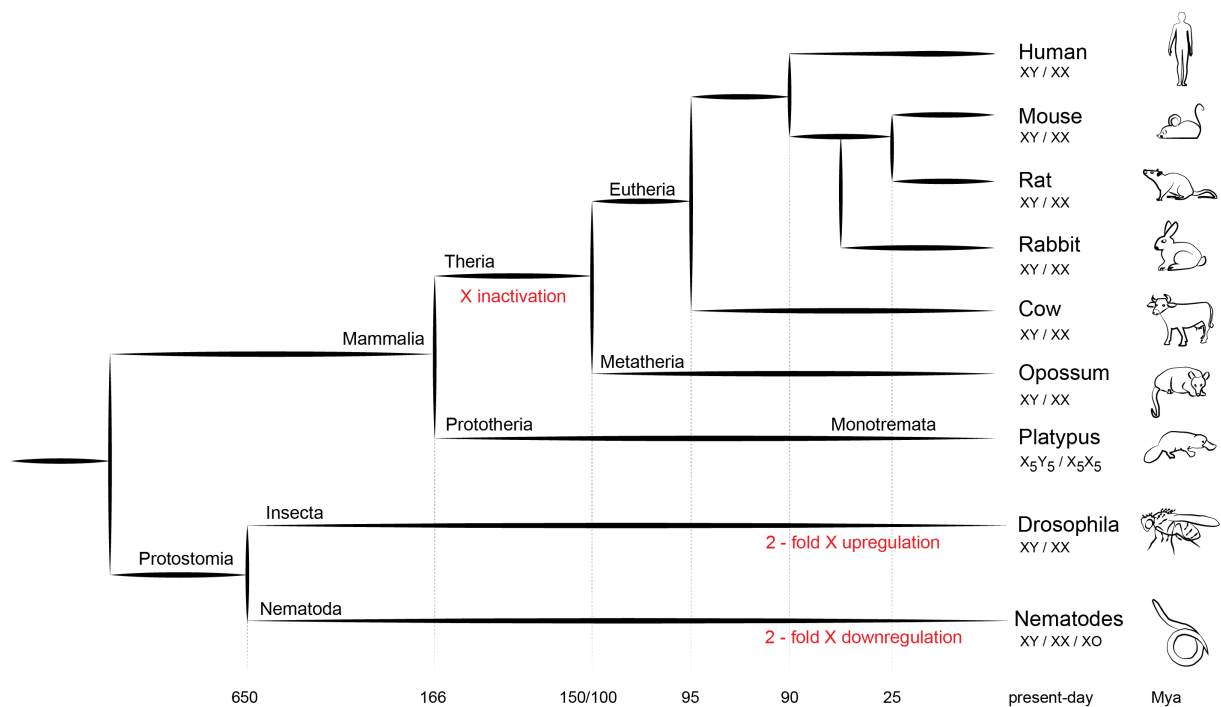


Figure . Dosage compensation strategies in XY/XX systems

Phylogenetic tree highlighting the emergence and the nature of dosage compensation strategies in XY/XX systems.

2.1 Dosage compensation in Mammals

Different mammals have developed different strategies to accomplish X chromosome inactivation.

In prototherian mammals, which possess multiple sex-chromosomes, X inactivation is incomplete and displays locus and tissue-specificity (similarly to birds). In therian mammals, dosage compensation for X-linked gene products between XX and XY individuals is generally achieved by fully silencing one of the two X chromosomes in females.

2.1.1 Dosage compensation in Prototheria (Monotremes)

Monotremes are unique mammals that display a mixture of reptilian and mammalian characteristics, as they lay eggs, yet have fur and produce milk. Represented only by the platypus and four species of echidna, they are distantly related to the other therian mammals, from which they diverged about 166 million years ago (Bininda-Emonds et al., 2007) (Figure 1).

The platypus (*Ornithorhynchus anatinus*) has five pairs of sex-chromosomes (Deakin et al., 2012). Males have five Xs and five Ys, and females have five pairs of Xs, all

genetically different (Grützner et al., 2004; Rens et al., 2004; Veyrunes et al., 2008). During male meiosis, the X and Y chromosomes pair at the ends to form an alternating chain of ten, with each chromosome linked to its neighbor *via* a pseudoautosomal region, from which the Xs segregate to one pole and the Ys to the opposite pole, forming female- and male-producing sperm (Rens et al., 2004). This system is thought to have arisen from sequential exchanges of DNA material between an ancient sex chromosome pair and four autosome pairs (Rens et al., 2007). Monotremes do not have *Sry*, and the actual candidate for sex determination is the *Amh* (anti-Müllerian hormone) gene, which is involved in the sex determination pathways of reptiles and fish (Cortez et al., 2014).

Necessity of dosage compensation in the female individuals of this mammalian lineage would be expected because of the complex sex chromosome system they possess, with the five (mostly unpaired) Xs accounting for 15% of the haploid genome. However, it appears that some genes on the sex chromosomes are not dosage-compensated, whereas others display varying degrees of compensation via stochastic, locus-specific transcriptional inhibition (Deakin et al., 2008a).

Arguing for the absence of dosage compensation, a study examining gene expression levels using RNA-seq data from different tissue types and cultured fibroblasts revealed that, for genes on platypus chromosome X5, median expression levels in females are nearly double than those in males (Julien et al., 2012). At the same time though, for any given X-linked gene, a specific proportion of fibroblast nuclei within a cell population was reported to silence one locus, with the percentage of cell inactivating that particular locus being gene-specific (Livernois et al., 2013). The fact that, in the same population, each gene is expressed from one or both loci at a reproducible frequency perhaps correlates with the fact that expression levels of X-linked compensated genes in females are almost double than in males (Julien et al., 2012; Livernois et al., 2013). Interestingly, for a given cell, the activity status of an X-linked gene is not clonally inherited (Livernois et al., 2013), unlike the situation in therians (section 2.1.2). Importantly, inactivated loci colocalize to the same X, suggesting that for each X chromosome pair, one chromosome might be fully active and the other partially inactive (Livernois et al., 2013).

It was postulated that the partial and variable dosage compensation system of monotremes represents an ancestral system that evolved to be more tightly

controlled in placental mammals. However, the X and Y chromosomes of monotremes are not homologous to the therian X and Y, and the two systems are proposed to have evolved concomitantly, yet independently (Cortez et al., 2014).

On the other hand, the platypus' five X and five Y chromosomes do share considerable genetic homology with the ZW sex chromosome pair of the chicken (Rens et al., 2007), where Z dosage compensation was also shown to be incomplete and gene-specific. This supports the hypothesis that the monotreme sex chromosome has evolved by repeated autosome translocation onto an original bird-like ZW pair (Gruetzner et al., 2006; Rens et al., 2004). It is thus possible that, like chickens, monotremes alter chromatin regionally, rather than chromosome-wide, to achieve gene-specific dosage compensation. This, together with the platypus' XY lack of homology with the therian XY chromosomes, suggests that monotremes might have retained an ancestral compensation system (Deakin et al., 2009), very different from the one later adopted by therian mammals.

2.1.2 Dosage compensation in Theria

Eutheria and Metatheria share orthologous sex chromosomes, that originated around 180 mya, 20 mya after the divergence of monotremes (Cortez et al., 2014). Despite their common origin, the separate evolution of the ancestrally common X and Y chromosomes led more recently to alterations in gene content along the eutherian and metatherian lineages.

The present-day eutherian X chromosome is enriched in genes with brain- and testis-specific functions, acquired via translocation of autosomal regions into the X (Delbridge et al., 2009; Pask and Graves, 1999, 2001; Pask et al., 1997, 2000) (Figure 3). As a consequence, the therian X is bigger than its marsupial counterpart. Interestingly, the autosomal regions translocated into the X chromosome in Eutheria contained genes that already had functions in brain and testes development (Hu et al., 2009, 2011; Pask et al., 1997).

The lack of autosomal addition in metatherian X evolution and the consequent absence of X-linked brain- and testis-specific genes suggests that the metatherian X chromosome may represent more closely the ancestral therian X chromosome.

Despite these differences in X-linked gene content, both Eutheria and Metatheria achieve dosage compensation for X-linked gene products between XX and XY

individuals by silencing one of the two Xs in females during early embryonic development. X chromosome inactivation (XCI) represents a specialized form of facultative heterochromatin formation whose status is epigenetically maintained through cell divisions and into the adult organism.

It is generally accepted that X inactivation is more stable and complete in Eutheria than in Metatheria (Piper et al., 1993). Eutherian XCI is a chromosome-wide silencing mechanism, with only a small proportion of genes (15% in humans and 3% in mice) escaping its effects (Berletch et al., 2015; Carrel and Willard, 2005; Yang et al., 2010). Interestingly, human escapees are concentrated on the short arm of the X chromosome (Carrel and Willard 2005, Cotton A 2013), reflecting the recent addition of this region to the human X and Y chromosomes (Veyrunes et al., 2010). On the contrary, marsupial XCI has been reported to be incomplete, with a large proportion of genes escaping XCI. In the South American grey short-tailed opossum (*Monodelphis domestica*) only around 14% of genes fail to be silenced (Wang et al., 2014), but in the Australian tammar wallaby (*Macropus eugenii*) the escapees represent 32% of genes (Al Nadaf et al., 2010). In this context, it is puzzling that the median expression levels of X-linked genes between male and female opossums are similar, with a ratio close to 1 (Pessia et al., 2012).

2.1.2.1 Developmental regulation of X chromosome inactivation in Theria

In both Eutheria and Metatheria, X inactivation is developmentally regulated. Two forms of XCI exist, imprinted (iXCI) and random (rXCI). These are mostly species-specific, but can also take place in different cell types/developmental stages of the same organism in particular cases (Figure 4).

In Metatheria, X inactivation is imprinted, resulting exclusively in the inactivation of the paternal X chromosome (Xp) in all tissues (Sharman, 1971). Imprinted X inactivation is also found in the mouse: in the pre-implantation embryo, the Xp is always inactivated, and remains silent in the extra-embryonic tissues (that give rise to placental tissue) and in the primitive endoderm (that forms the visceral endoderm and yolk sac) (Harper et al., 1982; Huynh and Lee, 2003; Latham, 1996; Mak et al., 2004; Okamoto et al., 2004; West et al., 1977). Imprinted XCI is also observed in extra-embryonic cell lineages of rat (Wake et al., 1976) and cow (Dindot et al., 2004a, 2004b; Xue et al., 2002). In the mouse, this imprinted X inactivation is

reverted in the epiblast lineage (Mak et al., 2004; Sheardown et al., 1997; Williams et al., 2011a), where the second form of silencing, random X inactivation, is subsequently established (McMahon et al., 1983; Rastan, 1982). Random XCI also takes place in humans and rabbits, where it seems to be the only form of X inactivation, both in embryonic and extra-embryonic cell types.

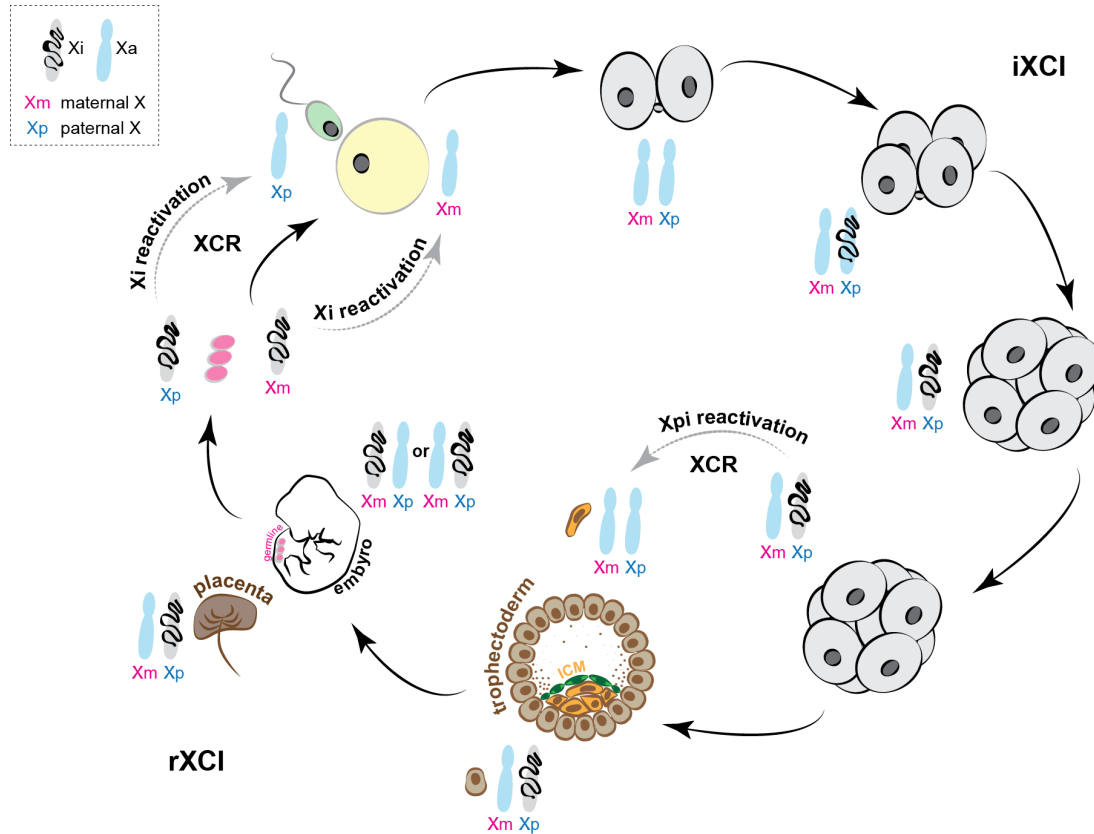


Figure 4. Developmental regulation of XCI in the mouse

In mice, X chromosome inactivation (XCI) begins at the four-cell stage and is initially imprinted (iXCI), with the preferential inactivation of the paternal X chromosome (Xp). In the inner cell mass (ICM) of the blastocyst, the Xp is reactivated (XCR), while in the extra-embryonic tissues it is maintained silenced. A second wave of inactivation takes place in the ICM (random XCI, rXCI), whereby one of the two chromosomes is randomly chosen for inactivation. The inactive state is maintained through cell divisions in the soma, but the inactive X (Xi) is reactivated during the formation of the female germ line, and all the gametes carry active X chromosomes.

It was suggested that marsupial X inactivation might exemplify an ancestral form of imprinted X inactivation, hypothesis that would be supported by the existence of imprinted X inactivation in the extra-embryonic tissues of rodents, and by the fact that fewer *cis*- and *trans*-acting factors seem to be required for inactivation, compared to rXCI. For example, an autosomal 250 kb transgene containing the XCI master

regulator *Xist* is able to recapitulate *cis*-inactivation with paternal iXCI kinetics in mice (Okamoto et al., 2005), while rXCI recapitulation fails even with the introduction of a much longer region (Heard et al., 1996). Moreover, imprinted XCI does not require a complex counting mechanism such as that of rXCI (section 3.2.1), since the paternal X is always the one that will be inactivated.

At the molecular level, iXCI implies the existence of an epigenetic difference between paternal and maternal X chromosomes. Indeed, the life cycle of the X chromosome differs between males and females. The imperfect pairing of the X chromosome with the Y chromosome during male gametogenesis (in both eutherians and metatherians) results in the inactivation of both sex chromosomes during the pachytene stage of meiosis, in a process called meiotic sex chromosome inactivation (MSCI) (Cloutier and Turner, 2010; Lifschytz and Lindsley, 1972). One hypothesis suggests that the Xp retains the epigenetic memory of MSCI silencing and enters fertilization in a pre-inactivated state (Huynh and Lee, 2003; Namekawa and Lee, 2009; Sun et al., 2015). However, *cis*-inactivation-competent autosomal *Xist* transgenes are not subjected to MSCI in the male germ line of mice (Okamoto et al., 2005), suggesting that MSCI is not necessary for iXCI in mice. Moreover, the Xp, like the autosomes, is transcribed at the time of zygotic genome activation (ZGA) (Kalantry et al., 2009; Namekawa et al., 2010; Okamoto et al., 2005; Patrat et al., 2009). Nevertheless, since repeat silencing is proposed to precede genic silencing in iXCI (Namekawa et al., 2010), it is still possible that the Xp is treated differently according to the genomic content, with repeat-rich (LINEs, SINEs)-regions being inherited in a pre-inactivated state through the paternal germ line and gene-rich regions being inactivated *de novo* after ZGA.

Imprinted XCI in mice has also been proposed to rely on a maternal imprint, established early on in the female germ line, that prevents the maternal X chromosome (Xm) from being silenced during early embryogenesis (Kay et al., 1994; Lyon and Rastan, 1984; Tada et al., 2000). Supporting this hypothesis, mouse embryos carrying two Xm chromosomes die during early embryogenesis, possibly due to the lack of dosage compensation (Shao and Takagi, 1990; Tada et al., 1993a). This maternal imprint seems to be acquired during oocyte maturation, and could be there to repress *Xist* until the late morula stage (Tada et al., 2000). Also, the imprint has to be contained into the 210 kb segment that recapitulates imprinted

inactivation (Okamoto et al., 2005) and it has to be independent of X chromosome sensing (section 3.2.1), since X chromosomes not protected by the imprint are inactivated during iXCI regardless of the number of X chromosomes in the cell (Kay et al., 1994; Okamoto et al., 2000).

Using parthenogenetic embryos (that carry two maternal genomes), it has recently been proposed that the nature of the imprint is the H3K9me3 mark at the *Xist* locus, which prevents *Xist* expression from the Xm during preimplantation development (Fukuda et al., 2014). However, during oogenesis, chromatin condensation, rather than H3K9me3, is associated with imprinted Xm *Xist* repression (Fukuda et al., 2015). Regardless of its nature (imprinted or random) XCI in Theria relies on a complex regulatory network involving several long non-coding RNAs (section 3) and, once established, it is clonally propagated, such that females are functionally mosaic for X-linked traits.

2.1.2.2 *Ex-vivo* cellular systems

In nature, X inactivation is initiated almost exclusively in the cells of the early embryo (Mak et al., 2004; Okamoto et al., 2004), and it seems thus dependent on specific pathways that operate under developmental control within this restricted window of time. It follows that the best system in which to investigate XCI mechanisms has to reflect the embryo at early stages of development (Figure 5).

Following fertilization of the egg and formation of a diploid zygote, multiple mitotic cell divisions give rise to the blastocyst. This structure consists of an inner layer of cells, the inner cell mass (ICM, or embryoblast) and an outer layer of cells, the trophoblast. Upon gastrulation, the ICM develops into the epiblast and then into the embryo, while the trophoblast becomes the trophectoderm, which ultimately forms the extra-embryonic tissues.

As revealed by early mouse studies, XCI dynamics are linked to the progressive restriction of pluripotency that takes place in the developing embryo and extends into the adult organism: pluripotent cell types generally have two active X chromosomes (XaXa), while somatic cell types have one active and one inactive X chromosome (XaXi).

The different states of embryonic pluripotency can be exported *ex vivo* in the form of several pluripotent cell types, stabilized by artificial culture conditions, that mimic the

natural development of the zygote from blastocyst to post-implantation embryo *in vivo*, and therefore represent different contexts for XCI establishment and propagation.

Mouse embryonic stem (ES) cells (mESCs) are derived from the ICM (Evans and Kaufman, 1981; Martin, 1981), and are the *ex vivo* equivalent of the epiblast lineage of the blastocyst, with which they share the same potential to differentiate in any of the three primary germ layers (mesoderm, definitive endoderm and ectoderm). Mouse ES cells have a high capacity for self-renewal *ex vivo*, and female cells display two active Xs (XaXa) and can recapitulate random XCI upon differentiation (Paterno and McBurney, 1985; Tada et al., 1993b). They can be genetically manipulated, enabling the study of XCI *ex vivo*, and they can also contribute to blastocyst chimeras (James et al., 2006) and be transmitted through the germ line, allowing the study of XCI in the embryo as well.

Pluripotent stem cells can also be derived from the late epiblast layer of post-implantation mouse embryos (Brons et al., 2007). These epiblast stem cells (EpiSCs) are able to generate cell types of the three germ layers, both *in vivo* and *ex vivo*, but, unlike ES cells (Beddington and Robertson, 1989), they can also differentiate into trophoctoderm (Brons et al., 2007). EpiSCs are characterized by the presence of an inactive X chromosome (Xi) (Nichols and Smith, 2009; Tesar et al., 2007), which they can reactivate upon reprogramming into ES-like cells (Guo et al., 2009). All in all, mouse EpiSCs and ES cells represent two different states of pluripotency along the same developmental pathway, with two distinct states of XCI. These pluripotency states are now commonly referred to as primed pluripotency (in EpiSCs) and naive pluripotency (in ES cells) (Nichols and Smith, 2009).

Two other types of cell lines can be derived from the two extraembryonic cell lineages of the blastocyst, trophoctoderm and primitive endoderm. These are the trophoblast stem (TS) and extraembryonic endoderm (XEN) cells, respectively. Like their lineage of origin, female TS (Mak et al., 2002) and XEN (Kunath et al., 2005) cells display imprinted XCI of the Xp. TS cells can be converted into ES-like cells (Kuckenbergh et al., 2011), and they reactivate their Xi during the process. Symmetrically, ES cells can be converted into TS-like (Niwa et al., 2005) and XEN-like (Shimosato et al., 2007) cells, but in this case the associated XCI process is random, rather than imprinted (Murakami et al., 2011).

Mouse pluripotent cell lines are not the only ones to have been successfully derived. The derivation of non-human primate ES cell lines (Thomson et al., 1995) and the establishment of specific culture conditions favorable to their maintenance led to the successful derivation of human ES cells (hESCs) from human embryos originally produced by *in vitro* fertilization for clinical purposes (Reubinoff et al., 2000). However, despite being derived from pre-implantation embryos at supposedly the same stage of development, human ESCs and mouse ES cells differ in many ways, including morphology, growth factors dependence (Reubinoff et al., 2000), genomic stability (International Stem Cell Initiative et al., 2011), and, importantly, the epigenetic state of the X chromosome.

Like mouse EpiSCs, cultured hESCs display one active and one inactive X chromosome even before differentiation (Shen et al., 2008; Silva et al., 2008). These observations initially raised the hypothesis that hESCs and mouse ES cells are in fact in a different developmental stage. However, XCI kinetics during human pre-implantation development are still not fully elucidated, so that it is not known whether hESCs are indeed the *ex vivo* equivalent of the epiblast, or whether available culture conditions cause hESCs to transition *ex vivo* into an EpiSC-like state (Dvash et al., 2010; Hoffman and Carpenter, 2005). Other than diverse temporal origin and culturing methods, the distinctions between hESCs and mouse ES cells could also reflect species differences.

The fact that mouse EpiSCs could be reverted to a more naive, ES-like state prompted the search for derivation techniques and culture conditions that would allow the establishment of naive hESC lines. While different attempts resulted in variable outcomes regarding the reactivation of the X chromosome, it would be extremely beneficial for the study of XCI dynamics if "real" naive hESCs were in a pre-XCI state, like their murine counterpart. However, it is still not known whether this would represent the *in vivo* situation.

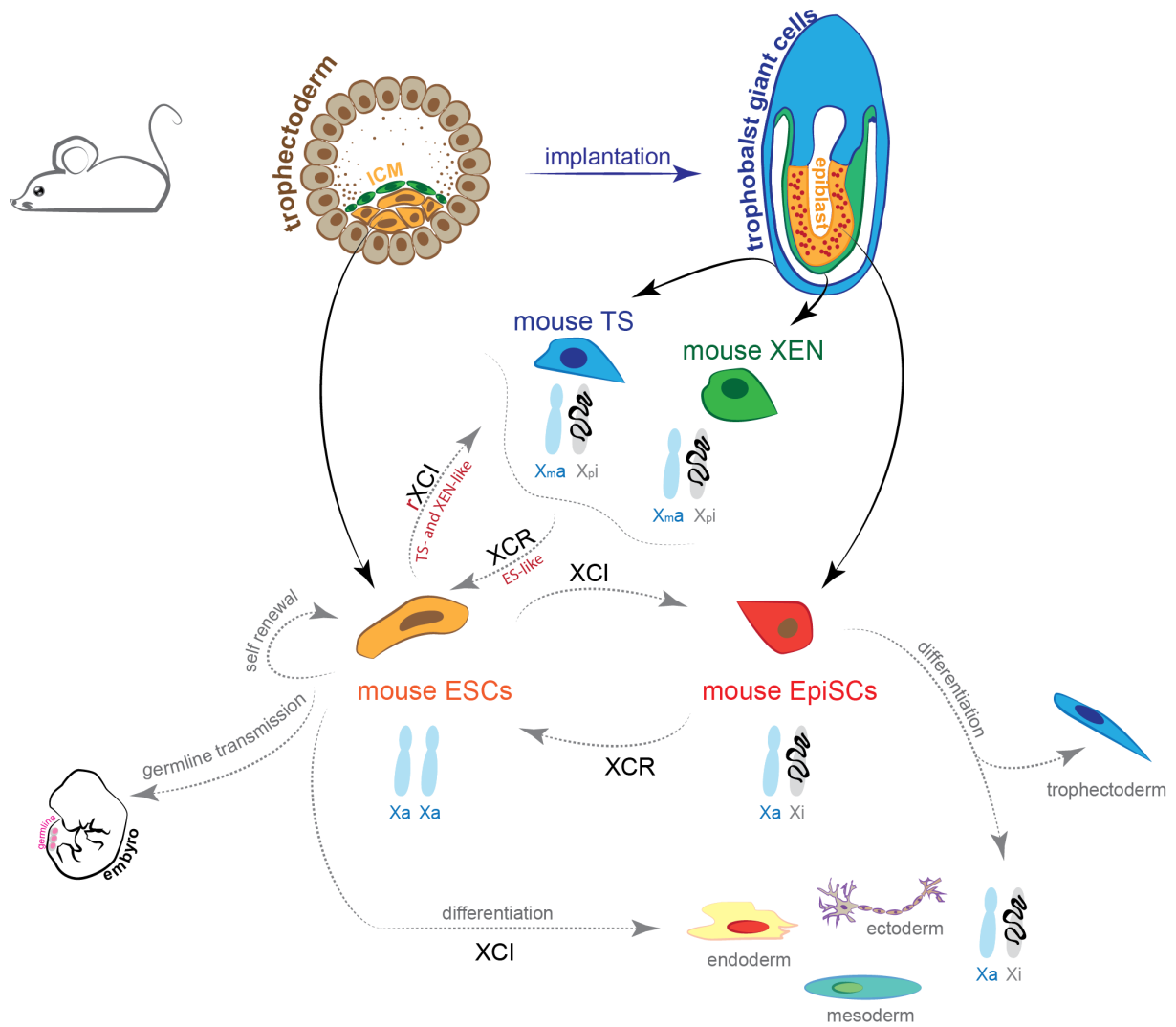
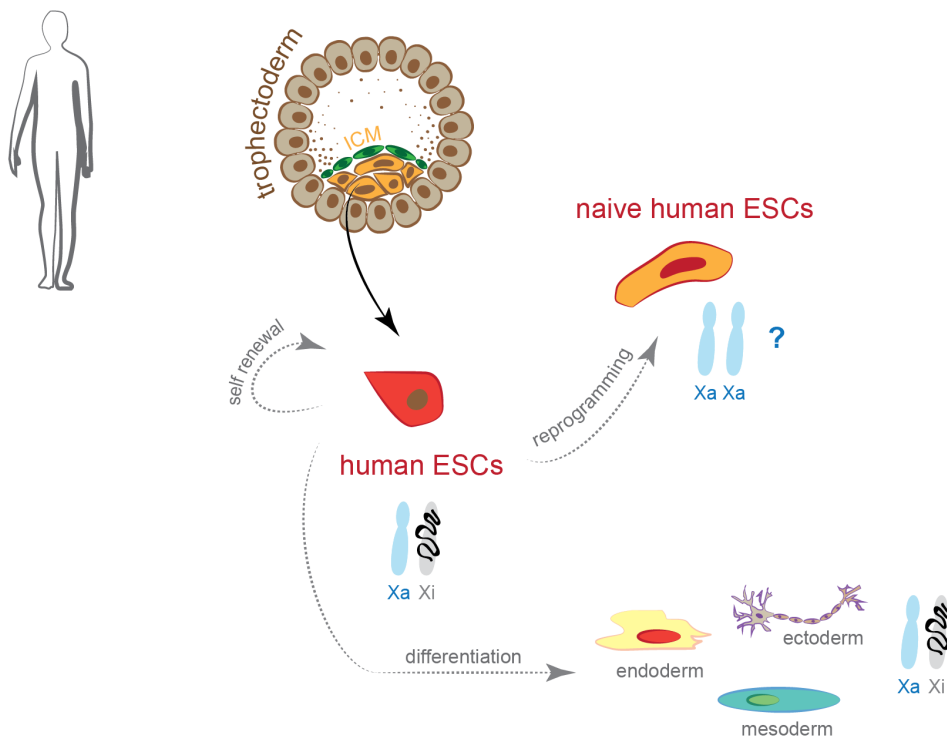


Figure 5. *Ex vivo* cellular systems

Different pluripotent cell types can be derived from embryos at early stages of development. In mice, mouse ES cells (ESCs), are derived from the inner cell mass of the blastocyst, while epiblast stem cells (EpiSCs), trophectoderm stem (TS) cells and extraembryonic endoderm (XEN) stem cells are derived from the epiblast. All these cell types have different potency states, can differentiate and can occasionally be converted into one another. In humans, human ES cells (hESCs) are derived from the inner cell mass of early embryo and they can differentiate or be reprogrammed into a more naive state of pluripotency.

3 Long non-coding RNA-based dosage compensation mechanisms in Theria

A striking feature of the therian dosage compensation systems is their dependence on long non-coding RNAs to achieve chromosome-wide transcriptional regulation.

3.1 Non-coding genome

A large proportion of the mammalian genome is transcribed into transcripts with no coding potential, whose proposed functions add a layer of complexity to the regulation of biological processes. Among those, long non-coding RNAs have gained widespread attention in the past years, and a small number of well-studied examples is now offering important information regarding the purpose of their existence.

3.1.1 The discovery of non-coding genome

Since the 1950s, the amount of DNA in the haploid genome, called C-value, has been counterintuitively shown to have little correlation with organism size or with the number of genes in the genome (Gall, 1981; Mirsky and Ris, 1951; Thomas, 1971). In fact, some single-celled protists have genomes that are larger than those of humans.

The discovery of non-coding DNA in the early 1970s resolved the main question of the C-value paradox: genome size does not reflect gene number in eukaryotes, since most of their DNA does not encode for proteins. The human genome, for example, comprises less than 2% protein-coding regions, with the remaining 98% being non-coding DNA (Elgar and Vavouri, 2008).

It also started to become clear that transcription occurs across genomic regions that are not protein-coding, and that RNA species include more than just the known rRNAs and tRNAs. These newly discovered RNA molecules, called "heterogeneous nuclear RNAs" (hnRNAs), are transcribed not only from repetitive and heterochromatic regions, but also from more than 20% of non-repetitive regions in mammalian genomes. Moreover, 50% of these hnRNAs are retained into the nucleus, and do not contain any coding sequences (Holmes et al., 1972; Lewin, 1980; Pierpont and Yunis, 1977).

In the 1980s, snRNAs and snoRNAs were shown as regulatory molecules needed for post-transcriptional RNA processing (Mattaj et al., 1993). Since then, accumulating

evidence revealed that RNA-based regulatory mechanisms could have a relevant role in the evolution of developmental complexity in eukaryotes. Indeed, the degree of complexity in different species positively correlates more with the portion of the genome that is transcribed into non-coding RNAs than in proteins (Taft et al., 2007).

It was not before the 1990s however, with the advent of high-throughput RNA sequencing techniques (RNA-Seq), that another class of regulatory RNAs, long non-coding RNAs (lncRNAs), began to be massively uncovered.

3.1.2 Long non-coding RNAs

Long non-coding RNAs are a wide and heterogeneous class of transcribed RNA molecules with a length of more than 200 nucleotides that do not encode proteins, or lack open reading frames producing more than 100 aminoacids. lncRNAs can be transcribed as whole or partial natural antisense transcripts (NATs) to protein-coding genes, or located between genes or within introns. Some of them also originate from pseudogenes (Milligan and Lipovich, 2014).

According to their genomic localization and orientation of transcription compared to other (often protein-coding) genes, they can be classified as antisense, intergenic, overlapping, bidirectional, and processed (Mattick and Rinn, 2015; Peschansky and Wahlestedt, 2014). lncRNAs often contain introns, and are (alternatively) spliced and polyadenylated.

Profiling of gene expression and *in situ* hybridization studies have shown that lncRNA expression is developmentally regulated, can be tissue- and cell type-specific, and can vary spatially, temporally, or in response to stimuli. A lot of lncRNAs are expressed in a more tissue-specific pattern and with greater variation among tissues compared to protein-coding genes (Derrien et al., 2012). At the cellular level, many lncRNAs are located exclusively in the nucleus, but some are cytoplasmic or are located both in the nucleus and in the cytoplasm (Ayupe et al., 2015).

Interestingly, analysis of RNA-seq data from 11 tetrapods species revealed that lncRNAs have emerged at high rate during mammalian evolution, with more than 100 new gene units per Myr in primates and rodents (Necsulea et al., 2014). lncRNAs can also diverge very rapidly, as shown by comparison studies in the human, the zebrafish and rodents (Cabili et al., 2011; Kutter et al., 2012; Ulitsky et al., 2011). A recent comparison of splice sites across 46 vertebrate genomes revealed that

lncRNAs are indeed fast evolving molecules, with a significant turnover of intron-exon structures (Nitsche et al., 2015).

At the same time, the expression levels of lncRNAs are at least one order of magnitude below that of mRNAs, observation that has questioned the biological relevance of lncRNA transcription. Because of this, and since lncRNA discovery is still at a preliminary stage, whether all these transcripts represent just "transcriptional noise" (Hüttenhofer et al., 2005) or instead serve specific biological functions is currently under investigation.

It is clear however, that at least some of them are important regulators of gene expression, and take part in a variety of cellular and developmental processes in several species, in which they act through several mechanisms (Figure 6). Among the best known examples of functional lncRNAs are those involved in the epigenetic regulation of allelic expression, from genomic imprinting (Bartolomei et al., 1991; Brannan et al., 1990; Cai and Cullen, 2007; Keniry et al., 2012; Lee et al., 1999c; Lyle et al., 2000; da Rocha et al., 2008; Schmidt et al., 2000; Williamson et al., 2006) to dosage compensation in *Theria* (following sections).

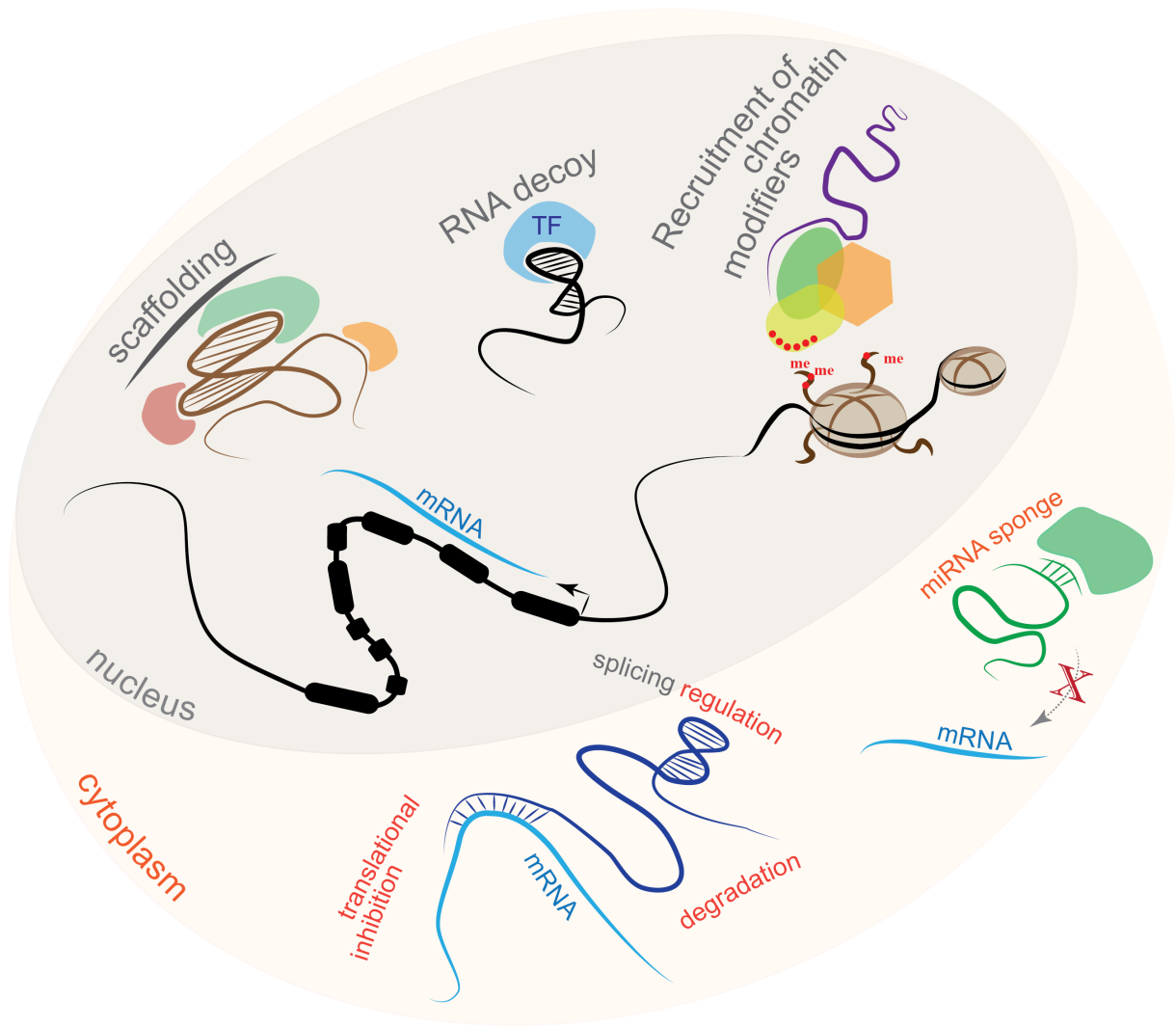


Figure 6. Mechanisms of action of lncRNAs

Long non-coding RNAs have been shown to perform several functions, in both the nucleus and the cytoplasm. In the nucleus, they can be part of ribonucleoprotein complexes, where they act as scaffold, they can regulate mRNA splicing and export, they can recruit chromatin modifiers to target loci, and act as decoys for proteins and transcription factors (TF). In the cytoplasm, they mostly regulate mRNA processing, including mRNA translation and mRNA degradation. They can also sequester microRNAs and prevent them from regulating target mRNAs.

3.2 Evolution of the long non-coding RNAs *Xist* and *Rsx*

In Eutheria, the long non-coding RNA *Xist* has been shown to be the master regulator of the X inactivation process, while the recently identified long non-coding RNA *Rsx* has been proposed to drive X inactivation in Metatheria. *Rsx* and *Xist* are not orthologous, which suggests an independent evolutionary origin of two RNAs with similar features (Figure 7).

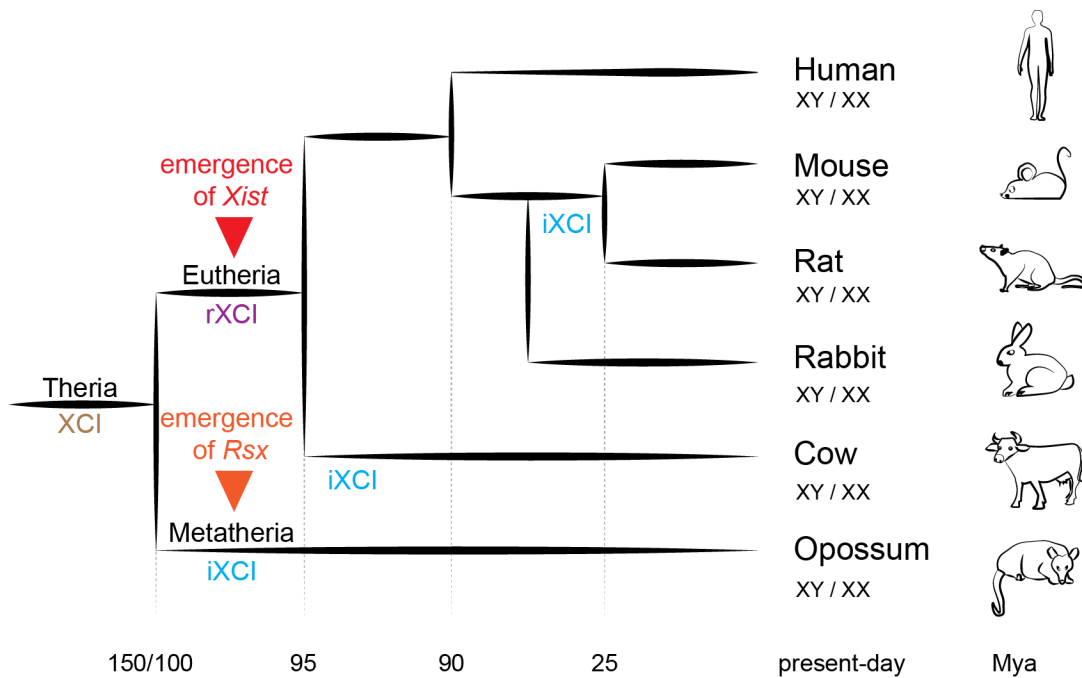


Figure 7. Emergence of *Xist* and *Rsx*

Phylogenetic tree indicating the emergence of the long non-coding RNAs *Xist* and *Rsx*, that are involved (or predicted to be involved in the case of *Rsx*) in X chromosome inactivation in Eutheria and Metatheria, respectively.

3.2.1 The eutherian *X inactivation center*

At the molecular level, X inactivation in present-day eutherians relies on a complex locus required in *cis*, called *X inactivation center* (*Xic*) (Figure 8), and described as being both necessary and sufficient to trigger X inactivation (Rastan and Robertson, 1985). The first genetic evidence for the *Xic* was obtained through studies of mouse embryos or embryonic stem cells carrying chromosomal translocations and truncations involving the X chromosome, as well as *via* the investigation of human disorders.

In mouse, cytological evidence on embryos carrying the T16H Searle's translocation showed that only one translocation product (16^X), but not the other (X^{16}) undergoes

inactivation (Rastan, 1983; Takagi, 1980). In ES cells, another mutation, the HD3, reduced the size of the X chromosome without preventing X inactivation (Rastan and Robertson, 1985). These observations enabled to define the *Xic* as the region included between the T16H and HD3 breakpoints. Further transgenesis experiments allowed to narrow down the candidate *Xic* region, but to date no exact boundaries have been precisely defined.

In humans, similar observations in patients displaying chromosomal abnormalities enabled to map the X inactivation center in a <1 Mb region within band Xq13.2, distal to the *AR*, *CCG-1*, *RPSX4* and *PHKA* loci, and proximal to *PGK1* (Brown, 1991).

Although it contains some protein-coding genes (*Nap1l2/NAP1L2*, *Ppnx/PPNX*, *Cdx4/CDX4*, *Chic1/CHIC1*, *Cnbp2/CNBP2*, *Xpct/XPCT*, *Rnf12/RNF12*), the *Xic/XIC* is especially enriched in genes producing long non-coding RNAs, the most iconic of which is the well known *Xist/XIST* (section 4). *Xist/XIST* is flanked, on both sides, by other lncRNA genes, including its antisense *Tsix/TSIX*, and the genes *Jpx/JPX*, *Ftx/FTX*, *Tsx*, *Xite* and *Linx*. All of them appear to play a function in XCI that is more or less documented: *Tsix*, *Xite* and possibly *Linx* are negative regulators of the process, while *Jpx* and *Ftx* have, like *Xist*, a positive effect on XCI (section 5 and section 6).

Interestingly the *Xic* is organized in two domains, both linearly and spatially, that separate activators from repressors (Nora et al., 2012). The boundary of these Topologically Associating Domains (TADs) maps in between *Xist* and *Tsix*, and their respective promoters thus lie in adjacent, oppositely regulated chromosomal neighborhoods (Figure 8). RS14, a conserved element located in between the two TADs, exerts a female-specific boundary function by binding the chromatin insulator CTCF. Deletion of this region results in impaired XCI and reduced differentiation efficiency in female but not in male cells (Spencer et al., 2011). Perhaps counterintuitively, TADs are present before and after differentiation, on the active and inactive Xs (Nora et al., 2013).

Differences between the murine and human *Xic/XIC* regions

Despite being globally homologous and syntenic, the minimal mouse *Xic* and human *XIC* regions show significant differences at the structural level: during evolution, multiple inversions originated shuffled gene order between the species, and the

present-day human *XIC* is more expanded compared to its murine orthologous region (Chureau et al., 2002). Overall, the portion of the *Xic* surrounding *Xist* seems to be undergoing a rapid evolutionary change: a small inversion including the *Xist/XIST* gene (Brockdorff et al., 1992; Lafrenière and Willard, 1993) and the *Cdx4*, *Bpx* and *Brx* genes (Horn and Ashworth, 1995; Rougeulle and Avner, 1996; Simmler et al., 1997) exists between human and mouse and is associated with the loss of genomic segments within the region. For example, *BPX* is located 600 kb downstream of *XIST* in human but only 300 kb 3' to *Xist* in mouse (Rougeulle and Avner, 1996). Cloning of the murine *Xpct* gene showed that the inversion also involved a region 5' to *Xist/XIST* (Debrand et al., 1999). Other differences include *Tsx*, which became a pseudogene in human (Duret et al., 2006; Elisaphenko et al., 2008), and *XPCT*, whose transcriptional direction is inverted compared to its murine counterpart (Duret et al., 2006; Elisaphenko et al., 2008).

The evolutionary divergence of the *Xic/XIC* not only involves gene composition, but also its three-dimensional organization. Hi-C analyses in hESCs suggested that the *TSIX* TAD is less prominent in humans than in mice, despite the general global conservation of the TADs in both species (Dixon et al., 2012).

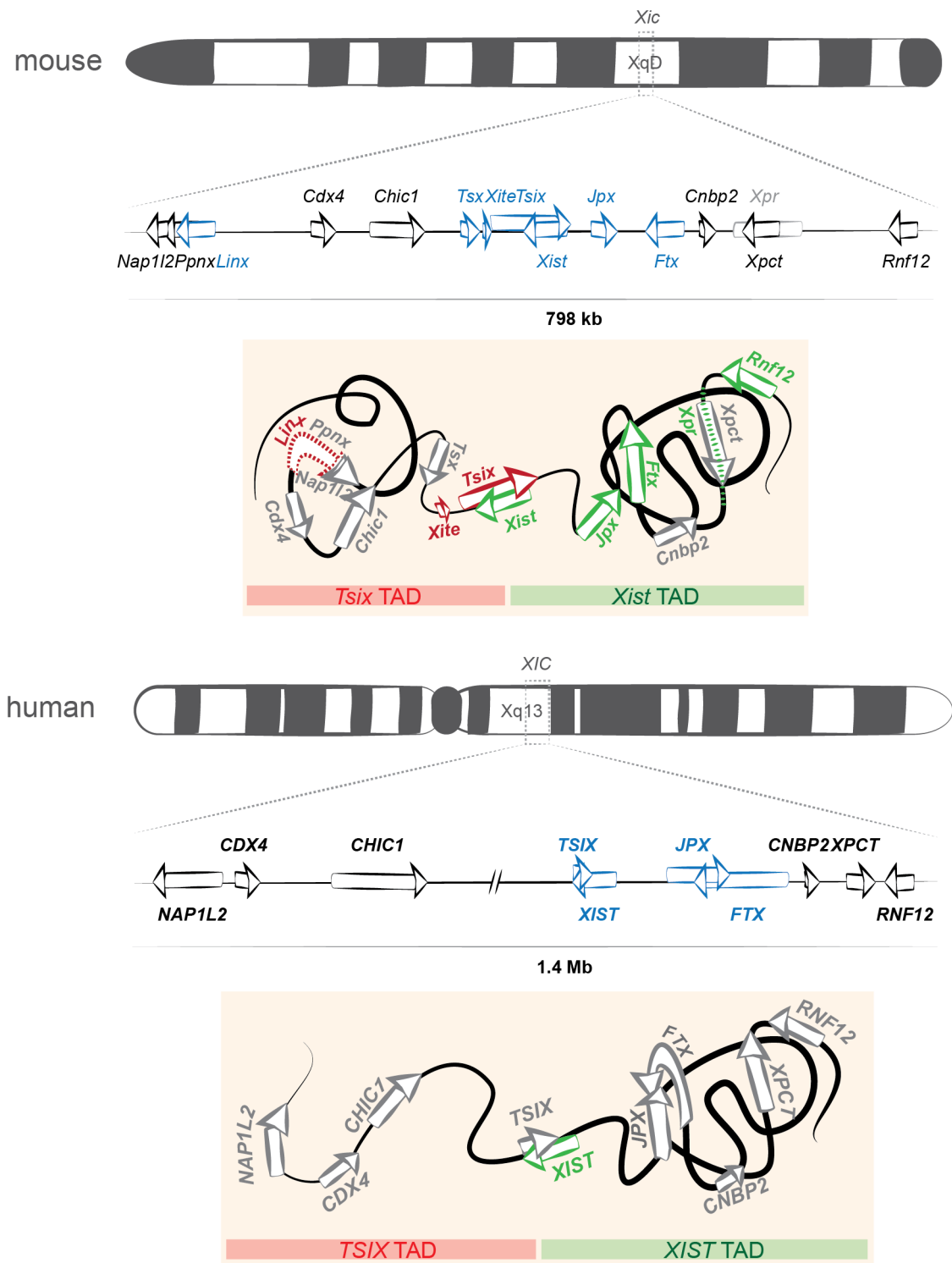


Figure 8. Murine and human X inactivation centers

Top panels: Representation of the mouse X chromosome, with a zoom of the X inactivation center (*Xic*) region, located in *XqD*, which harbors protein-coding genes (in black) as well as lncRNA genes (in blue). Below, schematic tridimensional structure of the *Xic*, partitioned in two topologically associated domains (TADs). The *Tsix* TAD includes repressors of XCI (in red), while the *Xist* TAD encompasses XCI activators (in green). *Linx* is represented by dotted lines as its function is yet to be determined. Bottom panels: Representation of the

human X chromosome, with a zoom of the X inactivation center (*XIC*) region, located in Xq13, which harbors protein-coding genes (in black) as well as lncRNA genes (in blue). Below, schematic tridimensional structure of the *XIC*. Within the *XIC*, the region upstream of *XIST* displays relative conservation compared to the mouse, both at the levels of gene content and for what concerns its 3D organization (TAD).

The *Xic* and the initiation of XCI

The *Xic* has been implicated in the initial steps that ensure correct silencing of X-linked genes: the counting of the number of X chromosomes, the sensing of the number of genomes present in the nucleus so that a correct X to autosome expression ratio is established, and the choice of which X chromosome to inactivate.

Indication of the existence of a counting step comes from observations in cells carrying abnormal complements of sex chromosomes. Inactivation partly relies on the number of X chromosomes present in the nucleus, as XXY "male" cells (Klinefelter syndrome in humans) inactivate one X chromosome (Iitsuka et al., 2001) while diploid female cells with more than one X (46, XXX super females) inactivate them all but one (Jacobs et al., 1959), and XO females (Turner syndrome in humans) do not undergo X inactivation at all (Ashworth et al., 1991).

Since at least two copies of the *Xic* must be present per diploid genome for the inactivation to occur (Lyon et al., 1962; Grumbach et al., 1963), this region was proposed to host all the necessary counting elements to ensure that only one copy of X-linked genes remains active. The study of X:autosome translocations lead to the formulation of the blocking hypothesis of X chromosome counting (Rastan, 1983). The hypothesis proposes that the state of the *Xic* is critical for this step: limited quantities of a blocking factor bind to only one *Xic* per diploid genome, leaving the other *Xic* free to receive the initiation signals. To present, it is still not clear which factor could be responsible for this role. However, since the *Xist* transcript can induce gene silencing in *cis* independently of flanking sequences but cannot induce inactivation of the endogenous X chromosome (Wutz and Jaenisch, 2000), the counting function of the *Xic* must be contained in regions outside *Xist*. The location of the binding site(s) for the blocking factor was proposed to be downstream *Xist*, since deletions 3' to the gene resulted in XCI in male cells (Clerc and Avner, 1998; Morey et al., 2004). Overall, in the last decade, several *Xic* regions outside *Xist* have been dissected as being involved in counting.

The inactivation mechanism also senses the autosomal complement: tetraploid cells keep two Xs active and two Xs inactive (Webb et al., 1992), suggesting that XCI depends not only on X-linked but also on autosomal factors. Using tetraploid XXXX, XXXY and XXY ES cells, it has been shown that each X chromosome has an independent probability of being inactivated, suggesting that counting and choice are directed by a stochastic mechanism. Moreover, this probability is proportional to the X chromosome:ploidy ratio, suggesting that a XCI activator must exist, that is X-linked (Lee et al., 1999a) and that is itself inactivated by the process (Monkhorst et al., 2008, 2009). If this model is true, mutations or deletions of the binding sites for this dose dependent X-linked activator will result in skewed XCI, because the probability of the mutated X to be inactivated will be lower.

The choice of which of the two X chromosomes to inactivate (Rastan and Robertson, 1985) seems to depend on several *Xic/XIC* regions, including, together with *Xist* (Gribnau et al., 2005; Marahrens et al., 1997), the *DXPas34/Tsix* segment (Clerc and Avner, 1998; Lee et al., 1999a; Luikenhuis et al., 2001). However, the fact that XCI does not occur in male cells with a *Tsix* promoter mutation lead to the hypothesis that a factor must exist outside *Tsix* that is required in two copies, and thus able to "count" the number of X chromosomes, to initiate XCI (Lee et al., 1999a). An additional genetic locus known to influence XCI choice in mouse is the *X controlling element* (*Xce*), whose allelic differences lead to biased inactivation (section 5.2.5).

Counting and choice have also been hypothesized to occur before the onset of XCI and depend on differences in sister chromatid cohesion (Mlynarczyk-Evans et al., 2006). Another model proposes that counting and choice are dictated by transient cross communication between the two Xs at the onset of XCI (Marahrens, 1999), idea supported by experiments showing that in a subset of female cells at the onset of XCI, the two X chromosomes transiently pair (Bacher et al., 2006; Xu et al., 2006) (section 5.2.1).

Evolution of the *Xic* from a protein-coding unit

Quest for homologs of murine *Xist* in vertebrates found *Xist* in all eutherians, including the most primitive placental mammals, elephant (*Loxodonta africana*) (Dementyeva et al., 2010; Duret et al., 2006) and armadillo (*Dasypus novemcinctus*) (Dementyeva et al., 2010). Despite being present in all of these species, *Xist* has

also been subjected to repeated modifications during evolution: *Xist* sequences in present-day species contain ancient and conserved repeats, as well as young species-specific repeats originating from mobile elements movements (Elisaphenko et al., 2008). Moreover, the fraction of repetitive sequences within *Xist* increased from 4.4% in the eutherian ancestor to 12.4% in the human (Elisaphenko et al., 2008).

To fully understand the evolutionary history of the *Xist* gene, it is necessary to look at the whole *Xic* region. Evidence shows that the present eutherian *Xic*, now dominated by long non-coding RNAs, was once protein-coding. In eutherian mammals, this ancient protein-coding unit was subjected to mutations that caused loss of protein-coding function of the embedded genes, as well as to the integration of several mobile elements, the combination of which contributed to the evolution and acquisition of function for the non-coding *Xist*, *Tsix*, *Jpx* and *Ftx* genes (Figure 9).

The unraveling of the evolutionary scenario of *Xic* emergence started with the observation that three human and murine *Xic* protein-coding genes (*CDX4/Cdx4*, *CHIC1/Chic1* and *Xpct/XPCT*) have orthologs in other classes of vertebrates, including chicken (*Gallus gallus*) and frog (*Xenopus tropicalis*) (Duret et al., 2006). The conserved linkage displayed by these protein-coding genes confidently enabled the labelling of this region as the *Xic* orthologous region (*XicHR*) in non-eutherian species (Duret et al., 2006). In addition to the three conserved protein-coding genes, five more protein-coding genes (*Fip112*, *Ln timer*, *Ras111c*, *UspL* and *Wave4*) are contained in the *XicHR*, both in chicken and frog, but absent in the eutherian *Xic* (Duret et al., 2006). Genomic sequence comparisons between the eutherian *Xic* and the chicken *XicHR* showed that the *Xic/XIC* long non-coding RNA genes *Xist*, *Tsx*, *Jpx* and *Ftx* evolved from the ancient protein-coding genes *Ln timer*, *Fip112*, *Usp1* and *Wave4* respectively, though pseudogenization and integration of different classes of mobile elements (Dementyeva et al., 2010; Duret et al., 2006; Elisaphenko et al., 2008).

Despite the conservation of *Xist* in Eutheria, the search for *Xist* homologs in other vertebrate lineages failed to retrieve any conserved sequences in non eutherians, with other *Xic* genes, including *Tsix*, *Jpx* and *Ftx* also lacking detectable orthologs in non eutherian genomes (Duret et al., 2006). Moreover, genes flanking the *XIC* in humans were found to map to two different regions of the X chromosome in two

marsupial species, the South and North American opossums, indicating the occurrence of rearrangements in this region since the divergence of marsupials and eutherians (Davidow et al., 2007; Deakin et al., 2008b; Hore et al., 2007; Shevchenko et al., 2007). It is now accepted that the same region that gave origin to the eutherian *Xic* was disrupted in metatherian mammals, and that a *XicHR* region is not present in these species.

Figure 9. *Xic* lncRNAs evolution from protein-coding genes

3.2.2 *Rsx*, a metatherian X-linked long non-coding RNA

The absence of *Xist* and the disruption of flanking markers suggests that X inactivation in marsupials cannot depend on the same *Xic* unit that drives XCI in eutherians, yet the presence of eutherian-like features on the Xi, such as a vast array of chromatin modifications including H3K27me3 (Chaumeil et al., 2011), indicates that a *Xist*-like regulator could exist in Metatheria. RNA-FISH analysis of XCI in the female brain of the short-tailed opossum *M. domestica* revealed a cloud signal reminiscent of the *Xist* signal (section 4.2) when a BAC containing the X-linked gene *Hprt1* and 131 kb of total flanking sequences was used (Grant et al., 2012). The same cloud signal was observed with a modified BAC carrying a deletion of the *Hprt1* gene, indicating that some other unknown transcript must exist in that region. RT-PCR experiments identified a 47 kb transcription unit that showed female-specific expression in a broad range of opossum tissues and was transcribed exclusively from the inactive X, hence called *Rsx* (RNA-on-the-silent X) (Grant et al., 2012). *Rsx* encodes a repeat-containing four-exon RNA, of approximately 27 kb in length, spliced and polyadenylated (Grant et al., 2012).

Rsx is conserved in both the two metatherian groups, the South American and Australasian groups, which diverged 75-80 Mya, indicating that *Rsx* originated before this major split and must be therefore common among all Metatheria (Grant et al., 2012). Profiling of parent-of-origin allelic-specific expression, DNA methylation and chromatin modifications across the metatherian X chromosome in fetal brain and extra-embryonic membranes further showed that XCI indeed depends on a *Xist*-like regulating mechanism relying on H3K27me3 and possibly *Rsx*, but that, contrary to the murine situation (section 0), histone modification and promoter DNA methylation are uncoupled (Wang et al., 2014). *Rsx* is the only metatherian X-linked gene that is, like *Xist*, differentially methylated at the promoter, with methylation on the active X (Xm) only (Wang et al., 2014). Interestingly, an *Rsx* transgene was capable of coating and inducing gene silencing in *cis* to some extent when ectopically inserted in mouse ES cells (Grant et al., 2012), suggesting that *Rsx* and *Xist* may share additional properties.

Whether *Rsx* is regulated by neighboring lncRNA genes is currently unknown, but inspection of the ancestral *Sox3* region might reveal important information. Interestingly, the protein-coding ancestor of *Xist* (*Lnx3* (Duret et al., 2006;

Elisaphenko et al., 2008)), is separated from *Sox3* in present-day eutherians but located in its proximity in birds (Gribnau and Grootegoed, 2012). In the light of this, it is possible that a "metatherian *Xic* unit" was also ancestrally located in the vicinity of *Sox3*, and that genome rearrangements in metatherian lineages prevented the evolution of the *Xic* from the same region as in eutherians. In this scenario, parallel but independent evolution would have lead to the emergence of *Rsx* in marsupials from another X-linked unit, possibly originally located close to *Sox3* and the eutherian proto-*Xic*. If this were true, putative *Rsx* regulators might be found in metatherian genomic regions whose ancient counterparts were close to the *Sox3/Lnx3* region.

4 *Xist*, the master regulator of X chromosome inactivation in Eutheria

Xist (X-inactive specific transcript) is the master regulator of XCI in Eutheria. During XCI, the *Xist* transcript accumulates in *cis* on the future inactive X chromosome, triggering a cascade of epigenetic modifications that results in the stable silencing of the entire chromosome, with only few genes remaining active. *Xist* expression profiles and the mechanisms of *Xist* regulation during development differ across therian species to a noticeable extent.

4.1 *Xist* discovery and its structure in different eutherian species

The *XIST* sex-specific expression and its genomic localization were first discovered in 1991, while screening a library of human cDNA probes (Brown, 1991). The *XIST* cDNA probe hybridized to RNA prepared from female but not male samples, and, interestingly, it mapped to the human *XIC*. Moreover, the observation that the probe hybridized to RNA from somatic cell hybrids only if they contained an inactive X chromosome showed that *XIST* was expressed from the inactive X (Xi) but not the active X (Xa), suggesting a role in X inactivation. The isolation of the mouse homolog, *Xist*, soon followed (Borsani et al., 1991), and the transcript was shown to display, like in the human, Xi-specific expression (Brockdorff et al., 1991).

So far, the *Xist* gene has been isolated in all eutherian mammals analyzed. Analysis of rodent, human, pig and cow sequences has enabled the identification of both conserved and species-specific regions of the *Xist* transcript (Elisaphenko et al., 2008).

The genomic loci of murine *Xist* and human *XIST* are approximately 23 kb and 35 kb in length, respectively. Both genes produce transcripts that are spliced and polyadenylated (Hong et al., 1999; Memili et al., 2001) and not translated into proteins, as the absence of significant open reading frames in their sequence suggests. Mouse *Xist* is transcribed from two alternative promoters, named P1 and P2, as far as two alternative polyadenylation sites, that generate RNA isoforms of 17 and 17.9 kb in length. Human *XIST* isoforms generally lack exon 4, part of exon 6, or exon 7 (Brown, 1991; Brown et al., 1992), resulting in an RNA transcript up to 19.3 kb in length (Hong et al., 2000). The overall sequence homology between mouse *Xist*

and human *XIST* only reaches 49%, but the global structure of the gene, including the first transcription start site (P1) and the exon/intron boundaries, is conserved, with six repeated elements (A to F) showing the highest degree of similarity (Nesterova et al., 2001) (Figure 10).

A putative porcine *XIST* sequence has been identified from assembly analysis using known sequences from GenBank, RNA-Seq data and PCR/sequencing (Hwang et al., 2013). The porcine *XIST* encodes a 25 kb transcript consisting of seven exons, of which two are conserved with human and mouse and two are porcine-specific (Hwang et al., 2013). Eight transcription start sites have been identified and transcription appears to span the locus only in female cells (Hwang et al., 2013). Similar analyses retrieved data from ESTs and cDNA databases to predict and compare the structure of the bovine *Xist*, which contains some exons conserved with mouse *Xist* and other exons conserved with human *XIST* (Chureau et al., 2002). Overall, these data indicate that the *Xist* locus is undergoing a constant remodelling during evolution.

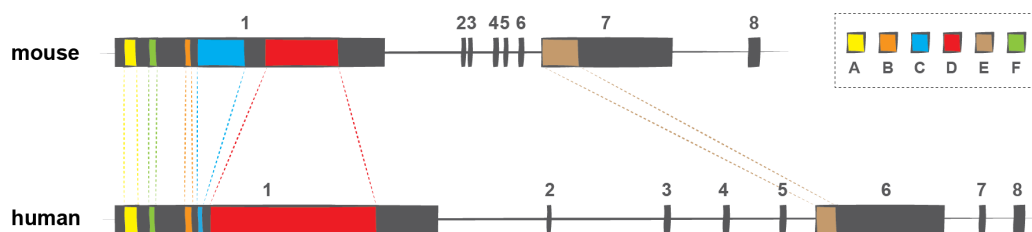


Figure 10. *Xist/XIST* structure in mouse and human

Structure of the mouse and human *Xist/XIST* genes, with exons represented as boxes and introns as lines. Conserved repeat elements (A–F) are drawn in color, in scale.

4.2 *Xist* cellular localization and expression profiles in different eutherian species

Not long after its initial discovery, the 17-kb long *Xist* RNA was found to be retained in the nucleus of female somatic cells (Brockdorff et al., 1992; Brown et al., 1992) where it coated the inactive X chromosome (Brown et al., 1992; Clemson et al., 1996; Panning and Jaenisch, 1996). This evidence suggested that *Xist* might be a structural element of the nuclear architecture of the Xi, making it the ideal candidate to explain the epigenetic transmission of silencing. Pioneer *Xist* RNA-FISH analyses of metaphase X chromosomes in voles also showed that *Xist* does not bind along the entire X-chromosome, but exhibits a particular banded pattern of localization, with

Xist-depleted regions being constitutively heterochromatic (Duthie et al., 1999). *Xist* banding was also detected in mouse and rat cell lines.

The 19-kb long human *XIST* RNA was shown to behave similarly to mouse *Xist*, also coating the Xi and dissociating from it during mitosis (Clemson et al., 1996; Hall and Lawrence, 2003)(Clemson 96, hall and lawrence 2003). Digestion of chromosomal DNA and chromatin showed that *XIST* RNA remained within the nucleus, probably bound to insoluble protein components of the nuclear matrix (Clemson 1996). Initial *XIST* RNA-FISH experiments also showed that the number of *XIST* accumulation foci ("clouds") coincided with the number of Xi in the cell (Brown 1992). *Xist* RNA clouds were also detected in cow and rabbit cells (Okamoto et al., 2011).

More recently, super-resolution three dimensional structured illumination microscopy (3D-SIM) analyses of the topology of the Xi territory challenged the idea that the *Xist* transcript is broadly distributed across the Xi (homogeneous "cloud"), and showed that it rather binds a limited number of genomic sites at a given time (Cerase et al., 2014; Smeets et al., 2014). In differentiating and fully differentiated female ESCs, a discrete number of foci (approximately 100) is visible, per cell. Each focus could represent multiple *Xist* molecules, consistent with the estimates of 300 to 1000 *Xist* molecules per cell (Buzin et al., 1994; Sun et al., 2006a).

Despite the conservation of *Xist* coating properties across different eutherian mammals, a remarkable diversity in the dynamics of *Xist* expression and X chromosome silencing exist (Figure 11). In mouse, *Xist* is transcribed at low levels in the pluripotent epiblast lineage prior to differentiation into embryonic tissues both in female and male embryos *in vivo* (Panning and Jaenisch, 1996; Panning et al., 1997). Similarly, in undifferentiated mouse ES cells, punctuate *Xist* RNA-FISH signals are detectable from both Xs in females (two pinpoints) and from the single X in males (one pinpoint) (Panning et al., 1997). On a side note, these pinpointed signals could also correspond to transcription of the *Xist* antisense *Tsix*, since the vector used to produce the *Xist* RNA-FISH probe (pXist3K) contains *Xist* exon 1, fully spanned by *Tsix* on the sense strand of the *Xist/Tsix* genomic locus. Both the epiblast cells *in vivo* and the ES cells *ex vivo* exhibit dynamic *Xist* expression patterns and undergo X inactivation upon differentiation (Panning et al., 1997; Paterno and McBurney, 1985; Sheardown et al., 1997; Tada et al., 1993b). In differentiating cells, *Xist* RNA coats the future Xi, while the Xa retains a low level of

Xist expression, which is subsequently turned off (Panning et al., 1997; Sheardown et al., 1997).

In human, *XIST* expression is detectable by qRT-PCR in both sexes at the blastocyst stage (Daniels et al., 1997; Ray et al., 1997), raising the intriguing possibility that it is uncoupled from XCI in early human embryos. Consistent with this, *XIST* RNA coating without the corresponding transcriptional silencing is found in the trophectoderm and the inner cell mass of female and male human preimplantation blastocysts (Okamoto et al., 2011). Adult human females, however, display *XIST* coating and gene silencing exclusively on one X chromosome. Human placentas, unlike mice ones, undergo random XCI (Daniels et al., 1997; Moreira de Mello et al., 2010; Ray et al., 1997).

In rabbits, eutherian mammals closely related to rodents, *Xist* is not subjected to imprinting, and XCI, like in human, begins later than in mice. Punctuate *Xist* RNA-FISH signals are detectable from both Xs in female and from the single X in male rabbit embryos from the 8-cell stage, which corresponds to the embryonic genome activation in rabbits (Okamoto et al., 2011). At the morula stage, *Xist* accumulation is visible in female but not male embryos, albeit in a low percentage of cells, and by the early blastocyst stage almost all trophectoderm cells display at least one *Xist* domain, and surprisingly, sometimes two (Okamoto et al., 2011). Thus, XCI initiates on both X chromosomes in a significant proportion of blastomeres, but the situation is resolved in the late blastocyst, when the majority of cells display a single *Xist* cloud (Okamoto et al., 2011).

Overall, the diversity in the timing of *Xist* upregulation across eutherians, as well as the differences observed in the structure of the *Xist* gene, might reflect the need to adapt to species-specific requirements during early development.

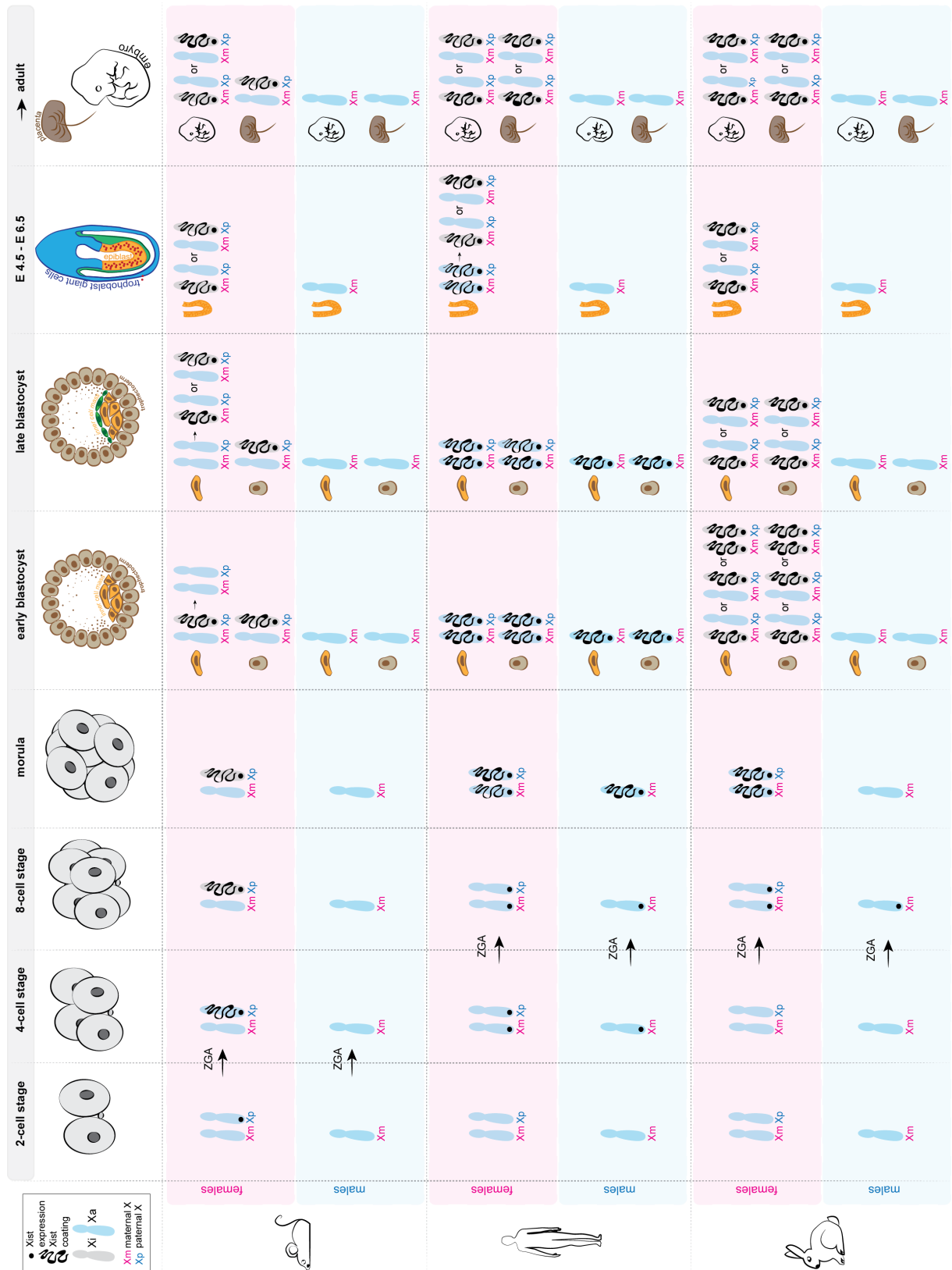


Figure 11. Developmental patterns of *Xist* expression

Developmental pattern of *Xist* expression, coating, and X-silencing in males and females of three eutherian species (mouse, human, and rabbit). In female mice, *Xist* imprinted expression from the Xp starts at the 2 to 4-cell stage transition, with *Xist* transcripts coating the future (paternal) Xi. From the 8-cell to the morula stage, the paternal X is progressively inactivated. At the blastocyst stage, the paternal X undergoes reactivation in the inner cell mass (ICM), so that both X chromosomes are now active. By contrast, the Xp remains inactivated in the trophoblast. Random XCI follows in the ICM, whereby the paternal and the maternal Xs have the same probability of being inactivated. This random pattern of inactivation is epigenetically maintained through cell divisions in the postimplantation embryo and in the adult. In male mice, *Xist* is never expressed and the Xm is active at all stages. In female humans, *XIST* is not imprinted and its expression is detectable from both X chromosomes from the 4- to 8-cell stage. It is unknown whether *XIST* coats the chromosomes at these stages. It does, however, at the morula stage, but the chromosomes remain transcriptionally active. This configuration persists in the blastocyst, in embryonic as well as extraembryonic lineages. Later on, *XIST* expression and coating are restricted to a single X chromosome and accompanied by X transcriptional repression. In male humans, *XIST* is thought to be expressed from the Xm at the 4-to 8-cell stage and to coat the chromosome by the morula stage and until the late blastocyst. Despite *XIST* expression and coating, the Xm remains active. *XIST* coating of the Xm disappears at later developmental stages. In female rabbits, like in humans, the two X chromosomes are active in the early cleavage stages. Punctuate *Xist* expression is detectable from both Xs from the 8-cell stage. *Xist* accumulation initiates in the morula and is more prominent in the blastocyst, coinciding with transcriptional inactivation. Importantly, a significant portion of blastocyst cells show biallelic *Xist* coating, which is accompanied in this case by the silencing of both X chromosomes. Only later is the situation resolved, finally resulting in inactivation of a single X in the adult. This inactivation is believed to be random, but this requires further investigation. In male rabbits, *Xist* could be expressed around the 8-cell stage, but the chromosome is not inactivated. The timing of zygotic genome activation (ZGA) is indicated for each species.

4.3 *Xist* requirement in X chromosome inactivation

Initial analyses of *Xist/XIST* RNA expression in different tissues during development showed a close correlation between *Xist/XIST* expression and the timing of XCI, suggesting that *Xist* expression is not just a consequence of X inactivation but rather an element involved in its control. For instance, *Xist* is expressed from the Xp just prior to imprinted XpCI in the trophectoderm and it is expressed concomitantly with the onset of random XCI before primitive streak formation *in vivo* (Kay et al., 1993).

Although correlation studies suggest an important role for *Xist* in XCI, the essential function of *Xist* in silencing was clearly demonstrated through loss and gain of function approaches. In mice, targeted deletions in the *Xist* gene have demonstrated that *Xist* expression is required for both imprinted and random inactivation. Female embryos that paternally inherit a mutated *Xist* gene die during early embryogenesis due to iXCI defects in the extraembryonic tissues, while female embryos that inherit the same *Xist* mutation maternally are healthy and fertile (Kalantry et al., 2009; Marahrens et al., 1997). Random X inactivation defects were initially observed in differentiating female ES cells and in the embryonic tissues of mice heterozygous for targeted deletions of the *Xist* gene (Marahrens et al., 1997; Penny et al., 1996). Both *ex vivo* and *in vivo*, the X chromosome carrying the mutant *Xist* allele failed to be silenced, and only the wild-type X ever became the Xi.

In addition to that, treatment with *Xist* antisense oligonucleotide analogs that block the coating ability of the *Xist* transcript prevent the production of a silenced X chromosome (Beard et al., 1995), indicating that not only the genomic sequence, but also the RNA product is needed for silencing. Similar results are obtained when using antisense oligos directed against a repeat region (repeat C) of the *Xist* RNA, which results in the disruption of the inactive X following the loss of *Xist* coating in the mouse (Beletskii et al., 2001).

4.3.1 *Xist* and the initiation of X chromosome inactivation

In the mouse, experiments with *Xist* transgenes concluded that *Xist* is sufficient for the initiation of silencing in *cis* (Plath et al., 2002; Wutz et al., 2002), and defined a critical time window during which *Xist* RNA is required for inactivation, both *ex vivo* (mESCs) and *in vivo*.

Single-copy inducible *Xist* transgenes targeted to the *Hprt* locus on the X

chromosome or on autosomes successfully trigger silencing of neighboring genes (Jiang et al., 2013; Wutz and Jaenisch, 2000; Wutz et al., 2002). An inducible *Xist* transgene was also shown to be able to initiate silencing in immature hematopoietic precursor cells in adult mice *in vivo*, suggesting the existence of a developmental time window, related to the pluripotency state of the cell, in which *Xist* is competent for inactivation (Savarese et al., 2006). The introduction of a multi-copy 450 kb long YAC transgene including *Xist* onto autosomes in male ES cells is also sufficient to activate *Xist* expression and affect the expression of genes in *cis* (Lee et al., 1996), suggesting that high-level *Xist* RNA expression is sufficient for chromosome coating and silencing, even in a male context.

However, while short single-copy *Xist*-expressing transgenes and large multi-copy *Xist/Xic* YACs are sufficient to trigger inactivation, large single-copy YACs seem unable to. Transgenic mice created by ectopically inserting a YAC containing *Xist* and portions of the *Xic* in an autosome lack any detectable *Xist* expression when the integration is single-copy (Matsuura et al., 1996). Similarly, the integration of a single-copy 460 kb YAC spanning *Xist* and two silent genes in ES cells, *Cdx4* and *Bpx*, results in the absence of detectable YAC-induced inactivation both in heterozygous and homozygous animals (Heard et al., 1996). In these animals, however, *Xist* is expressed in the early embryo and induced only when the transgene is paternally-inherited (Okamoto et al., 2005).

Aside from the need for a transgene to be integrated in multi-copy, it could be possible that the regions contained in these YACs cover only part and not all the functional *Xic*, or that the chromatin context at the site of integration is not permissive to *Xist* expression. Indeed, multi-copy integration of a shorter form of one of the YACs in the heterochromatic long arm of the Y chromosome shows detectable *Xist* expression, indicating that proximal heterochromatic sequences may be necessary for *Xist* induction *in vivo* (Matsuura et al., 1996). Moreover, the discrepancies in the silencing abilities of short single-copy inducible *Xist* transgenes versus long single-copy *Xist*-containing YACs might also be due to the extremely high levels of *Xist* expression from the former, which cannot be compared to those obtained from wild type promoter sequences.

Manipulation of *Xist* expression has shown that not only in the embryo, but also in the adult mouse, *Xist* is able to initiate silencing in specific cell types. These include the

lineage-restricted progenitors of the hematopoietic system (Savarese et al., 2006), even if hematopoietic stem cells (HSCs) and mature blood cells do not seem to have an appropriate cellular context to allow *Xist*-mediated silencing. *Xist* is able to trigger gene silencing also in some tumor cells (Agrelo et al., 2009), as demonstrated by using a mouse thymic lymphoma model triggered by an oncogenic lymphoma fusion kinase inside the T-cell compartment (Chiarle et al., 2008). In this context, forced *Xist* induction causes ectopic X inactivation and can stop tumor development (Agrelo et al., 2009).

Human *XIST* role in silencing was also confirmed in several studies using transgenes. Six tandem copies of a 480 kb transgene containing the human *XIC* region are recognised as a *XIC* and able to induce random X inactivation in some cells of a male chimeric mice when inserted into murine chromosome 11 (Migeon et al., 1999). However, similarly to the murine situation, a low-copy 300 kb transgene inserted into murine chromosome 6 is recognized as *XIC* but cannot induce inactivation in mice (Migeon et al., 2001).

4.3.1.1 *Xist* upregulation and stabilization

Run-on experiments and measurements of steady-state levels have shown that *Xist* upregulation and stabilization are associated with an increase of the half-life of the transcript, proposed to be mediated by factors produced only during cell differentiation and the onset of XCI (Panning et al., 1997; Sheardown et al., 1997). Post-transcriptional stabilization mediated by these factors could explain the observation that the expression of the unspliced (immature) *Xist* transcript is constant while the spliced (mature) *Xist* form accumulates upon differentiation of female ES cells (Claudio et al., 2006), as well as studies showing that *Xist* transcription rate is similar between ES cells and fibroblasts (Panning and Jaenisch, 1996; Sheardown et al., 1997).

However, the failure of *Xist* stabilization in undifferentiated ES cells does not seem to be due to the lack of stabilizing factors in these cells, as a full-length *Xist* cDNA transgene under control of a strong inducible promoter is able to generate high levels of *Xist* transcripts that are stable in the undifferentiated state (Wutz and Jaenisch, 2000). Studies involving human *XIST* transgenes further supported this observation (Heard et al., 1999).

Alternatively, another mechanism that could be responsible for monoallelic *Xist* accumulation is transcriptional upregulation. It was initially proposed that *Xist* transcriptional upregulation could be mediated by a developmentally regulated promoter switch by which *Xist* transcription from a putative third promoter (P0) initially generates an unstable transcript in ES cells, and subsequent transcription from P1 upon differentiation and in somatic cells gives rise to stable transcripts (Johnston et al., 1998). The existence of the P0 promoter was however questioned, since a P0-deleted *Xist* transgene in ES cells still produced an unstable *Xist* transcript initiating from P1 (Warshawsky et al., 1999). The P0 promoter was disregarded as an artifact of cross-amplification caused by a pseudogene of the ribosomal protein S12 gene *Rps12*, and the transcription upstream P1/P2 identified as coming from the 3' end of the *Tsix* RNA (Warshawsky et al., 1999). Nevertheless, the transgenes used in these experiments were truncated just upstream *Xist* and their expression could have been influenced by the integration site, thus leaving open the possibility of the existence of an alternative *Xist* promoter, like P0, upstream P1/P2. Another model based on *Xist* stabilization by transcriptional upregulation implicates *Tsix* in the induction of transient heterochromatic state at the *Xist* promoter (Sun et al., 2006a).

4.3.1.2 *Xist* spreading

Upon upregulation and stabilization, *Xist* initiates X chromosome inactivation by spreading in *cis* across the future inactive X chromosome, where it tethers to chromatin and gives rise to the characteristic, cytologically visible, cloud. Considerable effort has been put into the understanding of the mechanism of such spreading at the molecular resolution. In theory, the silencing apparatus could initially form at the *Xic* and diffuse serially from nucleosome to nucleosome (Duthie et al., 1999; Smith et al., 2004), or it could spread outwardly through defined "way stations" across the chromosome (Gartler and Riggs, 1983).

The fact that all autosomes are capable of inactivation when hosting an ectopic *Xist* transgene (Lee et al., 1999a; Lyon, 1998; Wutz and Jaenisch, 2000), but yet XCI travels less successfully in autosomal than in X chromosomal chromatin, initially led to the hypothesis that "boosters" must exist that are present genome-wide but enriched on the X chromosome. Intersperse repeat elements were identified as regions highly likely to fulfill this requirement and, among these, long intersperse

repeat elements (LINE1s, L1s) were suggested to be the most suitable candidates, due to their overrepresentation on the X chromosome (Mary Lyon's "repeat hypothesis" and (Bailey et al., 2000)). In accord with this hypothesis, microscopy studies have shown that *Xist*/*XIST* RNA on the X chromosome overlaps with regions enriched for non-coding genomic repeats, which are located inside the X-chromosome territory and physically separated from the gene-containing regions of the X chromosome (Chaumeil, 2006; Clemson et al., 2006). Moreover, *Xist* spreading in translocation products between the X chromosome and chromosome 4 has been shown to be attenuated at the translocation border, where the autosome displayed a low LINE1s density (Lee and Jaenisch, 1997). In addition, a DNA methylation profiling assay in X;autosome translocations showed that inactive autosomal genes (defined by having an hypermethylated promoter region) were located in LINEs-dense regions on the translocation product, indeed suggesting that these repetitive elements might favor the spreading of silencing (Bala Tannan et al., 2014).

Xist chromosomal localization during initiation and maintenance of X chromosome inactivation was recently explored by RNA Antisense Purification (RAP) experiments, which enabled to purify the endogenous *Xist* RNA in male ES cells engineered to carry an inducible *Xist* promoter or lacking the repeat A region (Engreitz et al., 2013) as well as by allele-specific CHART-seq in female samples at four different developmental stages (Simon et al., 2013). These studies described how *Xist* spreading occurs in a proximity transfer fashion following a two-step mechanism (Figure 12).

In the initiation phase, *Xist* exploits the three-dimensional genome architecture to spread from its locus to spatially proximal sites on the periphery of active gene-dense and short interspersed nuclear elements (SINEs)-rich regions (Engreitz et al., 2013; Simon et al., 2013). Gene-poor regions do not contact the *Xist* locus as often and are thus less accessible to *Xist* spreading (Engreitz et al., 2013). These observations are in line with initial cytological evidence suggesting a significant association of *Xist* localization with G-light bands (gene-rich regions), and a less prominent association of *Xist* with gene-poor region (Duthie et al., 1999; Smith et al., 2004). At this step, *Xist* is also depleted from escapees, but could concentrate near their borders (Simon et al., 2013).

Upon encounter with a new region, *Xist* RNA spreads over active genes, possibly

using the repeat A region (section 4.3.1.3) and through the interaction with the matrix protein HNRNPU (Engreitz et al., 2013). Subsequently, interaction with other proteins triggers the chromatin changes that compact and reposition the new silent regions into the *Xist* RNA compartment. It may be possible that these structural changes participate to *Xist* spreading by pulling new regions into their vicinity. In contrast with the classic model of linear *Xist* spreading (Duthie et al., 1999; Smith et al., 2004), the observation that the spreading on proximal sites depends on the relative positioning of the *Xic* (Engreitz et al., 2013; Simon et al., 2013) suggests that the affinity of the first sites does not necessarily have to be high, and implies that the secondary spreading could be a consequence of the initial contacts, chromosome structure or site affinity.

In the maintenance phase *Xist* shows enrichment over the entire X chromosome, without anchoring to particular focal binding sites (Engreitz et al., 2013). Interestingly, this whole-chromosome localization pattern is different from that observed in *Drosophila* for the roX2 RNA, which binds to specific sites (Chu et al., 2011; Simon et al., 2011). Nevertheless, *Xist* enrichment strongly correlates with H3K27me3 across the whole chromosome (Engreitz et al., 2013; Simon et al., 2013), in agreement with the proposed role for *Xist* in PRC2 recruitment (Plath et al., 2003; Silva et al., 2003; Zhao et al., 2008) and somehow in correlation with the suggested role of Polycomb stations in the spreading of XCI (Pinter et al., 2012). Additionally, when *Xist* is transiently removed from post-XCI cells, it recovers faster compared to the time course of silencing establishment during initiation, suggesting that the chromatin state may provide the epigenetic memory facilitating *Xist* positioning in the maintenance phase (Simon et al., 2013).

In contrast with what previously proposed (Chow et al., 2010; Lyon, 1998, 2006), both these studies failed to find a correlation between *Xist* enrichment on the chromosome and LINE1 sequence elements, supporting the idea that LINE-dense regions of the Xi are most frequently located next to the *Xist* coat rather than in its center (Chadwick et al., 2006; Duthie et al., 1999; Mak et al., 2004). Whether LINE1s are part of the mechanism of XCI or a result of it remains thus still enigmatic.

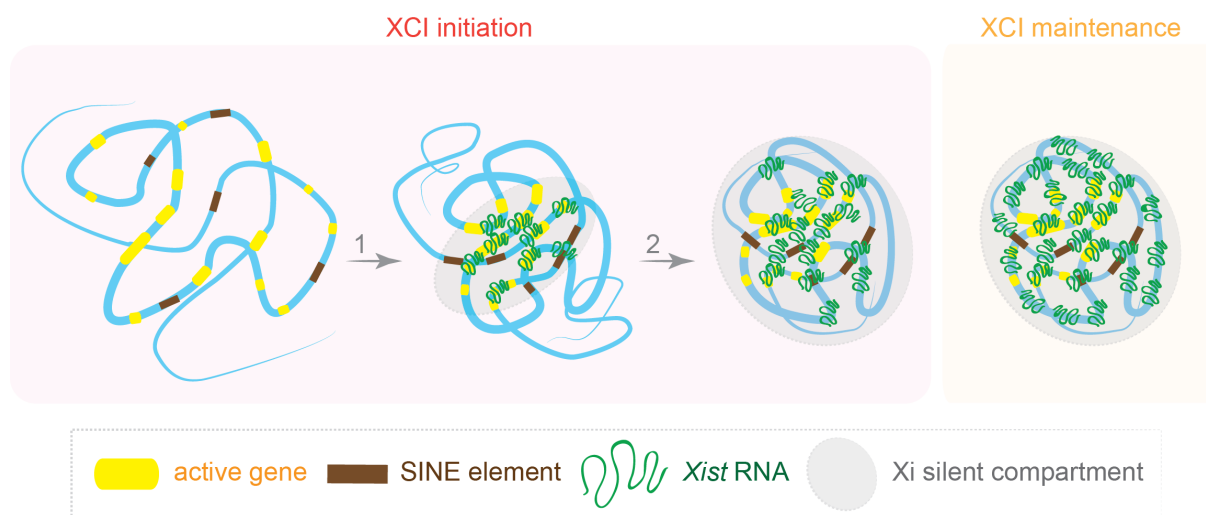


Figure 12. *Xist* spreading

During the initiation phase of X chromosome inactivation, *Xist* RNA spreads from its locus to spatially proximal gene- and SINEs-rich regions, pulling them progressively into a silent compartment. At this step, *Xist* RNA contacts focal binding sites. Gene-poor regions are recruited into the silent compartment in a second step. During the maintenance phase, *Xist* RNA is enriched over the entire X chromosome.

4.3.1.3 *Xist* functional regions and protein interactors

Several studies have dissected functional regions of the *Xist* transcript and identified several candidate protein interactors that could directly mediate *Xist* silencing effect (Figure 13).

A study with a series of inducible *Xist* cDNA transgenes containing defined deletions has pinpointed the stretch of conserved repeats at the 5' end of the transcript (known as the repeat A) as being necessary for *Xist* silencing function, but dispensable for chromosomal association (Wutz et al., 2002). The repeat A is conserved among all placental mammals and it has been shown to form a stable hairpin that adopts a stable tetraloop fold in human (Duszczuk et al., 2011). The repeat A has been proposed to recruit PRC2 to the Xi, consistent with its binding to PRC2 components, including Ezh2 and Suz12, *in vivo* and *in vitro* (Kanhare et al., 2010; Maenner et al., 2010; Zhao et al., 2008). It was also reported that the repeat A is required for *Xist* RNA processing through binding to the alternative splicing factor ASF/SF2, although this interaction is not considered to influence *Xist*'s silencing function (Royce-Tolland et al., 2010). More recently, the repeat A was reported as implicated in the spreading of *Xist* RNA across the inactivating X chromosome (Engreitz et al., 2013) (section 4.3.1.3).

The role of the repeat C is controversial. Experiments with deleted inducible *Xist* cDNA transgenes excluded that this region is required for coating in undifferentiated cells (Wutz et al., 2002), while a (PNA)-Interference Mapping study (Beletskii et al., 2001) and experiments with locked nucleic acids (LNAs) (Sarma et al., 2010) showed that it is in fact important for chromosomal localization, at least in differentiated female cells, where targeting of the region results in visible *Xist* dissociation from the Xi. Of course these discrepancies could be due to the different pluripotency status of the cells at the moment of analysis and thus relate to the need for the repeat C in the establishment versus maintenance of silencing.

The repeat E-containing *Xist* exon 7 was recently reported to be involved in the silencing of X-linked genes, but not required for *Xist* RNA expression level and stability (Yamada et al., 2015). Exon 7 resulted also important for the localization of *Xist* RNA to chromatin, possibly through the interaction with the nuclear matrix-associated protein HNRNPU/SAF-A (Yamada et al., 2015), whose requirement in *Xist* chromosomal localization has been previously reported (Agrelo et al., 2009; Fackelmayer, 2005; Hasegawa et al., 2010). These observations are consistent with initial reports of *Xist* retention in the nuclear scaffold after chromatin has been removed (Clemson et al., 1996), and with recent observations of *Xist* RNA localization within perichromatic spaces (Smeets et al., 2014).

Other functional sequences within *Xist* are required for coating and spreading on the chromosome but do not affect transcriptional repression (Wutz et al., 2002). The action of these regions appears to be cooperative and redundant, as no specific region is absolutely required. Moreover, contrary to the well-conserved repeat A region, these sequences share little or no homology across species (Brockdorff et al., 1992; Brown et al., 1992; Nesterova et al., 2001); this sequence divergence could thus suggest the presence of low affinity individual binding sites.

In addition to HNRNPU, other individual *Xist* protein interactors have been found. A screening approach identified SATB1 as critical factor in the initiation of *Xist*-mediated silencing (Agrelo et al., 2009). SATB1 has been shown to be involved in the formation of chromatin loops, by binding to AT-rich stretches at nuclear matrix attachment regions (de Belle et al., 1998), once again linking *Xist* coating function to the nuclear matrix. Because SATB1 localizes in the area around *Xist* and the Xi, rather than over the chromosome (Agrelo et al., 2009), it has been implicated in the

anchoring of gene-poor, LINE-dense regions of the Xi (Tattermusch and Brockdorff, 2011), that are located adjacent to the *Xist* cloud and gene-rich regions of the Xi (Calabrese et al., 2012).

The zinc-finger transcription factor YY1 was proposed to interact directly with the *Xist* transcript through the repeat C and to tether it to a nucleation center on the Xi, located just upstream of the repeat F in the *Xist* locus (Jeon and Lee, 2011a). Since immunofluorescence analyses indicated that YY1 does not form a microscopically visible cloud over the Xi, it was also suggested that the factor is not directly involved in the spread of *Xist* beyond the repeat F genomic region (Jeon and Lee, 2011).

The Polycomb repressive complex 2 (PRC2) was suggested as a candidate for the silencing role (Plath et al., 2003), shown as directly binding to *Xist* RNA via the EZH2 subunit (Zhao et al., 2008), and later proposed to be loaded onto *Xist* RNA via the ATRX protein (Sarma et al., 2014).

More recently, four studies provided a genome-wide, *in vivo* comprehensive identification of proteins directly bound to the *Xist* RNA (Chu et al., 2015; McHugh et al., 2015; Minajigi et al., 2015; Moindrot et al., 2015). Chromatin isolation by RNA Purification followed by mass spectrometry (ChIRP-MS) identified 81 proteins that are able to bind the *Xist* RNA. These experiments were performed in murine male cells carrying an inducible *Xist* cDNA transgene into chromosome 11 (Wutz and Jaenisch, 2000), as well as in a female epiblast stem cell line (EpiSC), representing cells that have just undergone random XCI, and in trophoblast stem cells (TSCs), a model for imprinted XCI. 30 of these proteins, mainly involved in RNA processing, splicing and stability were enriched also in samples of other ncRNAs, hence non *Xist*-specific. The other proteins, *Xist*-specific, enriched for gene repressors and could be separated into two sets, according to the developmental timing of binding to *Xist* RNA. 62 of them (termed Set 1) are found as bound to *Xist* in all the samples/developmental conditions, while an additional group of 19 (termed Set 2) are only found in differentiated cell types (differentiated ES, EpiSCs and TSCs). Set 1 proteins remain associated to *Xist* RNA from the pluripotent state on, and include, among others, PRC1 components, the known interactor HNRNPU, HNRNPK, and the *Drosophila* Split end homolog Spen. Knockdown of HNRNPK interferes with *Xist*-mediated silencing of the imprinted *Grb10* gene, located close to the *Xist* transgene integration region, and impairs PRC1 and PRC2-dependent histone modification but

not *Xist* localization. However, HNRNPK recovery is left unchanged by the deletion of the repeat A from the transgene, indicating that its silencing function is independent from the repeat A. Spen interaction with *Xist* via the repeat A was increased during development and presumably important for *Xist*-mediated silencing, as ChIRP-MS in the repeat A mutant fails to recover the protein product and knockdown of its mRNA results in failure of X-linked silencing. Set 2 proteins interact with *Xist* RNA only upon differentiation and include WTAP, SAP18 and RNF20. Both WTAP and RNF20 are thought to bind *Xist* via the repeat A, since neither is retrieved in the repeat A mutant cells.

Experiments of RNA antisense purification coupled to quantitative mass spectrometry (RAP-MS) in murine male ES cells harboring an inducible *Xist* promoter at the endogenous locus resulted in the identification of 10 proteins with high enrichment for *Xist* binding, and mostly included in the aforementioned study (colleen). Among those, SAF-A (HNRNPU), LBR and SHARP (SPEN) appear required for *Xist*-mediated silencing, as knockdown of their mRNAs results in failure of X-linked gene silencing upon *Xist* induction. SAF-A depletion also results in a diffuse localization pattern of the *Xist* RNA, indicating the involvement of this protein in *Xist* chromosomal localization, in agreement with previous observations (Hasegawa et al., 2010). SHARP knockdown also shows higher levels of PolII over the *Xist*-coated chromosome territory, suggesting that this protein could be required to create the initial silent compartment upon *Xist* binding to chromatin. Further manipulations enabled to propose a role for SHARP in the recruitment of the SMRT co-repressor, that in turn activates HDAC3, finally leading to histone deacetylation and gene silencing (McHugh et al., 2015).

Another independent confirmation of the role of SPEN/SHARP, WTAP, and additionally RBM15 (a component of the m6A RNA methyltransferase complex), in *Xist*-mediated silencing came from a pooled shRNA screen aimed at the identification of silencing factors primarily involved in the critical developmental window of opportunity for XCI (Moindrot et al., 2015). Once again, a male murine ES cell line with an inducible *Xist* transgene (on chromosome 17) was used, and knockdown of the candidate proteins resulted in defects in the *Xist*-mediated silencing of the autosome. Confirmation of these observations was obtained also in differentiating female ES cells through RNA-FISH analysis of *Xist* RNA and nascent mRNA of the

X-linked genes *Pgk1* and *Rnf12*: upon independent RBM15, SPEN and WTAP knockdown, the frequency of nascent mRNA foci associated with the *Xist* cloud increases, indicating a failure of silencing. In addition, RBM15 is important for efficient deposition of H3K27me3 on the Xi, as indicated by the reduction in the size of H3K27me3 domains upon RBM15 (but not WTAP or SPEN) knockdown. All the three proteins are also shown to co-localize with *Xist* RNA in the perichromatin spaces or nuclear matrix compartment (super resolution 3D-SIM microscopy), consistent with the studies mentioned above (Chu et al., 2015; McHugh et al., 2015). SPEN was demonstrated as required for *Xist*-mediated silencing also in an additional study, in which *Xist*-inducible haploid mouse ES cells were genetically screened for silencing factors involved in XCI (Monfort et al., 2015). A viral trap mutagenesis strategy was used to generate genome-wide insertional mutations in the genome of haploid ES cells harboring an inducible *Xist* promoter, and the cells were then screened for the rescue of the *Xist*-induced lethality upon *Xist* promoter induction. Cells harboring mutations in either *Xist*, the chromatin assembly factors *Hira* and *Ubn2*, the kinase *Cdk8*, or *Spen* recover the phenotype, abrogating the silencing pathway of *Xist* and thus preventing *Xist*-dependent cell death. Consistent with this, CRISPR-mediated deletion of *Spen* also restores the viability of *Xist*-expressing haploid cells. Due to the nature of the screen, it is however not possible to distinguish whether the mechanism of *Spen*-dependent silencing involves direct interaction of *Spen* with *Xist*, or indirect interaction via the regulation of *Xist* expression or the global destabilization of the cell state. Further electrophoretic mobility shift assays with recombinant *Spen* showed that the protein can directly bind to *Xist* RNA, at least *in vitro* (Monfort et al., 2015). Moreover, consistent with the proposed role for *Spen* in silencing, mutations in *Spen* do not affect *Xist* localization but reduce polycomb recruitment (Monfort et al., 2015).

A large *Xist* interactome (80 to 200 proteins) was reported via iDRiP (identification of direct RNA-interacting proteins) experiments (Minajigi et al., 2015). The most highly enriched proteins retrieved in this study belong to multiple classes and included cohesins, condensins, topoisomerases, RNA helicases, chromatin remodelers, histone modifiers, DNA methyltransferases, nucleoskeletal factors and nuclear matrix proteins. Known and proposed *Xist* interactors were recovered, such as PRC2 (zhao 2008), ATRX (Sarma et al., 2014), HNRNPU (Hasegawa et al., 2010), macrohistone

H2A (mH2A), RING1 (PRC1 component), CTCF, and SmcHD1, a condensin component. Consistent with the studies discussed above, SPEN was also identified. Curiously, PRC2 was not captured as factor directly bound to *Xist* RNA in the first three studies (Chu et al., 2015; McHugh et al., 2015; Moindrot et al., 2015), contrasting with the fourth (Minajigi et al., 2015) and with what previously reported through *in vitro* binding and purifications in non-denaturing conditions (Zhao et al., 2008), but consistent with the result that PRC2 is dispensable for the initiation of gene silencing during XCI (Kalantry and Magnuson, 2006) and with recent evidence indicating that PRC2 and *Xist* are in fact spatially separated in interphase nuclei (Cerase et al., 2014). Moreover, the specificity of the PRC2/*Xist* binding has been recently questioned due to the apparent promiscuity of PRC2 binding to many RNAs, including bacterial ones, in *in vitro* conditions (Davidovich et al., 2013). This does not exclude that further PRC2 recruitment could be mediated by *Xist* in an indirect manner. Indeed, indirect PRC2 recruitment was proposed to be mediated by a region of the *Xist* transcript which is not the repeat A (Cerase et al., 2014), consistent with the persistence of PRC2 and H3K27me3 on the *Xist*-coated chromosome in the repeat A deletion *Xist* mutant (Plath et al., 2003; da Rocha et al., 2014), and proposed to occur in an HDAC3-dependent fashion (McHugh et al., 2015), and possibly with the help of Jarid2 and Pcl2 cofactors (Casanova et al., 2011; da Rocha et al., 2014). Alternatively, PRC2 recruitment could be mediated by PRC1, whose components directly bind to *Xist* RNA (Chu et al., 2015), in accord with what previously reported (Blackledge et al., 2014; Cooper et al., 2014; Kalb et al., 2014). YY1 was also not recovered in these large-scale studies, and at the same time shown to rather bind to the *Xist* 5' genomic region, where it plays a role in the transcriptional activation of *Xist* (Makhlouf et al., 2014).

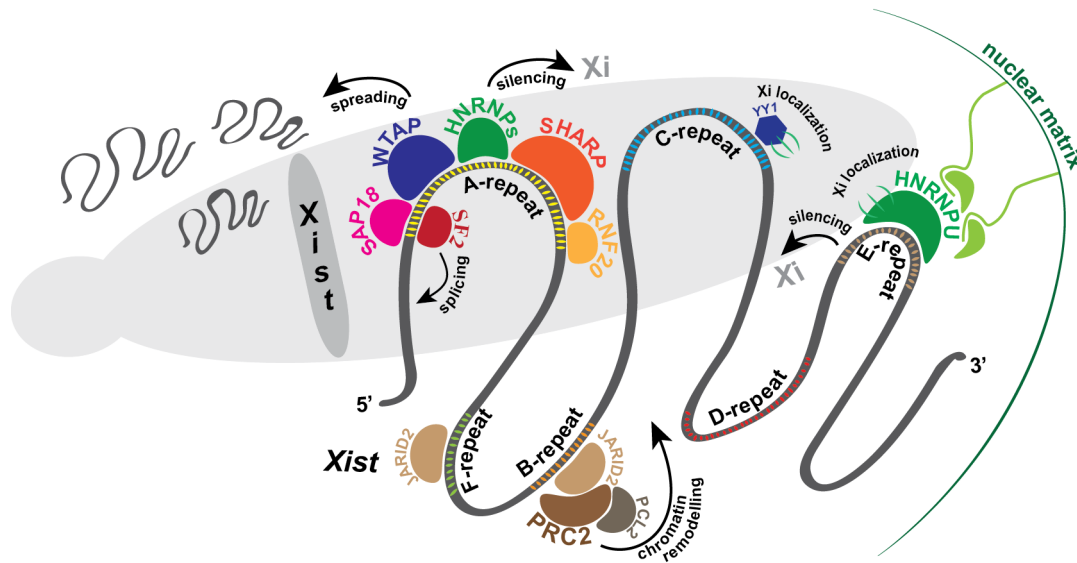


Figure 13. Functions, activities and interactors of *Xist* RNA repeat elements

The A-repeat recruits several proteins involved in X-chromosome silencing and is also required for spreading along the chromosome; The B-repeat interacts with JARID2, a PRC2 cofactor, thus participating to Xi chromatin remodeling. The C-repeat could be important for *Xist* tethering to the Xi, possibly via its interaction with the transcription factor YY1. The E-repeat also mediates *Xist* tethering to the Xi by interacting with HNRNP (whose interaction with *Xist* also depends on the A-repeat). The F-repeat is, like the B-repeat, is also critical for JARID2 recruitment.

4.3.1.4 Features of *Xist*-mediated transcriptional silencing

Following *Xist* RNA coating in *cis*, a cascade of events is triggered to establish the transcriptional inhibition of the X chromosome. Multiple epigenetic changes contribute to lock-in the inactive state over several cell cycles and provide cellular memory of silencing during differentiation and through cell divisions.

Epigenetic changes lock in the inactive state

In 1949, it was reported via cytological observations that a densely stained body ("sex-chromatin", later named "Barr body") was present in the nuclear material of female, but not male, cat neurons (Barr and Bertram, 1949). This element, clearly visible also in interphase nuclei of rat cells (Ohno et al., 1959), was later hypothesized to correspond to the heterochromatic (inactive) X chromosome (Ohno et al., 1959; Lyon, 1961), with its silenced state clonally propagated through cell divisions. The compact structure of the Barr body led to the assumption of the contribution of this particular high-order chromatin structure to the establishment and

maintenance of the silenced state of the inactive X chromosome. *Xist* RNA, always found to be associated with the Barr body (Brown, 1991; Clemson et al., 1996), represented a valid candidate for the recruitment of the epigenetic modifications constituting it.

Many years after these initial observations, it is now known that initiation of *Xist* transcription is followed by a change in the high order chromatin structure of the Xi, and results in the Xi differing from its active homolog, the Xa, in both composition and configuration. *Xist* expression, spreading and coating are followed by epigenetic changes that include RNA Polymerase II exclusion (Chaumeil, 2006), a shift to asynchronous replication timing (Takagi et al., 1982), repositioning of active genes into a transcriptionally silent nuclear compartment (Chaumeil, 2006), incorporation of the histone variant macro H2A (Costanzi et al., 2000; Mermoud et al., 1999), DNA (CpG) methylation (Norris et al., 1991), and a variety of histone modifications (Chadwick and Willard, 2003; Chaumeil et al., 2002; Heard et al., 2004), such as H3 and H4 hypoacetylation (Boggs et al., 1996; Jeppesen and Turner, 1993; Keohane et al., 1996a), hypomethylation of H3K4, dimethylation of H3K9 (Boggs et al., 2002; Heard et al., 2001; Mermoud et al., 2002; Peters et al., 2002), and the PRC2-dependent trimethylation of H3K27 (Plath et al., 2003; Rougeulle et al., 2004; Silva et al., 2003). Conversely, euchromatin marks such as H3K4me2/3 and H3 and H4 lysine acetylation are depleted from the Xi (Figure 14).

The modifications here listed appear very early in the X inactivation process, suggesting a major role in the initiation and in the transition to the maintenance of the inactive state. However, the molecular mechanisms that underlie the ability of the *Xist* RNA to trigger this series of events in the first place are not completely clear, and it is currently debated whether *Xist* RNA directly and continuously recruits silencing factors or transcriptional silencing is promoted via establishment and maintenance of repressive chromatin.

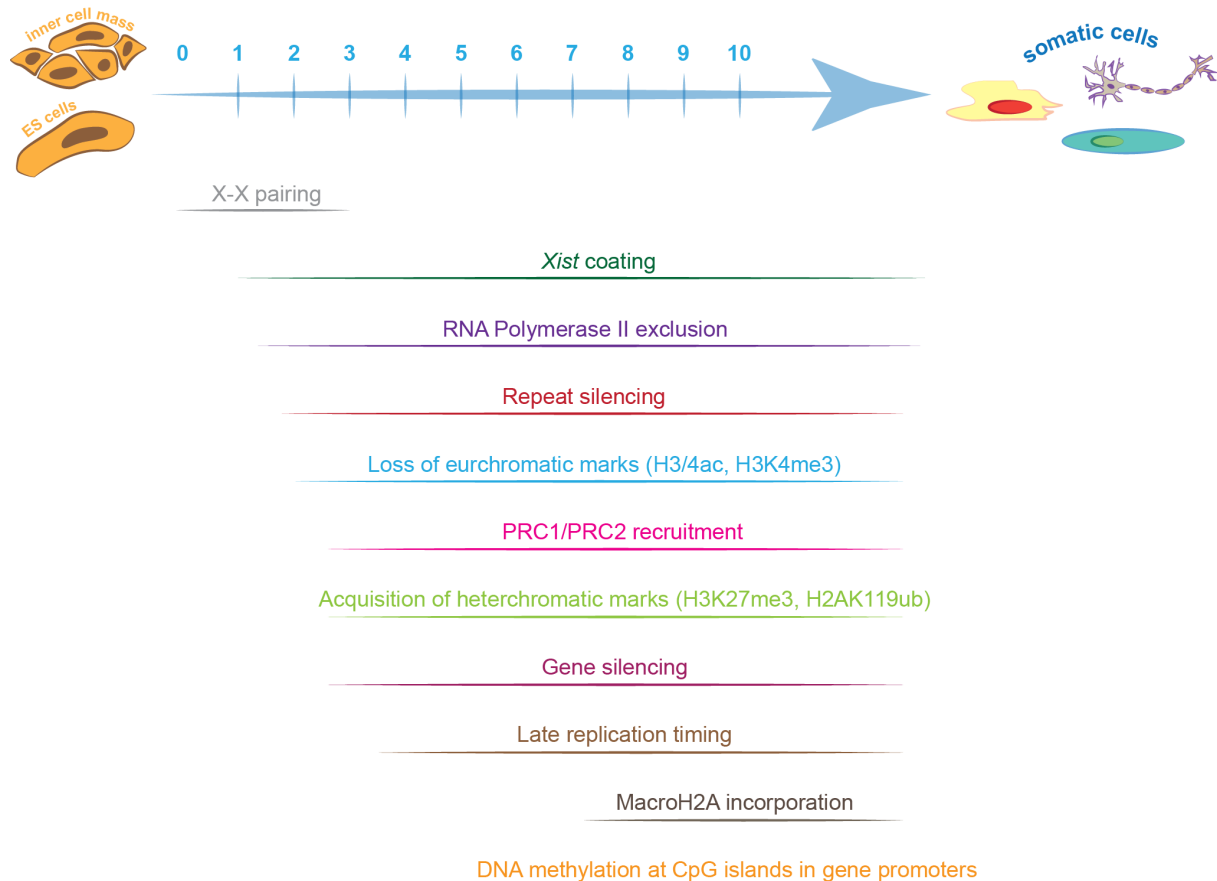


Figure 14. Features of *Xist*-mediated silencing

Approximate timeline of the events following *Xist* expression and upregulation. The arrow represents the time-course of ES cells differentiation. Vertical bars indicate each differentiation day.

Early silencing marks

While the heterochromatic marks deposited on the Xi are required to propagate the inactive state, little is known regarding the precise mechanisms of *Xist*-induced initiation of silencing before inactivation is locked in. In fact, the enrichment of epigenetic marks on the Xi is detected only after *Xist* coating and after many X-linked genes are silenced. For example, the PRC2 component EED, responsible for H3K27me3, is only required for XCI maintenance in differentiated extra-embryonic derivatives: trophoblast stem (TS) cells that lack *Eed* lose enrichment of heterochromatic marks on the Xi, but maintain silencing of X-linked loci (Kalantry and Magnuson, 2006).

Since the *Xist*-bound protein SHARP has been proposed to interact with HDAC3, it is possible that one of the early mechanisms of silencing involves histone deacetylation

and chromatin compaction. Other *Xist*-bound proteins may also play silencing roles that precede the deposition of epigenetic marks. For instance, *Xist*-bound scaffold proteins could mediate a chromosome-wide remodelling of the Xi. Facultative scaffold/matrix attachment regions (S/MARs) are indeed enriched in gene bodies, where *Xist* first accumulates (Engreitz et al., 2013; McHugh et al., 2015) and from where *Xist* is depleted in the case of escapees (Engreitz et al., 2013). *Xist* could for example accumulate on S/MAR-containing loci, interact with HNRNPU and other matrix-related proteins, and trigger the relocation of active genes into the silent compartment, the compaction of which would in turn preclude the access of the transcription machinery to their promoters. This would be consistent with the fact that one of the earliest detectable events during XCI initiation is the exclusion of RNA Polymerase II and general transcription factors from the *Xist* nuclear domain (Chaumeil, 2006; Clemson et al., 1996, 2006).

4.3.2 *Xist* and the maintenance of X chromosome inactivation

The participation of *Xist* RNA to XCI maintenance is correlated with specific cellular contexts and it is thus not found in all cell types.

4.3.2.1 XCI maintenance in the embryonic lineages: *Xist* and other epigenetic pathways

In the embryonic lineages of the developing embryo, DNA methylation pathways are required for locking-in the repressive state. DNA methylation at X-linked gene promoters is a hallmark of the Xi and the DNA methyltransferase *Dnmt1* has been shown to be necessary for XCI maintenance and to result in embryonic lethality when mutated (Sado et al., 2000). Similarly, the Xi-enriched protein *Smchd1* was identified as required for the maintenance of gene repression and DNA methylation on the Xi during early mouse development, since reactivation of genes on the Xi and female-specific lethality is observed when *Smchd1* is mutated (Blewitt et al., 2008).

While both *Dnmt1* and *Smchd1* are essential for XCI maintenance in the embryo, mutations in either gene do not affect (imprinted) XCI maintenance in the extra-embryonic tissues (Sado et al., 2000; Blewitt et al., 2008). Consistent with this, the Xi is hypomethylated in the placenta (Cotton et al., 2009), indicating that mechanisms other than DNA methylation are required for XCI maintenance in this lineage. A candidate for the role is the PRC2 component *Eed*, whose mutation results in a

female-specific trophoblast development defect (Wang et al., 2001).

Apart from DNA methylation, other epigenetic mechanisms are put in place to ensure XCI maintenance. For instance, the protein product of *Orc2* impacts Xi silencing through stabilization of HP1 α on the Xi and *Lrwd1* takes part in the same process possibly by interacting with the ORC complex (Kholmaier et al., 2004). Other candidates are genes involved in RNA splicing, including the previously identified ASF/SF2 protein (Royce-Tolland et al., 2010).

Not only chromatin structure but also the position of the X chromosome within the nuclear space changes during initiation and maintenance of XCI. In the mouse, it has been shown that prior to XCI the two active X chromosomes are randomly positioned in the nucleus, where they pair transiently upon initiation of silencing (Bacher et al., 2006; Xu et al., 2006) (section 5.2.1). Once silenced, the Xi occupies a discrete and condensed compartment close to the nuclear periphery (Barr and Bertram, 1949; Zhang et al., 2007). Here, the Xi is physically associated with the nuclear matrix, possibly with the help of the nuclear matrix-associated proteins that interact with *Xist* (section 4.3.1.3). The Xi is also tethered to the nucleolus via a lncRNA, *Firre* (Yang et al., 2015). *Firre* represents an X-linked macrosatellite repeat located outside the *Xic*. *Firre* escapes XCI and the transcript, like *Xist*, is proposed to bind the nuclear matrix protein HNRNPU (Yang et al., 2015). Knockdown of *Firre* results in decreased Xi perinucleolar association and reduced H3K27me3 enrichment. However, loss of *Firre* does not affect *Xist* expression nor causes re-activation of X-linked genes, and it does not affect XCI initiation.

Together with DNA methylation and heterochromatin proteins, continuous *Xist* expression is required for correct XCI maintenance in both the embryonic and extraembryonic lineages of the developing embryo. For example, in embryonic fibroblasts, *Xist* loss leads to erosion of some markers of XCI, such as H3K27me3 and macroH2A, and to limited gene re-expression (Nechanitzky et al., 2012). Similarly, perturbation of the normal pattern of *Xist* expression and localization via antisense LNA oligonucleotides leads to reactivation of the X-linked gene *Mecp2* (Brown and Willard, 1994). The inactive state can be further destabilized when DNA methylation and H4 hypoacetylation are perturbed, indicating a cooperative role of *Xist* RNA, DNA methylation and histone hypoacetylation in the maintenance of XCI (Csankovszki et al., 1999). Likewise, combined inhibition of the Mbd1-Atf7pi-Setdb1

network and DNA demethylation with *Xist* depletion lead to some gene reactivation in MEFs (Zhang et al., 2007). Deletion of *Xist* also disrupts the nucleolar association of the Xi that takes place during the S phase, but it is yet unclear whether the partial chromosome reactivation that occurs in this context is causally linked to nuclear repositioning. These observations suggest a synergistic action of *Xist* RNA and of other epigenetic pathways in the maintenance of the inactive state, so that removal of *Xist* alone only has mild effects on Xi-reactivation.

In the cells of the extra-embryonic tissues, *Xist* alone seems to have an even more important contribution. Mutant embryos that inherit a paternal *Xist* mutation properly establish silencing at the 8- and 16-cell stages but reactivate an Xp-linked GFP transgene in a significant number of extra-embryonic cells at the time of implantation (Kalantry et al., 2009). Similarly, a subset of extra-embryonic cells frequently reactivate Xp-linked genes in tissues carrying a paternal hypomorphic *Xist* mutation (Hoki et al., 2011).

4.3.2.2 *Xist*-independent XCI maintenance in adult tissues

Like in the embryonic lineage of the developing embryo, DNA methylation is required in the somatic tissues of the adult to maintain gene repression on the Xi, which is, relative to the Xa, hypermethylated in gene-rich regions (Cotton et al., 2011; Hellman and Chess, 2007). Unlike in the early embryo though, in adult somatic tissues only few and very specific cell types possess the right context for *Xist* silencing competence. For example, *Xist* induction initiates silencing in hematopoietic precursor cells but not in hematopoietic stem cells or fully differentiated blood cells (Agrelo et al., 2009), indicating that transient reactivation of the pathways sensitive to *Xist*-dependent silencing is present in cells such as pre-B cells and T-cell precursors but not in the other cell types of the same lineage.

5 *Xist* regulation

Selective inactivation of one of the two Xs in female cells implies a differential treatment of two homologous chromosomes in the same nucleus in a way that ensures *Xist* expression from one chromosome and *Xist* repression on the other. *Xist* expression is ultimately controlled by a balanced combination of repressors and activators, including *trans*-acting factors and *cis*-acting sequences. In eutherians, the *Xic/XIC* can be seen as the *Xist* gene and the *cis* regulatory elements that ensure *Xist* correct expression on both the Xa and Xi during development.

In the mouse, most of the genes of the *Xic* have been shown to play a role in *Xist* regulation, with *Xist* antisense *Tsix* mediating *Xist* repression on the Xa, *Jpx* and *Rnf12* activating *Xist* expression on the Xi in a dose-dependent manner, and *Ftx* having a positive effect on the expression of *Xist* and other neighboring *Xic* genes. The involvement of these genes in XCI will be detailed in the following sections.

In humans, a *XIST* antisense transcript, *TSIX*, has been identified, but does not seem to share the same functions as its murine homolog. A study aimed at the characterization of the human *XIC* region identified a >30 kb non translated nuclear RNA transcribed antisense to *XIST* and expressed in embryo-derived cells, but not in human adult somatic cells (Migeon et al., 2001). *TSIX* shares little homology with the murine *Tsix* and does not seem to act as antagonist to *XIST* expression; in fact *TSIX* is transcribed only from the inactive X in human fetal cells, and co-expressed with *XIST* (Migeon et al., 2002). Moreover, antisense transcription does not span the whole *XIST* gene and it does not reach the *XIST* promoter (Migeon et al., 2001). *TSIX* was possibly truncated during evolution, as hypothesised after 5' RACE experiments in mouse embryonic fibroblasts carrying the human *XIC* transgene (Migeon et al., 2001). Consistent with the possible lack of function of the human *TSIX*, the *TSIX* TAD in human ESCs is not pronounced as the mouse one (Dixon et al., 2012). The fact that *Tsix* is poorly conserved suggests that it might have evolved to fulfill specific needs during mouse development and raises the possibility that a *TSIX*-independent *XIST* regulatory mechanism exists in human. Interestingly, a human-specific X-linked lncRNA, termed *XACT*, has been identified in human ES cells (Vallot et al., 2013). *XACT* appears to be transcribed from a genomic region that spans more than 250 kb and to produce a polyadenylated transcript. Contrary to

XIST and similarly to *Tsix*, *XACT* RNA is expressed exclusively from an active X chromosome. Strikingly, *XACT* is also able to coat the active X in *cis* in both XX and XY human ES cells. Moreover, *XACT* expression seems to be restricted to human pluripotent cells, as analysis of RNA-seq libraries from female fibroblasts failed to detect *XACT* reads and RT-qPCR experiments showed absence of *XACT* expression in several tissues (Vallot et al., 2013). While it is tempting to speculate that *XACT* might have a *Tsix*-like function in the regulation of *XIST* expression, the fact that in female hESCs one X chromosome is, in most cases, already inactivated, makes it difficult to investigate the involvement of *XACT* in the control of XCI initiation.

5.1 *Xist* repression

In the mouse, X chromosome inactivation is prevented in the pluripotent state and established when the loss of pluripotency occurs. In pluripotent cells, *Xist* is repressed on both X chromosomes.

5.1.1 Undifferentiated mESCs: X chromosome inactivation and pluripotency

In the mouse, random XCI is intimately linked to differentiation, as initially proposed by a model envisioning that XCI occurs at different times and in different cell populations differentiating from a pluripotent state (Monk, 1978). Consistent with this model, ES cells (Rastan and Robertson, 1985) (Figure 15) and stem cells of teratocarcinomas (Martin et al., 1978) display transcriptional activity of both X chromosomes until initiation of differentiation, when also XCI is established. Reprogramming experiments also strengthened the idea that XCI and differentiation are tightly linked. Female nuclei of somatic cells fused with ES cells (Tada et al., 2001) or male embryonic carcinoma (EC) cells (Takagi et al., 1983) are known to reactivate the silent X chromosome. Xi reactivation is also observed upon somatic nuclear transfer into the enucleated egg (Eggan et al., 2000). Similarly, the generation of induced pluripotent stem (iPS) cells by forced expression of ES cells regulators (Takahashi and Yamanaka, 2006), is also accompanied by Xi reactivation (Stadtfield et al., 2008).

The link between XCI and pluripotency/differentiation is much less evident in the case of imprinted XCI. Imprinted XCI takes place in all the cells of the preimplantation embryo, prior the formation of the blastocyst and when the cells are still pluripotent (Mak et al., 2004; Okamoto et al., 2004). However, since the inactive Xp is

reactivated (X chromosome reactivation, XCR) in the ICM (Mak et al., 2004; Okamoto et al., 2004), it is perhaps in this transition that a critical layer of iXCI regulation exists, and coupling of inactivation to differentiation is, in this instance, just one aspect of a complex multilayered regulation system. Other than transiently in the ICM, reactivation of the (randomly inactivated) Xi also takes place in the pluripotent primordial germ cells (PGCs) *in vivo* (Tam et al., 1994). Hence, what it is clear is that, while iXCI is not *per se* associated to differentiation, rXCI is, both *in vivo* and *ex vivo*.

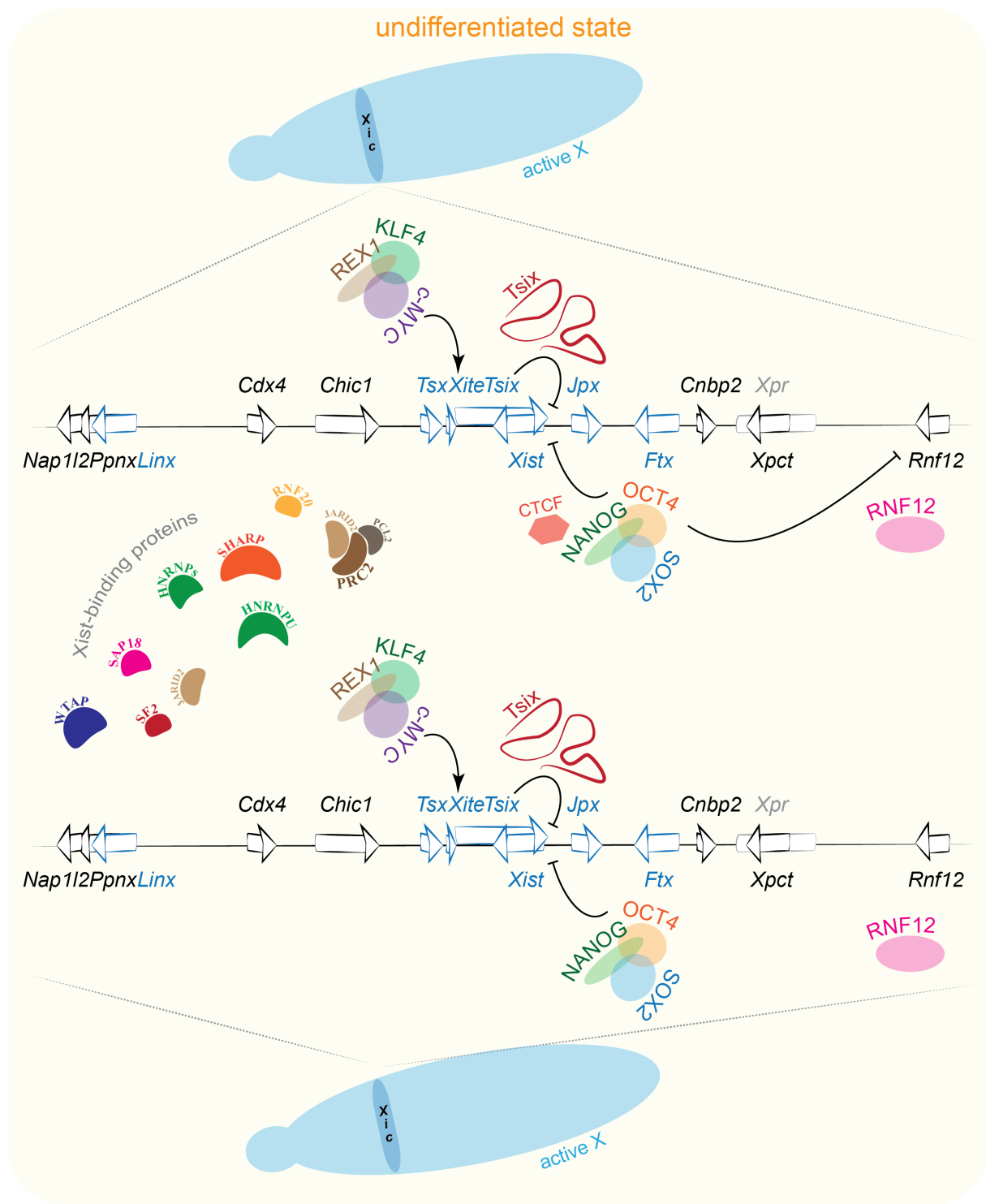


Figure 15. Absence of X chromosome inactivation in undifferentiated mESCs

In undifferentiated pluripotent mESCs, XCI is prevented by multiple mechanisms that ensure *Xist* repression from both X chromosomes. The pluripotency factors OCT4, NANOG and SOX2 bind to the *Xist* locus to repress transcription. At the same time, the pluripotency factors REX1, KLF4 and c-MYC stimulate *Tsix* expression and the production of *Tsix* RNA. By yet unclear mechanisms, *Tsix* prevents *Xist* upregulation in *cis*. At this stage, *Rnf12*, also bound by OCT4, NANOG and SOX2, is lowly transcribed and the pool of RNF12 protein is not sufficient to trigger *Xist* upregulation. Also, *Xist* RNA is not bound to its protein interactors.

5.1.2 XaXa status and differentiation potential

In many mammals, including mice (Seller and Perkins-Cole, 1987; Valdivia et al., 1993), rats (Hemet and Rabineau, 1966; Scott and Holson, 1977), cows (Avery et al., 1989; Xu et al., 1992; Yadav et al., 1993) and humans (Pedersen, 1980; Pergament et al., 1994), female embryos develop more slowly than males during early embryogenesis. This difference is unlikely to be mediated by hormones, since none is produced at these early stages of development. The disparity seems to be rather caused by the differential expression of sex-linked genes between males and females. It has been initially suggested that the faster development of male embryos could be a consequence of the expression of Y-linked genes, such as *Sry* and *Zfy*, as early as the pre-implantation stages (Zwingman et al., 1993). Indeed, chromosome constitution has been shown to play a relevant role in mouse, where both the presence of the Y chromosome and the absence of the second X chromosome tend to accelerate pre-implantation development (Burgoyne, 1993; Burgoyne et al., 1995). For example, mice with an XO karyotype (maternal X) are larger than XX mice at the same gestational age, consistent with faster development during pre-implantation development (Burgoyne, 1993). Indeed, the Y chromosome has been shown to carry a factor that accelerates development (Burgoyne, 1993), but this element is unlikely to be *Sry* (Burgoyne, 1995).

Identification of the molecular pathways responsible for chromosome composition sensing has been obtained through studies in mouse ES cells (Schulz et al., 2014). Genome-wide transcriptome comparison of ES cell lines bearing distinct chromosome constitutions revealed that double X dosage in undifferentiated cells stabilizes the pluripotent state by inhibiting the MAPK and *Gsk3* signalling pathways and by activating the *Akt* pathway. As a result, differentiation is blocked and remains on hold until dosage compensation has been achieved. According to the proposed model, by sensing the activity of the X chromosomes via the balance of two distinct signalling pathways, the female cell can coordinate X inactivation and differentiation. Whether the same mechanisms are put in place also *in vivo* remains to be tested.

Mechanistically, the fact that, in XaXa pluripotent cells, *Xist* RNA steady levels are extremely low and the locus is inefficiently transcribed suggests either the presence of a transcriptional repressor (or repressors) or the lack of a specific activator (or activators) in these cells, or a combination of both scenarios.

5.1.3 *Xist* repression and *Tsix* activation

The XaXa configuration is a key attribute of the pluripotent state and partly depends on *Xist* transcriptional repression. OCT4, SOX2 and NANOG are key transcription factors in the maintenance of ES cells identity, where they activate genes of the self-renewal program and repress lineage commitment genes. All three factors were shown to bind *Xist* intron 1 in male and female undifferentiated ES cells and to lose the binding upon differentiation (Navarro et al., 2008). Consistent with the role of these pluripotency factors in *Xist* transcriptional repression, depletion of NANOG and OCT4 results in *Xist* upregulation from the single X in male ES cells and from both Xs in differentiating female ES cells (Donohoe et al., 2009; Navarro et al., 2008). The mechanism of pluripotency factor-mediated *Xist* repression is proposed to be modification of the three-dimensional configuration at the *Xist* locus (Tsai et al., 2008).

However, while they are clearly needed for *Xist* repression, these factors might also not act directly, as they bind to thousands of loci on the genome to maintain pluripotency. In line with this, a genetic deletion of OCT4, SOX2 and NANOG binding sites within *Xist* intron 1 does not result in a strong change of *Xist* expression (Barakat et al., 2011), suggesting that other regulators might be involved. Indeed, a combined deletion of OCT4, SOX2 and NANOG binding sites on *Xist* intron 1 and *Tsix* enhances *Xist* upregulation compared to the deletion of *Xist* intron 1 binding sites alone (Nesterova et al., 2011).

Nanog appears to be needed for *Xist* upregulation also in the pre-implantation embryo, since its genetic overexpression causes rapid *Xist* RNA loss, even if without affecting the timing of XCR (Williams et al., 2011b). Conversely, pre-implantation embryos that lack NANOG are unable to specify the epiblast lineage and do not induce the loss of *Xist* RNA coating on the Xp upon XCR (Silva et al., 2009). Once again, *Nanog* alone may not be the only cause of this effect, since NANOG is already present in the cells of the late morula that display an Xi (Williams et al., 2011b; Silva et al., 2009), indicating that other mechanisms are coupled to *Nanog* to control *Xist* repression.

Another *Xist* repressing element is the lncRNA-encoding gene *Tsix*, identified following the observation that the region 3' to *Xist* influences X-chromosome counting at the onset of XCI. In a pioneer study, XO female cells harboring a 65 kb deletion 3'

to *Xist* induced *Xist* expression and triggered inactivation of their sole X chromosome, indicating that something must be important in this region for the suppression of ectopic XCI (Clerc and Avner, 1998). Further assessment of this 65 kb region by RNA-FISH in female and male ES cells, as well as deletion and sequence reinsertion experiments, led to the identification of an RNA transcribed antisense to *Xist*, expressed in female cells from both chromosomes prior to X inactivation (Lee et al., 1999b; Morey et al., 2001, 2004). Similarly to ES cells, *Tsix* is expressed in embryo of both sexes in pre-implantation development (Lee, 2000; Sado et al., 2001), when *Xist* is not yet activated.

There are several hypotheses regarding the mechanism by which *Tsix* prevents *Xist* upregulation and the debate is still open to whether it is the act of transcription or the *Tsix* RNA product that is responsible for *Tsix*-mediated *Xist* regulation. Antisense transcription across the *Xist* promoter is followed by DNA methylation and accumulation of repressive histone modifications (such as H3K9me3) and generally by a global pseudo-euchromatinization of a CTCF-flanked region of the *Xist* promoter (Navarro et al., 2005, 2006; Sado et al., 2005). However, DNA methylation of the *Xist* promoter might not be a direct mechanism for *Xist* regulation, since depletion of both DNMT3a and DNMT3b *de novo* methyltransferases does not lead to aberrant XCI (Sado et al., 2004). *Tsix* transcription was also proposed to block the spreading of the H3K27me3-rich domain flanking the *Xist* locus to the adjacent H3K4me2-rich *Xist* region, by a mechanism involving RNA Polymerase II accumulation at the 3' end of *Tsix* (Navarro et al., 2009).

Expression of *Tsix* transcripts was detected in a gradient along the *Xist* gene, with the highest proportion of transcript in the 5' region of *Tsix* relative to the 3' portion, leading to the hypothesis that, other than histone modifications, transcriptional interference might be the principal mechanism for the suppression of *Xist* expression (Luikenhuis et al., 2001; Shibata and Lee, 2003; Sun et al., 2006a). A report indicating the presence of *Tsix/Xist* duplexes matching *Xist* repeat A and the *Xist* promoter in ES cells, mouse embryonic fibroblasts (MEFs) and *in vivo* strengthened the hypothesis of the involvement of *Tsix* in an RNAi-like mechanism (Ogawa et al., 2008). Moreover, the deletion of *Dicer*, one of the main players of the RNAi pathway, was shown to cause *Xist* de-repression and loss of H3K27me3 accumulation on the Xi ((Ogawa et al., 2008). However, increased *Xist* levels in *Dicer* mutant ES cells

were later proposed to be a secondary consequence of the microRNA-mediated misregulation of *de novo* methyltransferases in these cells (Nesterova et al., 2008). In agreement with this, *Dicer*-deficient male and female embryos display normal *Xist* expression patterns, suggesting that *Xist* regulation is not affected in this model (Nesterova et al., 2008).

Tsix has also been proposed to act, in the form of transcript, as *Xist* competitor for the binding to *Xist*-activating factors. It was shown that *Tsix* inhibits the interaction between an independent 1.6 kb non-coding RNA transcribed across *Xist* repeat A (and termed RepA) and PRC2, which is required for H3K27 trimethylation and full-length *Xist* induction (Zhao et al., 2008).

Tsix expression in undifferentiated pluripotent ES cells relies on the recruitment of the pluripotency factor REX1 and the reprogramming-associated factors KLF4 and C-MYC at the *DXPas34* minisatellite, an enhancer associated with the *Tsix* promoter (Navarro et al., 2010). When *DXPas34* is deleted, binding of REX1, KLF4 and C-MYC is no longer occurring, and the transcriptional machinery cannot be efficiently recruited to the *Tsix* promoter. Symmetrically, knockdown of REX1 results in impaired *Tsix* transcriptional elongation (Navarro et al., 2010). *Rex1* expression decreases upon ES cell differentiation, in anti-correlation with XCI initiation, suggesting that loss of REX1 results in the downregulation of *Tsix* expression required to promote *Xist* upregulation from the future Xi. An additional link between XCI, *Tsix* and pluripotency is given by the observation that REX1 protein levels are regulated by the X-linked dose-dependent XCI activator *Rnf12* (section 5.2.3) (Gontan et al., 2012).

Consistent with the suggested role for *Tsix* as *Xist* repressor, *Tsix* was proposed to play a role in XCR, when epiblast precursors lose the *Xist* coating (Mak et al., 2004; Navarro et al., 2009; Nesterova et al., 2011; Sheardown et al., 1997), but actual evidence for this hypothesis is still missing, and reactivation is normal in embryos carrying paternally-inherited *Tsix* mutations (Kalantry and Magnuson, 2006). Overall, two different pluripotency-associated pathways couple XCI and differentiation: *Nanog*-, *Sox2*- and *Oct4*- mediated *Xist* repression and *Rex1*-, *Klf4*- and *c-Myc*-mediated *Tsix* activation.

However, while *Tsix* plays a fundamental role in *Xist* repression, the observation that abrogating *Tsix* transcription does not lead to *Xist* derepression prior to differentiation

(Lee and Lu, 1999) indicates that other complementary mechanisms, such as the pluripotency factor pathways discussed above and maybe others, must exist to prevent *Xist* upregulation at the undifferentiated state.

5.1.4 Absence of interaction with *Xist* binding proteins

Combined with the action of multiple mechanisms that ensure *Xist* transcriptional repression, the lack of *Xist* RNA binding to developmentally-regulated and differentiation-specific *Xist* interactors participates to prevent *Xist* upregulation in undifferentiated cells. While some proteins (Set 1, (Chu et al., 2015)) are bound to *Xist* RNA already in the undifferentiated state and thus probably represent the ground-state of *Xist* interactome, others (Set 2, (Chu et al., 2015)) interact with *Xist* only at the onset of differentiation or in XaXi cell types (EpiSCs and TSCs). It could be thanks to Set2 proteins that *Xist* acquires its silencing function in a developmentally-regulated manner. Along this line, the absence of *Xist*/Set2 protein interactions in undifferentiated cell types might provide safety against premature *Xist* activation during pluripotency.

5.2 *Xist* activation

When differentiation is induced, ES cells exit the ground state of pluripotency, transition through a primed state during which some pluripotency factors are downregulated, and then engage into differentiation towards a specific cell type (Nichols and Smith, 2009). With the decrease in pluripotency factors abundance *Xist* repression is released, and female-specific, monoallelic, *Xist* upregulation can trigger XCI (Figure 16). At the same time, the XCI-induced presence of only a single dose of X-linked genes contributes to the downregulation of pluripotency factors and facilitates the engagement into differentiation (Schulz et al., 2014).

While pluripotency factors are undoubtedly important for XCI regulation, they are unlikely to be the only developmental trigger of *Xist* upregulation, as experimental analysis of the relationship between *Xist* and pluripotency factors levels only found a minor anti-correlation between the two entities (Schulz, 2013). This means that, in the course of differentiation, changes in *Xist* expression are due to additional factors that stabilize *Xist* on the Xi and prevent *Xist* stabilization on the Xa.

interactors. On the Xa, *Tsix* continued expression prevents *Xist* upregulation in *cis* and DNA methylation at the *Xist* promoter also contributes to *Xist* repression.

5.2.1 X-X pairing and the initiation of XCI: *Tsix* and *Xpr*

The position of the X chromosomes within the nucleus of female cells changes during initiation and maintenance of XCI. The two *Xics* come into close proximity during the initiation of random inactivation in differentiating female ES cells (Bacher et al., 2006) and remain so only transiently (Xu et al., 2006). Since the pairing is detectable before *Xist* monoallelic upregulation from only one X chromosome, *Xic-Xic* interactions were suggested as being important for the sensing step, and, although not sufficient, also necessary for correct counting and choice. To date, two *cis*-acting pairing elements have been described: the *Tsix* promoter region and a genomic element located around the *Xpct* gene upstream of *Xist*.

Autosomal insertion of the *Tsix* promoter region (including *Tsix* enhancer *Xite*) in mouse ES cells induces *de novo* X-autosome pairing and inhibits endogenous X-X pairing (Xu et al., 2006). Importantly, pairing is disrupted when *Tsix* is deleted (Bacher et al., 2006, Xu et al., 2006). These results indicate that the *Tsix* promoter region is somehow implicated in counting by allowing the two X chromosomes to contact each other.

More generally, it has been observed that different deletions around the *Tsix* locus cause ectopic *Xist* induction in differentiating XO female and male ES cells, indeed suggesting that the cell is unable to sense correctly the actual number of X chromosomes when *Tsix* is absent (Lee, 2005). Since the deletions also encompass a repetitive region, called *DXPas34*, located close to *Tsix* exon 3 and needed to augment *Tsix* expression (Clerc and Avner, 1998; Cohen et al., 2007; Navarro et al., 2010; Vigneau et al., 2006) it is also possible that *Tsix* has a role in counting via its regulation by *DXPas34*.

Despite containing counting and choice elements 3' to *Xist*, including the *Tsix* promoter, a 460kb *Xic* transgene cannot induce *cis*-silencing in ES cells unless present as multicopy insertion (Heard et al., 1996). This likely indicates that the *Xic-Xic* colocalization requires long-range sequences that are partly absent from the transgene. In agreement with this, *Tsix* transcription might not be sufficient for the pairing, as *Tsix* is expressed in ES cells but *Xic* colocalization is not observed in this state (Bacher et al., 2006; Xu et al., 2006). Additionally, although ectopic multicopy

arrays of *Tsix* and *Xite* can pair with the endogenous *Xic* (Xu et al., 2006), single copy *Xist/Tsix/Xite* transgenes are not able to (Bacher et al., 2006), indicating that other sequences might be required for this *trans* interaction. A screen for homologous *trans*-interacting regions in differentiating female ES cells enabled the identification of a segment of the *Xic* necessary to bring the two chromosomes together and able to induce pairing even when present as a single copy transgene (Augui et al., 2007). This segment, called *X pairing region* (*Xpr*), is located 5' to *Xist* within an heterochromatic hotspot (Heard et al., 2001; Rougeulle et al., 2004) and overlaps the *Xpct* gene (Augui et al., 2007). *Xpr*-driven homologous association of the two X chromosomes is thus thought to be needed to ensure the female-specific correct *trans*-regulation between the two *Xist/Tsix* loci.

Even if the mechanism of pairing is not yet clear, CTCF, OCT4 and transcription are thought to participate in it (Donohoe et al., 2009; Xu et al., 2007). *Xic-Xic trans* interactions were also reported to occur in the distal sub-region of the *Xic*, mediated by the *Tsix*-induced H3K27me3 deposition at the *Xpct* locus (Navarro et al., 2009).

Although *Xic-Xic* pairing is an attractive model to explain XCI sensing, it is unclear how it can accommodate the observations made in diploid XXXX, tetraploid XXXX, tetraploid XXXY, and tetraploid XXYX cells, that inactivate respectively three, two, one or no X chromosomes (Monkhorst et al., 2008). Perhaps in these cells the mechanism that assesses the X chromosome:ploidy ratio exerts a stronger effect than *Xic-Xic* cross-talk on sensing and counting. Moreover, in female-male XX-YY heterokaryons, in which direct interactions between the "male" X chromosome and the "female" X chromosomes are precluded, any one of the three Xs (one of the "female" Xs or the single "male" X) can become the Xi upon differentiation (Barakat et al., 2014). This demonstrates that direct physical contact between the two X chromosomes is dispensable for XCI counting and initiation, and that the XCI is rather regulated by *trans*-acting cytoplasmic factors (Barakat et al., 2014). In this eventuality, X-X pairing would be a consequence rather than a cause of the XCI process.

5.2.2 *Tsix* and XCI choice

Other than in preventing *Xist* upregulation at the undifferentiated state and possibly in X-X pairing and counting, *Tsix* is involved in the regulation of *Xist* monoallelic expression. Before XCI, *Tsix* is expressed in female cells from both chromosomes, but its expression is gradually restricted to the Xa upon differentiation (Lee et al., 1999b; Morey et al., 2001, 2004), suggesting its involvement in the choice of which X chromosome to inactivate. Interestingly, a homozygous deletion of the *Tsix* promoter region indeed results in a chaotic choice mechanism, in which one, two or no X chromosomes are chosen for inactivation (Lee, 2005).

Consistent with *Tsix* role in choice, surviving female mice harboring a heterozygous *Tsix* mutation show skewed *Xist* expression toward the mutant X in fibroblasts, independently of the parent-of-origin inheritance of the mutation (Lee, 2000). Similarly, a targeted heterozygous *Tsix* deletion in female and male mouse ES cells results in skewed *Xist* expression toward the mutated X, without affecting the overall expression level of *Xist* (Lee and Lu, 1999; Morey et al., 2001, 2004; Rougeulle and Avner, 2004; Sado et al., 2001). Conversely, the insertion of an inducible promoter to force *Tsix* expression during female ES cells differentiation abrogates the possibility of *Xist* upregulation from the mutated chromosome (Stavropoulos et al., 2001).

Two models could explain this skew in XCI: a primary non-random X chromosome choice model, in which ectopic *Xist* induction from the *Tsix*-mutated chromosome is the leading cause of the preferential inactivation of the mutant X, or a secondary random X chromosome choice model, in which inactivation is normal and the cells that choose to inactivate the wild-type X chromosome are selected against due to death following the presence of two inactive Xs. The first model is more likely to be the correct one, since inactivation of the wild-type X chromosome in differentiating *Tsix* mutant ES cells is seldom observed (Lee and Lu, 1999; Morey et al., 2001; Sado et al., 2001).

5.2.3 Dose-dependent XCI regulators: *Rnf12* and *Jpx*

Rnf12

Pluripotency factors and *Xist* interactors are not the only proteins involved in the regulation of XCI. The X-linked *Xic*-encoded gene *Rnf12* produces the E3 ubiquitin

ligase RNF12, whose levels influence the probability of *Xist* upregulation (Jonkers et al., 2009). Heterozygous deletion of *Rnf12* strongly reduces XCI initiation at the start of differentiation, with decreased *Xist* expression from either X chromosome, suggesting a role for RNF12 in XCI activation in *trans* (Jonkers et al., 2009). This effect is independent of *Tsix* (Barakat et al., 2011) and mediated by RNF12 interacting partner REX1 (Gontan et al., 2012). Consistent with *Rex1* involvement in RNF12-mediated XCI regulation, differentiating male ES cells with a heterozygous *Rex1* deletion undergo ectopic XCI, and female ES cells overexpressing *Rex1* display a strong XCI inhibition (Gontan et al., 2012).

The XCI defect observed in *Rnf12* heterozygous cells is recovered at later differentiation time points, when *Xist* is expressed only from the *Rnf12* deleted chromosome. This indicates that cells that inactivate the mutant chromosome are selected for, possibly to avoid functional RNF12 deficiency. Indeed, one study demonstrated that *Rnf12* null deletion in ES cells leads to almost complete abrogation of XCI (Barakat et al., 2011). Contrasting results were however obtained in an independent analysis, in which *Rnf12*-deleted ES cells showed delayed XCI kinetics similar to those observed in heterozygous *Rnf12* mutant cells (Shin J 2010). The discrepancy between these two studies might be explained by the different origin of the ES cells analyzed: while in the first case (Barakat 2011) the *Rnf12* deletion was generated in vitro, in the second case (Shin et al., 2010) the cells were derived from *Rnf12*-null embryos, in which they might have adapted or been selected for compensation of the deletion.

According to the current model, *Rnf12* is also proposed to participate in the regulation of monoallelic *Xist* expression by acting as a negative feedback regulator (Jonkers et al., 2009): XCI-induced *Rnf12* silencing on the Xi prevents *Xist* upregulation from the Xa because the pool of RNF12 protein falls below the threshold that is needed to induce *Xist* expression in *trans* (Figure 17).

Rnf12-dependent monoallelic *Xist* regulation is described also in the case of imprinted XCI: an *Rnf12* deletion transmitted maternally results in the loss of maternal RNF12 pool in the zygote, which in turn leads to compromised *Xist* upregulation and gene silencing on the Xp and subsequent female-specific embryonic lethality (Shin et al., 2010).

The RNF12/REX1 ES-based model is however challenged by *in vivo* studies: both

Rnf12-null (Shin et al., 2014) and *Rex1*-null (Masui et al., 2008) mice are viable with no apparent phenotype. Once again though, the apparent lack of phenotype could be caused by compensation phenomena put in place during early development.

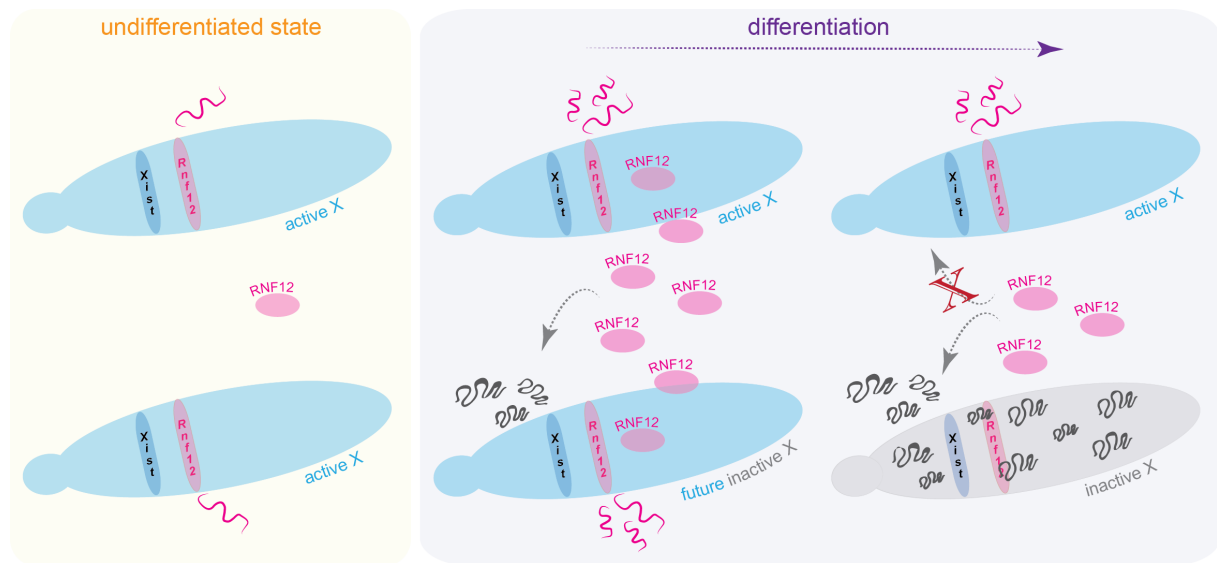


Figure 17. *Rnf12* activates *Xist* expression in a dose-dependent manner

In undifferentiated mESCs, *Rnf12* is lowly transcribed and RNF12 protein pool is too low to trigger *Xist* upregulation. In differentiating mESCs, *Rnf12* increase induces *Xist* upregulation, triggering the *cis*-inactivation that ultimately contributes to monoallelic *Rnf12* silencing and to the decrease of RNF12 protein levels. This feedback loop ensures that one X chromosome remains active.

Jpx

Since in cells heterozygous for an *Rnf12* deletion XCI is delayed but not prevented (Jonkers et al., 2009; Shin et al., 2010), additional X-linked *Xist* activators should exist. Investigation of EST databases and RT-qPCR coupled to 3'-RACE analyses led to the identification of *Jpx/Enox* RNA, a non-coding transcript mapping to an unmethylated CpG island and whose transcription initiates around 10 kb upstream of *Xist* P1 promoter and extends in the opposite orientation compared to *Xist* (Johnston et al., 2002). The initial observation that *Jpx* upregulation in differentiating female ES cells correlates with *Xist* induction and the fact that *Jpx* escapes XCI (Johnston et al., 2002; Tian et al., 2010), suggested a possible involvement of *Jpx* RNA in the regulation of *Xist*. However, *Jpx* upregulation in male ES cells (Tian et al., 2010) may also indicate an XCI-independent reason for *Jpx* upregulation. *Jpx* deletion in male ES cells does not result in deregulation of X-linked genes, while differentiating female ES cells heterozygous for *Jpx* deletion show a severe phenotype, with growth

defects, a high rate of cell death and a marked decrease in the number of cells displaying *Xist* coating (Tian et al., 2010). The effect is likely to be mediated by the transcript itself, as post-transcriptional *Jpx* knockdown recapitulates the phenotype and *Jpx* reintroduction in *trans* rescues it (Tian et al., 2010). When introduced in wild-type ES cells, the same transgene also results in a variable *Xist* induction concordant with the strength of *Jpx* promoter used, arguing in favor of a dose-dependent requirement for *Jpx* in *Xist* activation (Sun et al., 2013). The *Jpx* deletion is also rescued by a *Tsix* truncation, suggesting an antagonist role for these two long non-coding RNA genes and indicating that, when *Tsix* is absent, *Jpx* is not needed for *Xist* activation (Tian et al., 2010). Moreover, the introduction of a *Jpx* transgene in male *Tsix* mutant ES cells increases ectopic *Xist* expression (Sun et al., 2013).

Insights into *Jpx* mechanism of action were obtained with the observation that *Jpx* RNA is able to bind the insulator-binding protein CTCF *in vivo* (Sun et al., 2013). CTCF was previously shown to bind *Xist* P2 in male and female ES cells. While the binding remains unchanged in differentiating male ES cells, it is allelically-regulated and anticorrelated with *Xist* expression in differentiating female cells (Sun et al., 2013). Since CTCF binding on the active *Xist* promoter is decreased in differentiating female cells, *Jpx* upregulation is concomitant with *Xist* induction and *Jpx* RNA is able to bind to CTCF, it was proposed that *Jpx* activates *Xist* by physically removing CTCF from *Xist* P2 (Sun et al., 2013). However, this hypothesis was questioned by contradicting results showing that CTCF binds specifically to the active *Xist* allele (Makhlouf et al., 2014), thus arguing against a role for CTCF in *Xist* repression.

Evidence for a role for *Jpx* in *trans* in the activation of *Xist* was also challenged by a recent study in which the deletion of a large 416 kb genomic region including, among others, the full *Jpx* gene, resulted only in a mild phenotype in differentiating female ES cells, detectable exclusively when coupled to other mutations. Cells heterozygous for the 416 kb deletion display a significantly reduced number of nuclei with *Xist* coating only when also the trans-activator *Rnf12* is deleted (Barakat et al., 2014). In this study, *Jpx* is rather seen as a *cis*-activator needed to lower the threshold for RNF12-mediated *Xist* activation (Barakat et al., 2014). It is possible however that such a big deletion interferes with the *Xic* TAD structure and impacts XCI in an unexpected manner.

In all cases, since *Jpx* mutant female ES cells still display a low-level *Xist* expression

(Tian et al., 2010) and *Xist* coating is not drastically increased in male *Tsix* mutant ES cells with doubled *Jpx* RNA levels (Sun et al., 2013), it is clear that other *Xist* activators exist and that what drives *Xist* upregulation is ultimately a co-operation between all these factors.

5.2.4 Short non-coding transcripts

In addition to *Xic*-linked long non-coding RNAs, several short transcripts emanating from *Xist* first exon have been proposed to contribute to *Xist* activation. *RepA* is a 1.6 kb non-coding RNA overlapping the A-repeat and transcribed in the same orientation as *Xist* (Zhao et al., 2008). *RepA* is proposed to recruit PRC2 via its direct binding with the PRC2 subunit EZH2, interaction that appears to be counteracted by *Tsix* (Zhao et al., 2008). Depletion of *RepA* abolishes full-length *Xist* induction and H3K27me3 levels on the X and, similarly, absence of PRC2 compromises *Xist* upregulation (Zhao et al., 2008).

Another short transcript spanning the same region but in antisense orientation was recently described. *Xist AR* (*Xist activating RNA*) is expressed exclusively from the Xi *ex vivo* and *in vivo*, and it is required for robust *Xist* expression, as selective termination of *Xist AR* results in lack of *Xist* RNA expression in *cis* at the initiation of XCI (Sarkar et al., 2015).

5.2.5 *Xce* alleles and XCI skewing

In mice, a genetic component called *Xce* (*X controlling element*) modulates non-random skewing of XCI. Different mouse strains carry unique *Xce* alleles that result in skewing phenotypes in the progeny of hybrid crosses. The *Xce* locus has not been clearly defined, and conflicting results suggest that it is not just one locus, but may include different X-linked regions. However, the different attempts to map it (Calaway et al., 2013; Cattanach and Papworth, 1981; Chadwick et al., 2006; Simmler et al., 1993; Thorvaldsen et al., 2012) have located it close or overlapping with the *Xic* (Figure 18). Overall, six competing *Xce* alleles have been identified, with the order of strength being $a < f < e < b < c < d$, where *a* is the most likely to be inactivated and *d* the least likely. In *Xce* heterozygotes, the X chromosome carrying the weaker of the two alleles is more likely to be inactivated. The degree of skewing can vary a lot, with cases of mean X inactivation pattern as profound as ~25:75 in *Xce_d/Xce_a* hybrids (Calaway et al., 2013). Conversely, the choice is unbiased in *Xce* homozygotes

(Krietsch et al., 1986; Plenge et al., 2000). Interestingly, the parent-of-origin of the *Xce* allele plays a role in the degree of skewing, where maternal inheritance of an *Xce* allele increases its probability of remaining an *Xa* to a different extent depending on the strains used for the cross (Calaway et al., 2013; Chadwick et al., 2006; Thorvaldsen et al., 2012).

Interestingly, the smallest defined *Xce* region (Calaway et al., 2013) contains copies of the *Dmrt* genes, related to the *Drosophila melanogaster* *dsx* and *Caenorhabditis elegans* *mab-3* genes, which encode transcription factors with roles in sex determination.

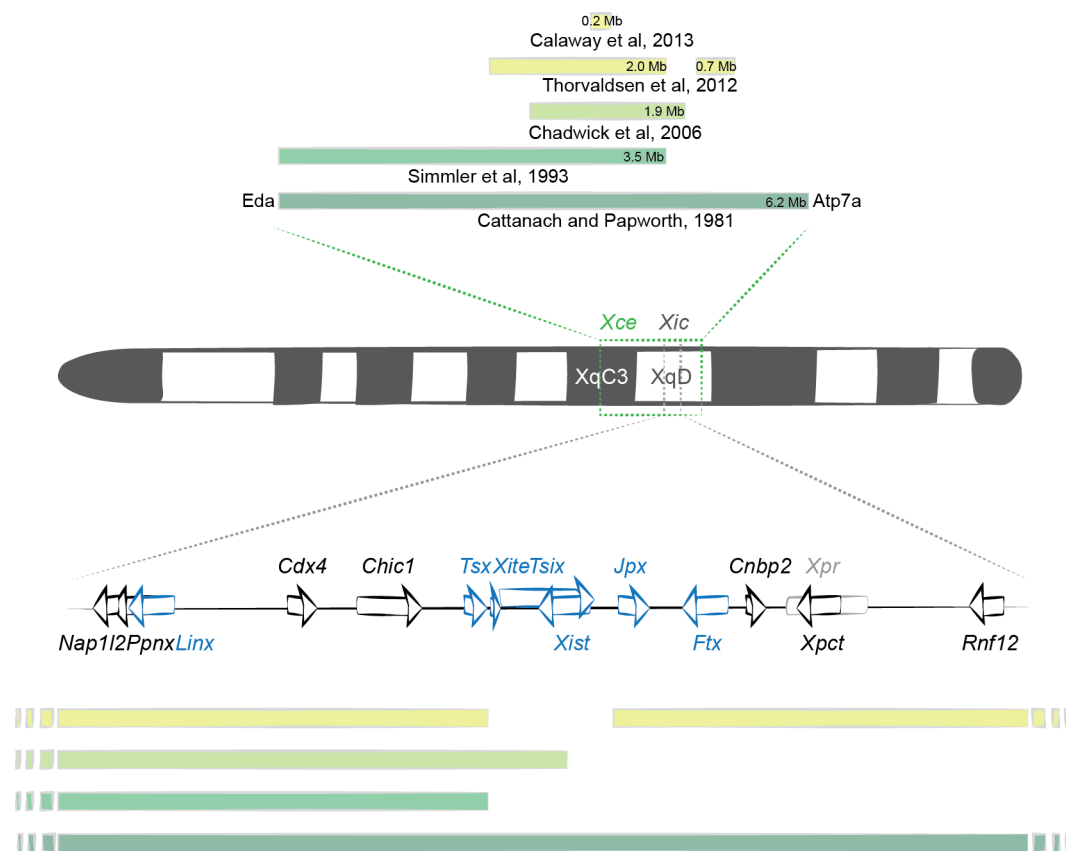


Figure 18. Delimitation of the *X* controlling element region

Different studies have attempted the precise delimitation of the *Xce*. The X chromosome is shown and horizontal bars represent the presumed location of the *Xce* according to each study. In the bottom part of the image, a zoom on the *Xic* is shown.

5.2.6 YY1-dependent *Xist* transcriptional activation

Not only X-linked, but also autosomal factors are involved in the monoallelic upregulation of *Xist*. The ubiquitous autosomal transcription factor YY1 plays a role in *Xist* activation, as shown in YY1 knock down experiments resulting in the loss of *Xist*

upregulation during XCI initiation and maintenance (Makhlouf et al., 2014). Prior to XCI, YY1 binds to the *Xist* 5' region, where it competes with the *Xist* repressor REX1. Upon differentiation, asymmetric *Xist* promoter methylation controls YY1 fixation, resulting in the exclusive binding of YY1 to the unmethylated *Xist* allele that induces *Xist* transcriptional expression (Makhlouf et al., 2014). Importantly, YY1 role in *Xist* activation is conserved also in humans (Makhlouf et al., 2014).

YY1, together with CTCF, was also reported as being a positive regulator of *Tsix* in undifferentiated ES cells via its binding to the *DxPas34* regulatory element (Donohoe et al., 2007)

5.2.7 *Xist* repression from the Xa: DNA methylation

Just like the stabilization of *Xist* on one chromosome is required to produce an Xi, *Xist* transcriptional repression is required on the other to keep it an Xa. At the beginning of differentiation, *Xist* silencing on the Xa is ensured by maintained *Tsix* expression (Lee, 2000; Sado et al., 2001). In differentiated cell types, *Tsix* expression is completely shut off, and *Xist* is transcriptionally repressed on the Xa by DNA methylation, as demonstrated by analyses showing that the *Xist* promoter is differentially methylated in somatic female cells (Norris et al., 1991). Consistent with this, loss of methylation on the Xa results in *Xist* transcriptional activation (Beard et al., 1995). Because of these observations, it would be reasonable to propose that DNA methylation at the *Xist* promoter is required as soon as *Tsix* is downregulated in differentiating cells. In agreement with these proposed dynamics, *Xist* promoter methylation is detected earlier in cells with a heterozygous mutation of *Tsix*, on the *Tsix*-deleted X (active) chromosome (Sun et al., 2006b). DNA methylation at the *Xist* promoter might however be functional only during XCI, since in undifferentiated female cells the two promoters are (weakly) methylated (Navarro et al., 2006; Penny et al., 1996), yet *Xist* is biallelically repressed.

6 *Ftx*, an additional non-coding gene in the X inactivation center

Ftx (*Five-prime-to-Xist*) was discovered in the framework of a comparative sequence analysis of the *X inactivation center* region in mouse, human and bovine (Chureau et al., 2002). An *in silico* approach was used to pinpoint conserved functional elements, and resulted in the identification of three novel genes mapping in the vicinity of *Xist*: *Cnbp2*, *Ppnx* and *Ftx*. *Cnbp2* and *Ppnx* encode putative peptides, while *Ftx* lacks protein-coding potential and produces a long non-coding RNA. The existence of the *Ftx* gene was hypothesized following detection of bidirectional trancription events across a heterochromatic hotspot (called B region) located within the *Xic*, a hundred kilobases upstream the *Xist* locus (Heard et al., 2001). The B region is marked by H3K9 dimethylation and H3K27 thrimethylation already before the spreading of repressive chromatin marks at the onset of XCI. Because of these features, the hotspot was hypothesized to represent a “nucleation center” for the propagation of silecing marks upon *Xist* accumulation. Within the hotspot, a stretch enriched in H3K4 thrimethylation was found to correspond to the *Ftx* promoter region (Figure 19).

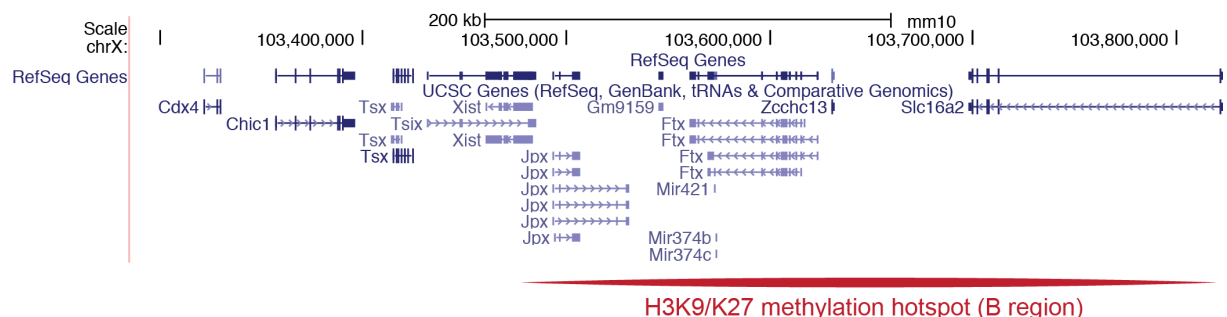


Figure 19. *Ftx* and the murine *Xic*

Localization of *Ftx* inside the murine *Xic* region. The location of the heterochromatin hotspot is indicated in red.

6.1 *Ftx* structure and conservation across eutherians

In the mouse, the *Ftx* genomic locus spans 62.8 kb of the X chromosome and maps 140 kb upstream to *Xist*, in the same orientation (Chureau et al., 2002, 2011). The gene contains fifteen exons lacking coding potential, all of them possessing perfect consensus splice sites. Exons 1, 3 and 4 are marked by CpG islands, hence possible candidates for alternative TSS positions (promoters P1, P3 and P4 respectively) (Chureau et al., 2002). Another putative promoter, P2, is located just upstream exon 2. The resulting transcript can be alternatively spliced and polyadenylated, giving rise to several isoforms, each containing a different combination of exons (Figure 20) (section 6.2).

Among all the genes of the *Xic*, *Ftx* is the sole to host a microRNA cluster (miR-374 and miR-421) in its locus, encoded inside *Ftx* intron 12 (Miska et al., 2004; Suh et al., 2004). Hosting a microRNA sequence is not a rare occurrence for non-coding genes, since insertion of miRNA-encoding mobile elements is often at the basis of lncRNA gene emergence. In this particular case, miR-374 and miR-421 arose from the ancient L2 family of retrotransposons.

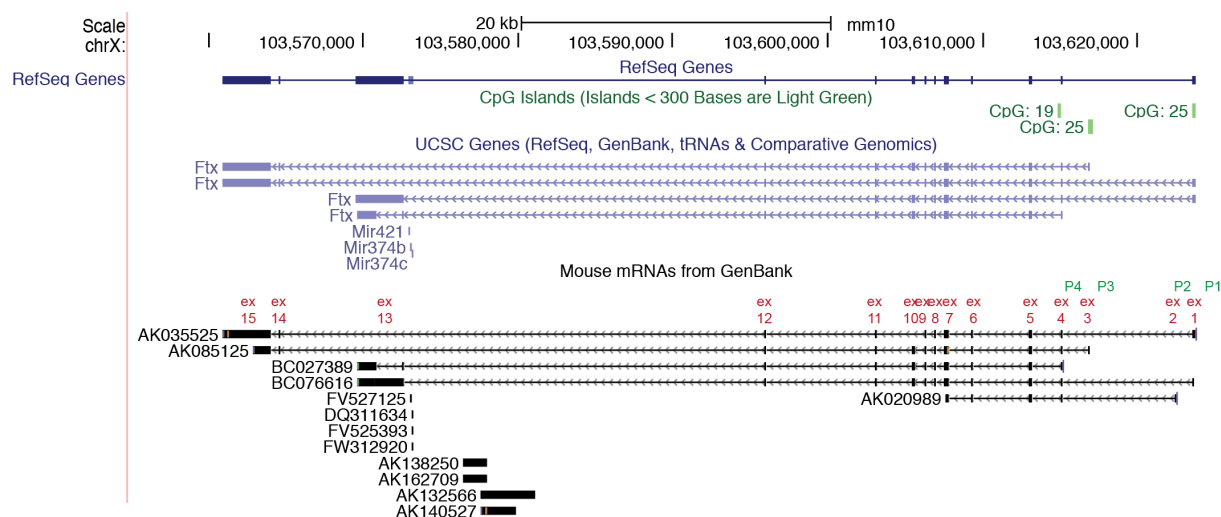


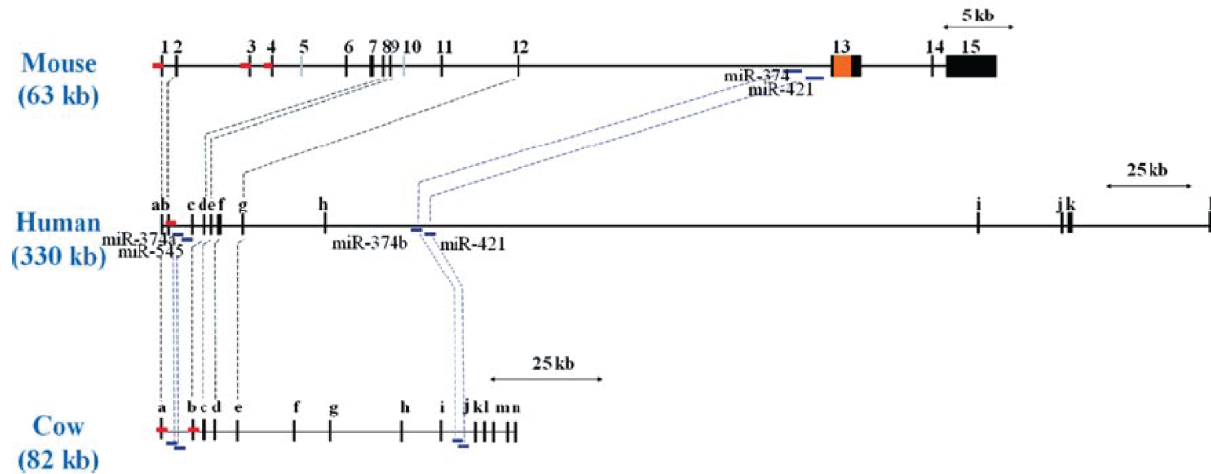
Figure 20. *Ftx* locus and transcript isoforms

UCSC depiction of the *Ftx* locus and the annotated isoforms. Exon numbering is indicated in red and positions of the putative promoters (P1, P2, P3, P4) are shown. Exons 1,3 and 4 are marked by CpG islands (green).

Like most of the genes mapping to the *Xic*, *Ftx* emerged as non-coding gene from a protein-coding gene, *Wave4* (*Wasf3*), exclusively in the eutherian lineage. Within

extant eutherian species, the *Ftx* locus is rather conserved, with the most conserved genomic segments being located in *Ftx* 5' region. The conservation interests the sequence, but does not include the exon/intron structure (Figure 21) (Chureau et al., 2011): in fact, only two exons (8 and 9 in the mouse annotation) are broadly conserved. Interestingly, *in silico* modelling of the RNA corresponding to these exons revealed that a conserved and stable secondary structure might exist across species (Figure 21), raising the hypothesis that this particular region could mediate the function of the *Ftx* transcript, for example by binding specific proteins or by interacting with other RNAs or DNA regions. Interestingly, like *Ftx* exons 8 and 9, the microRNA cluster is conserved across eutherian species and a second microRNA cluster (miR-374a and miR-545), possibly arisen by a duplication event at the miR-374/miR-421 cluster, is present in the second intron of *Ftx* in human and cow (but absent in rodents following pseudogenization) (Chureau et al., 2011) (Figure 21). The microRNA cluster seems to have appeared at the same time as *Ftx*, highlighting the possible co-evolution between the *Ftx* transcript and the microRNA elements.

A



B

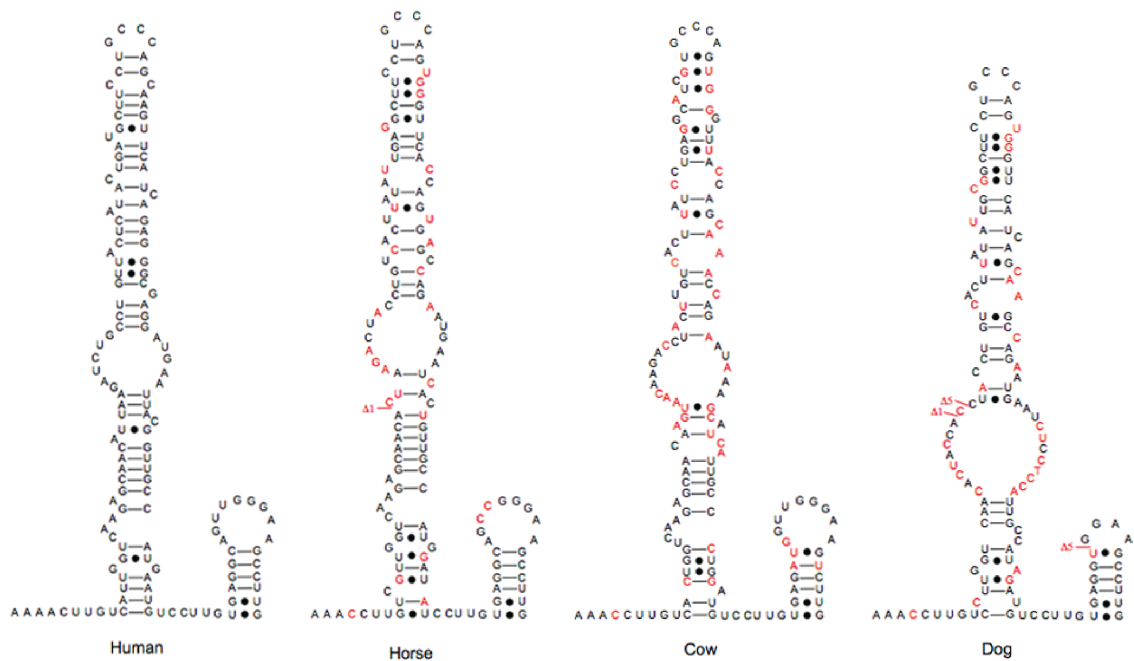


Figure 21. Conservation of *Ftx* across species

A. The genomic locus of *Ftx* is conserved across mammals, as shown by sequence alignments in representative species. The microRNA cluster is also conserved, and an additional cluster is present in the 5' region of the gene in human and cow. Similarity between exons is indicated by dotted lines. Image taken from (Chureau et al., 2011). B. *In silico* prediction of the secondary structure of the RNA transcript segment corresponding to the two best conserved exons (8 and 9 in the mouse) across species. This modelling was done by Sylvain Maenner.

6.2 *Ftx* cellular localization and expression patterns

In silico studies of the murine *Ftx* locus revealed that no similarities can be found between the genomic sequence and known peptides. Moreover, the longest open reading frame contains no more than 87 codons and it is not conserved across species, suggesting that the *Ftx* transcript is not translated. This conclusion is consistent with the fact that *Ftx* is located in a hotbed of lncRNA-producing genes (the *Xic*) and supported by evidence indicating that the transcript is mostly retained in the nucleus (Figure 22). Like the strictly nuclear *Xist* transcript, the *Ftx* transcript is enriched in the nuclear fraction of undifferentiated ES cells when compared to cytoplasmic RNAs like *Gapdh*. Unlike *Xist* molecules however, some *Ftx* transcripts can be found in small amounts also in the cytoplasm, indicating that the RNA can be exported. It remains to be tested if these proportions change in other cell types or during differentiation.

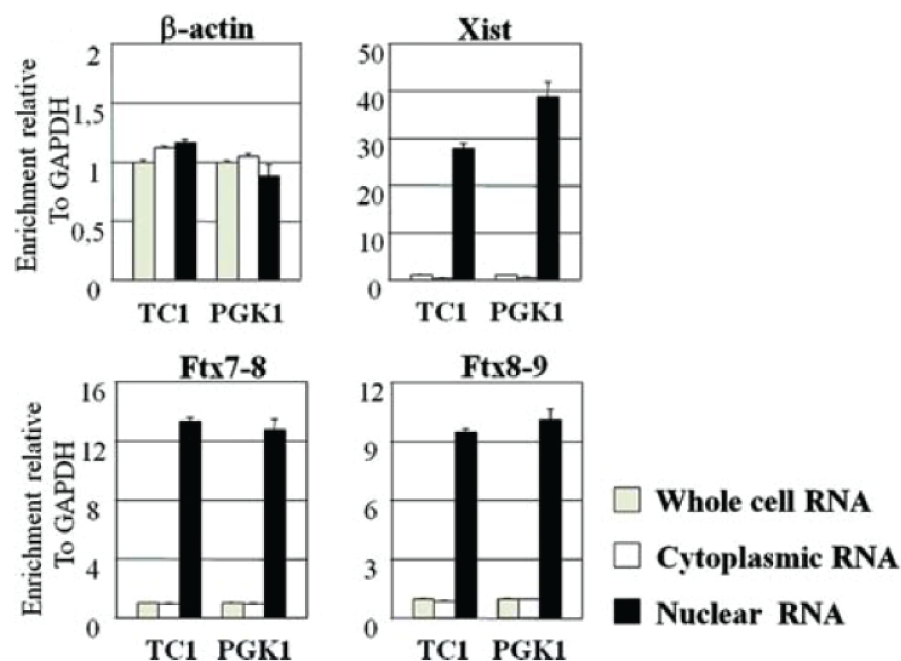


Figure 22. *Ftx* subcellular localization in undifferentiated mESCs

RNA from undifferentiated mouse ES cells was fractionated. TC1: male mESCs. PGK1: female mESCs. *Ftx* 7-8, *Ftx* 9-9: primer pairs overlapping exons 7 and 8, and exons 8 and 9 of *Ftx*, respectively. Values are normalized to *Gapdh* and whole-cell RNA is set to 1. Image taken from (Chureau et al., 2011).

According to published expressed sequence tags (ESTs), the *Ftx* gene generates a diversity of RNA isoforms in multiple tissues, produced through alternative promoter

usage, splicing and termination. Transcripts originate at exons 1, 2, 3 or 4 (Figure 23). Exons 5, 13 and 15 display more than one splice form, and exons 5 and 10 are skipped in some isoforms. Termination takes place after exon 12, which is either spliced to exon 13 or to exons 14 and 15. Consistent with ESTs data, a northern blot analysis using a probe hybridizing to *Ftx* exon 4 (common among all isoforms) showed two major bands of 2.3 and 4.2 kb, corresponding to group of isoforms terminating with the short versions of exon 13 and 15 or with their long counterparts, respectively (Figure 23) (Chureau et al., 2011). The size difference between isoforms within the two groups is too small to be detected by northern blot, and not all the isoforms can therefore be discriminated. This complex exon/intron structure is seldom seen in lncRNA genes, which globally display a strong bias towards two-exon products (Derrien et al., 2012); this perhaps indicates that the *Ftx* transcript represents more than just “transcriptional noise” and can have one or more functions. If this is true, careful investigation of *Ftx* expression profiles should give an indication on the context and modality of *Ftx* requirement.

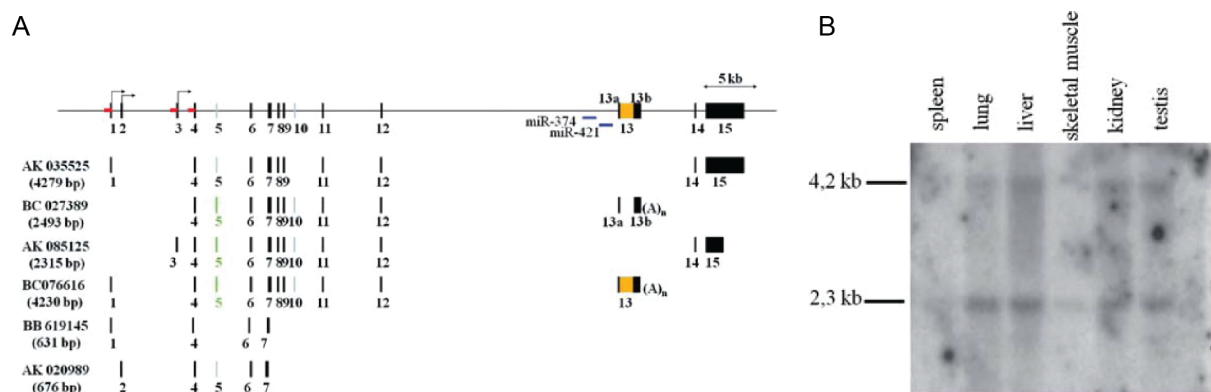


Figure 23. The *Ftx* locus generates multiple RNA isoforms

A. The *Ftx* locus and the corresponding published ESTs. Exon 5 is represented in green when a slightly longer form is present. B. Northern blot analysis of the *Ftx* transcript, using a probe against *Ftx* exon 4. RNA was extracted from different adult mouse organs. Two bands corresponding to the groups of isoforms with short versions of exons 13 and 15 (2.3 kb) and of isoforms with whole exons 13 and 15 (4.2 kb) can be seen. Image taken from (Chureau et al., 2011).

6.3 *Ftx* expression profiles

Ftx is expressed ubiquitously in adult mice tissues, where it is found at especially high levels in the brain and in some cell types of the hematopoietic system (Figure 24). Interestingly, dynamic *Ftx* expression is observed upon female ES cell differentiation at the onset of XCI (Chureau et al., 2011) (Figure 25). *Ftx* partially escapes X inactivation both in randomly and imprintedly inactivated cell lineages, but it is expressed at lower levels from the inactive X (Chureau et al., 2011; Kunath et al., 2005; Mak et al., 2002). This interesting expression profile suggests that *Ftx* might be implicated in multiple processes, among which, importantly, the regulation of XCI.

RT-qPCR analyses performed in the laboratory showed that the expression of the two *Ftx*-embedded miRNAs does not always systematically follow that of *Ftx* transcript. While, in undifferentiated ES cells, the expression of both the miRNAs is, as well as *Ftx* expression, very low, variable expression levels are observed compared to *Ftx* and remarkably also within the cluster itself during early cell differentiation. In particular, while *Ftx* and miR-374 expression levels increase sharply, miR-421 expression level remains almost steady (Figure 26). Differences are observed also in adult tissues: for example, *Ftx* and miR-421 expression levels are high in brain cells, where miR-374 expression is almost undetectable (not shown).

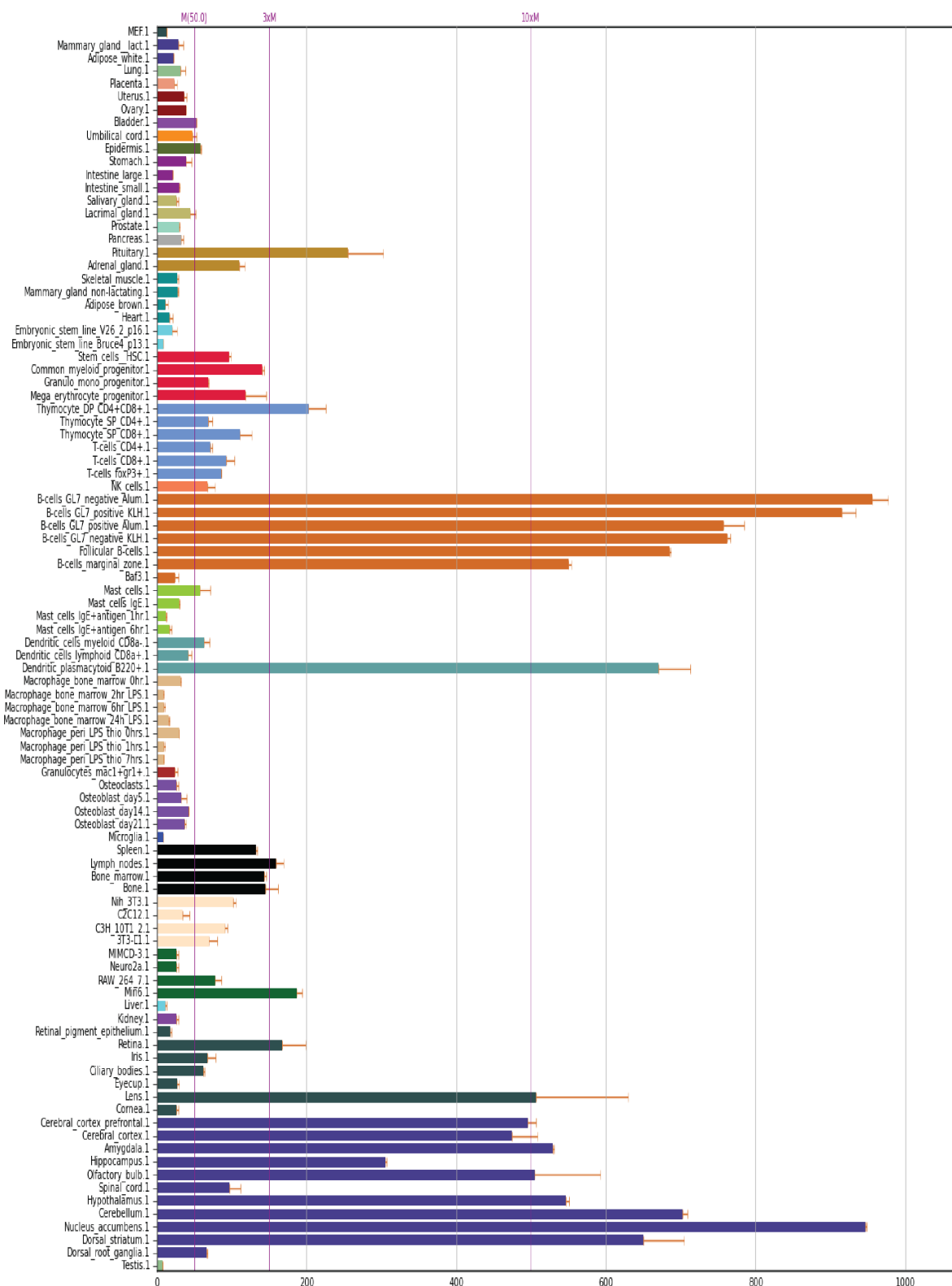


Figure 24. *Ftx* expression profiles in different tissues

Ftx expression profiles in different mouse tissues. Data from the BioGPS portal.

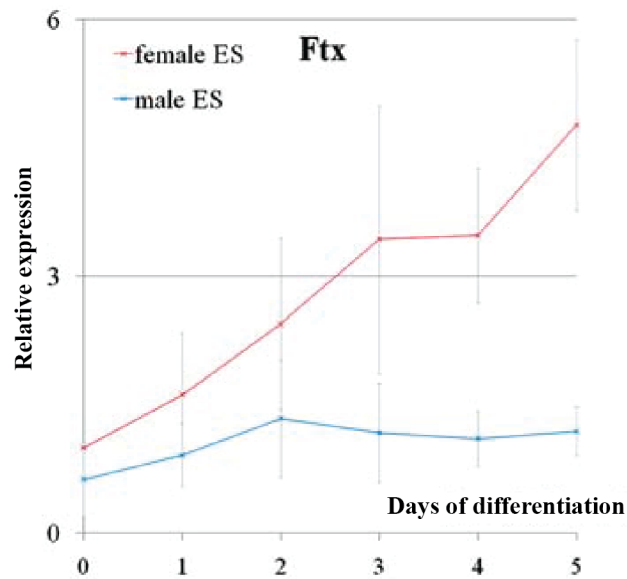


Figure 25. *Ftx* expression profiles in differentiating mESCs

RT-qPCR analyses of *Ftx* expression profiles in male and female differentiating mouse ES cells.

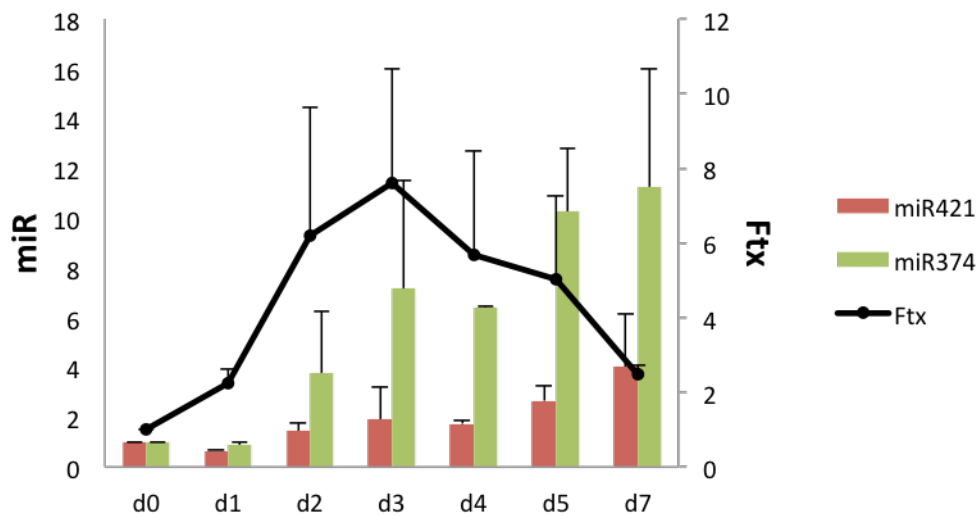


Figure 26. Expression profiles of *Ftx* and the miR cluster in differentiating female ESCs

RT-qPCR quantification of the expression levels of *Ftx*, miR-374 and miR-421 in differentiating female mouse ES cells.

6.4 *Ftx* and the hotspot

To evaluate the effects of *Ftx* in the maintenance of the hotspot, the *Ftx* promoter region was deleted in male ES cells. Loss of *Ftx* expression in these cells results in local alteration of the chromatin structure in the *Ftx* genomic region in *cis*, but intact chromatin structure is maintained inside and outside the hotspot (Figure 27) (Chureau et al., 2011). Strikingly, *Ftx* deletion leads to abnormal activity along the hotspot region, normally characterized by complex sense and antisense transcription. In particular, a gradient of downregulation extends from *Ftx* to *Tsix* and, importantly, *Xist* expression is significantly decreased. Accordingly, DNA methylation at a CpG island of *Xist* exon1 is increased, and H3K2me2 at the *Xist* promoter is decreased (Chureau et al., 2011). Although these results argue in favor of a role for *Ftx* in *Xist* activation, definitive experiments in female cells have not been published yet, and are the subject of this study (section RESULTS).

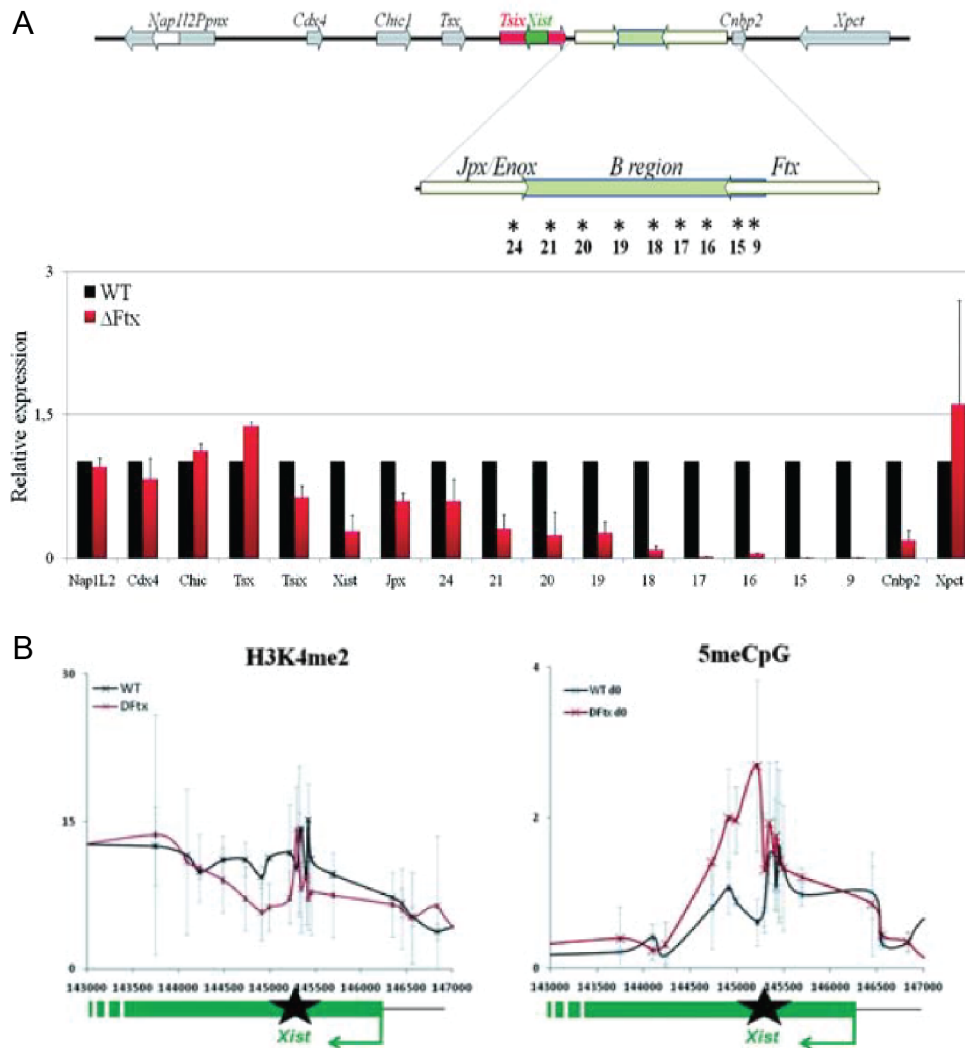


Figure 27. *Ftx* impacts on *Xist* expression levels and chromatin structure at the *Xist* locus

A. RT-qPCR analysis of Xic-linked gene expression in WT (black) and *Ftx* mutant (red) cells. Values are normalized to *Arpo* expression levels and value of WT is set to 1. B. Left panel: ChIP analysis of H3K4me2 in WT (black line) and *Ftx* mutant (red line) male ES cells ($n = 2$; for *Ftx* mutants, results obtained for the three clones were pooled). Graphs show the percentage of immunoprecipitation (%IP) obtained after normalization to the input. All values are mean+SD. The average %IP calculated for each position was plotted against the genomic location, corresponding to the *Xist* 5' region. The green box represents the start of *Xist* exon 1 and the black asterisks symbolizes CpG island. Note that only the first 3.2 kb of *Xist* exon 1 is shown. Right panel: DNA methylation analysis by MeDIP of the *Xist* 5' region. Images taken from (Chureau et al., 2011).

RESULTS

1 *Ftx* deletion strategy

While male cells contain one copy of the physical *cis*-environment needed for XCI and are thus suitable for the investigation of the *cis*-effect of X-linked genes, they do not possess the complicate network of X-encoded dose-dependent factors that ultimately contribute to coordinate XCI establishment. As a result, male cells do not upregulate *Xist* during differentiation, which makes it difficult to explore the *Xist* regulatory landscape, especially in the case of *Xist* activators. Because of this, we decided to delete *Ftx* in female cells and investigate its effects on *Xist* upregulation and XCI initiation. In order to completely abrogate *Ftx* transcription, we targeted the region encompassing all the putative promoters, which extends across 9.3 kb and include *Ftx* exons 1 to 4. We also paid attention to include in the targeted region the binding sites for the core pluripotency factors, NANOG, OCT4 and SOX2, as well as RNA Pol II and p300 fixation sites.

2 Generation of *Ftx* mutant female ES cell lines: LF2

In order to evaluate the contribution of *Ftx* to XCI dynamics in females, we set out to generate mutant mouse ES cell lines devoid of the region containing the four *Ftx* putative promoters (P1, P2 and P3). To do so, we selected a female *Mus musculus*, strain 129/Ola, ES cell line (LF2), known to properly establish XCI upon differentiation (Keohane et al., 1996b; Mermoud et al., 1999; Nichols et al., 1990), and we used the CRISPR/Cas technology (section METHODS) to delete the ~15 kb genomic sequence starting upstream *Ftx* P1 and ending after *Ftx* exon 4 (Figure 1A). We designed several combinations of sgRNAs pairs targeting two separate sites at the extremities of the region to be deleted and we tested them. Based on the intensity of a PCR-generated gel band, we selected the combination of guides that induced the two double strand breaks (DSBs) at the highest efficiency (Figure S1A and Figure S1B). We assessed for the presence of targeting events first in the population of transfected cells and subsequently in a series of independent clones. Out of the 282 clones analyzed, 34 were positive for the PCR assay detecting the

deletion, with 25 of them being heterozygous (wild-type, WT and knock-out, KO bands) and 9 being homozygous (KO band exclusively) for the deletion.

Mouse ES cell lines are known to display chromosomal abnormalities in culture, including chromosome duplications (Rebuzzini et al., 2008; Sugawara et al., 2006) and the loss of one X chromosome in females and the single X chromosome in males (Gaztelumendi and Nogués, 2014). To verify that our clones still retained both X chromosomes, we performed *Xist* DNA-FISH and quantified the percentage of cells showing one, two or more signals in each clonal population (not shown). Out of the 9 homozygous clones, 4 displayed either a high percentage of tetraploid cells (> 5%) or a high percentage of cells with a single X chromosome (> 5%), and were thus discarded, and 2 clones had almost 100% of cells with two X chromosomes. The additional 3 clones did not survive freezing and thawing. Out of the 25 heterozygous clones, 6 displayed either a high percentage of tetraploid cells (> 5%) or a high percentage of cells with a single X chromosome (> 5%), and were thus discarded, and at least 5 had almost 100% of cells with two X chromosomes. The remaining clones were not tested. To adapt to the number of available homozygous clones (2), we decided to focus on only 2 heterozygous clones as well.

Since, in order to obtain independent clones, our strategy required that the cells were plated at low density directly after transfection (Figure S1A), we could not exclude that the clones defined as heterozygous by the PCR assay were in fact mixed populations of *Ftx* deleted and *Ftx* WT cells, depending on the delay of the CRISPR/Cas9-dependent DSB induction (before or after the first mitosis upon plating). This scenario was not an issue for the homozygous clones, owing to the fact that they were selected on the basis of the absence of a WT band (and presence of a KO band) in the PCR assay.

To estimate the homogeneity of the clonal populations with heterozygous deletion, we performed *Tsix/Ftx* RNA-FISH at the undifferentiated state and observed that a significative proportion of cells showed a double *Ftx* pinpoint, co-localizing with both *Tsix* signals and thus with both X chromosomes (Figure S1C). Although strongly suggesting that the clones were heterogeneous, with *Ftx* being transcribed from both X chromosomes, this result could also indicate the ability of the *Ftx* transcript to move from its transcription site on one (WT) X chromosome to the *Xic* on the other (*Ftx* deleted) X chromosome, especially if *Ftx* acted, like the *Xist*-activating *Xic*-linked

lncRNA *Jpx* (Sun et al., 2013), in *trans*. The interpretation of these results was further complicated by the presence of *Ftx*-specific scattered signal in the totality of the nuclear space, indeed possibly indicating the ability of the *Ftx* transcript to move away from its transcription site (Figure S1C).

To circumvent this problem, we performed a real-time quantitative PCR (RT-qPCR) analysis that measured the genomic DNA (gDNA) amount of the region supposed to be deleted and the gDNA amount of a region outside the targeted sites. We reasoned that, in homogeneous heterozygous populations, the gDNA inside the deleted region would be half the amount of the gDNA outside the deleted region and, conversely, in heterogeneous heterozygous populations, the difference in amount would be smaller, to an extent depending on the percentage of heterogeneity. The results indicated that, while the homozygous clones, as expected, were pure populations, the heterozygous clones were not, with the proportion of heterogeneity estimated at 30%*Ftx*^{+/-}/70%WT. for both populations (Figure S1D). This result does not exclude, *per se*, that *Ftx* can act in *trans*, since the presence of scattered *Ftx* signals is still detectable by RNA-FISH in pure populations (following sections).

To obtain pure heterozygous populations while maintaining the independent origin of the clones, we subcloned the two initial heterogeneous populations. We analyzed 48 subclones from each of them, and, confirming our estimations, found that 37.5% (first population) or 26% (second population) of the subclones were real heterozygous, while the remaining subclones were wild-type, according to the PCR assay (not shown). Neither heterogeneous population was a mixture of WT and homozygous cells. We next verified the homogeneity of 20 subclones (10 for each initial heterogeneous population) by RT-qPCR on gDNA, and concluded that we were able to obtain pure subclonal populations (Figure S1E).

For further analysis, we finally selected a total of 4 independent clones, of which 2 were heterozygous (*Ftx*^{+/-} A and *Ftx*^{+/-} B) and 2 homozygous (*Ftx*^{-/-} A and *Ftx*^{-/-} B) for the deletion (Figure 1B and Figure 1E). To make sure that the final clones had two X chromosomes, we performed *Xist* DNA-FISH (Figure 1D). In order to locate the exact borders of the deletion and to verify the integrity of the homologous region on the WT X chromosome (in the heterozygous clones), we sequenced the PCR products encompassing the targeted sites. We noticed that one of the two sgRNAs (at the 3' junction) could direct the DSB in two different positions, resulting in deletions of

slightly different sizes, while the second (at the 5' junction) induced a DSB reproducibly at the same location (Figure 1C). In both heterozygous clones, the targeted regions on the WT *Ftx* allele were intact, except for a few missing nucleotides in the *Ftx*^{+/-} B clone at the 5' junction and for both *Ftx*^{+/-} A and *Ftx*^{+/-} B clones at the 3' junction (Figure 1C). These small deletions are however unlikely to perturb *Ftx* transcription or splicing, as they reside away from the TSSs and outside the exons.

Figure 1

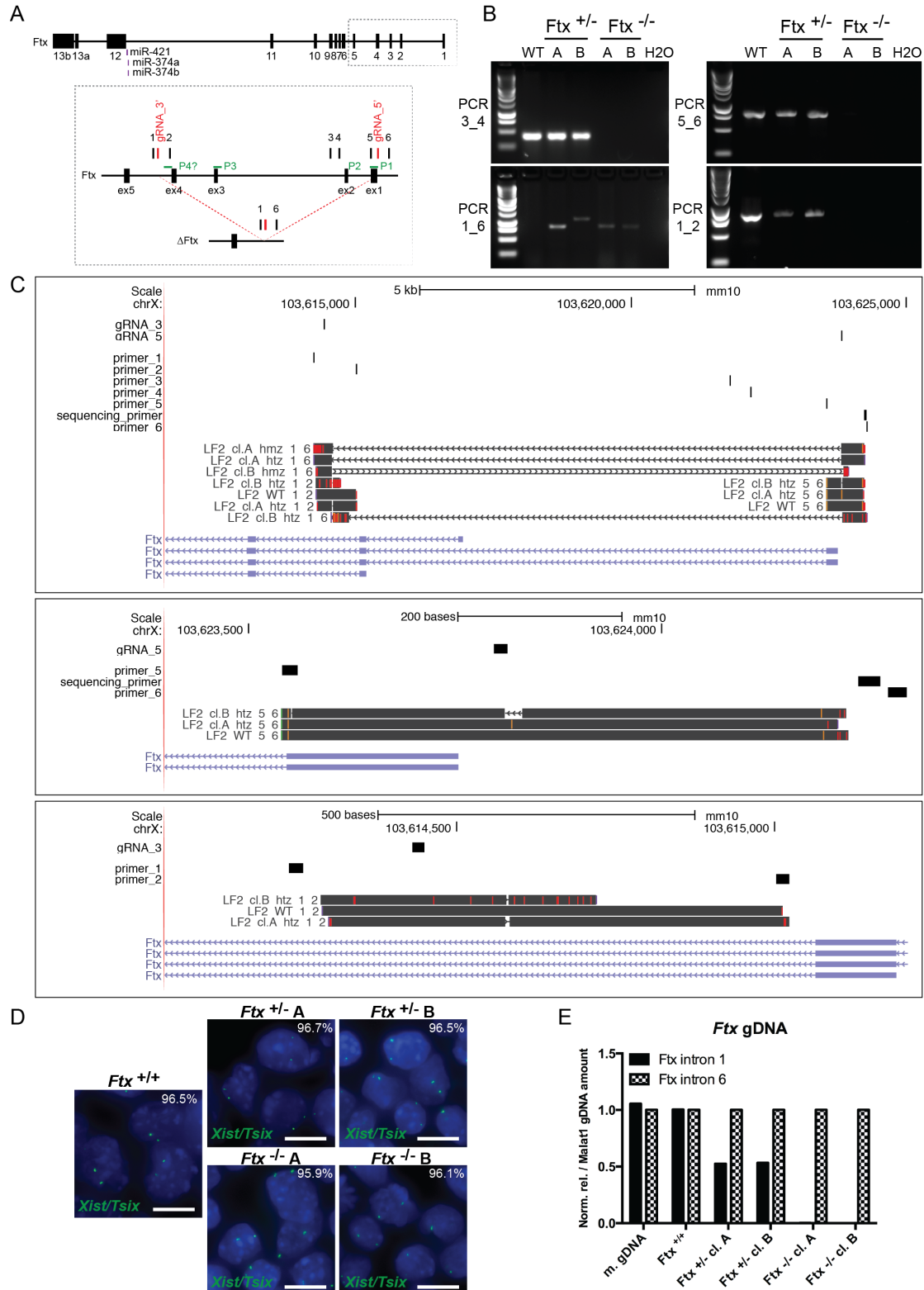


Figure 1. Generation of LF2 *Ftx* mutant cell lines

A. Schematic representation of the *Ftx* gene, including its 5' portion, with its putative promoters (P1, P2, P3 and P4) and CpG islands (horizontal green lines), the first five exons (black boxes) and four introns (black lines), in scale. Guide RNAs locations (vertical red lines) and PCR primers positions (vertical black lines) are shown.

B. Left panel: PCR assay to detect the wild-type allele (top) and deletion event (bottom) in WT cells and in the selected *Ftx* mutant clones. Right panel: PCR assay to assess the integrity of the 3' (top) and 5' (bottom) targeted regions on the WT *Ftx* allele in the selected *Ftx* mutant clones.

C. UCSC genome browser alignments of the sequencing results on the PCR products from B.

D. Representative *Xist/Tsix* DNA-FISH images in WT cells and in the selected *Ftx* mutant clones.

E. RT-qPCR quantification of the genomic DNA amount inside (*Ftx* intron 1) and outside (*Ftx* intron 6) the deleted region in mouse gDNA (control), WT and selected *Ftx* mutant clones. Values are normalized to *Ftx* intron 6 gDNA amount.

Figure S1

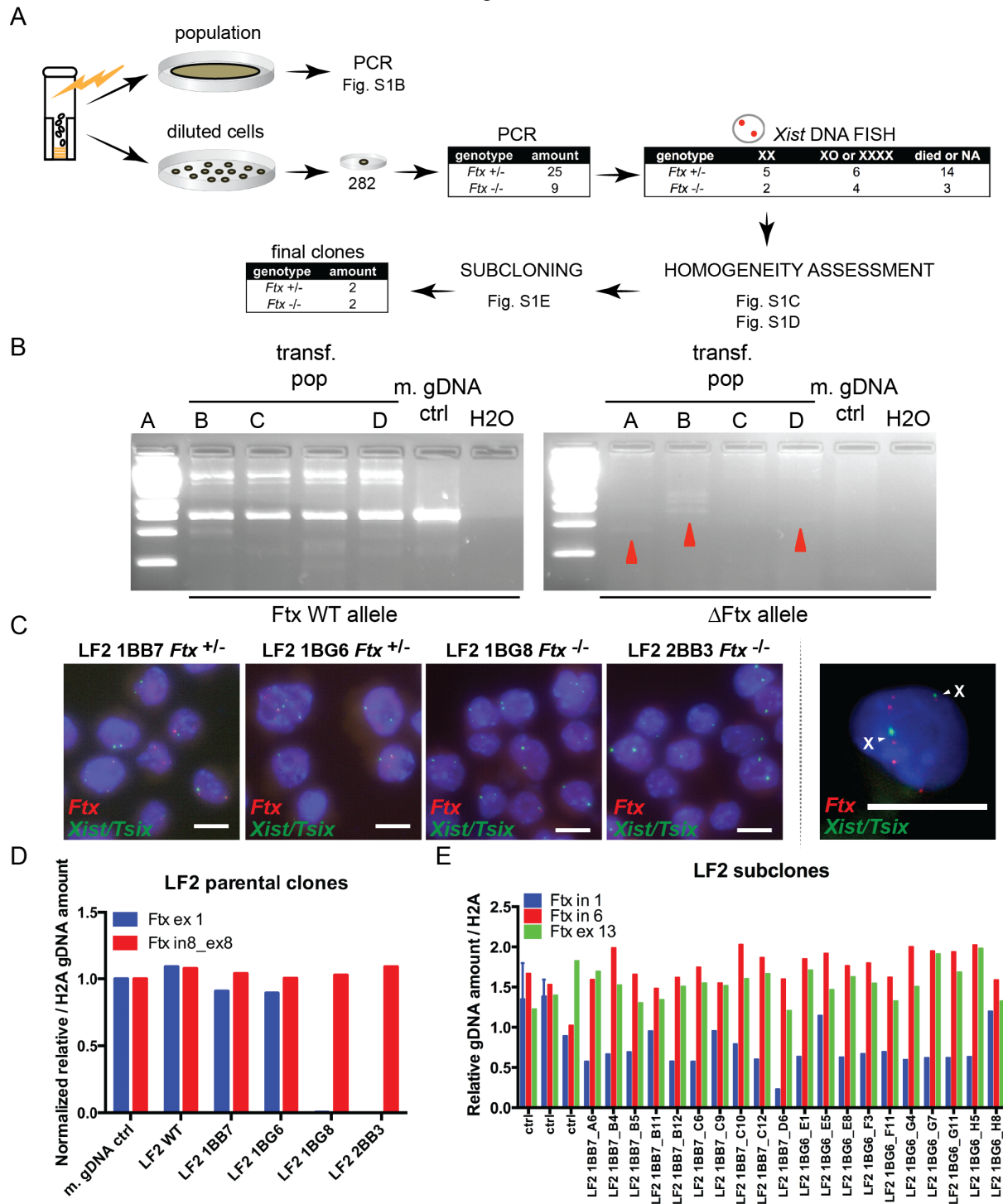


Figure S 1. Generation of LF2 *Ftx* mutant cell lines

A. Schematic representation of the procedure for the generation of LF2 *Ftx* mutant cell lines. After transfection, the cells were either plated at high density and checked for the presence of the deletion event at the population level, or directly seeded at low density to obtain independent clones. After ten days, clones were picked and individually screened for the presence of the targeting event. Positive clones were further checked for the presence of two X chromosomes and assessed for homogeneity. Non-homogeneous clonal population were sub-cloned.

B. PCR assay to detect the wild-type allele (left) and deletion event (right) in populations transfected with different sgRNAs combinations (A, B, C and D). Red arrowheads point towards the bands indicating the deletion event. Final clones were obtained from the combination B.

C. Left panel: Representative *Xist/Tsix-Ftx* RNA-FISH images of the parental clones. Scale bar = 10 μ m. Right panel: Representative *Xist/Tsix-Ftx* RNA-FISH image of cells with scattered *Ftx* signal. Scale bar = 10 μ m.

D. RT-qPCR quantification of the genomic DNA amount inside (*Ftx* intron 1) and outside (*Ftx* intron 8-exon 8) the deleted region in the mutant clonal populations before sub-cloning. Values are normalized to mouse gDNA amounts (control).

E. RT-qPCR quantification of the genomic DNA amount inside (*Ftx* intron 1) and outside (*Ftx* intron 6 and *Ftx* exon 13) the deleted region in two groups of independent clones after sub-cloning. Values are normalized to H2A amounts.

3 *Ftx* mutant cell lines display female-specific differentiation defects

Once the *Ftx* mutant clones were obtained and characterized, we proceeded with the investigation of *Ftx* expression profiles and differentiation dynamics of female and male (Chureau et al., 2011) mutant ES cells.

Visualization of *Ftx* expression at the single-cell level by RNA-FISH showed that, in undifferentiated wild-type female cells, *Ftx* is expressed in the majority (~90%) of the cells (Figure 2A), either biallelically or monoallelically (Figure 2B). In *Ftx*^{+/-} cells these proportions change, with the majority (~90%) of the cells still expressing *Ftx*, but without cells showing biallelic expression (consistent with the heterozygous deletion), and an increased number of cells with monoallelic expression and with the lack of *Ftx* expression (Figure 2B). In *Ftx*^{-/-} undifferentiated cells, *Ftx* is not expressed, indicating that deletion of the selected region is sufficient to completely abrogate *Ftx* transcription (Figure 2A and Figure 2B). These proportions are maintained also during differentiation (Figure 5A). In undifferentiated wild-type male cells *Ftx* is expressed only in 50% of the cells (Figure 2A and Figure 2B). Other than on the X chromosomes, *Ftx* RNA can be found, in both female and male cells, in the nuclear space, where it occupies multiple discrete positions in a high proportion of cells (Figure S2A). The signal appears to be *Ftx*-specific, as it is not present in *Ftx*^{-/-} or *Ftx*^Y cells. Whether these foci correspond to particular chromatin structures or are located inside specific subnuclear domains is currently unknown.

Time-course measurements of *Ftx* expression during monolayer undirected differentiation showed that, in both wild-type male and female differentiating ES cells, *Ftx* expression is dynamic (Figure 2C). Since the cell lines used do not share the same genetic background (female 129/Ola and male TC1 129/Sv lines) and they are not supported by the same culture conditions (feeder-dependent or feeder-independent, section METHODS), *Ftx* expression in males and females cannot be compared with regard to the level, which could be dependent on strain- and culture-specific features. Despite this, the profile of *Ftx* expression is similar between the two sexes, with an increase in the expression levels during the first days of differentiation. Contrasting with previous findings published by the laboratory (Chureau et al., 2011), these results indicate that the increase in *Ftx* expression in differentiating ES cells is

not sex-specific. The discrepancy between the results here described and the published data could be due to the fact that, in this study, two replicates performed with the same differentiation method were averaged, while, in the published study, many replicates of differentiation time-courses performed with distinct differentiation protocols (with or without the addition of retinoic acid) were pooled. Further investigations are however required to provide a definitive answer.

In order to understand if *Ftx* is required for correct differentiation kinetics, we analyzed female and male differentiating *Ftx* mutant ES cells. As expected, when we measured the residual *Ftx* levels in *Ftx*^{+/-}, *Ftx*^{-/-} and *Ftx*^{-Y} cells by RT-qPCR, we found about 50% of the RNA in the heterozygous cells compared to wild-type cells and no residual *Ftx* expression in the homozygous or hemizygous cells along the whole course of differentiation (Figure 2C and Figure S2B).

Although all female *Ftx*^{+/+}, *Ftx*^{+/-} and *Ftx*^{-/-} cell lines were morphologically identical in the undifferentiated state, they started to diverge upon differentiation (strong difference starting from D3), with the mutant lines showing a higher number of round-shaped, ES-like colonies compared to wild-type, and a lack of isolated, sharp-edged, differentiated cells (Figure 3A and Figure S2C). This was accompanied by a visible increase in cell death in *Ftx*^{+/-} and *Ftx*^{-/-} cell lines. Nevertheless, assessment of the expression of the stem cell markers *Nanog*, *Oct4*, *Klf4* and *Rex1* showed appropriate downregulation of all of them in all female cell lines during the course of differentiation (Figure 3B). It can be noted however that the extent of downregulation of these genes (especially of *Oct4*) during our differentiation experiments is not as striking as reported in other studies, including some of those investigating XCI mechanisms. This discrepancy could be due to differences in the method of differentiation, which relies solely on low-density plating of the cells and LIF withdrawal in our case (section METHODS), but that is performed via embryoid body formation and with the addition of retinoic acid (RA), known to repress pluripotency-associated genes and activate ectoderm markers (Okada et al., 2004; Wichterle et al., 2002), in the majority of the other studies.

In any case, the fact that a high proportion of cells die in the *Ftx* mutant lines could mask a mutant-specific increase in the percentage of cells still expressing pluripotency markers during differentiation.

In contrast to *Ftx*^{+/-} and *Ftx*^{-/-} cells, male *Ftx*^{-Y} cells displayed no obvious phenotype.

Consistent with this, no massive cell death was detected ([Figure 3A](#)).

All together, these results indicate that, while the lack of *Ftx* has no substantial effects in males, deleting even only one copy of *Ftx* results in marked abnormalities in females.

Figure 2

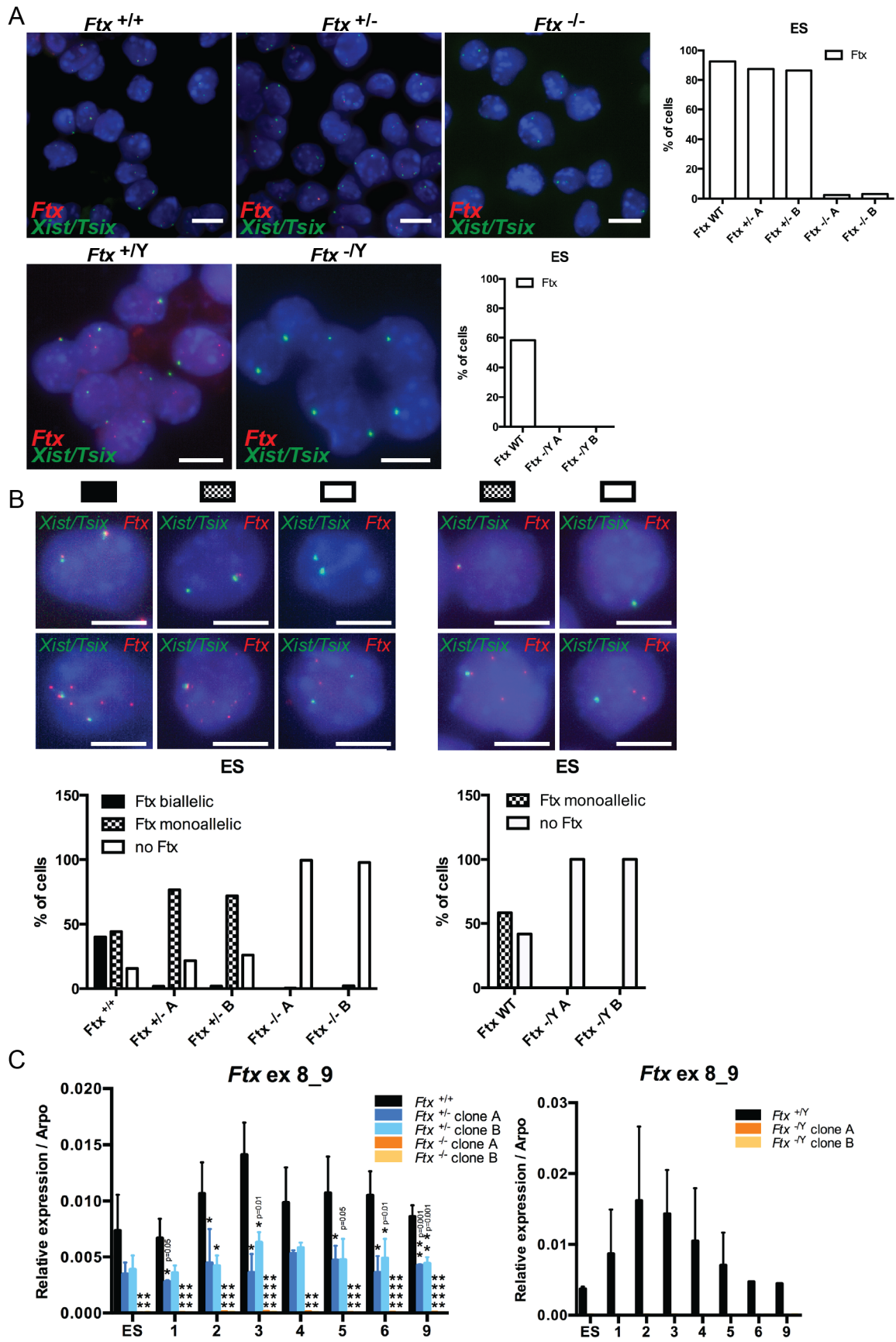


Figure 2. *Ftx* expression profiles in female and male WT and *Ftx* mutant cells

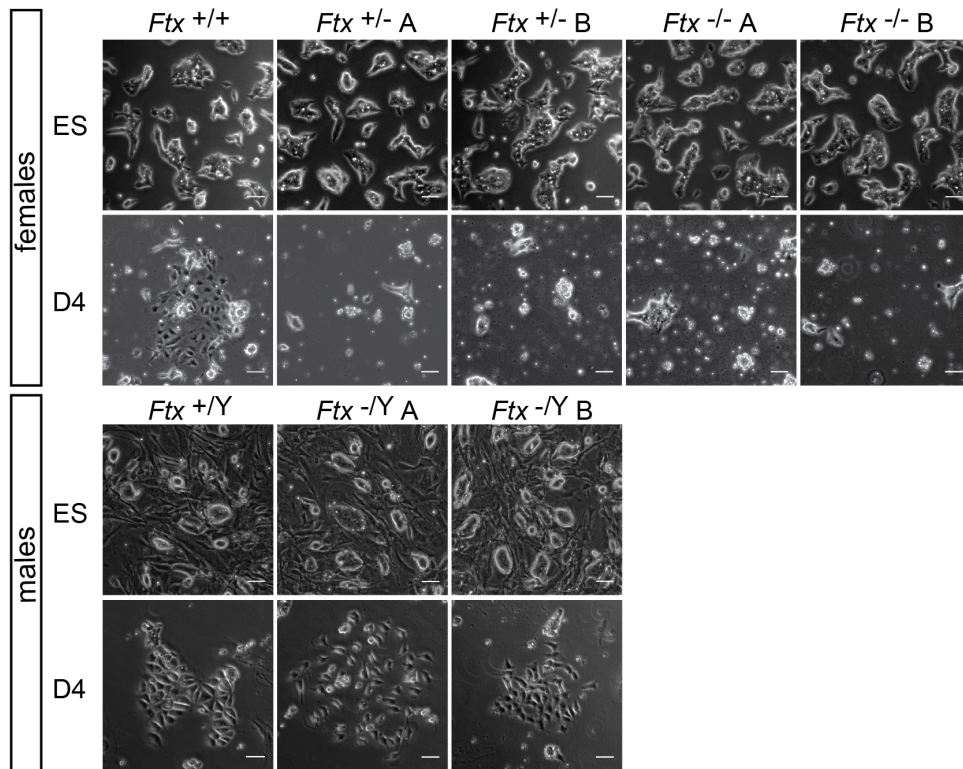
A. Top panel, left: Representative RNA-FISH images of female LF2 WT and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. *Ftx* probe, Cy3-labelled, red. Scale bar = 10 μ m. Top panel, right: Quantification of the percentage of *Ftx* expressing cells in undifferentiated (ES) female LF2 WT and *Ftx* mutant cells. Sample sizes (n) = 197; 613; 307; 243; 321. Bottom panel, left: Representative RNA-FISH images of male TC1 WT and *Ftx* mutant cells. *Xist/Tsix* probe (p510), FITC-labeled, green. *Ftx* probe, Cy3-labelled, red. Scale bar = 10 μ m. Bottom panel, right: Quantification of the percentage of *Ftx* expressing cells in undifferentiated (ES) male TC1 WT and *Ftx* mutant cells. Sample sizes (n) = 206; 100; 103.

B. Left panel, top: RNA-FISH images of each possible (*Xist/*)*Tsix/Ftx* configuration in undifferentiated (ES) female LF2 WT and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. *Ftx* probe, Cy3-labelled, red. Scale bar = 10 μ m. Left panel, bottom: Quantification of the percentage of each configuration in undifferentiated (ES) female LF2 WT and *Ftx* mutant cells. Sample sizes (n) = 197; 613; 307; 243; 321. Right panel, top: RNA-FISH images of each possible (*Xist/*)*Tsix/Ftx* configuration in undifferentiated (ES) male TC1 WT and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. *Ftx* probe, Cy3-labelled, red. Scale bar = 10 μ m. Right panel, bottom: Quantification of the percentage of each configuration in undifferentiated (ES) male TC1 WT and *Ftx* mutant cells. Sample sizes (n) = 206; 100; 103.

C. Time-course analyses of *Ftx* expression by RT-qPCR in differentiating female LF2 and male TC1 WT and *Ftx* mutant cell lines. Averages and standard deviation (S.D.) from at least three (female) or two (male) independent differentiation experiments are plotted. Values are normalized to *Arpo* RNA levels. Statistical significance was assessed with a two-tailed t-test.

Figure 3

A



B

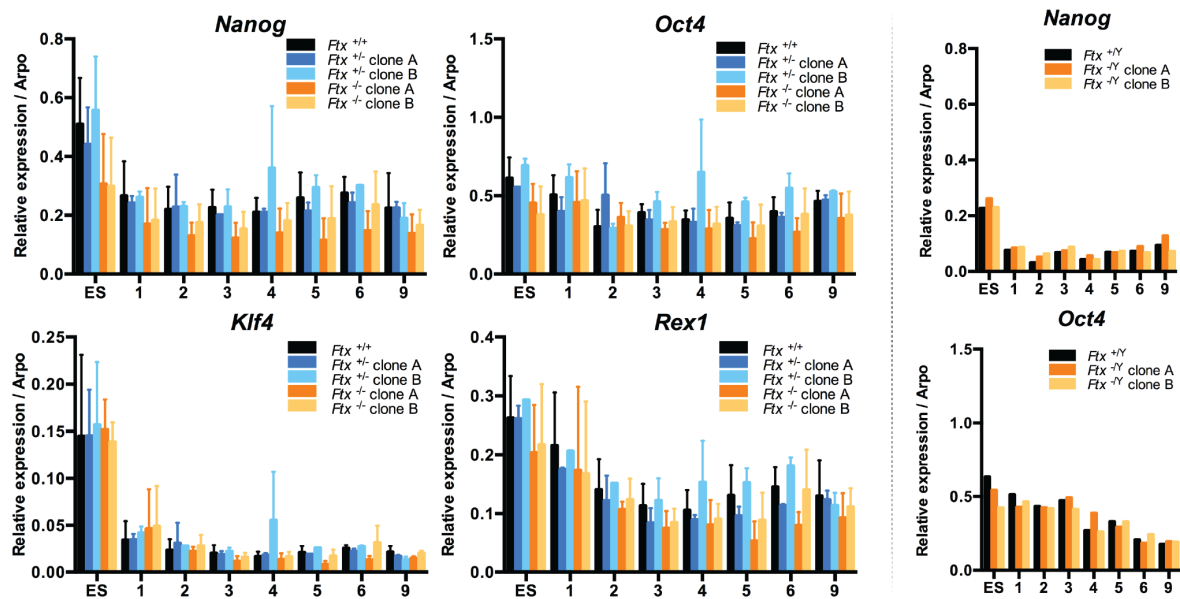


Figure 3. Female-specific *Ftx*-dependent differentiation defect

A. Brightfield photographs of undifferentiated (ES) and differentiating (D4) female LF2 and male TC1 WT and *Ftx* mutant cells. Scale bar = 50 μ m.

B. Left panel: Time-course analysis of the expression of the pluripotency markers *Nanog*, *Oct4*, *Klf4* and *Rex1* by RT-qPCR in differentiating female LF2 WT and *Ftx* mutant cell lines. Averages and standard deviation (S.D.) from at least three independent differentiation experiments are plotted. Values are normalized to *Arpo* RNA levels. Statistical significance was assessed with a two-tailed t-test. Right panel: Time-course analyses of the expression of the pluripotency markers *Nanog* and *Oct4* by RT-qPCR in differentiating male TC1 WT and *Ftx* mutant cell lines. Values from one differentiation experiment are plotted. Values are normalized to *Arpo* RNA levels.

Figure S2

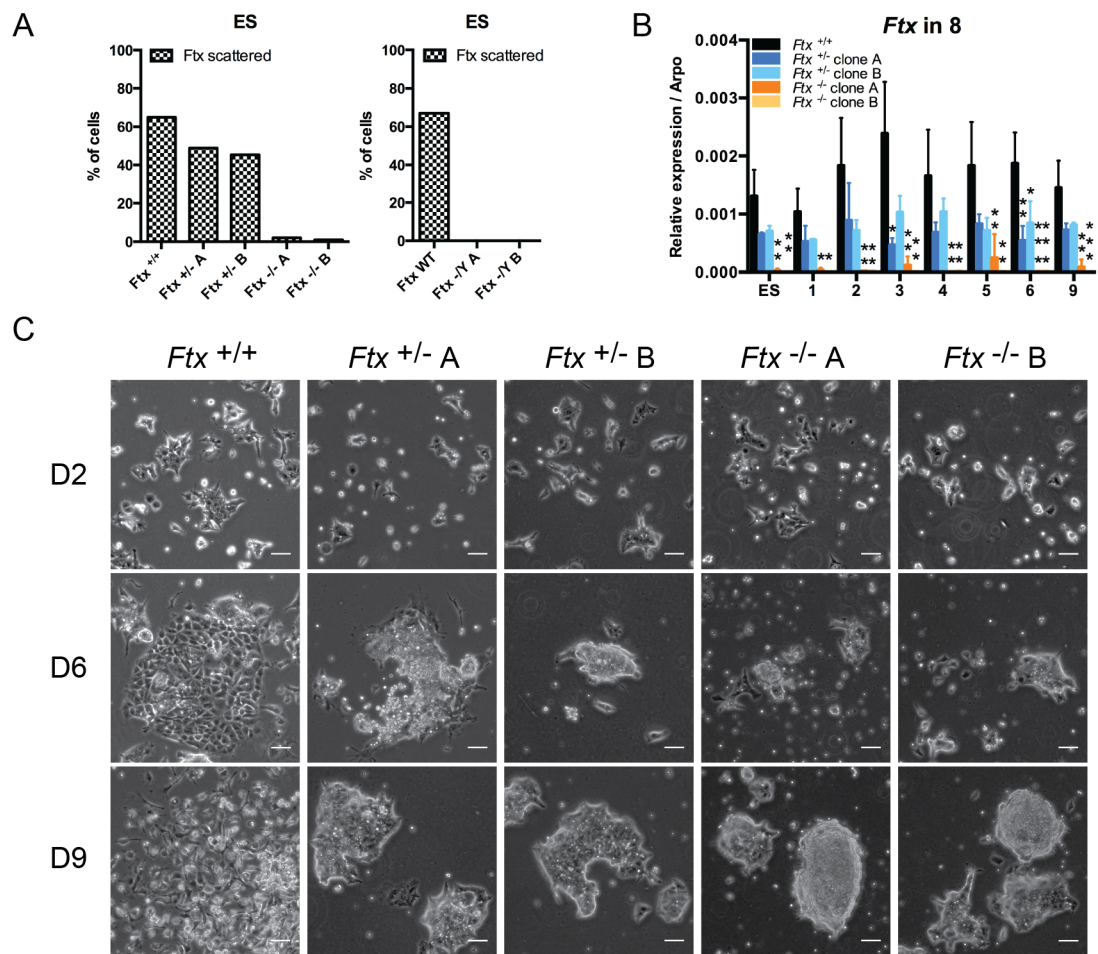


Figure S 2. *Ftx* expression profiles and female-specific *Ftx*-dependent differentiation defect

A. Left panel: Quantification of the percentage of cells showing *Ftx* scattered signal in undifferentiated (ES) female LF2 WT and *Ftx* mutant cells. Sample sizes (n) = 197; 613; 307; 243; 321. Right panel: Quantification of the percentage of cells showing *Ftx* scattered signal in undifferentiated (ES) male TC1 WT and *Ftx* mutant cells. Sample sizes (n) = 206; 100; 103.

B. Time-course analyses of *Ftx* expression by RT-qPCR in differentiating female LF2 WT and *Ftx* mutant cell lines. Averages and standard deviation (S.D.) from at least two independent differentiation experiments are plotted. Values are normalized to *Arpo* RNA levels. Statistical significance was assessed with a two-tailed t-test.

C. Brightfield photographs of differentiating (D2, D6 and D9) female LF2 and male TC1 WT and *Ftx* mutant cells. Scale bar = 50 μ m.

4 *Ftx* is required for *Xist* accumulation

The female-specific nature of the defect suggested a link to XCI, as this process is coupled to cell differentiation (Donohoe et al., 2009; Monk and Harper, 1979; Navarro et al., 2008) (section INTRODUCTION 5). Furthermore, the increase of *Ftx* expression during differentiation occurs in a window that corresponds to the onset of XCI and to *Xist* upregulation in females. We therefore measured steady-state *Xist* levels by RT-qPCR and visualized *Xist* accumulation by RNA-FISH. In *Ftx*^{+/+} cell lines, *Xist* was efficiently upregulated (Figure 4A and Figure S3A) and accumulated (Figure 4B) in ~32% of cells by day 4. In *Ftx*^{-/-} cell lines, *Xist* upregulation was severely compromised, with no major increase in *Xist* levels (Figure 4A and Figure S3A) and only 2% to 4% of cells with a *Xist* cloud by day 4 (Figure 4B). Given the low percentage of XCI even at later days, we concluded that XCI was blocked rather than just delayed, indicating that *Ftx* is required for proper *Xist* dynamics and accumulation. *Ftx*^{+/-} cell lines displayed inefficient *Xist* upregulation (Figure 4A and Figure S3A) and accumulation (Figure 4B), but the defect was not as marked as in *Ftx*^{-/-} cell lines (~16% of cells with a *Xist* cloud by day 4), indicating that *Ftx* acts in a copy-dependent manner (Figure 4A and Figure S3B).

Importantly, in both *Ftx*^{+/-} and *Ftx*^{-/-} cell lines, the actual phenotype is probably underestimated, due to the fact that dead cells detach from culture and are therefore not analyzed. Consistent with a role for *Ftx* in *Xist* activation, *Xist* was appropriately suppressed in undifferentiated and differentiating *Ftx*^{-Y} male cells (Figure S3A).

Although *Xist* can still be upregulated in a proportion of cells with the *Ftx* deletion, its accumulation pattern differs from that of wild-type cells. In both *Ftx*^{+/-} and *Ftx*^{-/-} cell lines, we noticed the presence of a significant number of cells displaying a delocalized *Xist* signal, with the percentage of "normal" *Xist* clouds dropping below 30% (on the totality of cells with *Xist* cloud) in *Ftx*^{-/-} lines (Figure 4C). In these lines, we carefully assessed the volume and the sphericity of the *Xist* clouds, and concluded that bigger but less compact *Xist* clouds were present in *Ftx*^{-/-} cells compared to wild-type (Figure 4C).

We next assessed the ability of the *Xist* clouds to efficiently silence the chromosome and found that in *Ftx*^{-/-} cells the X-linked gene *Taf1* was actively transcribed from the *Xist*-coated chromosome in a higher percentage of cells compared to the wild-type

population (Figure 4D). This indicates that the Xi may not be correctly inactivated, and suggests that the abnormal *Xist* cloud is not entirely functional.

To explore whether the *Xist* phenotype was mediated by other known *Xist* regulators, we measured the expression levels of the *Xic*-encoded genes *Tsix*, *Jpx* and *Rnf12* during differentiation (Figure 4E). The expression of *Tsix*, a negative *Xist* regulator, is known to decrease in differentiating cells, while the expression of *Jpx*, described as a *Xist* activator, has been shown to increase. In our analysis, the expression profiles of these two genes (in wild-type cells) behave as expected, but, like in the case of the pluripotency factors mentioned in the previous section, the extents of decrease (*Tsix*) and increase (*Jpx*) are not as marked. Once again, the discrepancies with other published studies can be explained by variations in the differentiation methods.

While *Tsix* and *Rnf12* did not show significant differences in their expression levels between wild-type and mutant cell lines during the whole time-course of differentiation, *Jpx* expression was decreased in both *Ftx*^{+/-} and *Ftx*^{-/-} cell lines. Because of this, we cannot exclude that *Jpx* is, at least in part, mediating the *Ftx* effect. However, since *Jpx*, as a transcript, was shown to act in *trans* (Tian et al., 2010), we would expect to see impaired *Xist* accumulation on either X chromosome in *Ftx*^{+/-} cell lines, and not just from the *Ftx*-deleted one. As for *Rnf12*, it would be important to measure the amounts of RNF12 protein in addition to the levels of mRNA, as it is the protein product that promotes *Xist* expression in differentiating cells (Barakat et al., 2011; Gontan et al., 2012; Jonkers et al., 2009).

To gain insight into the link between the *Xist* and *Ftx* transcripts, we analyzed the patterns of *Xist* and *Ftx* expression by RNA-FISH during differentiation. In differentiating *Ftx*^{+/+} cells (D4), *Ftx* is expressed in ~94% of the cells. In *Xist* positive cells, *Ftx* is present on the Xi on ~84%, with ~70% of the cells being biallelic for *Ftx* and ~14% monoallelic (Figure 5A). In the remaining ~14% of the *Xist* positive *Ftx* expressing cells, *Ftx* is found exclusively on the Xa (Figure 5A). Only in ~2% of *Xist* positive cells the *Ftx* signal is absent. These proportions are maintained along all the time course of differentiation (Figure S3D).

In differentiating *Ftx*^{+/-} cells (D4), *Ftx* is expressed in the majority of cells (~80%) and it is always monoallelic (Figure 5A). Here, we assumed that the *Ftx* pinpoint corresponded to the wild-type allele, and that the absence of pinpoint marked the *Ftx* deleted chromosome. In these lines, the majority (~80%) of *Xist* positive cells show

Xist coating on the wild-type, *Ftx*-expressing X chromosome, while only a small proportion (5% to 9%) display *Xist* coating on the mutant X chromosome. The remaining *Xist* positive cells (11% to 14%) lack *Ftx* expression. Similar to the wild-type population, these proportions are maintained also at later days (Figure S3D). Since in the LF2 cell line each X chromosome has a 50% probability of being chosen for inactivation, the skew towards a preferential inactivation of the wild-type X chromosome observed in *Ftx*^{+/-} populations suggests a role for *Ftx* in *cis*.

In *Ftx*^{-/-} populations, the presence of a low but not negligible number of *Xist* clouds in the absence of *Ftx* expression is consistent with the fact that *Xist* can indeed be upregulated from a non *Ftx* expressing X chromosome, as seen in a small proportion of cells in wild-type populations. In these non *Ftx* expressing cells, *Xist* upregulation and accumulation are probably driven by the activity of other *Xist* activators (*Rnf12*, *Jpx*, *Xist AR*, YY1 and possibly others).

We also quantified the association of *Ftx* expression (that identifies a wild-type X chromosome) with the type of *Xist* cloud (normal or abnormal) and we found that, in *Ftx*^{+/-} lines, the *Ftx*-deleted X chromosome, when chosen for inactivation, displayed almost exclusively an abnormal *Xist* cloud (Figure 5B). This was also the case in *Ftx*^{-/-} populations (Figure 5B). All together, these data indicate that, in most cases, *Xist* and *Ftx* are co-expressed, and, when *Ftx* is deleted, *Xist* cannot be properly upregulated and silence the chromosome.

Like in undifferentiated cells, *Ftx* foci were often found in the nuclear space of differentiating cells (Figure S3C).

Extended culture of ES cells has been related to epigenetic changes (Maitra et al., 2005), which can result in differences among independent subclones derived from the same original cell line (Humpherys et al., 2001). To exclude the possibility that the *Ftx*-dependend *Xist* phenotype was in fact merely due to the subcloning procedure, we analyzed two independent wild-type-derived subclones, and observed that, with respect to *Xic*-encoded and stem cell marker genes, they behaved like the parental population (Figure S3E).

Figure 4

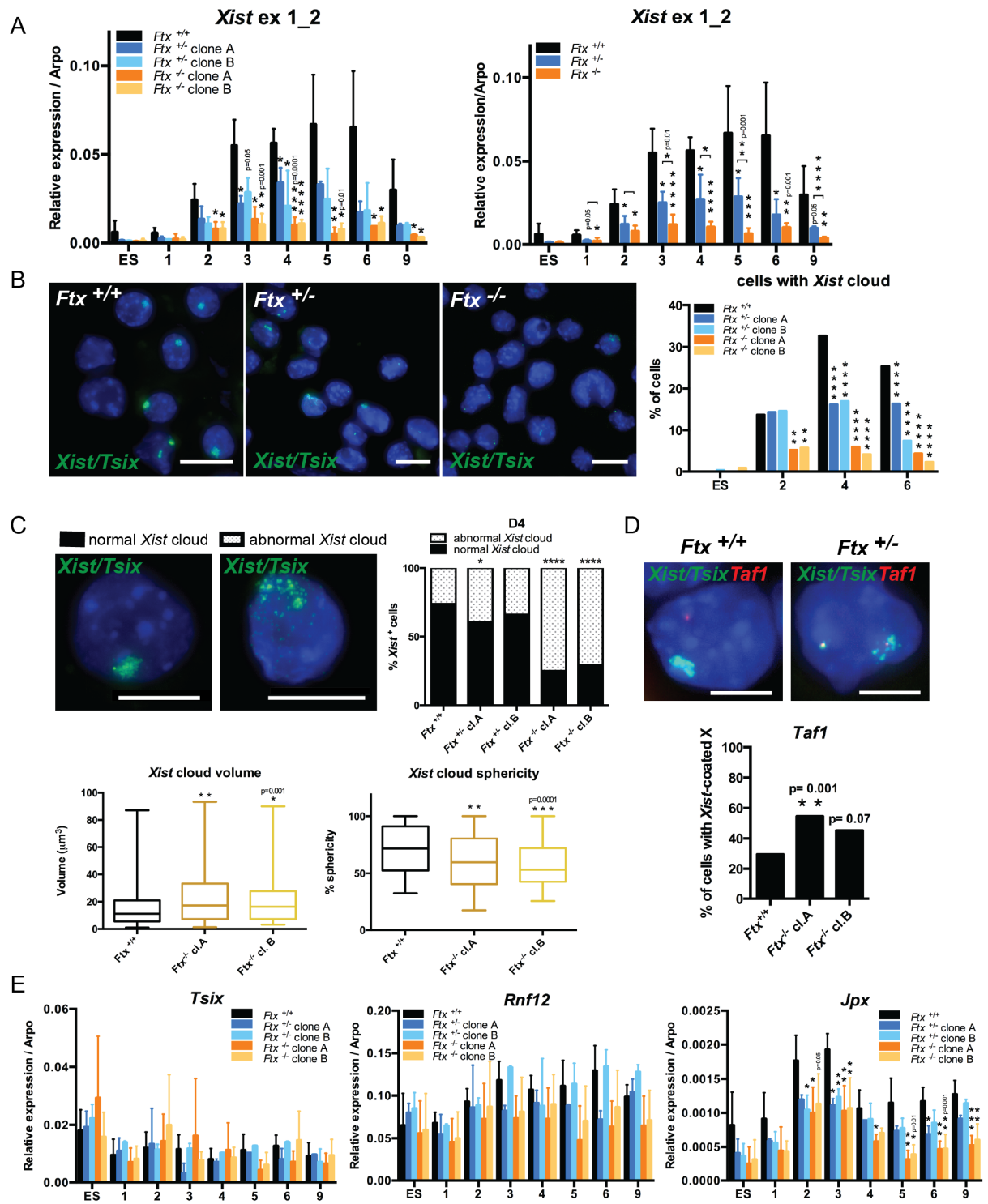


Figure 4. *Ftx* is required for *Xist* accumulation and proper XCI progression independently of other *Xic*-linked genes

A. Left panel: Time-course analysis of *Xist* expression by RT-qPCR in differentiating female LF2 WT and *Ftx* mutant cell lines. Averages and standard deviation (S.D.) from at least three independent differentiation experiments are plotted. Values are normalized to *Arpo* RNA levels. Statistical significance was assessed with a two-tailed t-test. Right panel: Time-course analysis of *Xist* expression by RT-qPCR in differentiating female LF2 WT and *Ftx* mutant cell lines. Expression values (same as in A) from samples with the same genotype are pooled and averaged. Statistical significance was assessed with a two-tailed t-test.

B. Left panel: Representative RNA-FISH images of differentiating (D4) female LF2 WT and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. Scale bar = 10 μ m. Right panel: Quantification of the percentage of cells with *Xist* cloud in differentiating (D2, D4 and D6) female LF2 WT and *Ftx* mutant cells. Statistical significance was assessed with Fisher's exact test. Sample sizes (n) = ES: 197-613; D2: 210-343; D4: 289-655; D6: 394-691.

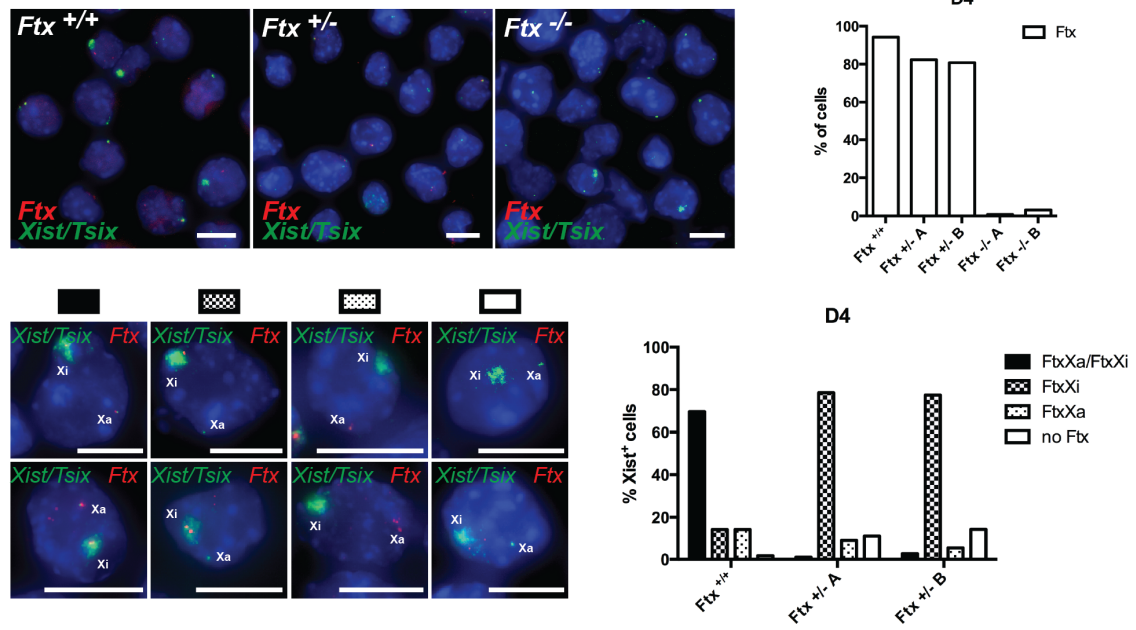
C. Top panels: *Xist/(Tsix)* RNA-FISH images of differentiating (D4) female LF2 cells (left) and quantification of the percentage of cells displaying either a normal or an abnormal *Xist* domain (right). Statistical significance was assessed with Fisher's exact test. Sample sizes (n) = 153; 94; 100; 88; 65. Bottom panels: Assessment of the *Xist* cloud volume (left) and sphericity (right) in female LF2 WT and *Ftx*^{-/-} cells. Statistical significance was assessed with Mann-Whitney test. Sample sizes (n) = 97; 87; 89.

D. Top: Representative *Xist/Taf1* RNA-FISH images of differentiating (D4) female LF2 wild-type and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. *Taf1* probe, Cy3-labeled, red. Bottom: Quantification of the percentage of cells with *Xist* cloud in which the *Taf1* signal is present. Statistical significance was assessed with Fisher's exact test. Sample sizes (n) = 88; 77; 53.

E. Time-course analyses of the expression of the *Xist*-regulating *Xic*-encoded genes *Tsix*, *Jpx*, and *Rnf12* in differentiating female LF2 WT and *Ftx* mutant cell lines. Averages and standard deviation (S.D.) from at least three independent differentiation experiments are plotted. Values are normalized to *Arpo* RNA levels. Statistical significance was assessed with a two-tailed t-test.

Figure 5

A



B

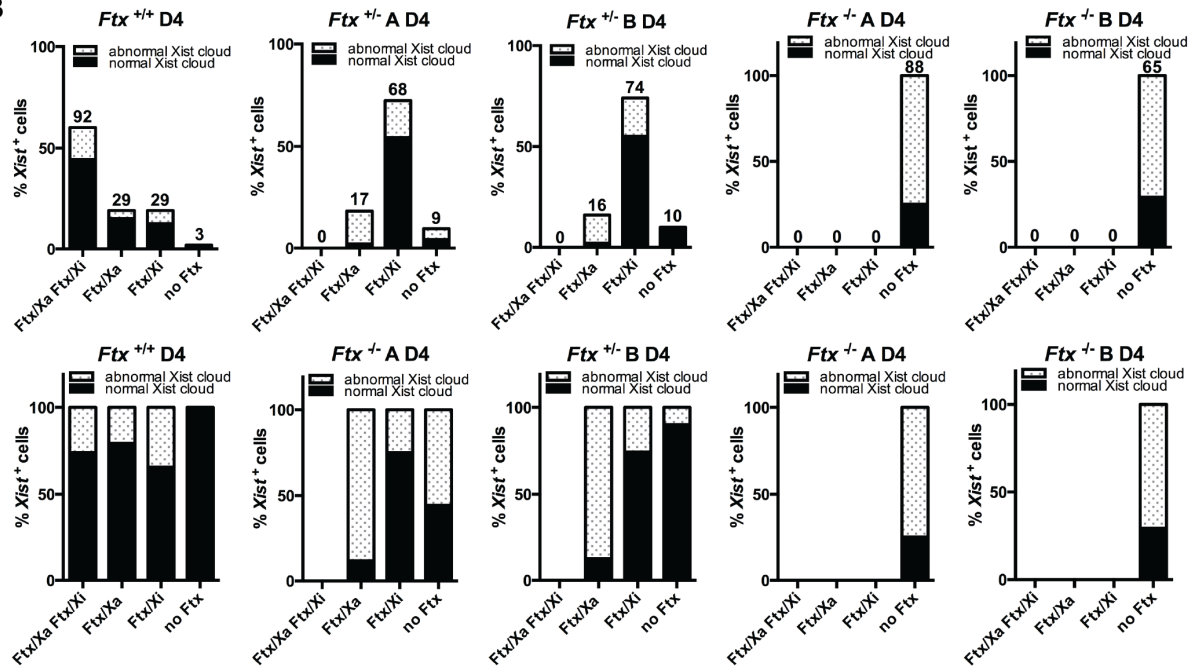


Figure 5. *Ftx* and *Xist* are mostly co-expressed

A. Top panel, left: Representative RNA-FISH images of differentiating (D4) female LF2 WT and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. *Ftx* probe, Cy3-labeled, red. Scale bar = 10 μ m. Top panel, right: Quantification of the percentage of *Ftx* expressing cells in differentiating (D4) female LF2 WT and *Ftx* mutant cells. Sample sizes (n) = 536; 607; 655; 289; 484. Bottom panel, left: RNA-FISH images of each possible *Xist/(Tsix)/Ftx* configuration in differentiating (D4) female LF2 WT and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. *Ftx* probe, Cy3-labeled, red. Scale bar = 10 μ m. Bottom panel, right: Quantification of the percentage of each configuration in differentiating (D4) female LF2 WT and *Ftx* mutant cells. Sample sizes (n) = *Ftx*^{+/+}: 175; *Ftx*^{+/-} A: 98; *Ftx*^{+/-} B: 111.

B. Top panel: *Xist/Tsix-Ftx* RNA-FISH quantification of the association between the *Xist* cloud (normal or abnormal) and *Ftx* transcripts in female LF2 WT and *Ftx* mutant cells. Sample sizes (n) = 153; 94; 100; 88; 65. Bottom panel: Same as in top panel, shown in percentages.

Figure S3

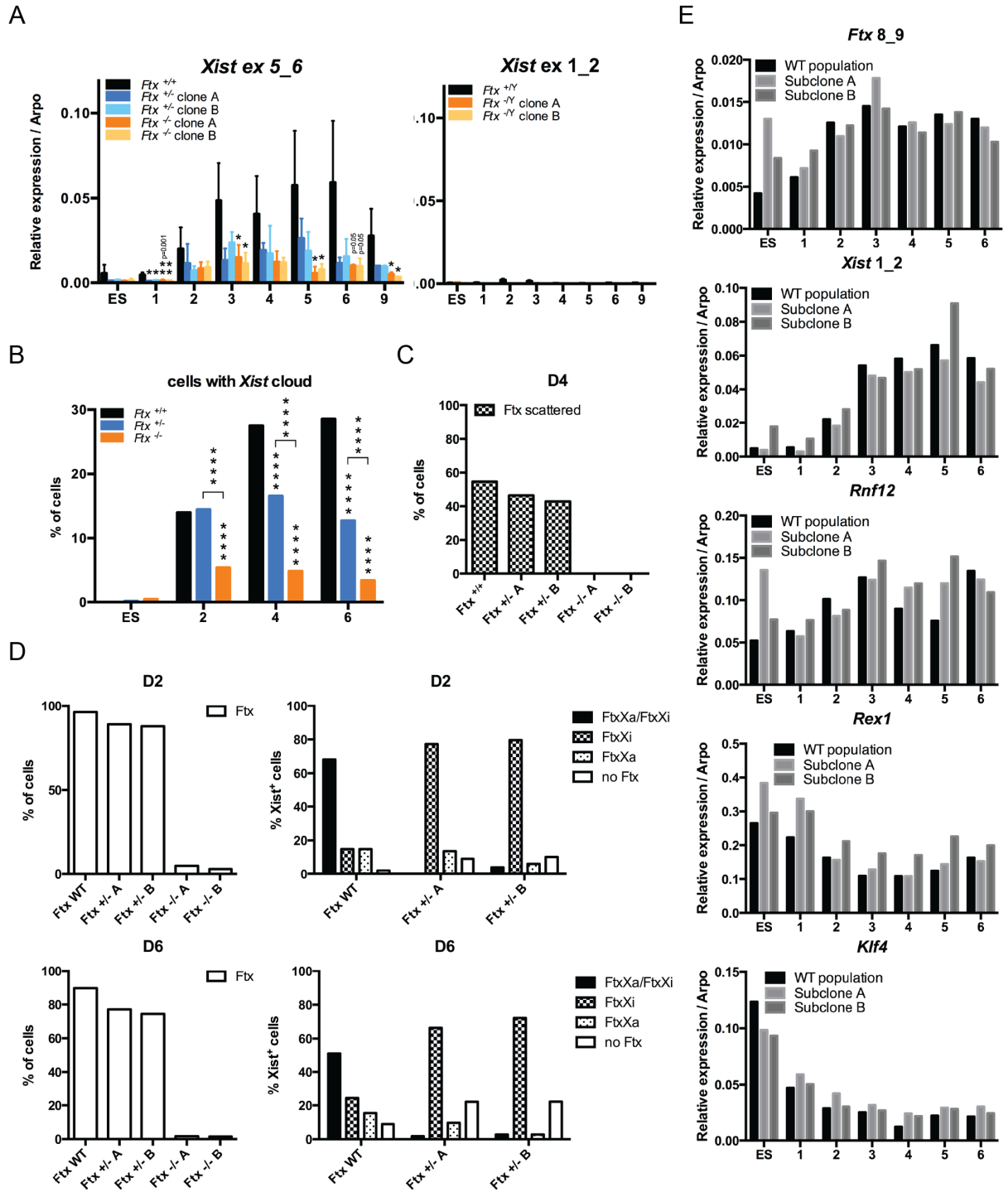


Figure S 3. *Ftx* is required for *Xist* accumulation and the effects is independent of the subcloning procedure

A. Left panel: Time-course analysis of *Xist* expression by RT-qPCR in differentiating female LF2 WT and mutant *Ftx* cell lines. Averages and standard deviation (S.D.) from at least three independent differentiation experiments are plotted. Values are normalized to *Arpo* RNA levels. Statistical significance was assessed with a two-tailed t-test. Right panel: Time-course analysis of *Xist* expression by RT-qPCR in differentiating male TC1 WT and *Ftx* mutant cell lines. Values from one differentiation experiment are plotted.

B. Quantification of the percentage of cells with *Xist* cloud in differentiating (D2, D4 and D6) female LF2 WT and mutant cells. Countings (same as in Figure 4B) from samples with the same genotype are pooled. Statistical significance was assessed with Fisher's exact test. Sample sizes (n) = ES: 197-613; D2: 210-343; D4: 289-655; D6: 394-691.

C. Quantification of the percentage of cells showing *Ftx* scattered signal in differentiating (D4) female LF2 WT and *Ftx* mutant cells. Sample sizes (n) = 338; 98; 111; 289; 484.

D. *Xist/Tsix-Ftx* RNA-FISH. Quantification of the percentage of *Ftx* expressing cells and of each *Xist/(Tsix)/Ftx* configuration (same as in Figure 5A) in differentiating (D2 and D6) female LF2 WT and *Ftx* mutant cells. Sample sizes (n) = D2, total cells: 468; 307; 336; 252; 210. D6, total cells: 485; 691; 482; 417; 394. D2, *Xist*⁺ cells: 47; 44; 49. D6, *Xist*⁺ cells: 123; 113; 36.

E. Time-course analysis of *Ftx*, *Xist*, *Rnf12*, *Rex1* and *Klf4* expression by RT-qPCR in differentiating female LF2 WT and WT subcloned cell lines. Values from one differentiation experiment are plotted. Values are normalized to *Arpo* RNA levels.

5 Generation of *Ftx* mutant female ES cell lines: PGK1

While the skewed *Xist* expression patterns in *Ftx*^{+/-} cells already suggest a role for *Ftx* in *cis*, a definitive answer can only be given by the analysis of cells bearing X-linked polymorphisms. To investigate this, we deleted the *Ftx* promoter region as described above (Figure 1A) in the female *Mus musculus*, strain 129/Pgk, ES cell line PGK1, in which the two X chromosomes can be told apart thanks to the presence of SNPs.

We assessed for the presence of targeting events first in the population of transfected cells and then in a series of independent clones (Figure S4). Out of the 192 clones tested, 6 were positive for the PCR assay detecting the presence of the deletion. All of them showed both WT and KO bands and were thus classified as heterozygous (not shown). No pure homozygous clones were found at this step. Taking advantage of the presence of a SNP inside the region to be deleted (*Ftx* intron 1), we designed an allelic RT-qPCR assay that would enable us to identify which X chromosome was targeted (*X*₁₂₉ or *X*_{Pgk}), check for the presence of both X chromosomes and assess the purity of the population all at the same time.

In homogeneous heterozygous populations, amplification inside the region to be deleted should correspond to either only the 129 or only the Pgk allele (WT allele). The presence of amplification signal from both alleles inside the deleted region indicated that none of the 6 clones was a pure population (Figure S4C). However, the difference in the amount of each allele compared to the other was compatible with the presence of *Ftx*-deleted cells and thus with a variable degree of heterogeneity in each population. To obtain pure populations of independent clones, we subcloned all 6 populations. We analyzed 16 subclones from each of them, and found that 19% (first population), 37% (second population), 12% (third and fifth populations), and 69% (sixth population) of the subclones were real heterozygous, while the remaining subclones were wild-type, according to the PCR assay. Interestingly, the fourth population contained two clones that displayed the KO band exclusively, and that could thus either have lost an X chromosome, or bear a homozygous *Ftx* deletion.

We then repeated the allelic RT-qPCR assay on the subclones and found that we had (at least) one heterozygous clone with ΔFtx_{Pgk} (*Ftx*^{+/-} A), one clone with ΔFtx_{129} (*Ftx*^{+/-} B) and one $\Delta Ftx_{Pgk}/\Delta Ftx_{129}$ clone (*Ftx*^{-/-}) (Figure S4D). Among these, we

selected 3 independent clones for further analysis ($Ftx^{+/-}$ A, $Ftx^{+/-}$ B and $Ftx^{-/-}$ A) (Figure 6A and Figure 6C). To make sure that all the cells in each selected clone were XX, we performed *Xist* DNA-FISH (Figure 6D). In order to precisely locate the borders of the deletion and to verify the integrity of the homologous region on the WT X chromosome (in the heterozygous clones), we sequenced the PCR products spanning the targeted sites. We observed that the DSBs were induced as expected in all the heterozygous and homozygous clones, but a 500bp region was deleted also on the X_{Pgk} in the ΔFtx_{129} ($Ftx^{+/-}$ B) clone at the 3' junction, and a small segment was deleted on the X_{129} in the ΔFtx_{Pgk} ($Ftx^{+/-}$ A) clone at the 3' junction (Figure 6B). The 500bp region deleted on the X_{Pgk} in the ΔFtx_{129} clone partially spans a mediator and a cohesin binding site, while no known binding sites are present inside the small deletion on the X_{129} in the ΔFtx_{Pgk} clone. Neither of these additional deletions prevents or affects *Ftx* transcription from that chromosome (Figure 8A).

Figure 6

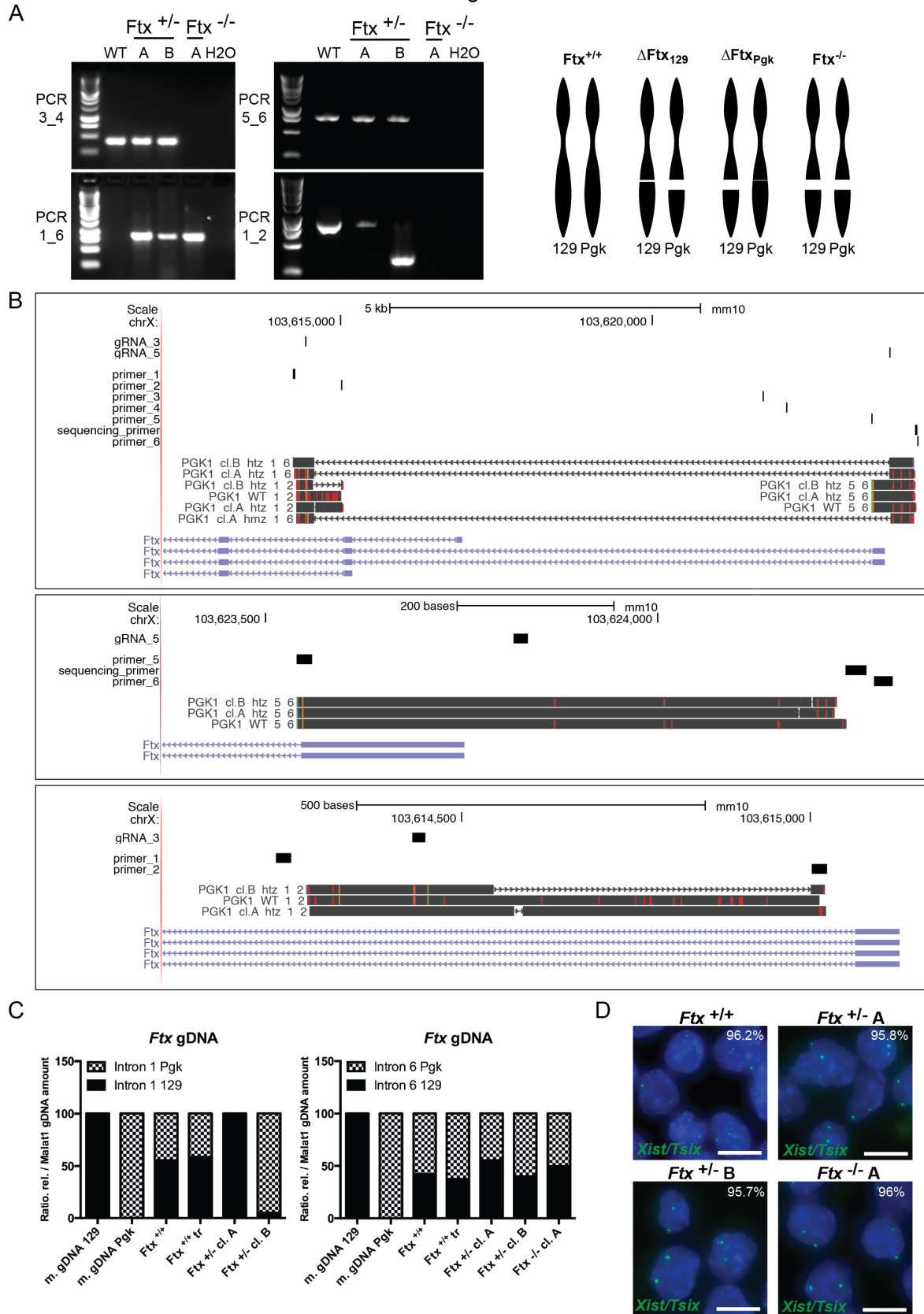


Figure 6. Generation of PGK1 *Ftx* mutant cell lines

A. Left panels, left: PCR assay to detect the wild-type allele (top) and deletion event (bottom) in WT cells and in the selected *Ftx* mutant clones. Left panels, right: PCR assay to assess the integrity of the 3' (top) and 5' (bottom) targeted regions on the WT *Ftx* allele in the selected *Ftx* mutant clones. Right panel: illustration of the cell lines used. White boxes indicate the deletions inside the *Ftx* locus.

B. UCSC genome browser alignments of the sequencing results on the PCR products from A.

C. Representative *Xist/Tsix* DNA-FISH images of WT and the selected *Ftx* mutant clones. Scale bar = 10 μ m.

D. Ratio of allelic (129 and Pgk) genomic DNA amount inside (*Ftx* intron 1, left panel) and outside (*Ftx* intron 6, right panel) the deleted region in mouse gDNA (control), WT and selected *Ftx* mutant clones, by RT-qPCR.

Figure S4

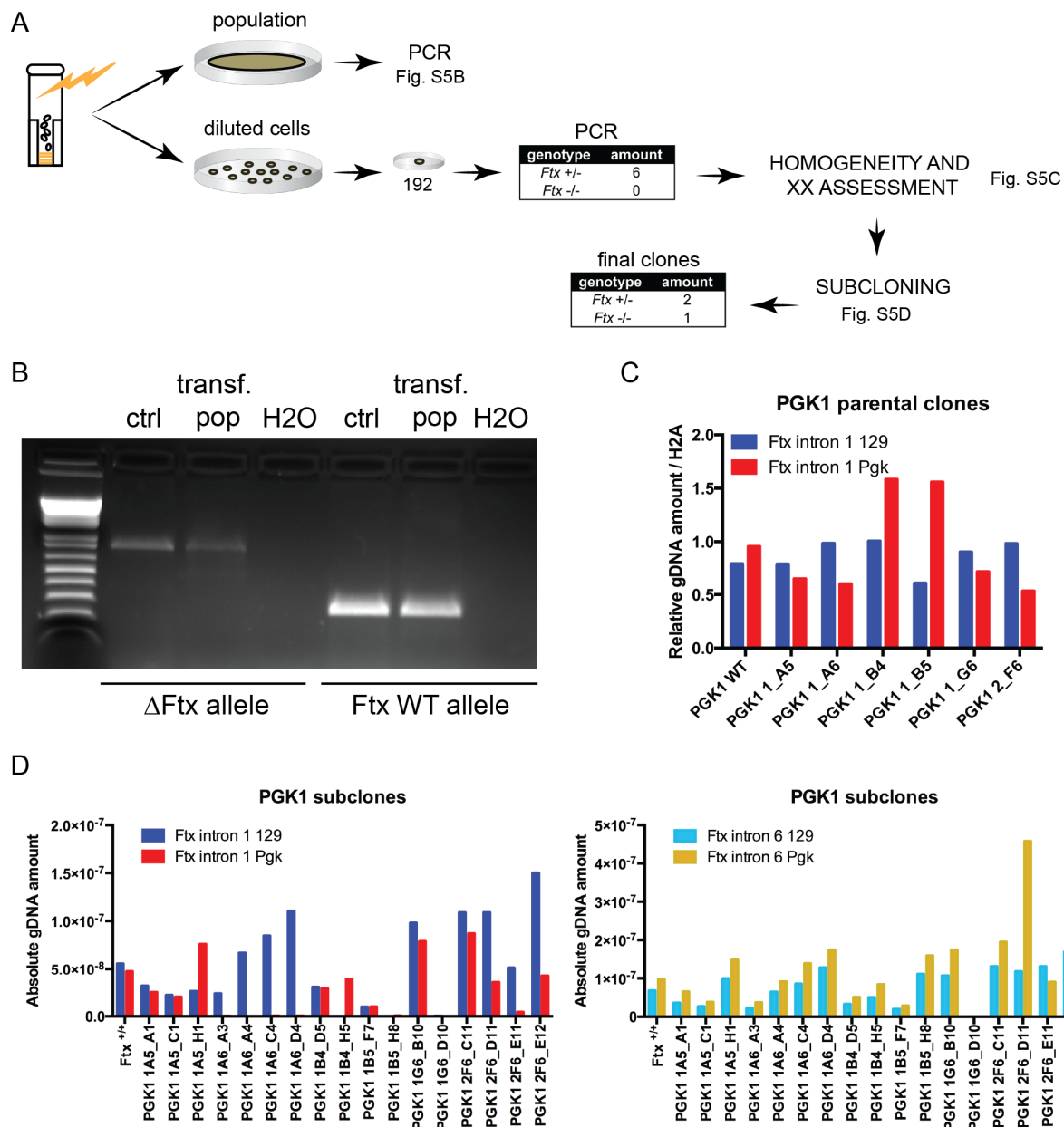


Figure S 4. Generation of PGK1 *Ftx* mutant cell lines

A. Schematic representation of the procedure for the generation of female PGK1 *Ftx* mutant cell lines. After transfection, the cells were either plated at high density and checked for the presence of the deletion event at the population level, or directly seeded at low density to obtain independent clones. After ten days, clones were picked and individually screened for the presence of the targeting event. Positive clones were further checked for the presence of two X chromosomes and assessed for homogeneity. Non-homogeneous clonal population were sub-cloned.

B. PCR assay to detect the wild-type allele and deletion event in the transfected population.

C. RT-qPCR quantification of the allelic (129 and P_{gk}) genomic DNA amount inside (*Ftx* intron 1) the deleted region in WT and *Ftx* mutant clonal populations before sub-cloning.

D. RT-qPCR quantification of the allelic (129 and P_{gk}) genomic DNA amount inside (*Ftx* intron 1, left panel) and outside (*Ftx* intron 6, right panel) the deleted region in five groups of independent clones after sub-cloning.

6 The lack of *Ftx* results in cell death and impaired *Xist* expression also in PGK1 cells

In order to investigate if the *Ftx*-induced phenotype was reproduced in this second cell line, we explored *Ftx* expression profiles and differentiation dynamics in PGK1 wild-type and *Ftx*-mutant cells. Like in the LF2 cell line, *Ftx* is expressed in the majority (~90%) of the cells in undifferentiated wild-type PGK1 cells, either biallelically or monoallelically (Figure S5A). In *Ftx*^{+/-} PGK1 cells *Ftx* is still expressed in the majority (~80%) of the cells, but exclusively in a monoallelic pattern (consistent with the heterozygous deletion), with a concomitant increase in the proportion of cells showing no *Ftx* expression (Figure S5A). Finally, *Ftx*^{-/-} PGK1 cells lack *Ftx* expression completely (Figure S5A). These proportions are maintained during differentiation (Figure S5B).

Differentiating PGK1 *Ftx* mutant cell lines recapitulated the differentiation defect observed in LF2 mutant cells, with *Ftx*^{+/-} and *Ftx*^{-/-} cells displaying a high mortality rate and an increased presence of round-shaped ES-like colonies compared to wild-type, along the whole time-course of differentiation. (Figure 7A and Figure S5C). When we measured the expression levels of the pluripotency markers *Nanog*, *Klf4* and *Rex1* by RT-qPCR, we did not observe any striking difference between mutant and wild-type lines (Figure 7B). Once again, the extent of downregulation of these genes is however not striking in our differentiation experiments.

We next measured steady-state *Ftx* and *Xist* levels by RT-qPCR and visualized *Xist* accumulation by RNA-FISH. In *Ftx*^{+/+} cell lines, *Ftx* is upregulated in the first days of differentiation. In *Ftx*^{+/-} cell lines, *Ftx* upregulation follows the same dynamics as in wild-type cells, but, consistent with the deletion, the levels are reduced to half (Figure 7C). In the *Ftx*^{-/-} line, *Ftx* is not expressed (Figure 7C). In *Ftx*^{+/+} cell lines, *Xist* is efficiently upregulated (Figure 7C and Figure S5D) and accumulated (Figure 7D), with ~30% of cells showing a *Xist* cloud at day 4. In the *Ftx*^{-/-} line, *Xist* upregulation and accumulation are strongly impaired, with low levels of *Xist* expression along the whole time-course (Figure 7C and Figure S5D) and a drop in the percentage of cells with *Xist* cloud below 10% at day 4 (Figure 7D). This result strengthens our conclusions for a role of *Ftx* in *Xist* upregulation. The situation is less clear in *Ftx*^{+/-} cell lines, where *Xist* upregulation (Figure 7C and Figure S5D) and accumulation

(Figure 7D) are more or less affected according to the clone and thus according to the type of deletion (ΔFtx_{Pgk} or ΔFtx_{129}): in ΔFtx_{Pgk} cells *Xist* levels are not different from those of wild-type cells and *Xist* accumulation at the single-cell level is not significantly reduced, but in ΔFtx_{129} cells *Xist* upregulation and accumulation are drastically reduced, to a similar extent than in $Ftx^{-/-}$ cells. This difference between the two clones could be due to the additional deletion of a small segment on the homologous X chromosome in the ΔFtx_{129} clone (Figure 6B), which could render X_{Pgk} more similar to X_{129} . It is also possible that X_{Pgk} is more “resistant” to effects of the *Ftx* deletion than X_{129} , so that the effect on *Xist* is subtler in ΔFtx_{Pgk} cells. In any case, additional biological replicates of the differentiation experiment are needed to draw a final conclusion.

As observed in LF2 cells, no striking differences in the transcription levels of *Tsix* and *Rnf12* were found in *Ftx*-mutant compared to wild-type lines (Figure 7E).

Taken together, these data demonstrate that *Ftx* is strongly and reproducibly involved in *Xist* upregulation.

Figure 7

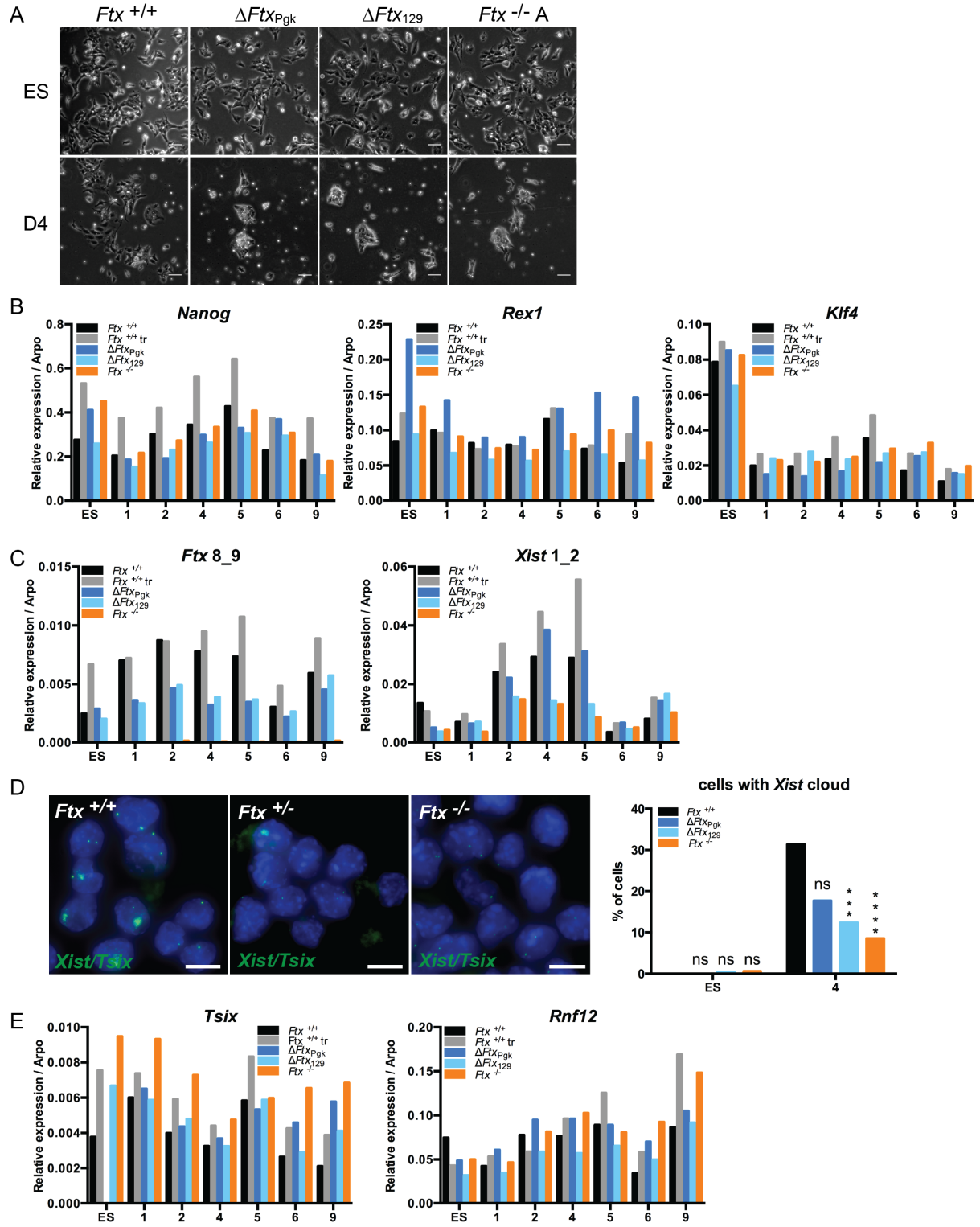


Figure 7. The lack of *Ftx* results in cell death and impaired *Xist* expression in PGK1 cells

A. Brightfield photographs of undifferentiated (ES) and differentiating (D4) female PGK1 WT and *Ftx* mutant cells. Scale bar = 50 μ m.

B. Time-course analyses of *Nanog*, *Klf4* and *Rex1* expression by RT-qPCR in differentiating female PGK1 WT and *Ftx* mutant cell lines. Values from one differentiation experiment are plotted. Values are normalized to *Arpo* RNA levels.

C. Time-course analyses of *Ftx* and *Xist* expression by RT-qPCR in differentiating female PGK1 WT and *Ftx* mutant cell lines. Values from one differentiation experiment are plotted. Values are normalized to *Arpo* RNA levels.

D. Left panel: Representative RNA-FISH images of differentiating female PGK1 WT and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. Scale bar = 10 μ m. Right panel: Quantification of the percentage of cells with *Xist* cloud in differentiating (D4) female PGK1 WT and *Ftx* mutant cells. Statistical significance was assessed with Fisher's exact test. Sample sizes (n) = ES: 160-299; D4: 220-514.

E. Time-course analyses of *Tsix* and *Rnf12* expression by RT-qPCR in differentiating female PGK1 WT and *Ftx* mutant cell lines. Values from one differentiation experiment are plotted. Values are normalized to *Arpo* RNA levels.

Figure S5

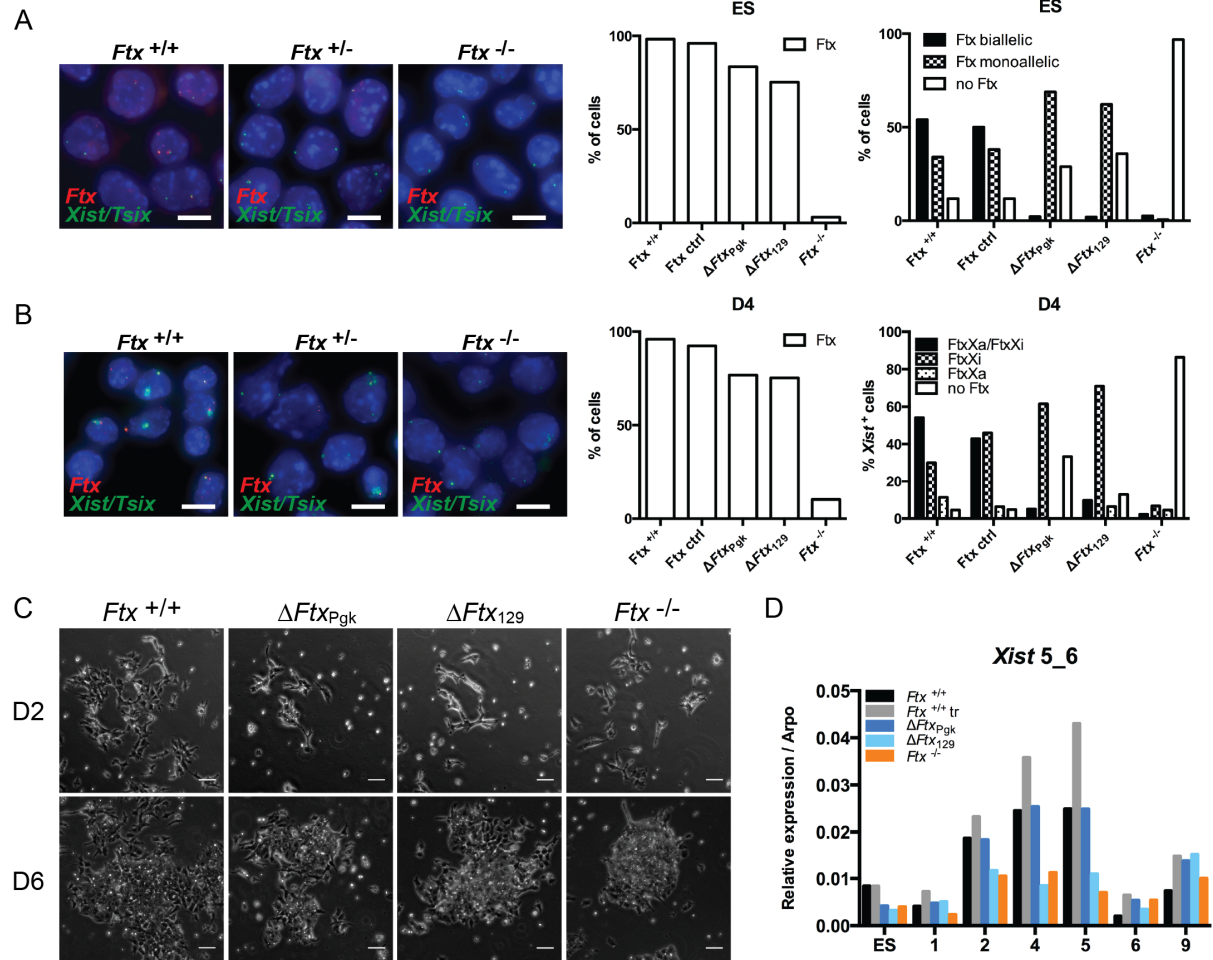


Figure S 5. The lack of *Ftx* results in cell death and impaired *Xist* expression in PGK1 cells

A. Left panel: Representative RNA-FISH images of undifferentiated (ES) female PGK1 WT and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. *Ftx* probe, Cy-3 labeled, red. Scale bar = 10 μ m. Quantification of the percentage of *Ftx* expressing cells in undifferentiated (ES) female PGK1 WT and *Ftx* mutant cells. Sample sizes (n) = 235; 202; 225; 299; 160. Right panel: Quantification of the percentage of each possible (*Xist/*)*Tsix/Ftx* configuration (as in Figure 2B) in undifferentiated (ES) female PGK1 WT and *Ftx* mutant cells. Sample sizes (n) = 235; 202; 225; 299; 160.

B. Left panel: Representative RNA-FISH images of differentiating (D4) female PGK1 WT and *Ftx* mutant cells. *Xist/Tsix* probe, FITC-labeled, green. *Ftx* probe, Cy-3 labeled, red. Scale bar = 10 μ m. Quantification of the percentage of *Ftx* expressing cells in differentiating (D4) female PGK1 WT and *Ftx* mutant ES cells. Sample sizes (n) = total cells: 424; 198; 220; 251; 514. *Xist*⁺ cells: 133; 63; 39; 31; 44. Right panel: Quantification of the percentage of each possible (*Xist/*)*Tsix/Ftx* configuration (as in Figure 2B) in differentiating (D4) female PGK1 WT and *Ftx* mutant cells.

C. Brightfield photographs of differentiating (D2 and D6) PGK1 female PGK1 WT and *Ftx* mutant cells. Scale bar = 50 μ m

D. Time-course analysis of the expression of *Xist* in differentiating female PGK1 WT and *Ftx* mutant cell lines. Values from one replicate are plotted. Values are normalized to Arpo RNA levels.

7 *Ftx* is required for *Xist* accumulation in *cis*

In order to investigate if *Ftx* is required for *Xist* upregulation in *trans* or in *cis*, we performed *Ftx* and *Xist* allele-specific RT-qPCR analyses. The presence of a single nucleotide polymorphism (SNP) in *Ftx* intron 6 and *Xist* exon 5 allowed us to distinguish the *Ftx* and *Xist* transcripts emanating from either the Pgk or the 129 alleles.

As suggested by the RNA-FISH experiments (Figure 5A, Figure S3D, Figure S5A and Figure S5B) *Ftx* is expressed from both X chromosomes in undifferentiated and differentiating *Ftx*^{+/+} cells, and it is expressed and upregulated exclusively from the wild-type X chromosome in *Ftx*^{+/-} cells (Figure 8A). In *Ftx*^{-/-} cells, only very low residual *Ftx* expression is detectable from X_{Pgk} (Figure 8A), possibly arising from spurious transcription across the locus.

While in *Ftx*^{-/-} cells *Xist* expression is decreased from either X chromosome (Figure 8B), the effect of the heterozygous *Ftx* deletion in ΔFtx_{Pgk} and ΔFtx_{129} cells is restricted in *cis*, with *Xist* upregulation being reduced exclusively from X₁₂₉ in the ΔFtx_{129} population and exclusively from X_{Pgk} in the ΔFtx_{Pgk} population (Figure 8B). To better represent this result, we calculated the ratios of allelic expression and indeed observed that, while in *Ftx*^{+/+} and *Ftx*^{-/-} cells *Xist* is expressed at the same level from either X chromosome (50% *Xist*_{Pgk} and 50% *Xist*₁₂₉), in *Ftx*^{+/-} cells *Xist* expression is skewed, with *Xist* being mostly transcribed from the WT X chromosome (X₁₂₉ in ΔFtx_{Pgk} cells (clone A) and X_{Pgk} in ΔFtx_{129} cells (clone B)) (Figure 8D).

However, the extent to which *Xist* expression is decreased is not the same in the two *Ftx*^{+/-} clones, with *Xist*_{Pgk} residual expression in ΔFtx_{Pgk} cells being higher than *Xist*₁₂₉ residual expression in ΔFtx_{129} cells (Figure 8B). This is reminiscent of the skewed XCI patterns observed in cells carrying two distinct *Xce* alleles, but in our case the effect seems to be confined to *Ftx*^{+/-} populations. Importantly, the ΔFtx_{Pgk} line, despite showing no decrease in the percentage of cells with a *Xist* cloud, does display skewed *Xist* expression. This, together with the fact that cell death is observed also in this line, could indicate that very few XaXa cells are surviving, and that the majority of the remaining cells are XaXi cells with skewed *Xist* expression.

The *cis* effect of the *Ftx* deletion is maintained also at intermediate and at later days during differentiation (Figure S6A). These results are unlikely to be due to the loss of

one X chromosome during differentiation, as allelic-specific quantification of the *Xist*_{Pgk} and *Xist*₁₂₉ loci indicates the presence of an equal amount of genomic DNA in the undifferentiated state and at day 4 of differentiation for all cell lines (Figure S6B). To evaluate the contribution of the lncRNA *Jpx*, shown to be significantly downregulated in *Ftx* mutant cells (Figure 5C), to the phenotype, we took advantage of the presence of a SNP in *Jpx* exon 4 and measured allelic *Jpx* expression. In this cell line (PGK1), the effects of the *Ftx* deletion on *Jpx* levels are not as drastic as in LF2 cells (Figure 8C). While a definitive conclusion is hard to provide with the low number of replicates that we have for the moment, *Jpx* does not seem to be the main effector of the *Ftx*-mediated *Xist* regulation mechanism, as *Jpx* is known to act in *trans*, yet *Xist* is decreased exclusively from the *Ftx*-deleted chromosome in *Ftx*^{+/-} cells. Moreover, the ratios of *Jpx* allelic expression (50% *Jpx*_{Pgk} and 50% *Jpx*₁₂₉) are not consistent even with a *Jpx*-mediated *cis* regulation of *Xist* upregulation (Figure 8D). This suggests that *Ftx* effect on *Xist* is either direct or depends on yet unidentified intermediate regulators or interactors.

All together, these results indicate that *Ftx* is required in *cis* for proper *Xist* accumulation in differentiating mESCs. Wild-type cells are able to upregulate and accumulate *Xist*, and induce random XCI of either X chromosome upon differentiation (Figure S6C). Cells bearing a homozygous *Ftx* promoter deletion are still able to upregulate and accumulate *Xist* and induce random XCI, but do so only in a very small percentage of the population. The remaining cells, unable to undergo correct XCI, either die or retain an ES-like pluripotent state (Figure S6C). Finally, cells bearing a heterozygous *Ftx* promoter deletion are able to upregulate and accumulate *Xist* in a reduced proportion of cells, and display a selective advantage of cells that have chosen to silence the *Ftx* wild-type allele, which leads to skewed random XCI patterns. Like in the case of *Ftx* knock out populations, the remaining cells, that have chosen to silence the *Ftx* mutant allele, either die or retain an ES-like pluripotent state (Figure S6C).

Figure 8

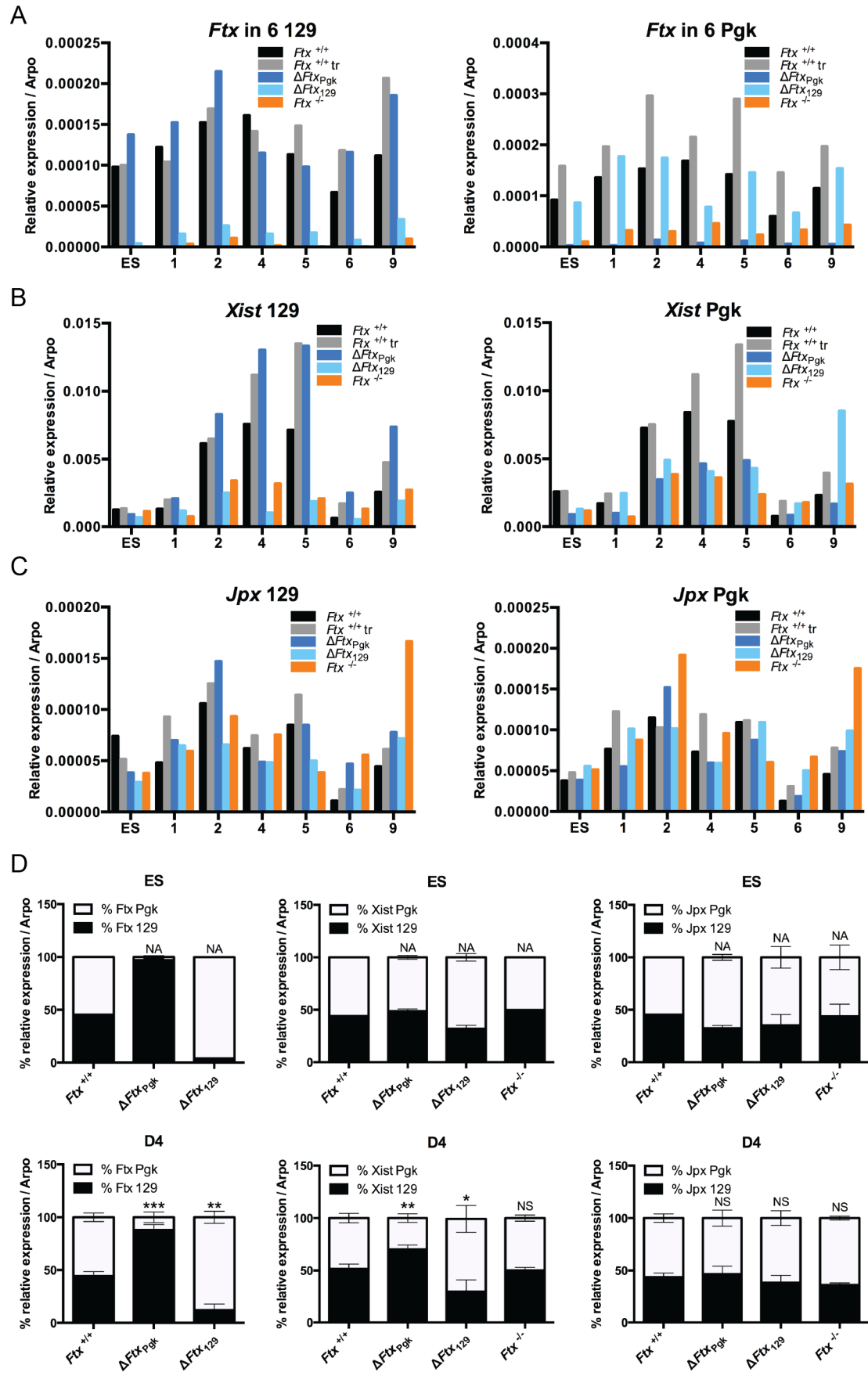


Figure 8. *Ftx* is required for *Xist* accumulation in *cis*

A. Time-course analyses of *Ftx* allelic expression by RT-qPCR in differentiating female PGK1 WT and *Ftx* mutant cell lines. Values are normalized to *Arpo* RNA levels. One replicate is shown.

B. Time-course analyses of *Xist* allelic expression by RT-qPCR in differentiating female PGK1 WT and *Ftx* mutant cell lines. Values are normalized to *Arpo* RNA levels. One replicate is shown.

C. Time-course analyses of *Jpx* allelic expression by RT-qPCR in differentiating female PGK1 WT and *Ftx* mutant cell lines. Values are normalized to *Arpo* RNA levels. One replicate is shown.

D. Ratios of *Ftx*, *Xist* and *Jpx* allelic expression in undifferentiated (ES) and differentiating (D4) female PGK1 WT and *Ftx* mutant cell lines. Averages and standard deviations (S.D.) from two (ES) or three (D4) independent differentiation experiments (monolayer, LIF removal, + retinoic acid) are plotted. Values are normalized to *Arpo* RNA levels. Statistical significance was assessed with a two-tailed t-test.

Figure S6

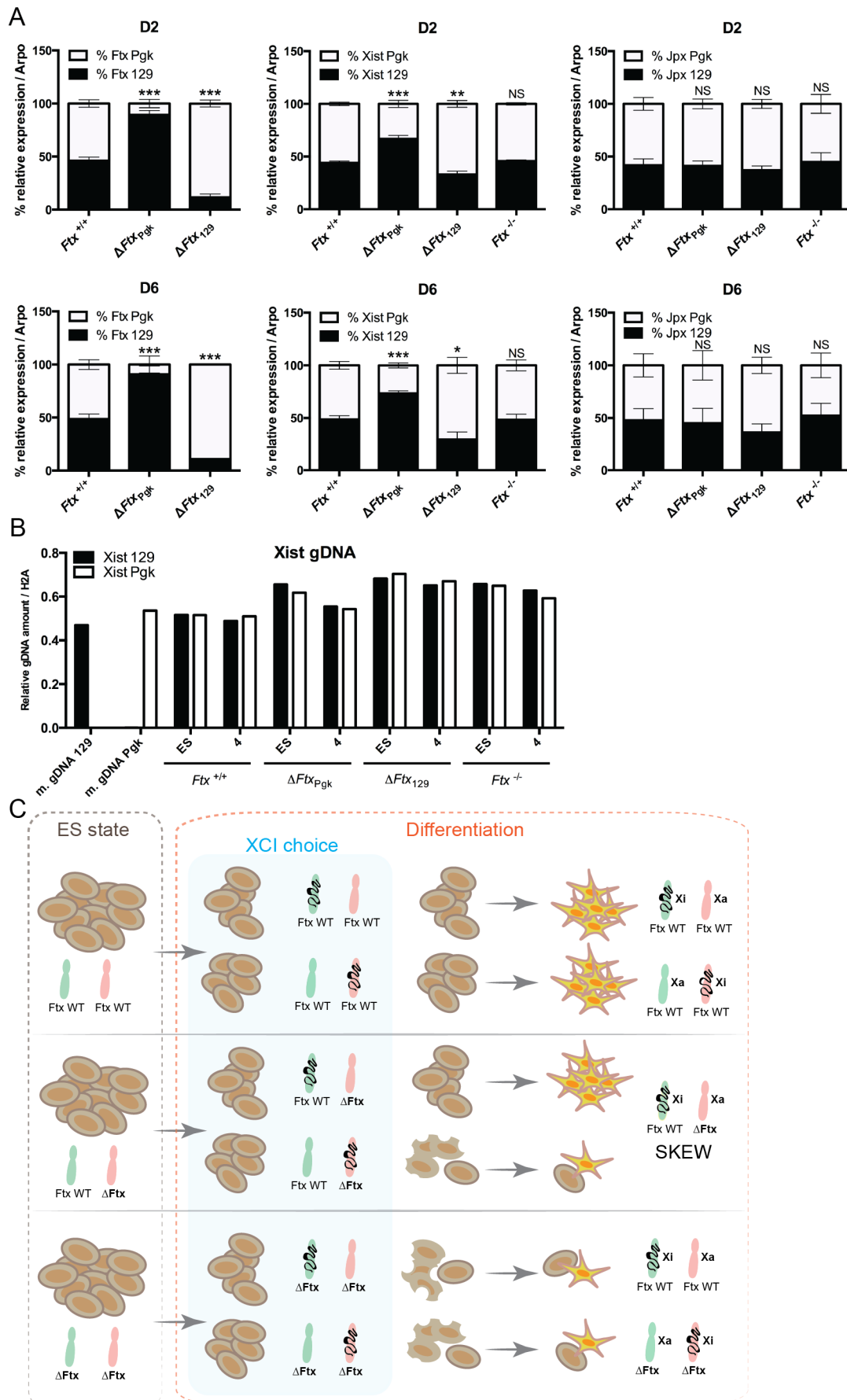


Figure S 6. *Ftx* is required for *Xist* accumulation in *cis*

A. Ratios of *Ftx*, *Xist* and *Jpx* allelic expression in differentiating (D2 and D6) female PGK1 WT and *Ftx* mutant cell lines. Averages and standard deviations (S.D.) from three independent differentiation experiments (monolayer, LIF removal, + retinoic acid) are plotted. Values are normalized to *Arpo* RNA levels. Statistical significance was assessed with a two-tailed t-test.

B. RT-qPCR quantification of the allelic genomic DNA amount of the *Xist* locus in undifferentiated (ES) and differentiating (D4) female PGK1 WT and *Ftx* mutant ES cells, with mouse genomic DNA highlighting the specificity of the primers. Values are normalized to H2A gDNA amount.

C. Schematic representation of the effects of the *Ftx* deletion in female ES cells over the course of differentiation.

8 Insights into the mechanisms of *Ftx*-mediated *Xist* regulation

8.1 At the level of *Xist*

Once demonstrated that *Ftx* is required for *Xist* upregulation in *cis*, we sought to examine how *Xist* is affected. Given the reduction in *Xist* levels that we observe in *Ftx* mutant cells and the elevated percentage of abnormal *Xist* clouds in the majority of *Ftx*^{-/-} cells and on the *Ftx*-deleted X chromosome in the majority of *Ftx*^{+/-} cells, *Ftx* could, in principle, regulate *Xist* at the transcriptional level, so that fewer *Xist* RNA molecules can be produced and accumulated on the Xi, or affect *Xist* accumulation in a post-transcriptional manner, by interfering with *Xist* RNA splicing, stabilization, tethering to the Xi or spreading. To discriminate between these possibilities, we quantified the amount of the pre-mature (unspliced) *Xist* transcript at the population level (Figure 9A) and visualized *Xist* ongoing transcription and *Xist* accumulation separately at the single-cell level (Figure 9B).

Since *Xist* antisense *Tsix* is still expressed (albeit at low and progressively decreasing levels) in differentiating mESCs (Figure 5C), we performed *Xist*-specific or *Tsix*-specific reverse transcription assays, to discriminate between the *Xist* and the *Tsix* RNAs, and we subsequently measured the amount of each product with a set of common RT-qPCR primers (Figure 9A). The results show that, both in undifferentiated and differentiating cells, the levels of unspliced *Xist* transcripts are similar between wild-type and *Ftx* mutant cell lines, suggesting that *Ftx* might act on *Xist* post-transcriptionally. Furthermore, no major differences can be seen regarding the levels of unspliced *Tsix* transcripts. Similar data were obtained also in male ES cells (not shown).

To assess *Xist* transcription at the single-cell level, we performed RNA-FISH using strand-specific single-stranded DNA probes that hybridize exclusively to the intronic regions of the *Xist* transcript (*Xist*IN), thus preferentially detecting the unspliced form. Complementary to this, we used a set of strand-specific single-stranded probes that hybridize exclusively to the exonic regions of the *Xist* transcript (*Xist*EX), thus detecting only the spliced form. As an additional control, we also used a set of strand-specific probes hybridizing to the *Tsix* transcript. We then calculated, within all the cells displaying at least one signal (*Xist*IN, *Xist*EX or *Tsix*), the percentage of cells displaying *Xist* transcription alone, *Xist* accumulation alone, a combination of

both, or a complete absence of *Xist* transcription and accumulation.

In *Ftx*^{+/+} populations, as observed with the traditional double-stranded *Xist* probe, around 30% of the cells display a *Xist* cloud by day 4. In the majority of these cells, *Xist* is actively transcribed, as indicated by the co-localization of the *XistEX* and *XistIN* signals, and only few cells display *Xist* accumulation without ongoing *Xist* transcription (*XistEX* signal exclusively). In the cells in which *Xist* is not accumulated, *Xist* is never even transcribed, as demonstrated by the absence of cells displaying exclusively a *XistIN* signal (Figure 9B).

In theory, if *Xist* was affected at the post-transcriptional level, we would observe, in *Ftx*^{-/-} populations, a percentage of cells showing *Xist* transcription without *Xist* accumulation (*XistIN* signal exclusively) comparable to the sum of the percentage of cells showing *Xist* accumulation alone (*XistEX* signal) and the percentage of cells showing *Xist* transcription and accumulation (*XistEX* and *XistIN* signals) in *Ftx*^{+/+} populations. Conversely, if *Xist* were affected at the transcriptional level, we would see, in *Ftx*^{-/-} populations, a decrease in the percentage of cells displaying ongoing *Xist* transcription (absence of *XistIN* signal and, most likely, also *XistEX* signal).

The fact that in *Ftx*^{-/-} populations all the cells that do not display any *Xist* accumulation do not display ongoing *Xist* transcription either and the number of cells in which *Xist* is actively transcribed is significantly decreased (Figure 9B) seems to suggest, contrary to the aforementioned RT-qPCR results, that *Xist* is affected transcriptionally, and that the decrease in mature (spliced) *Xist* transcript levels observed in *Ftx* mutant cells arises from the decrease in pre-mature (unspliced) *Xist* transcript levels. This discrepancy could be due to the different sensitivity and the nature of the two techniques, with the RT-qPCR being more sensitive and quantitative but only assessing the sample at the population level, and the RNA-FISH being less sensitive and offering only a qualitative answer, but at the single-cell level. Because of this, further experiments will be needed to clarify whether *Xist* is affected transcriptionally or post-transcriptionally.

8.2 At the level of *Ftx*

In both *Ftx*^{+/+} and *Ftx*^{+/-} populations, the fact that the presence of a *Xist* cloud is almost always accompanied by *Ftx* expression on the same X chromosome could indicate that the *Ftx* transcript is somehow involved in the *cis*-mediated monoallelic *Xist* upregulation. Alternatively, the *Ftx* transcript could be a byproduct of the possible interaction between the *Xist* and *Ftx* promoters, which would suggest a primary role for the genomic region or the act of transcription rather than for the transcript itself. This would be compatible with the observed skew in *Xist* expression and the inability of most of the *Ftx*^{+/-} cells to upregulate *Xist* from the chromosome lacking this region. In order to distinguish between these scenarios, we designed a locked nucleic acids (LNAs) assay, which enabled us to target the *Ftx* transcript while leaving the genomic region intact. LNA Gapmers were designed to target the conserved exons, 8 and 9. We first started by knocking down the *Ftx* transcript in mouse embryonic fibroblasts (MEFs), a context in which *Xist* accumulation is found in a high percentage of the cells. This enabled us to test the technique while overcoming the experimental difficulties arising from the heterogeneity of differentiating ES lines, in which the percentage of cells with *Xist* cloud rarely exceeds 60%.

Female MEFs were transfected with either a scramble LNA or an LNA targeting *Ftx*. MEFs from either condition were morphologically identical and no increase in cell death could be detected in the *Ftx*_{LNA} population (Figure 10A). In *Ftx*_{LNA} MEFs, the spliced form of the transcript was efficiently knocked down, but the levels of *Xist* RNA remained unchanged (Figure 10B), suggesting that *Xist* can be correctly accumulated on the Xi in the absence of *Ftx* transcript. We also analyzed *Ftx* and *Xist* expression patterns at the single cell level by RNA-FISH. In control MEFs, like in undifferentiated and differentiating ES cell lines, *Ftx* is expressed the majority (~90%) of the cells and the transcript can be found on the X chromosome as well as accumulating at discrete foci in the totality of the nuclear space (Figure 10C). In *Ftx*_{LNA} MEFs, the LNA mostly targeted the transcript localizing away from the *Ftx* transcription site on the X chromosome (“scattered” *Ftx* signal), which possibly corresponds to the spliced form of the *Ftx* transcript (Figure 10C). In these *Ftx*_{LNA} MEFs, no decrease in the percentage of cells with *Xist* cloud was observed in the whole population and, in each single cell, the *Xist* cloud appeared normal (Figure 10C). Collectively, these results suggest that the *Ftx* (spliced) transcript is dispensable

for proper *Xist* accumulation, at least in differentiated cells. However, since the levels of *Ftx* expression in MEFs are similar to those in undifferentiated ES cells but lower than those in differentiating cells, it is possible that *Ftx*- as locus- is not required for the maintenance phase of XCI in post-XCI cells.

In order to investigate if the *Ftx* transcript is required for XCI establishment, we performed the same experiment in differentiating ES cells. Unlike *Ftx*_{LNA} MEFs, *Ftx*_{LNA} differentiating ES cells appeared to die at high rate (Figure 10D), a phenotype reminiscent of that observed upon differentiation of *Ftx*^{+/-} and *Ftx*^{-/-} cell lines (Figure 3). In the surviving *Ftx*_{LNA} cells, the *Ftx* transcript was efficiently knocked down, but *Xist* levels did not change drastically (Figure 10D), hinting that *Ftx* RNA is dispensable also for XCI establishment. However, the fact that most of the *Ftx*_{LNA} cells die during differentiation and are therefore not analyzed might suggest that the transcript is in fact required and that the surviving cells represent just a fraction of the population, corresponding exactly to those cells in which *Xist* can be upregulated in the absence of *Ftx* transcription, phenomenon that is seen in a small percentage of differentiating wild-type cells (Figure 5). Nevertheless, since cell death was observed also in control cells (possibly linked to the transfection process), we cannot confirm that the mortality of *Ftx*_{LNA} cells is due exclusively to the fact that *Ftx* is knocked down. Careful comparison of the levels of cell death in the two populations would be required, but our current experimental settings (low numbers of cells can be transfected with a given amount of LNA solution) prevented this analysis so far.

To try to assess the relevance of the *Ftx* transcript without being biased by the elevated cell death, we took advantage of the *cis* phenotype observed in *Ftx* heterozygous populations. The *cis* phenotype can be detected independently of cell death because the surviving cells are (mostly) the ones that chose to inactivate the X chromosome carrying the *Ftx* wild-type allele, hence showing skewed XCI configurations. We reasoned that, by introducing the *Ftx* LNA in PGK1 *Ftx*^{+/-} cell lines we would specifically target the transcript produced by the wild-type *Ftx* allele. The cells would thus be equivalent to full *Ftx* knock out cells with regard to the transcript, but, in each single nucleus, one X chromosome would lack the whole genomic region (including the transcript), while the other X chromosome would keep the genomic region intact. If the transcript was required, we would observe a reversal of the skewed XCI patterns in *Ftx* heterozygous cells to the random XCI patterns observed

in wild-type and *Ftx*^{-/-} cells (Figure 8), and if the transcript was not required we would not observe such reversal and the skewed XCI patterns would be maintained. It is also possible that the transcript and the region are equally necessary, for example in case a particular chromatin loop is needed to bring the *Ftx* transcript close to the *Xist* locus. In this case, we would expect a mild reversal of skewed to random XCI patterns. In *Ftx*^{+/-} cells transfected with the *Ftx* LNA we observed that the levels of *Ftx* (spliced) transcript were reduced compared to the LNA scramble control, but to a lower extent than in wild-type cells transfected with the *Ftx* LNA: an equivalent amount of *Ftx* transcript was still present in wild-type and *Ftx*^{+/-} cells transfected with the *Ftx* LNA (Figure 10E). In these lines, we measured the ratios of *Xist* allelic expression and found that the skewed XCI patterns in *Ftx*^{+/-} *Ftx*_{LNA} lines were not completely reversed to random (Figure 10E), suggesting that the *Ftx* transcript is not the main player of the *Ftx*-dependent *Xist* regulation. However, to this day, we cannot exclude that a subtler effect is indeed mediated by the transcript but not detectable due to experimental limitations. For instance, these experiments were performed only once, and statistical significance cannot be assessed for. Additionally, as residual *Ftx* expression is present in *Ftx*_{LNA} *Ftx*^{+/-} samples, RNA-FISH will be needed to analyze the XCI status in each individual cell, including the assessment of *Xist* cloud shapes and volumes. Lastly, it seems that the *Ftx* LNA mostly targets the spliced *Ftx* transcript localizing at distinct nuclear foci, while it does not disrupt the *Ftx* transcript at its locus on the X chromosome in all the cells. It can be possible that this latter fraction of the transcripts is the one that is actually involved in *Xist* regulation. Overall, although these results seem to suggest that the *Ftx* transcript is not absolutely required for *Xist* accumulation during XCI establishment and maintenance, further experiments are needed to formally rule out the participation of the *Ftx* RNA to *Xist* regulation.

Figure 9

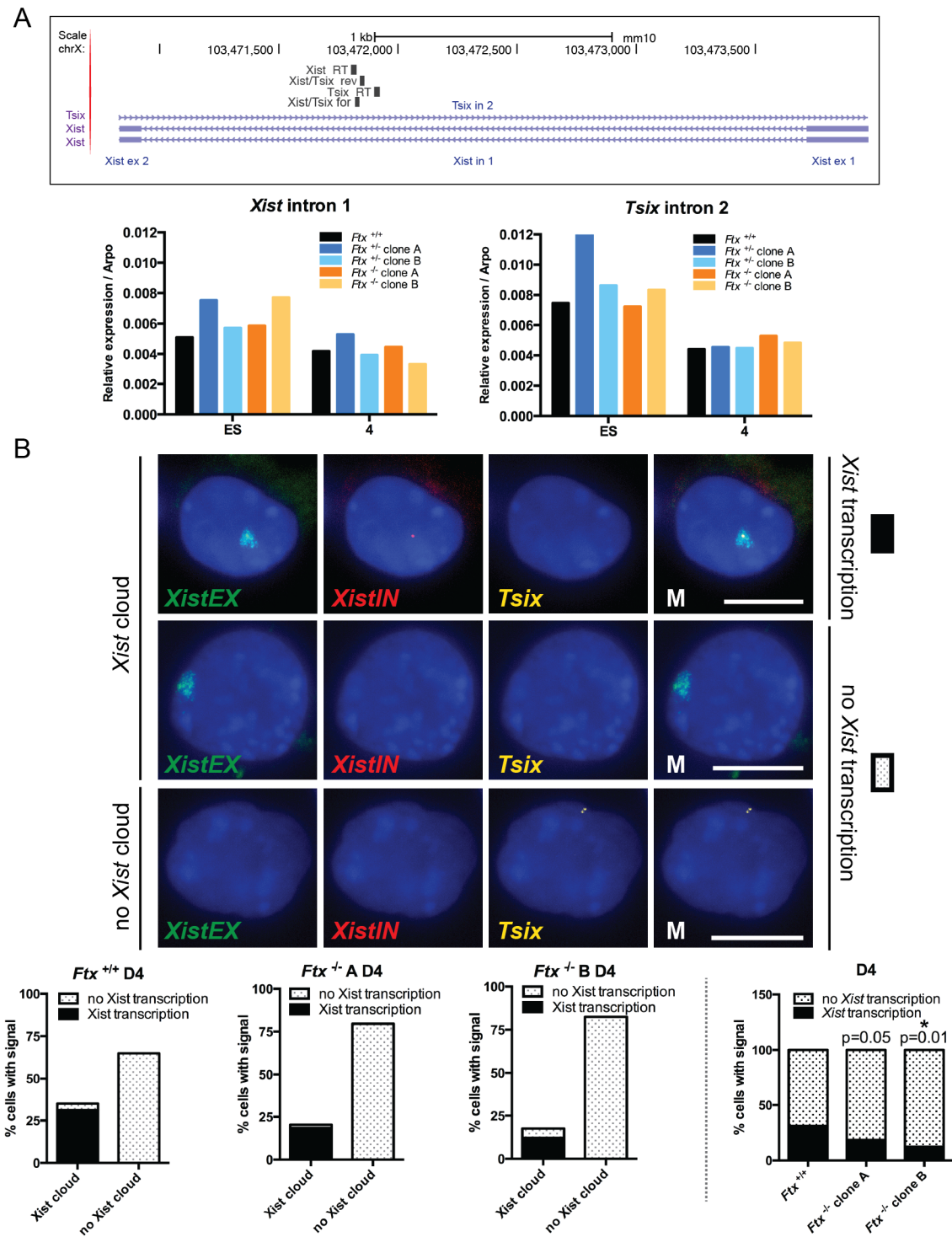


Figure 9. Insights into the mechanisms of *Ftx*-dependent *Xist* regulation at the level of *Xist*

A. Top panel: Schematic representation of the 5' portion of the *Xist* and *Tsix* transcripts, including *Xist* first exon and intron, and *Tsix* second intron, in scale. Reverse transcription (*Xist* RT and *Tsix* RT) and RT-qPCR (*Xist/Tsix* for/rev) primer positions are shown. Bottom panel: Strand-specific RT-qPCR quantification of the pre-mature (unspliced) *Xist* transcript levels and of the pre-mature *Tsix* transcript levels in undifferentiated (ES) and differentiating (D4) LF2 WT and *Ftx*^{-/-} cell lines. One replicate is shown.

B. Top panel: RNA-FISH images of each possible *XistEX/XistIN* configuration, with the *Tsix* signal used as control for positive efficiency of the technique, in differentiating (D4) female LF2 WT and *Ftx* mutant cells. *XistEX* single-stranded probe, FITC-labeled, green. *XistIN* single-stranded probe, Cy3-labelled, red. *Tsix* single-stranded probe, Cy5-labeled, yellow. Bottom panel, left: Quantification of the percentage of each configuration in differentiating (D4) LF2 female LF2 WT and *Ftx* mutant cells. Scale bar = 10 μm. Sample sizes (n) = *Ftx*^{+/+}: 77; *Ftx*^{-/-} A: 103; *Ftx*^{-/-} B: 74. Bottom panel, right: Quantification of the percentage of cells with ongoing *Xist* transcription (*XistIN*), within the cells with signal. Statistical significance was assessed with Fisher's exact test. Sample sizes (n) = *Ftx*^{+/+}: 77; *Ftx*^{-/-} A: 103; *Ftx*^{-/-} B: 74.

Figure 10

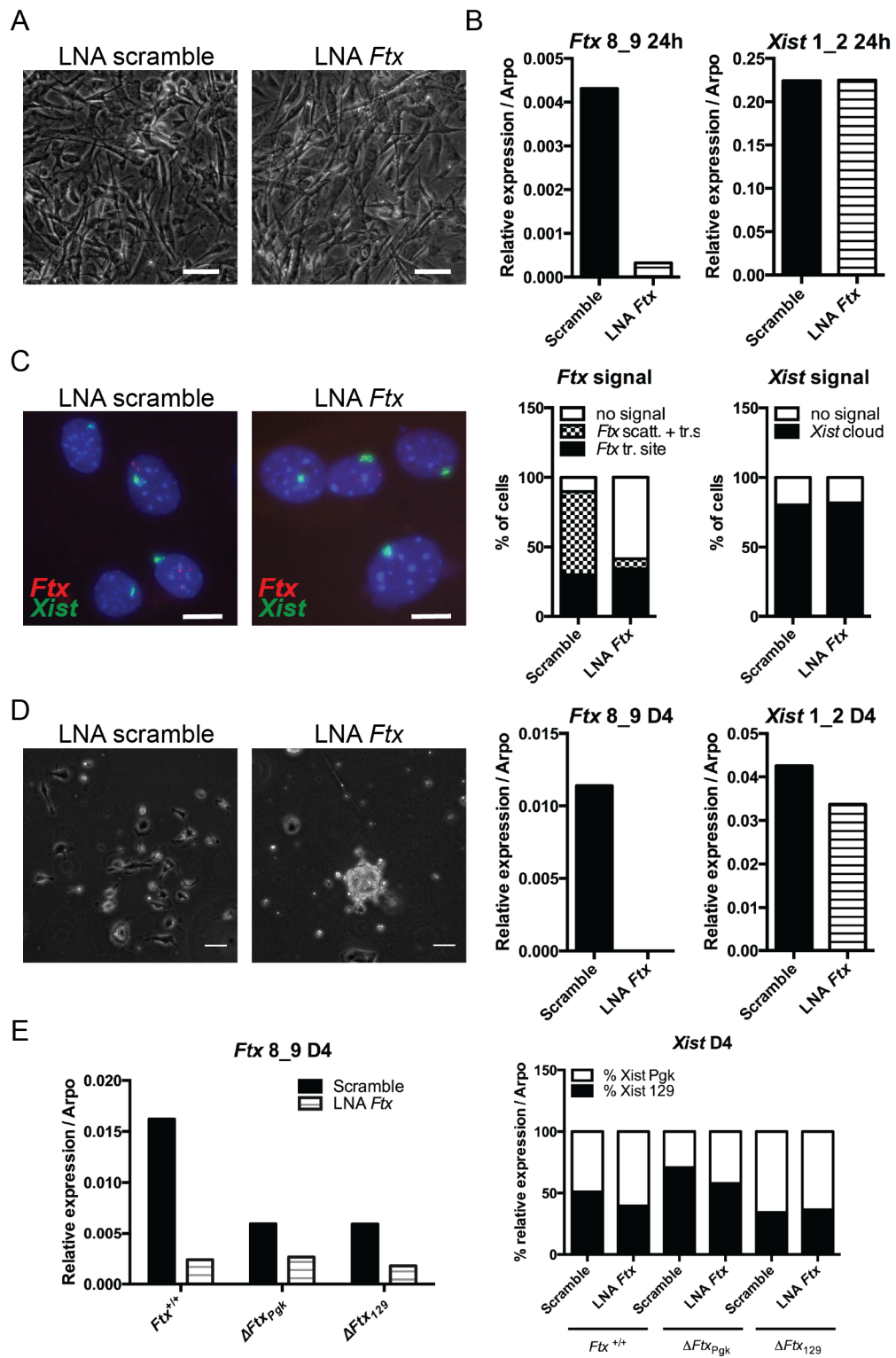


Figure 10. Knocking down the *Ftx* transcript

A. Brightfield photographs of female MEFs transfected with either LNA scramble or LNA *Ftx*. Scale bar = 50 μ m.

B. Analysis of *Ftx* and *Xist* expression by RT-qPCR in female MEFs transfected with either LNA scramble or LNA *Ftx*, 24 hours after transfection. One replicate is shown. Values are normalized to *Arpo* RNA levels.

C. Left panel: Representative RNA-FISH images of female MEFs transfected with either LNA scramble or LNA *Ftx*. *Xist/Tsix* probe, FITC-labeled, green. *Ftx* probe, Cy3-labeled, red. Scale bar = 10 μ m. Right panel: Quantification of the percentages of cells with *Ftx* (left) of *Xist* (right) signal in female MEFs transfected with either LNA scramble or LNA *Ftx*. Sample sizes (n) = LNA scramble: 117, LNA *Ftx*: 98.

D. Left panel: Brightfield photographs of differentiating (D4) female LF2 cells transfected with either LNA scramble or LNA *Ftx*. Scale bar = 50 μ m. Right panel: Analysis of *Ftx* and *Xist* expression by RT-qPCR in differentiating (D4) female LF2 cells transfected with either LNA scramble or LNA *Ftx*. One replicate is shown. Values are normalized to *Arpo* RNA levels.

E. Left panel: Analysis of *Ftx* expression by RT-qPCR in differentiating (D4) female PGK1 cells transfected with either LNA scramble or LNA *Ftx*. One replicate is shown. Values are normalized to *Arpo* RNA levels. Right panel: Ratios of *Xist* allelic expression in differentiating (D4) female PGK1 WT and *Ftx* mutant cells transfected with either LNA scramble or LNA *Ftx*. One replicate is shown. Values are normalized to *Arpo* RNA levels.

These experiments were performed by Nancy Gutierrez-Hernandez (Master 1 student) under the supervision of Dr. Céline Morey.

CONCLUSIONS

X chromosome inactivation is a complex process that relies on multiple complementary networks to orchestrate and maintain the transcriptional inactivation of a single X chromosome exclusively in females and in a developmental manner. In the mouse, proper XCI establishment depends on the interplay between *cis*- and *trans*-acting factors, several of which are non-protein coding, including the master switch, *Xist*. *Xist* is tightly regulated to ensure the female-specific, monoallelic expression that results in X chromosome-wide silencing. Here, we have characterized the contribution of an X-linked non-coding gene, *Ftx*, to *Xist* upregulation at the onset of XCI.

Ftx expression is not sex-specific, but, while differentiating male ES cells lacking *Ftx* behave normally, differentiating female ES cells lacking *Ftx* display a severe phenotype, with high levels of cell death and an increased number of ES-like cells still present in the population during the whole time course of differentiation. Since the *Ftx* deletion is female-lethal but has no functional consequence in male cells, we conclude that *Ftx* is not directly required to ensure proper differentiation kinetics.

The fact that the phenotype is female-specific and the fact that, in the mouse, X chromosome inactivation and the potency status of the cell are intertwined, indicates that the differentiation defects observed in female *Ftx* mutant cells are likely to result from impaired XCI dynamics. In agreement with this hypothesis, female *Ftx* mutant cells fail to correctly upregulate and accumulate *Xist* during differentiation. This effect is robust, as it is observed with different clones and reproduced in two cell lines with different genetic backgrounds. Interestingly, the extent of *Xist* misregulation is correlated with the genotype of the clone and ranges from mild to severe, with cells that lack *Ftx* on both X chromosomes displaying the most striking phenotype. This suggests that *Ftx* is required for correct *Xist* accumulation in a copy-dependent manner. Moreover, in most of the (few) *Ftx* mutant cells in which *Xist* is seen coating, the pattern of accumulation appears abnormal, with the *Xist* cloud being less compact than in wild-type cells and showing a reduced ability to silence X-linked genes.

Importantly, in cells with a heterozygous *Ftx* deletion, impaired *Xist* accumulation is

observed exclusively on the same X chromosome in which the *Ftx* deletion resides, indicating that *Ftx* is required for correct *Xist* upregulation in *cis*.

All together, these results indicate that *Ftx* is required for XCI establishment, where it functions as a strong *Xist* activator and possibly takes part in the choice step of XCI initiation. In this context, *Ftx* represents one important component of the delicate balance of activators and repressors that ultimately converge to dictate the precise *Xist* expression patterns that characterize the XCI process.

DISCUSSION

1 *Ftx* and cell differentiation

The observation, at the very beginning of this project, that *Ftx* expression increases during ES cell differentiation in both female and male samples led us to investigate the contribution of *Ftx* to this fundamental process using female as well as male ES cells depleted for *Ftx*. The absence of severe phenotype in differentiating male *Ftx* mutant cells indicates that *Ftx* does not play a direct role in differentiation. Conversely, in female cells, the lack of *Ftx* has serious consequences, among which a high rate of cell death and the persistence of cells with undifferentiated morphology even further in along the time-course of differentiation. Precise quantification of the rate of cell proliferation and cell death, as well as the characterization of the potency state of the ES-like surviving cells, is currently ongoing using fluorescence-activated cell sorting (FACS).

Surprisingly, quantification of the expression levels of the classical core pluripotency markers did not reveal any statistically significant differences between wild-type and *Ftx* mutant cells. It is however to be noted that downregulation of most of the chosen markers is subtle (*Nanog*, *Rex1*) or almost absent (*Oct4*) in our system: the fact that expression levels are comparable between undifferentiated and differentiating (wild-type) cells makes it difficult to detect potential lack of downregulation in *Ftx* mutant lines. Nevertheless, no differences could be detected also when the marker was correctly downregulated (*Klf4*).

A possibility is that, due to the substantial cell death, the actual phenotype is underestimated and cannot be detected in the surviving population, which is composed of a mixture of differentiated and ES-like cells. Alternatively, the distinct morphology of our *Ftx* mutant cells could be independent from the activity of the traditional pluripotency factors, and rather be the result of the perturbation of independent pluripotency-related pathways. Indeed, a recent analysis on ES cell lines with different chromosome compositions revealed that the MAPK and *Gsk3* cascades are involved in the female-specific coordination between XCI and

differentiation (Schulz et al., 2014). It would be interesting to analyze the expression of these genes in our system. Overall, it is clear that the prominent differentiation defect observed in female *Ftx* mutant cells is indirect, and possibly a consequence of misregulated XCI patterns.

2 *Ftx* and the regulation of *Xist*

Consistent with an indirect role for *Ftx* on differentiation via X chromosome inactivation, *Xist* upregulation is drastically impaired in female cells lacking *Ftx*. At the population level, the amount of spliced *Xist* transcripts is significantly decreased in both ΔFtx heterozygous and ΔFtx homozygous cell lines. Only few cells in ΔFtx populations show a *Xist* domain, indicating that, while *Xist* can be upregulated in the absence of *Ftx*, this is an extremely rare event. The XCI phenotype found in *Ftx* mutant cells is strong and reproducible, as it has been observed multiple times in several independent clones originating from two different mouse ES cell lines (LF2 and PGK1). The only exception to this was PGK1 *Ftx*^{+/-} clone A (ΔFtx_{Pgk}), in which *Xist* expression was not decreased compared to wild-type cells. Nevertheless, a high level of cell death was detected also in this clone, suggesting that, in the surviving cells, the XCI phenotype may just be subtler.

Similar analyses have also been performed using a third ES cell line (EL16.7), in which both ΔFtx heterozygous and ΔFtx homozygous clones were produced, by traditional homologous recombination and CRISPR/Cas9-mediated deletion, respectively. The results obtained with this third line are not included in the present manuscript due to difficulties in obtaining satisfactory differentiation efficiencies and to the fact that the ΔFtx heterozygous clones displayed genomic rearrangements at the *Ftx* locus. Despite these issues, cell mortality during differentiation and decreased *Xist* expression and accumulation were found in all the analyzed EL16.7 clones.

The mere fact that the XCI phenotype is detectable implies that cells unable to upregulate *Xist* can, in part, survive during differentiation. Since cells that do not undergo XCI tend to retain an undifferentiated state, this is compatible with the observed increase in the amount of ES-like colonies in differentiating *Ftx* mutant populations, and likely due to the permissive culture conditions (low-density, LIF-

removal, monolayer differentiation) of the experiment. Indeed, when cells are pushed to differentiate with faster kinetics by growth in LIF-depleted medium supplemented with retinoic acid, the cell death is massive and, in the few surviving cells, the XCI phenotype is not longer detectable (not shown).

In *Ftx* heterozygous cells, the *Ftx* deleted chromosome is almost never coated by *Xist*, indicating that *Ftx* is required in *cis* for *Xist* accumulation. In the light of this, it would be interesting to speculate that *Ftx* is, like the *cis*-acting non-coding gene *Tsix*, involved in the choice step of XCI. Indeed, both *Tsix*^{+/-} and *Ftx*^{+/-} populations display skewed XCI patterns. However, contrary to the heterozygous deletion of *Tsix*, the heterozygous deletion of *Ftx* results in increased cell death upon differentiation. This suggests that, while *Tsix* influences primary choice, *Ftx* is rather involved in secondary choice, whereby the cells in which the *Ftx* deleted chromosome is chosen for inactivation are selected against.

In *Ftx* mutant lines, a high percentage of cells showing an abnormal *Xist* domain can be found within *Xist*-positive cells, especially in ΔFtx homozygous lines. This abnormal *Xist* cloud could be part of the phenotype, representing cells in which *Xist* cannot be properly upregulated due to the lack of *Ftx*, but that manage to survive nevertheless, or are about to die and detach. These cells could be used to explore the molecular basis of *Ftx* contribution to *Xist* upregulation (see following sections for discussion). However, as a small percentage of cells with abnormal *Xist* cloud can be seen also in wild-type cells, the presence of many of these in *Ftx* mutant cell lines could also just be a result of the high levels of cell mortality in the population, that would skew the relative proportions of the different classes of cells towards an apparent increase in this particular one. In any case, the *Xist* cloud found in *Ftx*^{-/-} cells does not appear to be fully functional, as expression of the X-linked gene *Taf1* is visible on the *Xist*-coated chromosome in a high proportion of cells. In order to better assess X-chromosome activity, analysis of the expression of additional X-linked genes is needed.

3 Mechanisms of *Ftx* action

Our attempts at dissecting which *Xist* processing steps are affected by the *Ftx* deletion resulted in contradicting results. Quantification of the pre-mature *Xist* transcript by RT-qPCR showed no differences in the amount of this form between wild-type and *Ftx* mutant cell lines, suggesting that *Xist* is correctly transcribed in the absence of *Ftx*; on the contrary, strand-specific RNA-FISH using intronic probes indicated that the percentage of cells displaying ongoing *Xist* transcription is significantly reduced in *Ftx* mutant lines compared to wild-type lines, suggesting that *Xist* pre-mature transcript is not present in the totality of the cells.

Beside technical differences in sensitivity of the two assays, one reason for this discrepancy could be that, in *Ftx* mutant conditions, elongation is impaired and *Xist* transcription does not extend beyond the first intron, where the RT-qPCR primers are positioned. In this case, a full length pre-mature *Xist* would not be produced in all the cells in *Ftx* mutant populations. A truncated pre-mature *Xist* would likely be undetectable by the RNA-FISH assay, which relies on probes that are equally distributed along the entire transcript. Attempts have been made to design a gene-specific-RT/RT-qPCR assay to probe further downstream positions along the transcript, but, in all cases, the *Xist*- and *Tsix*- specific retro-transcription primers have so far shown a high level of cross-reactivity, making it impossible to distinguish between *Xist* and *Tsix* gene products. RNA-FISH experiments using a strand-specific probe located entirely within *Xist* intron 1 will be performed instead.

If a full-length *Xist* pre-mature transcript is indeed produced, the observed reduction in mature *Xist* levels could arise from defects in splicing. The lack of correctly spliced, mature transcript could perhaps also explain the formation of abnormal *Xist* clouds, in which the lack of the required number of mature *Xist* molecules prevents X chromosome silencing and compaction. Another, perhaps less likely, justification for the existence of such abnormal clouds (in presence of correctly spliced *Xist* transcripts) would be a defect in *Xist* tethering to the Xi, which would result in the partial delocalization of the mature *Xist* RNA without affecting the pre-mature form. However, it has been observed that when *Xist* tethering to the Xi is prevented (by YY1 knockdown or ectopic transgene squelching) the total amounts of *Xist* spliced transcripts remain stable, despite disappearance of the *Xist* cloud (Jeon and Lee,

2011b). To better assess the localization of the abnormal *Xist* cloud, visualization of the Xi territory by using X chromosome paint probes and assessment of its co-localization with the *Xist* signal will be necessary.

On a side note, our RT-qPCR results show that, at least in our system, *Xist* is upregulated at the post-transcriptional level: in wild-type cell the levels of pre-mature *Xist* are stable between the undifferentiated and differentiating states. It remains to be assessed whether this is due to the limited time-frame considered (4 days) or the phenomenon extends also to post-XCI cells. In the light of this and given our strong XCI phenotype, it would be not so surprising if *Ftx* indeed regulated *Xist* at the post-transcriptional level. Interestingly, another indication of a possible role for *Ftx* in the post-transcriptional regulation of RNAs in general is the localization of the *Ftx* transcript at specific nuclear foci. *Ftx* could, for instance, accumulate inside Cajal bodies, known to be the sites for RNA-related metabolic processes, or inside similar sub-organelles. To explore this possibility, staining of p80/Coilin (that anchors the Cajal bodies to the nucleolus) or of other components of the organelle combined with *Ftx* RNA-FISH could be performed.

4 Transcript, act of transcription and/or genomic region?

Even if the study of the intermediate steps of *Xist* processing can give insights into the mechanisms of *Ftx*-mediated *Xist* regulation, a refined answer can only be provided by the investigation of the contribution to the phenotype of each of the elements that are included in the deleted segment, namely the *Ftx* transcript, the act of transcription at the *Ftx* promoter, and the physical genomic region including the *Ftx* promoters, the first four exons and three introns.

Overall, given that the *Ftx* deletion results in defects of *Xist* upregulation in *cis*, it is highly possible that the structure of the genomic region encompassing the *Xist* and *Ftx* loci plays a major role in the establishment of their genetic interaction. Indeed, it is now widely accepted that genes can be configured into looped structures or chromatin hubs that juxtapose regulatory elements to activate or repress transcription. Examples of long-range interactions have been observed in mammalian systems for the α -globin gene cluster in erythroid cells (Vernimmen et al., 2007), the *Th2*, *IFN γ* and *MHC* loci in T-cells (Eivazova and Aune, 2004; Kumar et

al., 2007; Spilianakis and Flavell, 2004), the *IgH* and *IgK* loci in B-cells (Sayegh et al., 2005), *HoxB1* (Würtele and Chartrand, 2006), and at the imprinted gene clusters *Dx15*, *Dx16* and *H19-Igf2* (Horike et al., 2005; Kurukuti et al., 2006; Reik et al., 2004). Despite the precise mechanisms of loop formation are not yet fully understood, several transcription factors have been implicated in this process (Drissen et al., 2004; Mahmoudi et al., 2002; Petrascheck et al., 2005; Song et al., 2007), highlighting the possibility that the transcription machinery and the act of transcription (if not the transcript themselves) could be required for the coordination of gene expression at linearly distant loci or to establish the required chromatin environment for proper gene expression at specific individual loci. Looping interactions mediated by transcription factors before the onset of transcription can also exist, and have been called poised (and pre-poised), but their functional significance remains to be assessed (Spilianakis and Flavell, 2004). In the case of *Ftx* and *Xist*, a chromatin loop could be for example required to bring the two promoter regions in close proximity and favor the recruitment (either via the transcription machinery or independently) of the splicing/stabilizing factors needed for *Xist* RNA processing and accumulation. Considering all these scenarios, different but complementary working models can be proposed (Figure 1). Firstly, *Ftx* could act as a transcript and promote *Xist* accumulation in *cis* through the recruitment of nuclear factors that would mediate *Xist* processing. In this scenario, the activity of the transcript could be independent from any three-dimensional chromatin structure, but be favoured by a conformation able to bring the *Ftx* product close to the *Xist* locus on the same X chromosome. Secondly, if the *Ftx* product is not required for *Xist* accumulation but is present just as a result of the need for ongoing transcription, the main mediator of the *Ftx*-induced XCI phenotype could be the need to recruit the transcription machinery in the vicinity of the *Xist* promoter or the transcription-induced deposition of specific epigenetic marks at the *Xist* locus. An act-of-transcription-based mechanism has for example been proposed for *Tsix* (section INTRODUCTION 5.1.3). However, while *Tsix* localization as *Xist* overlapping antisense gene is enough to explain the local influence of *Tsix* on *Xist*, in the case of *Ftx* specific chromatin loops would be needed to bring *Ftx* and *Xist* promoter regions in close contact. An important role for chromatin structure and the existence of a particular three-dimensional organization

at the *Ftx* and *Xist* loci is envisioned also in the third model, in which chromatin looping via anchoring factors such as CTCF and cohesins is proposed to stabilize the *Xist* TAD and allow *Xist* expression in *cis*. Preliminary, ongoing and future experiments are aimed at testing these different hypotheses.

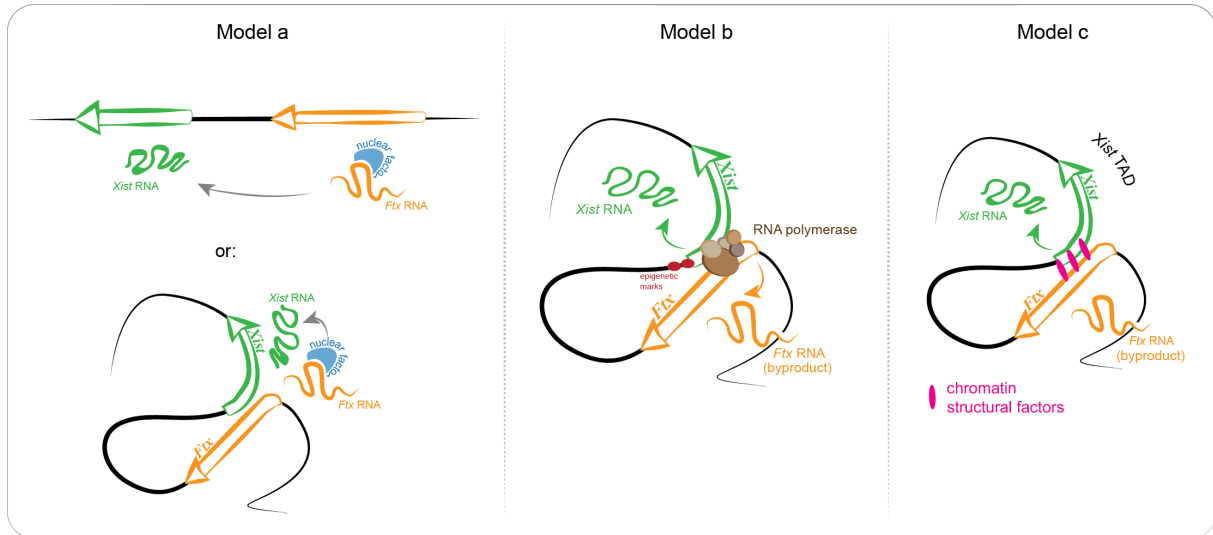


Figure 1. Working models

Three working models explaining the *Ftx* mechanism of action. **Model a.** *Ftx* is required as transcript to promote *Xist* upregulation in *cis*, possibly by recruiting the nuclear factor(s) needed for *Xist* processing. This mechanism of action could be favoured by a specific three-dimensional chromatin structure (bottom). **Model b.** The act of transcription across the *Ftx* locus contributes to the positioning of the transcription machinery in the vicinity of the *Xist* gene, thereby allowing *Xist* transcription or the deposition of activating chromatin marks at the *Xist* promoter. The *Ftx* transcript is produced just as a result of the need for active transcription in the region. **Model c.** Structural elements bind at the *Ftx* and *Xist* promoter regions, structuring a specific loop needed to stabilize the *Xist* TAD and allow *Xist* expression. Once again, the *Ftx* transcript is dispensable, and produced just as result of a transcription in this particularly active chromatin region.

4.1 Exploring the role of the *Ftx* transcript

Preliminary experiments have been carried out to test the first model. The *Ftx* transcript was knocked down using LNA Gapmer oligonucleotides in MEFs and in differentiating ES cells and, while the results are not yet conclusive, it seems that the lack of the *Ftx* transcript is not the main cause of impaired XCI dynamics in females. However, the mechanisms by which LNAs induce the degradation of their targets are not fully unravelled, so that it is challenging to adapt the technique to specific needs. In our case, the LNA efficiently downregulates the transcript localizing at nuclear foci

(scattered signal), but does not seem to affect equally the *Ftx* molecules at the transcription site on the X chromosome: if this latter fraction of transcripts is the one mediating the *cis*-regulation of *Xist*, using LNAs to downregulate it may not be the best option. An additional possibility would be to exploit another mediator of RNAi pathway, siRNAs, and test if this method is able to induce rapid degradation of the transcript being generated at the transcription site.

4.2 Exploring the role of ongoing transcription

Investigating the role for active transcription at the *Ftx* locus is challenging because of the complex promoter usage and isoform generation. Since it is commonly known that promoter usage and alternative splice forms can be tissue-specific and vary with cell-type, we measured the activity of all promoters by quantifying the amount of the first exon for each TSS in a context that is relevant to our study, differentiating ES cells (Figure 2). The results show that, in undifferentiated ES cells, P4 is the major promoter, followed by P1. The situation is maintained also in differentiating ES cells (D2 to D6), where the activity of P1 increases slightly but does not reach that of P4. The same tendency is found also in *Ftx* heterozygous cells. In order to efficiently prevent transcription at the *Ftx* locus in differentiating ES cells, disruption of P4 and P1 is thus required. On a side note, the usage of promoter P3 could not be assessed equally, since exon 3 is relatively short (96 bp) and partially spans a repeated stretch, making it difficult to design specific RT-qPCR primers fully contained within the exon. Expression of intron 3 was quantified instead, but alternative methods, such as 5' RACE and quantification of the distribution of RNA Pol II occupancy at the *Ftx* locus might be needed to refine the analysis.

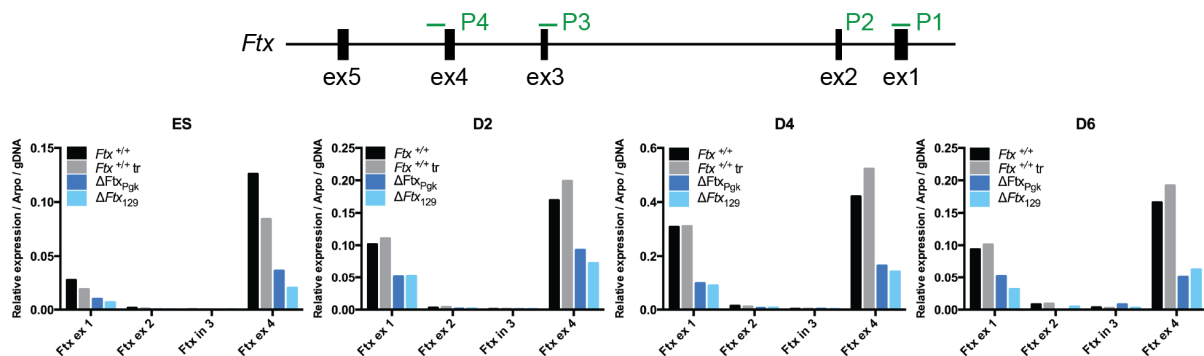


Figure 2. *Ftx* promoter usage in differentiating female ES cells

RT-qPCR analysis of *Ftx* promoters usage in differentiating female PGK1 WT and *Ftx* mutant cell lines. Values are normalized to the amount of gDNA to adjust for primer efficiency and to

Arpo RNA levels.

Inhibition of transcription at promoters can be achieved in a couple of ways. Recently, the CRISPR technology has been adapted to provide gene repression by transcriptional interference (CRISPRi) (Gilbert et al., 2013). This method has been successfully used to recruit a transcriptional repressor at target loci, thereby preventing transcriptional activation. While it would be interesting to design a CRISPRi assay to target the *Ftx* promoters, it is still not clear if the dCas9 fusion proteins cause steric effects that could affect chromatin looping. An alternative to this method is the disruption of the core promoters by CRISPR Cas9 nickase (Wen et al., 2016), which would prevent promoter recognition by the transcription machinery. Finally, an elegant experiment to perturb transcription without drastically affecting the genomic region would be to induce the inversion of the segment containing the promoter(s). This can be achieved by inducing DSBs at the extremities of the selected region: just like deletion, inversion is one of the possible outcomes of the excision of a DNA segment. It is possible that we already produced such mutations while generating *Ftx* mutant cell lines, but we did not screen the clones for the presence of this event.

4.3 Exploring the role of the genomic region

Other than for correct establishment of expression patterns in general, the three-dimensional architecture of the genome, especially on the X chromosome, has been shown to be important for the initial steps of *Xist* spreading, which exploits genomic proximity and topology to spread in *cis*. All these considerations imply that, for example, the formation of an abnormal *Xist* cloud as observed in our *Ftx* mutant cells could result from decompaction of the chromatin structure within the *Xist* TAD prior to *Xist* upregulation due to missing structural segments inside the *Ftx* locus, or from the loss of some of the spatially proximal focal binding sites needed for *Xist* spreading during XCI initiation.

Using chromosome conformation capture it is now possible to investigate looping interactions between enhancers, silencers, boundary elements and promoters at individual loci. The original 3C method (Dekker et al., 2002) and some of its adaptations (4C, 5C, Hi-C) have contributed to the understanding of the three-dimensional structure of the X chromosome in general and of the *Xic* in particular

(Nora et al., 2012). However, these techniques (except some versions of 4C) are not suitable to probe specific short-range interactions in detail. To this regard, a novel method, called Capture-C, has been recently developed (Hughes et al., 2014). Capture-C enables the identification of *cis*-acting elements interacting with a chosen region of interest at high resolution.

By using the *Xist* promoter as target region, Capture-C could be used in our case to explore if a looping is created between the *Xist* and *Ftx* promoters, and to assess if an existing contact is disrupted in *Ftx* mutant cells. A caveat of the technique is that it only provides averages of the interactions in question, which hinders the analysis of individual loci in rare or less abundant cell populations, such as *Xist*-positive cells in differentiating ES populations. An additional complication for X-linked loci in female cells is given by the fact that it is impossible to know if the target interaction is occurring on the Xa or on the Xi. The direct visualization of an interaction occurring on a specific chromosome can be obtained by high resolution 3D FISH with DNA probes (coupled to RNA-FISH to identify the Xi in the case of the X chromosome). However, the current resolution limits of fluorescence microscopy preclude a clear distinction between looped and linear conformation of closely spaced sites: given that the distance between the *Xist* and *Ftx* promoters does not exceed 140kb, this could be a major impairment in the context of our study. Another general limitation of 3D FISH methods is that, just as co-expression does not mean reciprocal regulation, juxtaposition or co-localization does not necessarily imply that the two regions are functionally interacting.

An alternative and complementary way to investigate the contribution of the genomic sequence to the phenotype that we observe in *Ftx* mutant clones includes the identification of possible contact sites based on the information that CTCF and cohesins are usually the mediators of such interactions. By mapping the binding sites of these factors it is possible to pinpoint small regions putatively involved in the generation of chromatin loops, and by deleting them it is possible to determine their importance in the context of our *Ftx*-mediated XCI phenotype. Interestingly, in PGK1 *Ftx*^{+/-} clone B (ΔFtx_{129}), the small deletion present also on X_{Pgk} partially removes a cohesin and a mediator binding-sites (Figure 3): the fact that the XCI phenotype in this clone is almost as strong as in the *Ftx*^{-/-} clone suggests that these two elements

could be involved in the formation of some kind of three-dimensional structure. Future experiments will be aimed at the precise mapping and deletion of such binding sites, followed by the investigation of *Xist* expression and XCI dynamics. We obtained a preliminary visualization of multiple binding sites in undifferentiated ES cells by analyzing published ChIP-sequencing data (Creyghton et al., 2010; Kagey et al., 2010) (Figure 4). Our next goal will be to collect similar data for differentiating cells, perform ChIP-qPCR experiments to investigate if such factors are bound to candidate sites in our system and finally proceed with the deletion of relevant ones.

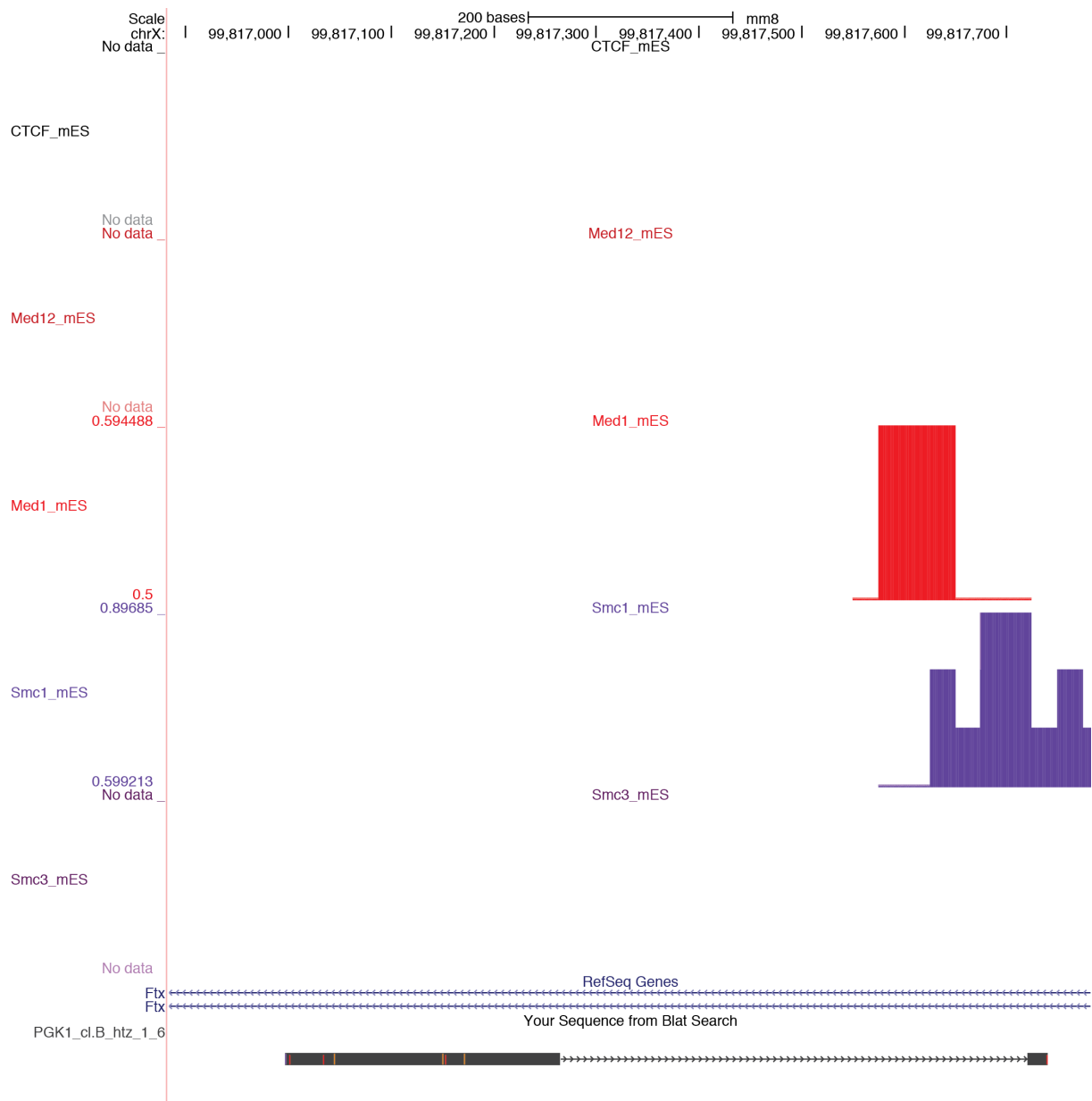


Figure 3. Binding sites mapping to the X_{Pgk} deletion in PGK1 clone B (ΔFtx_{129}).

Alignments of ChIP-seq reads for CTCF, cohesins and mediator in mouse undifferentiated ES cells. The black sequence corresponds to the aligned PCR product generated from amplification of the 3' junction (CRISPR site) at the *Ftx* locus on X_{Pgk} in PGK1 clone B (ΔFtx_{129}). Data from *Kagey et al., 2010* and *Creyghton et al., 2010*.

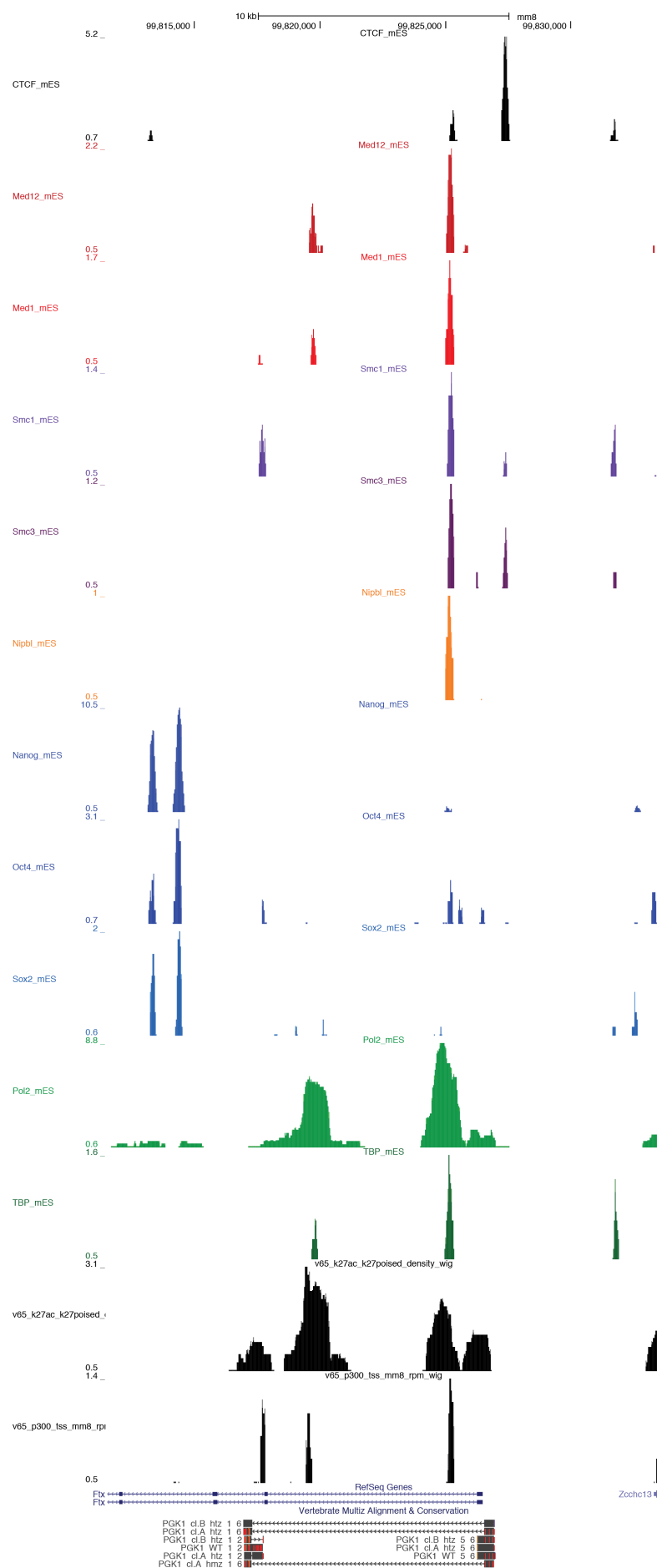


Figure 4. Protein factors binding sites mapping to the *Ftx* deleted region

Alignments of ChIP-seq reads for multiple nuclear factors, including CTCF and cohesins in mouse undifferentiated ES cells. Data from *Kagey et al., 2010* and *Creyghton et al., 2010*.

5 *Ftx* and other XCI regulators

Multiple layers of regulation provide coordination between XCI, pluripotency and differentiation, and the complexity is such that affecting this network partially might have detectable effects in some contexts but not in others. Because of this, contradicting data are often generated and it is not always easy to integrate results from independent studies.

Results from our experiments indicate that *Ftx* is required for *Xist* upregulation in *cis*, and that removal of *Ftx* leads to female-restricted lethality upon differentiation due to XCI failure. Other XCI regulators such as *Tsix* and *Rnf12* behave normally in ΔFtx conditions, at least with regard to their expression profiles, while others, such as *Jpx*, are misregulated, but unlikely to represent the main cause of the phenotype. This is partially in disagreement with what reported from studies of *Jpx*^{+/-} ES cell lines, in which the reduction in the levels of the *Jpx* transcript are enough to induce cell death and result in impaired *Xist* expression from either X chromosome (in *trans*) in female differentiating cell populations (Tian et al., 2010). Perhaps in our system *Jpx* expression does not drop below the threshold levels required for *Jpx*-mediated *Xist* upregulation. This could explain for example why some ΔFtx cells can upregulate *Xist* on *Ftx*-deleted chromosomes. While we cannot state for sure what the mechanism of *Jpx* downregulation in ΔFtx cells is, it is possible that the genomic region and its three-dimensional architecture play a role: *Jpx* and *Ftx* reside in the same TAD and have positive effects on *Xist*, so that it is conceivable that they are somehow brought in close proximity and co-regulated.

While the independent heterozygous deletions of *Ftx* (our study) and *Jpx* (Tian et al., 2010) both result in impaired *Xist* upregulation and have strong consequences for XCI patterns, their heterozygous co-deletion barely affects *Xist* expression in differentiating female ES cells and does not result in cell mortality (Barakat et al., 2014). An explanation for this discrepancy could reside once again in how the whole region is three-dimensionally structured: in the latter example it is possible that the

deletion of such a large domain (which includes, other than the full *Ftx* and *Jpx* loci, also the *Xpr* region) impacts on the *Xist* TAD in a manner that does not compromise *Xist* expression, while removal of smaller but critical stretches destabilizes intra-TAD contacts critical for *Xist* transcription, stabilization or spreading in the former examples. Another consideration is that the ES cell lines used in these three studies (ours, Tian et al., 2010 and Barakat et al., 2014) have different genetic backgrounds and the differentiation protocols used are not exactly the same.

Despite the reported differences in the extent of the contribution of *Ftx* to XCI, one conclusion that our study has in common with that of Barakat et al., 2014 is the implication of *Ftx* in *Xist* cis-regulation. Just as deleting the *Ftx* promoter region results in impaired *Xist* upregulation from the same X chromosome (our results), combining the deletion of the large *Ftx-Jpx-Xpr* segment with the deletion of *Rnf12* impairs *Xist* upregulation to a higher extent than the *Rnf12* heterozygous deletion alone and partially reverts the skew towards the *Rnf12*-deleted chromosome observed in *Rnf12*^{+/-}/*Ftx*^{+/-}/*Jpx*^{+/-} cells (Barakat et al., 2014).

All in all, despite differences and sometime contradicting data can be found in the literature, we can at least suggest that *Ftx* cooperates with *Jpx* and *Rnf12* in the activation of *Xist* in differentiating female ES cells.

6 *Ftx* and XCI in other eutherian species

An interesting open question regards the conservation of the function of *Ftx* in XCI across eutherians. Within this lineage, there are substantial variations in the mechanisms of X chromosome inactivation, both in the developmental timing of its establishment and, especially, in its regulation. In mice and in humans, the *Xic* and *XIC* regions, despite being homologous and syntenic, differ at the structural level and in their three-dimensional organization. Within the *Xic/XIC*, there are examples of genes whose function is conserved, and genes whose function is not: for instance, *Xist/XIST* is absolutely required for XCI in the two species, while *Tsix* is not conserved.

Hints of functional conservation usually rely on the degree of structural/linear conservation of the genomic region and of the gene product. While this reasoning might help to identify conserved functions for protein-coding genes, the situation is

more complicated in the case of lncRNA loci, whose products evolve rapidly, can adopt various three-dimensional structure or possibly even exist without having a role *per se*. In the case of *Ftx*, the RNA transcript is fairly conserved, and predicted to adopt a stable loop structure (corresponding to exons 8 and 9) across species. However, our data suggest that the transcript itself may not be involved in the regulation of *Xist*, hence being dispensable for XCI. The study of the functional conservation of *Ftx* in the XCI process is hindered by another issue, the current lack of easily manipulable *ex vivo* systems that can pose as counterpart of differentiating murine ES cells in the recapitulation of XCI establishment. For example, human ES cells in culture display an XaXi status even before differentiation and, while an increasing number of studies are focusing on the generation of naive XaXa hESCs, routine manipulation of these cells is not yet in place. Surely, progress in this field will be made in the upcoming years, and the evaluation of the functional conservation of *Ftx* will be extended to the human and, eventually, to other species as well.

7 From an *ex vivo* to an *in vivo* model

While *ex vivo* cellular models are useful and practical tools to investigate the mechanisms of gene regulation in general, it is also important to evaluate to what extent the conclusions drawn in these systems hold true in *in vivo* contexts. In particular, *in vivo* systems represent perhaps the best context into which assess the physiological relevance of molecules like lncRNAs, whose functionality- or lack thereof- is still debated. At the same time though, if lncRNAs are mostly needed to fine-tune the expression of other genes, their lack could be compensated for by other genes, or by the action of other lncRNAs.

In order to explore the contribution of *Ftx* to cellular processes *in vivo*, the generation of *Ftx* deficient mice lines (C57BL/6 genetic background) is currently ongoing in the laboratory. Given the ubiquitous expression of *Ftx* (section INTRODUCTION 6.3), the experimental design envisioned the generation of an inducible *Ftx* deletion, which could be triggered in a spatio-temporal manner, either in specific organs and tissues of the adult organism or constitutively during the first steps of embryonic development. Importantly, this strategy would allow to bypass critical steps in case of embryonic lethality, a possibility if the XCI phenotype is reproduced *in vivo*. Initial

experiment performed in the laboratory showed that conditional animals of both sexes are healthy and fertile, and that it is possible to obtain fertile constitutive *Ftx* hemizygous and *Ftx* heterozygous mice.

Analyses of male mice are ongoing to explore the consequences of the *Ftx* deletion in the brain and in the blood compartment, while crosses are being generated to assess the relevance of the *Ftx* deletion in iXCI and rXCI *in vivo*. In order to evaluate the parental effect of the *Ftx* deletion inheritance in both iXCI and rXCI, crosses between *Ftx* heterozygous females and *Ftx* wild-type males (maternal inheritance of the deletion) and between *Ftx* wild-type females and *Ftx* knock-out males (paternal inheritance of the deletion) have been initiated (Table 1) or planned. Crosses between *Ftx* heterozygous females and *Ftx* knockout males are also in progress to evaluate if *Ftx* knockout females can be viable (Table 1).

So far, we assessed the distribution of the offspring genotypes generated by the cross between *Ftx* heterozygous females and *Ftx* wild-type males and found that it was significantly different from expected (1/4 of the total for each possible genotype) ($p=0.02$, Chi-square test). We also tested the proportion of genotypes within each sex and found that, while the proportion of *Ftx* mutant males was not significantly different from the proportion of wild-type males (and than expected, $p=0.1$, Chi-square test), the proportion of *Ftx* mutant females was significantly smaller than the proportion of wild-type females, (and than expected, $p=0.01$, Chi-square test). This suggests that maternal inheritance of the *Ftx* deletion results in reduced survival exclusively in female individuals. Assuming that *Ftx* is required in *cis* for *Xist* upregulation also *in vivo*, the observed embryonic lethality is unlikely to be due to defects in imprinted XCI, since the paternal X chromosome carries a wild-type copy of *Ftx* and can therefore theoretically be inactivated. To confirm this hypothesis, analysis of embryos at different embryonic stages of development is needed: iXCI defects usually result in death at embryonic day E6.5/8.5 following collapse of placental tissues due to lack of dosage compensation, but the embryo proper can develop, albeit to a variable extent (Kalantry et al., 2009; Marahrens et al., 1997). Complementary information on the link between *Ftx* and iXCI will be obtained with the analysis of the effects of the paternal inheritance of the *Ftx* deletion: if *Ftx* was required for iXCI, paternal inheritance of the *Ftx* deletion would likely result, as in

$X_p^{\Delta Xist}/X_m$ mice, in iXCI failure in the extra-embryonic tissues. It is however to be noticed that the *Ftx* gene is not entirely included in the 210 kb region (Tg80 transgene) shown to harbor sufficient factors to trigger imprinted XCI in autosomes in preimplantation embryos when inherited paternally, implying that full length genes other than *Ftx* are more likely to perform this function (Okamoto et al., 2005).

The female-specific embryonic lethality observed upon maternal inheritance of the *Ftx* deletion could rather derive from defects in random XCI. To explore this possibility, careful analysis of the developing embryo proper should be performed. It is reasonable to speculate that defects in rXCI would result in embryonic lethality due to failure of dosage compensation in half of the cells constituting the embryonic tissues, or give rise to smaller/developmentally-retarded embryos if the surviving cells can compensate for such loss. In both cases, the extra-embryonic tissues should undergo normal iXCI. However, two pioneer studies revealed that the maternal inheritance of a *Xist* deletion does not cause female-specific lethality, but only results in complete non-random X inactivation (Penny et al., 1996; Marahrens et al., 1997). Similar conclusions were also obtained from studies of ES cells carrying a heterozygous *Xist* deletion (Penny et al., 1996). Interestingly, allelic expression data for X-linked genes in these cells are compatible with a model envisioning secondary non-random X inactivation, in which the counting and choice functions of XCI are unaffected by the *Xist* deletion and, as a consequence, the cells that chose to inactivate the *Xist* deleted chromosome are counterselected and do not engage into differentiation. Curiously, an increase in cell death has not been reported in this system, implying that *Xist* mutant XaXa cells can survive to some extent in culture. While this is not exceptional in *ex vivo* systems (we observe the same tendency), it is easy to imagine that the XaXa state would be deleterious *in vivo*. The fact that embryonic lethality of $X_m^{\Delta Xist}/X_p$ individuals has not been observed *in vivo* in these studies (Penny et al., 1996; Marahrens et al., 1997) could be due to the nature of the experiment (in both cases, embryo aggregates were produced) and to the fact that only very small numbers of individuals were considered.

In any case, it will be informative to investigate if non-random XCI is a feature of our *Ftx*^{+/-} animals. To this end, *Xist/Ftx* RNA-FISH of adult or embryonic cell types and assessment of allele-specific expression of X-linked genes in the female offspring of

crosses between *Ftx*^{+/-} females and males with a different genetic background could be done.

Looking at the numbers of individuals from our crosses, it can be noticed that *Ftx* wild-type females are more abundant than *Ftx* wild-type males. Even if rodents, like most mammals, tend to produce roughly equivalent numbers of males and females, litters with imbalances in sex ratio can occur spontaneously. However, because of this, we cannot completely exclude that more male individuals are needed for the statistical test to be powerful enough to detect a subtle lethality phenotype, in this case be XCI-independent, affecting also the male sex. Crosses between wild-type individuals are now ongoing and will be used as reference to test crosses between *Ftx* mutant lines for the sex ratio of the offspring in this genetic background under the environmental conditions at our facility.

Finally, crosses to generate *Ftx* knockout females are currently under way and, while some female *Ftx* knockout individuals are born and are viable, the number of pups is still too small to allow proper statistical testing of the relative proportions (Table 1).

Cross	Male pups		Female pups		Total pups
(♀ X ♂)	<i>Ftx</i> +/Y	<i>Ftx</i> -/Y	<i>Ftx</i> +/+	<i>Ftx</i> +/-	
<i>Ftx</i> +/- X <i>Ftx</i> +/Y	64	50	80	53	
	114	p=0.1	133	p=0.01*	247
Cross	Male pups		Female pups		Total pups
(♀ X ♂)	<i>Ftx</i> +/Y	<i>Ftx</i> -/Y	<i>Ftx</i> +/-	<i>Ftx</i> -/-	
<i>Ftx</i> +/- X <i>Ftx</i> -/Y	4	4	4	3	
	8		7		15

Table 1. Mating of ΔFtx mice

Results from crosses between *Ftx* heterozygous females and *Ftx* wild-type males (top) and between *Ftx* heterozygous females and *Ftx* knockout males (bottom). Within each sex, statistical significance was assessed with a Chi-square test.

This analysis was performed by Dr. Jean-François Ouimette.

While our analyses were ongoing, the contribution of *Ftx* to imprinted and random XCI was evaluated *in vivo* also by another group (Kobayashi et al., 2013; Soma et al., 2014). According to their studies, in preimplantation mouse embryos, *Ftx* and the embedded miRNAs are expressed predominantly from the paternal X chromosome,

together with *Tsix*, *Jpx* and the protein-coding gene *Zcchc13* (Kobayashi et al., 2013). This initial observation and the fact that *Ftx* is the most abundant transcript at the preimplantation stage (together with *Xist*) led the authors to formulate the hypothesis that *Ftx* might be important for the regulation of *Xist* expression in iXCI. However, the *Ftx*-deficient female blastocysts analyzed in their study displayed normal *Xist* expression levels and localization, as well as invariant X-linked gene expression, suggesting that *Ftx* is in fact dispensable for iXCI in mice pre-implantation development (Soma et al., 2014). It has to be noted however that iXCI defects are usually detectable in the extra-embryonic tissues after implantation rather than in the early blastocyst. Because of this, the authors could have missed a potential role for *Ftx* in iXCI after implantation.

Viability of newborn pups from crosses between *Ftx*-deficient individuals was also assessed, but no statistically significant differences were detected in the ratio of expected individuals of female sex, suggesting that not only iXCI, but also rXCI was functional in these mice (Soma et al., 2014). An explanation for the discrepancy between this finding and ours could perhaps reside in the size of the cohorts analyzed: 101 individuals from the *Ftx*^{+/-} X *Ftx*^{+/-} cross were analyzed in their report, while more than 200 are tested in our study. In their case, since only a relatively low number of individuals are considered, the power of the statistical test used might not allow the detection of a low-penetrance phenotype. Overall, this indicates that the *Ftx* deletion has detectable but subtle effects on female survival, at least when maternally inherited.

8 *Ftx* expression patterns: a role for the RNA transcript?

Despite our results indicate that the *Ftx* transcript is not the main effector of the *Ftx* mechanism of action at the onset of XCI, the fact that *Ftx* expression is modulated and dynamic, together with the fact that some *Ftx* molecules localize at distinct nuclear positions and the existence of a remarkable complexity of the alternative promoter usage and splicing patterns suggest that the transcript could be functional in some contexts.

8.1 *Ftx* as a miRNA precursor

The remarkable conservation of the *Ftx* embedded microRNA cluster (miR-347 and miR-421) across species raises the hypothesis that the *Ftx* transcript could be to serve as precursor for the production of the corresponding pri- and/or pre-miRNAs. This scenario is seen for example in the case of the imprinted *H19* locus in humans, which, in some instances, produces a lncRNA transcript serving exclusively as miRNA precursor (Cai and Cullen, 2007; Keniry et al., 2012). In the case of the *Ftx* transcript, the hypothesis of its requirement for miRNA production would be consistent with the lack of mature miR-374 and miR-471 molecules in male *Ftx* mutant cells (Chureau et al., 2011) despite the integrity of the miRNA genomic segment (within *Ftx* intron 12).

Interestingly, the scattered *Ftx* signal that we observe by RNA-FISH in the nucleus of undifferentiated, differentiating and differentiated cells could correspond to the pre-/pri-/mature miRNAs or represent the *Ftx* host transcript undergoing miRNA processing at specific nuclear sites before export into the cytoplasm. Indeed, some miRNAs have been shown to co-localize with specific substructures within the nucleus. For example, miR-206, a regulator of myogenic differentiation, is found not only in the cytoplasm but also in the granular component of the nucleolus in rat myoblasts, where it accumulates in ribosome-rich regions (Politz et al., 2006). Variations in the nucleolar localization of miRNAs can be also found between cells in the same cultures, indicating that their subnuclear localization might represent a transient step of their regulatory process (Politz et al., 2009).

The fact that the microRNA cluster is not expressed when the *Ftx* promoter region is deleted implies that the mature microRNA molecules could be involved in the *Ftx*-mediated *Xist* regulation. Interestingly, a predicted target for miR-421 in human is the ATM protein, whose disruption or inhibition in a murine cell line leads to heterochromatin and silencing defects on reporter genes on the Xi (Ouyang et al., 2005). Although these observations might provide a molecular link between *Ftx* and XCI and thus explain part of the phenotype observed in our female *Ftx* mutant cells, human miR-421 is believed to suppress ATM expression by targeting the 3'UTR of its mRNA, sequence that is not conserved in the murine *Atm* and thus unlikely to represent an equivalent target for murine miR-421. Moreover, the fact that *Xist*

misregulation is observed exclusively in *cis* in *Ftx* mutant cells makes it unlikely for the microRNAs, *trans*-acting, to be the main mediators of the *Ftx*-dependent *Xist* regulation. In any case, a definitive answer regarding the possible contribution of such microRNAs to XCI can only be provided by the deletion of the genomic sequence corresponding to each mature microRNA in female cells.

8.2 A role for the *Ftx* transcript in other contexts

Ftx is expressed also in the hematopoietic compartment of adult mice and in brain tissues. Interestingly, inside the hematopoietic lineages, *Ftx* is highly expressed in the B-cells, one of the rare adult cell types in which *Xist* is competent for silencing. While the fact that these two lncRNAs are co-expressed does not imply that one is required to regulate the other, their co-existence could indicate that both genes are somehow required for the establishment or maintenance of the inactive state in this cell type: perhaps, it is in this context that the *Ftx* transcript plays a role.

As for the brain, albeit most lncRNAs tend to be expressed either exclusively or at their highest levels in this organ (Lipovich et al., 2010; Mercer et al., 2008), raising the hypothesis that they represent merely an output of a permissive open chromatin state, they could also be needed to precisely fine-tune the regulation of important, brain-specific genes. Whether the *Ftx* transcript is produced in the brain as a result of spurious transcription or it is directly involved in the establishment of a precise neural identity remains to be thoroughly assessed, but its peculiar expression patterns seem to suggest that *Ftx* has a role in this context. In the brain, *Ftx* is expressed ubiquitously, but higher levels are registered in the olfactory bulb, dentate gyrus and cerebellum (Sophia Hafner, unpublished results). Strikingly, *Ftx* expression in the brain increases post-natally (Sophia Hafner, unpublished results), hinting that a functional role for *Ftx* in this organ may indeed be a possibility: many brain-related processes, such as the integration of environmental stimuli and the reshaping of neuronal morphology, are initiated only after birth. Moreover, male *Ftx* mutant ES cells display an important deregulation of some *Hox* genes (Sophia Hafner, unpublished results). Aside from specifying body plan axes during embryonic development, *Hox* genes play an important role in the determination of cellular identity of several neuronal subtypes of the mammalian nervous system, from selective muscle innervation to the establishment of synaptic specificity (Philippidou

and Dasen, 2013). In the light of this, it would be fascinating to speculate that a synergistic action of *Hox* genes and *Ftx* is needed to guide the correct maturation of a subset of neurons. Importantly, since the misregulated *Hox* genes are autosome-encoded it is likely that in the brain it is the *Ftx* transcript, rather than the genomic region, that exerts a regulatory function, either directly or indirectly.

These observations bring about the intriguing possibility that the *Ftx* gene has been co-opted by several processes (XCI, neural differentiation and maybe others) to perform different functions via distinct mechanisms (chromatin looping/transcription and via the transcript, and maybe others), and open up the way to future investigation on the conservation of such functions across species.

MATERIALS AND METHODS

Overview of the CRISPR/Cas system

Not long ago, the type II bacterial CRISPR/Cas system has been demonstrated as an efficient gene-targeting technology. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) sequences and Cas (CRISPR-associated) proteins are the two elements of an ancient viral and plasmid defence prokaryotic adaptive restriction system conserved in bacterial genomes (Horvath and Barrangou 2010, Wiedenheft 2012). Cas proteins, CRISPR RNAs (crRNAs), and *trans*-activating crRNA (tracrRNA) form ribonucleoprotein complexes, which target and degrade foreign nucleic acids (Gasiunas 2012, Jinek 2012). CRISPRs represent the memory of the infection, a group of short, directly repeating nucleotide sequences flanked by short, unique DNA fragments from previous infections. Cas proteins are the effectors of the system, able to process the CRISPR sequences into small RNAs and to cleave the infectious DNA elements that match the CRISPR-derived sequences. It was shown that the Cas9 endonuclease from *Streptococcus pyogenes* type II CRISPR/Cas system can be used to produce sequence specific double strand breaks (DSB) in vitro by providing a single guide RNA (sgRNA), obtained by the fusion of the crRNA (a 20 bp sequence homologous to the target DNA) and a tracrRNA-resembling hairpin RNA structure (Jinek 2012).

Cell culture and cell differentiation

Female LF2 and PGK1 ES cell lines were grown at 37°C, 8% CO₂ in Dulbecco's modified Eagle medium (DMEM, Life Technologies), 15% Fetal calf serum (FCS, GIBCO) and 1000 U/mL LIF (Millipore), on gelatin-coated dishes and in absence of feeder cells. Male TC1 cells were grown at 37°C, 8% CO₂ in Dulbecco's modified Eagle medium (DMEM, Life Technologies), 15% Fetal calf serum (FCS, GIBCO) and 1000 U/mL LIF (Millipore), on gelatin-coated dishes and in presence of feeder cells. Cell passaging (1:6 split) was done by enzymatic treatment with trypsin (Trypsin-EDTA 0.05%, GIBCO) for 4 minutes.

Differentiation experiments were carried out on gelatin-coated dishes at a density of 10,000 cells/ cm², in Dulbecco's modified Eagle medium (DMEM, Life Technologies)

and 10% Fetal calf serum (FCS, GIBCO). For the differentiation of the PGK1 cell line, the medium was supplemented with retinoic acid (RA) at a final concentration of 10^{-6} M or 10^{-7} M. RNAs were extracted at day 0, 1, 2, 3, 4, 5, 6 and 9 of differentiation from biological triplicates (independent differentiation experiments). Cells were fixed at day 0, 2, 4 and 6 for RNA-FISH.

CRISPR/Cas9 deletion of *Ftx* promoter region

The CRISPR/Cas9 system was used to delete the *Ftx* promoter region and first four exons in female mouse ES cells. Two target regions were selected: the first positioned just upstream of the first *Ftx* promoter and the second downstream *Ftx* exon 4. DNA oligonucleotides harboring the sgRNA target sequences were designed using the ZiFiT Targeter program (<http://zifit.partners.org/ZiFiT/Disclaimer.aspx>), paying special attention to select the ones with the lowest possible number of predicted off-target sites (0 or 1). Oligonucleotide pairs were annealed to generate short double-stranded DNA fragments with overhangs compatible with ligation into the BsmBI-digested plasmid pMLM3636. The Cas9 expression plasmid pJDS246 and the sgRNA expression plasmid pMLM3636 were obtained from Addgene (Addgene #43860 and #47754). SgRNA guide sequences are listed in Table 1.

5 million cells at approximately 70% confluency were transfected with a total of 3µg of DNA using a Nucleofector 4D System (Lonza) and directly plated at a density of 10cells/cm² for clone picking. One week later, clones were screened for the presence of the deletion event by PCR and real-time quantitative PCR on extracted genomic DNA. Mixed clonal populations were subcloned. Screening primers are listed in Table 1.

Sequencing of PCR products

PCR products were sequenced at GATC-Biotech by Sanger sequencing. Sequencing primers are listed in Table 1.

RNA extraction and real-time quantitative PCR

Total RNA was extracted using Trizol reagent and treated with Turbo DNase (Invitrogen) for 20min at 37°C (Life Technologies). 1µg of RNA was random primed reverse-transcribed at 42°C using SuperScript II reverse-transcriptase (Life Technologies) and Random Primers (Promega). Control reactions lacking enzyme

were systematically run and verified negative for representative samples.

Strand specific RT-PCR was performed to detect the pre-mature (unspliced) *Xist* transcript using primers detecting *Xist* (sense orientation) or *Tsix* (antisense orientation) together with *Arpo*-specific primer for normalization. 250ng of DNase-treated RNA extract were reverse transcribed with SuperScript III reverse-transcriptase (Life Technologies) at 55°C. Control reactions lacking enzyme or primers were verified negative. Control reactions testing each primer individually were also carried out to exclude cross reaction between primers.

To assess RNA levels by real-time quantitative PCR (RT-qPCR), the SYBR Green kit (Life Technologies) was used and plates were run on a Viia7 real-time PCR system (Thermo Fisher Scientific). All samples were run in duplicates, or in quadruplicates for strand-specific RT-qPCR. Expression levels were normalized to the reference gene *Arpo* transcript levels according to the $2^{\Delta Ct}$ method. To assess gDNA amounts, RT-qPCR was performed and sample values were normalized to the reference gene H2A DNA levels according to the $2^{\Delta Ct}$ method. Primer sequences are listed in Table 1.

RNA-FISH

Differentiating ES cells were centrifuged onto SuperfrostPlus slides (VWR) with the Cytospin 3 Cyto centrifuge (Shandon). Cells were fixed with 3% paraformaldehyde (Electron Microscopy Science)/1X PBS for 10 min at room temperature (RT), permeabilized in ice-cold cytoskeleton (CSK) buffer (NaCl, MgCl₂, sucrose, and piperazine-N,N'-bis(2-ethanesulfonic acid [pH 6.8]), freshly supplemented with 1mM EGTA (Sigma Aldrich)/0.5% Triton X-100 (Sigma Aldrich)/2mM Vanadyl Ribonucleoside Complex (VRC, New England Biolabs) for 5 min, and progressively dehydrated in 70%-90%-100% ethanol.

For *Xist* RNA detection, a 19 kb genomic lambda clone probe (p510) was labeled with fluorescent dUTPs (SpectrumGreen-dUTP, Abbott Molecular) by nick translation for 3 hours at 15°C using 1µg of DNA.

For *Ftx* RNA detection, a 39.1 kb fosmid probe (WI1-1177B13, BACPAC) was labeled with fluorescent dUTPs (Orange-dUTP, Abbott Molecular) by nick translation for 3 hours at 15°C using 1µg of DNA.

All the probes generated from fosmids were precipitated with mouse Cot1 DNA

(Invitrogen) using 1/10 3M NaOAc and 2.5 volumes of ethanol, and the *Xist* probe with salmon sperm DNA (Invitrogen), also using 1/10 3M NaOAc and 2.5 volumes of ethanol, for 3 hours at -20 °C. Precipitated probes were washed with 70% ethanol and resuspended in formamide (deionized formamide >99.5%, Sigma Aldrich), then denatured for 7 minutes at 75 °C and pre-incubated for 10 minutes at 37 °C when precipitated with Cot1 DNA.

Slide hybridization was carried out overnight at 37°C, in hybridization buffer (4× saline sodium citrate (SSC, Ambion), 20% dextran sulfate, 2mg/ml BSA (New England Biolabs), and 2mM Vanadyl Ribonucleoside Complex (VRC, New England Biolabs). The slides were subsequently washed three times in 50% formamide/2X SSC [pH 7.2] at 42°C, and three times in 2X SSC at 42°C. Slides were mounted in Vectashield containing DAPI (Vector Laboratories) and visualized under a fluorescence microscope.

For the analysis of *Xist* unspliced transcripts, strand-specific single-stranded RNA-FISH probes were used. For each target (unspliced *Xist*, spliced *Xist*) a mix of multiple 20-mer oligonucleotides, each labelled with a single Quasar 570 or 670 fluorophore, was designed (Stellaris probe designer: www.biosearchtech.com/stellaris) and synthesized. All probe sets contained ≥ 36 oligonucleotides each. Hybridizations were carried out on progressively dehydrated slides overnight at 37°C in hybridization buffer (10% dextran sulfate, 10% formamide/2X SSC). Samples were then washed for 30 minutes in 10% formamide/2X SSC [pH 7.2], and the slides mounted in Vectashield containing DAPI (Vector Laboratories) and visualized under a fluorescence microscope.

DNA-FISH

Cells were prepared as for RNA-FISH and treated with 100µg/ml RNase A (Macherey-Nagel)/2X SSC at 37°C for 30 minutes before dehydration. Cells were then dehydrated in 70%-90%-100% ethanol and denatured in 70%formamide/2X SSC for 7 minutes at 80°C after adaptation for 2 minutes at 60°C in 2X SSC. Another dehydration series in 70%-90%-100% ethanol was performed, at room temperature. Probes were prepared and slide hybridization carried out as for RNA-FISH. The slides were subsequently washed three times in 50% formamide/2X SSC [pH 7.2] at 42°C, and three times in 2X SSC at 42°C. Slides were mounted in

Vectashield containing DAPI (Vector Laboratories) and visualized under a fluorescence microscope.

Microscopy and image analysis

All images were taken on an Axioplan 2 Imaging fluorescence microscope (Zeiss) with a cooled Coolsnap camera (Roper Scientifics) and controlled by the MetaMorph 7.04 software (Roper Scientifics), using a Plan-NEOFLUAR 100X/1.30 oil objective (Zeiss), except for the countings, where images were taken with a Plan-NEOFLUAR 63X/1.25 oil objective (Zeiss). Optical sections were collected at 0.2 μ m steps through each nucleus at different wavelengths (DAPI [345,455], FITC [488,507], CY3 [(625)650,670]). Approximately 30 optical sections per nucleus were collected. Stacks were processed using Icy (<http://icy.bioimageanalysis.org>), and the images are represented as 2D projections of the stacks (maximum projection). 3D modelization of the volume of the signals generated by the *Xist* probe was achieved using the Icy plugin Thresholder, for around n=90 nuclei for each sample.

Knockdown using LNA gapmers

Three different Locked Nucleic Acid (LNA) gapmers targeting *Ftx* exons 8-9 and a control scrambled LNA gapmer devoid of target RNA were designed using the Exiqon online tool. LNAs were delivered via lipofection using the transfection reagent RNAimax (Invitrogen), according to manufacturer recommendations. LNAs were used at a concentration of 50 nM for each transfection. As positive control, an LNA Gapmer targeting the murine Malat1 (metastasis associated lung adenocarcinoma transcript 1) lncRNA was used. A human MALAT1 LNA carrying a FAM fluorophore was used as a control for transfection efficiency.

Generation of <i>Ftx</i> mutant cells		
sgRNA_5'_oligo1	ACACCGACCTGAGTACTGGACTCAGG	CRISPR/Cas9 sgRNA
sgRNA_5'_oligo2	AAAACCTGAGTCCAGTACTCAGGTCC	CRISPR/Cas9 sgRNA
sgRNA_3'_oligo1	ACACCGGGGCGGAAGGTGTGGATGGG	CRISPR/Cas9 sgRNA
sgRNA_3'_oligo2	AAAACCCATCCACACCTTCCGCCCCG	CRISPR/Cas9 sgRNA
1	CAGGAAGTACTACATTGCATCAG	screening of deletion event
6	TTTCTTCAGGCTTTGTAACCTGGG	screening of deletion event
3	GTAAGCCCCGGGACACTT	detection WT non deleted sequence
4	CCAGAATCTTAGTCTCTCCCTGG	detection WT non deleted sequence
seq	AAATCTACATACGTATATTTCCATTCC	deletion event sequencing primer
5	CTTTACCTCCTGCTGAAGCG	CRISPR 5' junction integrity
2	GCCCATGCATATCAATCAGTGAA	CRISPR 3' junction integrity
Strand-specific RT primers		
<i>Xist</i>	CCGGCTGAAAGCTCATTCTT	<i>Xist</i> -specific RT
<i>Tsix</i>	CAAATTAGTGTCTGACCCAGC	<i>Tsix</i> -specific RT
<i>Arpo</i>	TGTCGAGCACTTCAGGGTTA	<i>Arpo</i> -specific RT
RT-qPCR primers		
<i>Ftx</i> Ex 8_9	GCCGCTTTTGAAGAGATAAACT	TGATTATCCTCATGTGTTGCTG
<i>Ftx</i> In 8	GCCGCTTTTGAAGAGATAAACT	GCAAGCCAAGTAGAGTCACTGAG
<i>Ftx</i> In 1 129	AATGCTAAATTACTAGTTTTAAGCAATTAAC	AGATTGGAGGTTGGTTGTTAGC
<i>Ftx</i> In 1 <i>Pgk</i>	ACTAGCAATCCTCAAAGGCACA	AGATTGGAGGTTGGTTGTTAGG
<i>Ftx</i> In 6 129	CCTACCTCCAAGACGCTGAGTT	TTGCTGGATTTGAGGGTTACTTCTAAT
<i>Ftx</i> In 6 <i>Pgk</i>	CTAGTATCCTACCTCCAAGACaCTGAGTA	TGCTGGATTTGAGGGTTACTTCTTAC
<i>Xist</i> Ex 1_2	GGATCCTGCTTGAACACTACTGC	CAGGCAATCCTTCTTCTTGAG
<i>Xist</i> Ex 5_6	GCACTGCTACCCTACTCATGC	GCTGCCTCCCTTTTGTGTGT
<i>Xist</i> 129 Ex 5	GGTTCTCTCTCCAGAAGCTAGGAAGG	TGGTAGATGGCATTGTGTATTATATGG
<i>Xist</i> <i>Pgk</i> Ex 5	GGTTCTCTCTCCAGAAGCTAGGAGAA	TGGTAGATGGCATTGTGTATTATATGG
<i>Xist</i> 129 Ex 5_In 6	GGTTCTCTCTCCAGAAGCTAGGAAGG	GCTACTCACTTGTGTATTATATGGCATGA
<i>Xist</i> <i>Pgk</i> Ex 5_In 6	GGTTCTCTCTCCAGAAGCTAGGAGAA	GCTACTCACTTGTGTATTATATGGCATGA
<i>Tsix</i>	CGCATAGCTGGCAAGCATTT	AAGAAGAGCGTGATAGCCAGCT
<i>Jpx</i>	ATACGTCCTTCCCGGGATTA	GGATTAATGATCAGGGCATGTT
<i>Jpx</i> 129 Ex 4	CGTGGGTCATTTCCAGTGT	GGAAATTTGTTCTCCCTTCTGG
<i>Jpx</i> <i>Pgk</i> Ex 4	CGTGGGTCATTTCCAGTGC	GGAAATTTGTTCTCCCTTCTGG
<i>Rex1</i>	CAGCTCCTGCACACAGAAGA	ACTGATCCGCAAACACCTG
<i>Nanog</i>	AGGATGAAGTGCAAGCGGTG	TGCTGAGCCCTTCTGAATCAG
<i>Klf4</i>	ACCCACACTTGTGACTATGCAG	CCCAGTCACAGTGGTAAGGTTT
<i>Rnf12</i>	ACAAATTAAGAGGGTCCACCA	AGAATCACTGTTGGTCACATCATC
<i>Oct4</i>	CCCCAATGCCGTGAAGTTG	TCAGCAGCTTGGCAAACCTGTT
<i>Arpo</i>	TCCAGAGGCACCATTGAAATT	TCGCTGGCTCCACCTT
<i>Xist/Tsix</i>	GTCCTGACCCAGCCCTCA	AGGCAGCCTGAAGATGGTGA

Table 1. List of primers

Primers used for the construction of sgRNAs; PCR screening primers; sequencing primers; primers for gene-specific reverse-transcription; RT-qPCR primers for amplification of cDNA and gDNA.

REFERENCES

- Agrelo, R., Souabni, A., Novatchkova, M., Haslinger, C., Leeb, M., Komnenovic, V., Kishimoto, H., Gresh, L., Kohwi-Shigematsu, T., Kenner, L., et al. (2009). SATB1 Defines the Developmental Context for Gene Silencing by Xist in Lymphoma and Embryonic Cells. *Dev. Cell* 16, 507–516.
- Ashworth, A., Rastan, S., Lovell-Badge, R., and Kay, G. (1991). X-chromosome inactivation may explain the difference in viability of XO humans and mice. *Nature* 351, 406–408.
- Augui, S., Fillion, G.J., Huart, S., Nora, E., Guggiari, M., Maresca, M., Stewart, A.F., and Heard, E. (2007). Sensing X chromosome pairs before X inactivation via a novel X-pairing region of the Xic. *Science* 318, 1632–1636.
- Avery, B., Bak, A., and Schmidt, M. (1989). Differential cleavage rates and sex determination in bovine embryos. *Theriogenology* 32, 139–147.
- Ayupe, A.C., Tahira, A.C., Camargo, L., Beckedorff, F.C., Verjovski-Almeida, S., and Reis, E.M. (2015). Global analysis of biogenesis, stability and sub-cellular localization of lncRNAs mapping to intragenic regions of the human genome. *RNA Biol.* 12, 877–892.
- Bacher, C.P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., and Heard, E. (2006). Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nat. Cell Biol.* 8, 293–299.
- Bachtrog, D. (2013). Y chromosome evolution: emerging insights into processes of Y chromosome degeneration. *Nat. Rev. Genet.* 14, 113–124.
- Bachtrog, D., Mank, J.E., Peichel, C.L., Kirkpatrick, M., Otto, S.P., Ashman, T.-L., Hahn, M.W., Kitano, J., Mayrose, I., Ming, R., et al. (2014). Sex Determination: Why So Many Ways of Doing It? *PLoS Biol.* 12.
- Bailey, J.A., Carrel, L., Chakravarti, A., and Eichler, E.E. (2000). Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proc. Natl. Acad. Sci.* 97, 6634–6639.
- Bala Tannan, N., Brahmachary, M., Garg, P., Borel, C., Alnefaie, R., Watson, C.T., Thomas, N.S., and Sharp, A.J. (2014). DNA methylation profiling in X;autosome

translocations supports a role for L1 repeats in the spread of X chromosome inactivation. *Hum. Mol. Genet.* 23, 1224–1236.

Barakat, T.S., Gunhanlar, N., Pardo, C.G., Achame, E.M., Ghazvini, M., Boers, R., Kenter, A., Rentmeester, E., Grootegoed, J.A., and Gribnau, J. (2011). RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS Genet.* 7, e1002001.

Barakat, T.S., Loos, F., van Staveren, S., Myronova, E., Ghazvini, M., Grootegoed, J.A., and Gribnau, J. (2014). The Trans-Activator RNF12 and Cis-Acting Elements Effectuate X Chromosome Inactivation Independent of X-Pairing. *Mol. Cell* 53, 965–978.

Barr, M.L., and Bertram, E.G. (1949). A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature* 163, 676.

Bartolomei, M.S., Zemel, S., and Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. *Nature* 351, 153–155.

Beard, C., Li, E., and Jaenisch, R. (1995). Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev.* 9, 2325–2334.

Beddington, R.S., and Robertson, E.J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Dev. Camb. Engl.* 105, 733–737.

Beletskii, A., Hong, Y.-K., Pehrson, J., Egholm, M., and Strauss, W.M. (2001). PNA interference mapping demonstrates functional domains in the noncoding RNA Xist. *Proc. Natl. Acad. Sci.* 98, 9215–9220.

De Belle, I., Cai, S., and Kohwi-Shigematsu, T. (1998). The genomic sequences bound to special AT-rich sequence-binding protein 1 (SATB1) in vivo in Jurkat T cells are tightly associated with the nuclear matrix at the bases of the chromatin loops. *J. Cell Biol.* 141, 335–348.

Belote, J.M., and Lucchesi, J.C. (1980). Male-specific lethal mutations of *Drosophila melanogaster*. *Genetics* 96, 165–186.

Berletch, J.B., Ma, W., Yang, F., Shendure, J., Noble, W.S., Disteche, C.M., and Deng, X. (2015). Escape from X inactivation varies in mouse tissues. *PLoS Genet.* 11, e1005079.

Bininda-Emonds, O.R.P., Cardillo, M., Jones, K.E., MacPhee, R.D.E., Beck, R.M.D., Grenyer, R., Price, S.A., Vos, R.A., Gittleman, J.L., and Purvis, A. (2007). The delayed rise of present-day mammals. *Nature* **446**, 507–512.

Blackledge, N.P., Farcas, A.M., Kondo, T., King, H.W., McGouran, J.F., Hanssen, L.L.P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., et al. (2014). Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, 1445–1459.

Blewitt, M.E., Gendrel, A.-V., Pang, Z., Sparrow, D.B., Whitelaw, N., Craig, J.M., Apedaile, A., Hilton, D.J., Dunwoodie, S.L., Brockdorff, N., et al. (2008). SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat. Genet.* **40**, 663–669.

Boggs, B.A., Connors, B., Sobel, R.E., Chinault, A.C., and Allis, C.D. (1996). Reduced levels of histone H3 acetylation on the inactive X chromosome in human females. *Chromosoma* **105**, 303–309.

Boggs, B.A., Cheung, P., Heard, E., Spector, D.L., Chinault, A.C., and Allis, C.D. (2002). Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes. *Nat. Genet.* **30**, 73–76.

Borsani, G., Tonlorenzi, R., Simmler, M.C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C., et al. (1991). Characterization of a murine gene expressed from the inactive X chromosome. *Nature* **351**, 325–329.

Brannan, C.I., Dees, E.C., Ingram, R.S., and Tilghman, S.M. (1990). The product of the H19 gene may function as an RNA. *Mol. Cell. Biol.* **10**, 28–36.

Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D., and Rastan, S. (1991). Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **351**, 329–331.

Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., and Rastan, S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71**, 515–526.

Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A.,

et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195.

Brown, S.D. (1991). XIST and the mapping of the X chromosome inactivation centre. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **13**, 607–612.

Brown, C.J., and Willard, H.F. (1994). The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. *Nature* **368**, 154–156.

Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafrenière, R.G., Xing, Y., Lawrence, J., and Willard, H.F. (1992). The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71**, 527–542.

Burgoyne, P.S. (1993). A Y-chromosomal effect on blastocyst cell number in mice. *Dev. Camb. Engl.* **117**, 341–345.

Burgoyne, P.S., Thornhill, A.R., Boudrean, S.K., Darling, S.M., Bishop, C.E., and Evans, E.P. (1995). The genetic basis of XX-XY differences present before gonadal sex differentiation in the mouse. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **350**, 253–260 discussion 260–261.

Buzin, C.H., Mann, J.R., and Singer-Sam, J. (1994). Quantitative RT-PCR assays show Xist RNA levels are low in mouse female adult tissue, embryos and embryoid bodies. *Dev. Camb. Engl.* **120**, 3529–3536.

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25**, 1915–1927.

Cai, X., and Cullen, B.R. (2007). The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA N. Y. N* **13**, 313–316.

Calabrese, J.M., Sun, W., Song, L., Mugford, J.W., Williams, L., Yee, D., Starmer, J., Mieczkowski, P., Crawford, G.E., and Magnuson, T. (2012). Site-specific silencing of regulatory elements as a mechanism of X inactivation. *Cell* **151**, 951–963.

Calaway, J.D., Lenarcic, A.B., Didion, J.P., Wang, J.R., Searle, J.B., McMillan, L., Valdar, W., and Pardo-Manuel de Villena, F. (2013). Genetic architecture of skewed X inactivation in the laboratory mouse. *PLoS Genet.* **9**, e1003853.

Carrel, L., and Willard, H.F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* **434**, 400–404.

Casanova, M., Preissner, T., Cerase, A., Poot, R., Yamada, D., Li, X., Appanah, R., Bezstarosti, K., Demmers, J., Koseki, H., et al. (2011). Polycomblike 2 facilitates the recruitment of PRC2 Polycomb group complexes to the inactive X chromosome and to target loci in embryonic stem cells. *Dev. Camb. Engl.* 138, 1471–1482.

Cattanach, B.M., and Papworth, D. (1981). Controlling elements in the mouse. V. Linkage tests with X-linked genes. *Genet. Res.* 38, 57–70.

Cerase, A., Smeets, D., Tang, Y.A., Gdula, M., Kraus, F., Spivakov, M., Moindrot, B., Leleu, M., Tattermusch, A., Demmerle, J., et al. (2014). Spatial separation of Xist RNA and polycomb proteins revealed by superresolution microscopy. *Proc. Natl. Acad. Sci. U. S. A.* 111, 2235–2240.

Chadwick, B.P., and Willard, H.F. (2003). Chromatin of the Barr body: histone and non-histone proteins associated with or excluded from the inactive X chromosome. *Hum. Mol. Genet.* 12, 2167–2178.

Chadwick, L.H., Pertz, L.M., Broman, K.W., Bartolomei, M.S., and Willard, H.F. (2006). Genetic control of X chromosome inactivation in mice: definition of the Xce candidate interval. *Genetics* 173, 2103–2110.

Charlesworth, B. (1991). The evolution of sex chromosomes. *Science* 251, 1030–1033.

Charlesworth, B. (1996). The evolution of chromosomal sex determination and dosage compensation. *Curr. Biol. CB* 6, 149–162.

Chaumeil, J. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev.* 20, 2223–2237.

Chaumeil, J., Okamoto, I., Guggiari, M., and Heard, E. (2002). Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells. *Cytogenet. Genome Res.* 99, 75–84.

Chaumeil, J., Waters, P.D., Koina, E., Gilbert, C., Robinson, T.J., and Graves, J.A.M. (2011). Evolution from XIST-independent to XIST-controlled X-chromosome inactivation: epigenetic modifications in distantly related mammals. *PloS One* 6, e19040.

Chiarle, R., Voena, C., Ambrogio, C., Piva, R., and Inghirami, G. (2008). The anaplastic lymphoma kinase in the pathogenesis of cancer. *Nat. Rev. Cancer* 8, 11–

23.

Chow, J.C., Ciaudo, C., Fazzari, M.J., Mise, N., Servant, N., Glass, J.L., Attreed, M., Avner, P., Wutz, A., Barillot, E., et al. (2010). LINE-1 Activity in Facultative Heterochromatin Formation during X Chromosome Inactivation. *Cell* **141**, 956–969.

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., and Chang, H.Y. (2011). Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44**, 667–678.

Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., and Chang, H.Y. (2015). Systematic Discovery of Xist RNA Binding Proteins. *Cell* **161**, 404–416.

Chureau, C., Prissette, M., Bourdet, A., Barbe, V., Cattolico, L., Jones, L., Eggen, A., Avner, P., and Duret, L. (2002). Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. *Genome Res.* **12**, 894–908.

Chureau, C., Chantalat, S., Romito, A., Galvani, A., Duret, L., Avner, P., and Rougeulle, C. (2011). Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. *Hum. Mol. Genet.* **20**, 705–718.

Ciaudo, C., Bourdet, A., Cohen-Tannoudji, M., Dietz, H.C., Rougeulle, C., and Avner, P. (2006). Nuclear mRNA degradation pathway(s) are implicated in Xist regulation and X chromosome inactivation. *PLoS Genet.* **2**, e94.

Clemson, C.M., McNeil, J.A., Willard, H.F., and Lawrence, J.B. (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* **132**, 259–275.

Clemson, C.M., Hall, L.L., Byron, M., McNeil, J., and Lawrence, J.B. (2006). The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 7688–7693.

Clerc, P., and Avner, P. (1998). Role of the region 3' to Xist exon 6 in the counting process of X-chromosome inactivation. *Nat. Genet.* **19**, 249–253.

Cloutier, J.M., and Turner, J.M.A. (2010). Meiotic sex chromosome inactivation. *Curr. Biol. CB* **20**, R962–R963.

Cohen, D.E., Davidow, L.S., Erwin, J.A., Xu, N., Warshawsky, D., and Lee, J.T. (2007). The DXPas34 Repeat Regulates Random and Imprinted X Inactivation. *Dev.*

Cell 12, 57–71.

Conrad, T., and Akhtar, A. (2011). Dosage compensation in *Drosophila melanogaster*: epigenetic fine-tuning of chromosome-wide transcription. *Nat. Rev. Genet.* 13, 123–134.

Cooper, S., Dienstbier, M., Hassan, R., Schermelleh, L., Sharif, J., Blackledge, N.P., De Marco, V., Elderkin, S., Koseki, H., Klose, R., et al. (2014). Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment. *Cell Rep.* 7, 1456–1470.

Cortez, D., Marin, R., Toledo-Flores, D., Froidevaux, L., Liechti, A., Waters, P.D., Grützner, F., and Kaessmann, H. (2014). Origins and functional evolution of Y chromosomes across mammals. *Nature* 508, 488–493.

Costanzi, C., Stein, P., Worrall, D.M., Schultz, R.M., and Pehrson, J.R. (2000). Histone macroH2A1 is concentrated in the inactive X chromosome of female preimplantation mouse embryos. *Dev. Camb. Engl.* 127, 2283–2289.

Cotton, A.M., Avila, L., Penaherrera, M.S., Affleck, J.G., Robinson, W.P., and Brown, C.J. (2009). Inactive X chromosome-specific reduction in placental DNA methylation. *Hum. Mol. Genet.* 18, 3544–3552.

Cotton, A.M., Lam, L., Affleck, J.G., Wilson, I.M., Peñaherrera, M.S., McFadden, D.E., Kobor, M.S., Lam, W.L., Robinson, W.P., and Brown, C.J. (2011). Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation. *Hum. Genet.* 130, 187–201.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U. S. A.* 107, 21931–21936.

Csankovszki, G., Panning, B., Bates, B., Pehrson, J.R., and Jaenisch, R. (1999). Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. *Nat. Genet.* 22, 323–324.

Csankovszki, G., McDonel, P., and Meyer, B.J. (2004). Recruitment and spreading of the *C. elegans* dosage compensation complex along X chromosomes. *Science* 303, 1182–1185.

Daniels, R., Zuccotti, M., Kinis, T., Serhal, P., and Monk, M. (1997). XIST expression

in human oocytes and preimplantation embryos. *Am. J. Hum. Genet.* **61**, 33–39.

Davidovich, C., Zheng, L., Goodrich, K.J., and Cech, T.R. (2013). Promiscuous RNA binding by Polycomb repressive complex 2. *Nat. Struct. Mol. Biol.* **20**, 1250–1257.

Davidow, L.S., Breen, M., Duke, S.E., Samollow, P.B., McCarrey, J.R., and Lee, J.T. (2007). The search for a marsupial XIC reveals a break with vertebrate synteny. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* **15**, 137–146.

Deakin, J.E., Hore, T.A., Koina, E., and Marshall Graves, J.A. (2008a). The status of dosage compensation in the multiple X chromosomes of the platypus. *PLoS Genet.* **4**, e1000140.

Deakin, J.E., Koina, E., Waters, P.D., Doherty, R., Patel, V.S., Delbridge, M.L., Dobson, B., Fong, J., Hu, Y., van den Hurk, C., et al. (2008b). Physical map of two tammar wallaby chromosomes: A strategy for mapping in non-model mammals. *Chromosome Res.* **16**, 1159–1175.

Deakin, J.E., Chaumeil, J., Hore, T.A., and Marshall Graves, J.A. (2009). Unravelling the evolutionary origins of X chromosome inactivation in mammals: insights from marsupials and monotremes. *Chromosome Res.* **17**, 671–685.

Deakin, J.E., Graves, J. a. M., and Rens, W. (2012). The evolution of marsupial and monotreme chromosomes. *Cytogenet. Genome Res.* **137**, 113–129.

Debrand, E., Chureau, C., Arnaud, D., Avner, P., and Heard, E. (1999). Functional Analysis of the DXPas34Locus, a 3' Regulator of Xist Expression. *Mol. Cell. Biol.* **19**, 8513–8525.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science* **295**, 1306–1311.

Delbridge, M.L., Patel, H.R., Waters, P.D., McMillan, D.A., and Marshall Graves, J.A. (2009). Does the human X contain a third evolutionary block? Origin of genes on human Xp11 and Xq28. *Genome Res.* **19**, 1350–1360.

Dementyeva, E.V., Shevchenko, A.I., Anopriyenko, O.V., Mazurok, N.A., Elisaphenko, E.A., Nesterova, T.B., Brockdorff, N., and Zakian, S.M. (2010). Difference between random and imprinted X inactivation in common voles. *Chromosoma* **119**, 541–552.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The GENCODE v7 catalog of

human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789.

Dindot, S.V., Farin, P.W., Farin, C.E., Romano, J., Walker, S., Long, C., and Piedrahita, J.A. (2004a). Epigenetic and genomic imprinting analysis in nuclear transfer derived *Bos gaurus*/*Bos taurus* hybrid fetuses. *Biol. Reprod.* 71, 470–478.

Dindot, S.V., Kent, K.C., Evers, B., Loskutoff, N., Womack, J., and Piedrahita, J.A. (2004b). Conservation of genomic imprinting at the *XIST*, *IGF2*, and *GTL2* loci in the bovine. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* 15, 966–974.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380.

Donohoe, M.E., Zhang, L.-F., Xu, N., Shi, Y., and Lee, J.T. (2007). Identification of a Ctfc Cofactor, Yy1, for the X Chromosome Binary Switch. *Mol. Cell* 25, 43–56.

Donohoe, M.E., Silva, S.S., Pinter, S.F., Xu, N., and Lee, J.T. (2009). The pluripotency factor Oct4 interacts with Ctfc and also controls X-chromosome pairing and counting. *Nature* 460, 128–132.

Drissen, R., Palstra, R.-J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., and de Laat, W. (2004). The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes Dev.* 18, 2485–2490.

Duret, L., Chureau, C., Samain, S., Weissenbach, J., and Avner, P. (2006). The *Xist* RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science* 312, 1653–1655.

Duszczuk, M.M., Wutz, A., Rybin, V., and Sattler, M. (2011). The *Xist* RNA A-repeat comprises a novel AUCG tetraloop fold and a platform for multimerization. *RNA* 17, 1973–1982.

Duthie, S.M., Nesterova, T.B., Formstone, E.J., Keohane, A.M., Turner, B.M., Zakian, S.M., and Brockdorff, N. (1999). *Xist* RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material in cis. *Hum. Mol. Genet.* 8, 195–204.

Dvash, T., Lavon, N., and Fan, G. (2010). Variations of X chromosome inactivation occur in early passages of female human embryonic stem cells. *PloS One* 5, e11330.

Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W., Yanagimachi, R., and

- Jaenisch, R. (2000). X-Chromosome inactivation in cloned mouse embryos. *Science* 290, 1578–1581.
- Eivazova, E.R., and Aune, T.M. (2004). Dynamic alterations in the conformation of the *lfn* gene region during T helper cell differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 101, 251–256.
- Elgar, G., and Vavouri, T. (2008). Tuning in to the signals: noncoding sequence conservation in vertebrate genomes. *Trends Genet.* TIG 24, 344–352.
- Elisaphenko, E.A., Kolesnikov, N.N., Shevchenko, A.I., Rogozin, I.B., Nesterova, T.B., Brockdorff, N., and Zakian, S.M. (2008). A Dual Origin of the Xist Gene from a Protein-Coding Gene and a Set of Transposable Elements. *PLoS ONE* 3, e2521.
- Engreitz, J.M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E.S., et al. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341, 1237973.
- Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154–156.
- Fackelmayer, F.O. (2005). A stable proteinaceous structure in the territory of inactive X chromosomes. *J. Biol. Chem.* 280, 1720–1723.
- Fukuda, A., Tomikawa, J., Miura, T., Hata, K., Nakabayashi, K., Eggen, K., Akutsu, H., and Umezawa, A. (2014). The role of maternal-specific H3K9me3 modification in establishing imprinted X-chromosome inactivation and embryogenesis in mice. *Nat. Commun.* 5, 5464.
- Fukuda, A., Mitani, A., Miyashita, T., Umezawa, A., and Akutsu, H. (2015). Chromatin condensation of Xist genomic loci during oogenesis in mice. *Dev. Camb. Engl.* 142, 4049–4055.
- Gall, J.G. (1981). Chromosome structure and the C-value paradox. *J. Cell Biol.* 91, 3s – 14s.
- Gartler, S.M., and Riggs, A.D. (1983). Mammalian X-chromosome inactivation. *Annu. Rev. Genet.* 17, 155–190.
- Gaztelumendi, N., and Nogués, C. (2014). Chromosome instability in mouse embryonic stem cells. *Sci. Rep.* 4, 5324.
- Gilbert, S.F. (2000). Environmental Sex Determination.

- Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451.
- Gontan, C., Achame, E.M., Demmers, J., Barakat, T.S., Rentmeester, E., van IJcken, W., Grootegoed, J.A., and Gribnau, J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature* 485, 386–390.
- Grant, J., Mahadevaiah, S.K., Khil, P., Sangrithi, M.N., Royo, H., Duckworth, J., McCarrey, J.R., VandeBerg, J.L., Renfree, M.B., Taylor, W., et al. (2012). Rrx is a metatherian RNA with Xist-like properties in X-chromosome inactivation. *Nature* 487, 254–258.
- Gribnau, J., and Grootegoed, J.A. (2012). Origin and evolution of X chromosome inactivation. *Curr. Opin. Cell Biol.* 24, 397–404.
- Gribnau, J., Luikenhuis, S., Hochedlinger, K., Monkhorst, K., and Jaenisch, R. (2005). X chromosome choice occurs independently of asynchronous replication timing. *J. Cell Biol.* 168, 365–373.
- Gruetzner, F., Ashley, T., Rowell, D.M., and Marshall Graves, J.A. (2006). How did the platypus get its sex chromosome chain? A comparison of meiotic multiples and sex chromosomes in plants and animals. *Chromosoma* 115, 75–88.
- Grumbach, M.M., Morishima, A., and Taylor, J.H. (1963). HUMAN SEX CHROMOSOME ABNORMALITIES IN RELATION TO DNA REPLICATION AND HETEROCHROMATINIZATION. *Proc. Natl. Acad. Sci. U. S. A.* 49, 581–589.
- Grützner, F., Rens, W., Tsend-Ayush, E., El-Mogharbel, N., O'Brien, P.C.M., Jones, R.C., Ferguson-Smith, M.A., and Marshall Graves, J.A. (2004). In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature* 432, 913–917.
- Guo, J., Li, Z., and Feng, Y. (2009). Expression and activation of the reprogramming transcription factors. *Biochem. Biophys. Res. Commun.* 390, 1081–1086.
- Hall, L.L., and Lawrence, J.B. (2003). The cell biology of a novel chromosomal RNA: chromosome painting by XIST/Xist RNA initiates a remodeling cascade. *Semin. Cell Dev. Biol.* 14, 369–378.
- Harper, M.I., Fosten, M., and Monk, M. (1982). Preferential paternal X inactivation in

extraembryonic tissues of early mouse embryos. *J. Embryol. Exp. Morphol.* 67, 127–135.

Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., Tsutui, K., and Nakagawa, S. (2010). The Matrix Protein hnRNP U Is Required for Chromosomal Localization of Xist RNA. *Dev. Cell* 19, 469–476.

Heard, E., Kress, C., Mongelard, F., Courtier, B., Rougeulle, C., Ashworth, A., Vourc'h, C., Babinet, C., and Avner, P. (1996). Transgenic mice carrying an Xist-containing YAC. *Hum. Mol. Genet.* 5, 441–450.

Heard, E., Mongelard, F., Arnaud, D., Chureau, C., Vourc'h, C., and Avner, P. (1999). Human XIST yeast artificial chromosome transgenes show partial X inactivation center function in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 96, 6841–6846.

Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C.D., and Spector, D.L. (2001). Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell* 107, 727–738.

Heard, E., Chaumeil, J., Masui, O., and Okamoto, I. (2004). Mammalian X-chromosome inactivation: an epigenetics paradigm. *Cold Spring Harb. Symp. Quant. Biol.* 69, 89–102.

Hellman, A., and Chess, A. (2007). Gene body-specific methylation on the active X chromosome. *Science* 315, 1141–1143.

Hemet, J., and Rabineau, D. (1966). [Position of the primordial gonocytes in relation to the genital ridges and its variations according to sex]. *Arch. Anat. Histol. Embryol. Norm. Expérimentales* 49, 499–510.

Hoffman, L.M., and Carpenter, M.K. (2005). Human embryonic stem cell stability. *Stem Cell Rev.* 1, 139–144.

Hoki, Y., Ikeda, R., Mise, N., Sakata, Y., Ohhata, T., Sasaki, H., Abe, K., and Sado, T. (2011). Incomplete X-inactivation initiated by a hypomorphic Xist allele in the mouse. *Dev. Camb. Engl.* 138, 2649–2659.

Holmes, D.S., Mayfield, J.E., Sander, G., and Bonner, J. (1972). Chromosomal RNA: its properties. *Science* 177, 72–74.

Hong, Y.K., Ontiveros, S.D., Chen, C., and Strauss, W.M. (1999). A new structure for the murine Xist gene and its relationship to chromosome choice/counting during X-

chromosome inactivation. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 6829–6834.

Hong, Y.K., Ontiveros, S.D., and Strauss, W.M. (2000). A revision of the human XIST gene organization and structural comparison with mouse Xist. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* **11**, 220–224.

Hore, T.A., Koina, E., Wakefield, M.J., and Marshall Graves, J.A. (2007). The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* **15**, 147–161.

Horike, S., Cai, S., Miyano, M., Cheng, J.-F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet.* **37**, 31–40.

Horn, J.M., and Ashworth, A. (1995). A member of the caudal family of homeobox genes maps to the X-inactivation centre region of the mouse and human X chromosomes. *Hum. Mol. Genet.* **4**, 1041–1047.

Hu, Y., Yu, H., Pask, A.J., O'Brien, D.A., Shaw, G., and Renfree, M.B. (2009). A-kinase anchoring protein 4 has a conserved role in mammalian spermatogenesis. *Reprod. Camb. Engl.* **137**, 645–653.

Hu, Y., Yu, H., Shaw, G., Renfree, M.B., and Pask, A.J. (2011). Differential roles of TGIF family genes in mammalian reproduction. *BMC Dev. Biol.* **11**, 58.

Hughes, J.R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R., and Higgs, D.R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* **46**, 205–212.

Humpherys, D., Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W.M., Binizskiewicz, D., Yanagimachi, R., and Jaenisch, R. (2001). Epigenetic instability in ES cells and cloned mice. *Science* **293**, 95–97.

Hüttenhofer, A., Schattner, P., and Polacek, N. (2005). Non-coding RNAs: hope or hype? *Trends Genet. TIG* **21**, 289–297.

Huynh, K.D., and Lee, J.T. (2003). Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature* **426**, 857–862.

Hwang, J.Y., Kim, E.B., Ka, H., and Lee, C.-K. (2013). Identification of the Porcine XIST Gene and Its Differential CpG Methylation Status in Male and Female Pig Cells.

PLoS ONE 8, e73677.

Iitsuka, Y., Bock, A., Nguyen, D.D., Samango-Sprouse, C.A., Simpson, J.L., and Bischoff, F.Z. (2001). Evidence of skewed X-chromosome inactivation in 47,XXY and 48,XXYY Klinefelter patients. *Am. J. Med. Genet.* 98, 25–31.

International Stem Cell Initiative, Amps, K., Andrews, P.W., Anyfantis, G., Armstrong, L., Avery, S., Baharvand, H., Baker, J., Baker, D., Munoz, M.B., et al. (2011). Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat. Biotechnol.* 29, 1132–1144.

Itoh, Y., Replogle, K., Kim, Y.-H., Wade, J., Clayton, D.F., and Arnold, A.P. (2010). Sex bias and dosage compensation in the zebra finch versus chicken genomes: general and specialized patterns among birds. *Genome Res.* 20, 512–518.

Jacobs, P.A., Baikie, A.G., Brown, W.M., Macgregor, T.N., Maclean, N., and Harnden, D.G. (1959). Evidence for the existence of the human “super female.” *Lancet Lond. Engl.* 2, 423–425.

James, D., Noggle, S.A., Swigut, T., and Brivanlou, A.H. (2006). Contribution of human embryonic stem cells to mouse blastocysts. *Dev. Biol.* 295, 90–102.

Jeon, Y., and Lee, J.T. (2011a). YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* 146, 119–133.

Jeon, Y., and Lee, J.T. (2011b). YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* 146, 119–133.

Jeppesen, P., and Turner, B.M. (1993). The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* 74, 281–289.

Jiang, J., Jing, Y., Cost, G.J., Chiang, J.-C., Kolpa, H.J., Cotton, A.M., Carone, D.M., Carone, B.R., Shivak, D.A., Guschin, D.Y., et al. (2013). Translating dosage compensation to trisomy 21. *Nature* 500, 296–300.

Johnston, C.M., Nesterova, T.B., Formstone, E.J., Newall, A.E., Duthie, S.M., Sheardown, S.A., and Brockdorff, N. (1998). Developmentally regulated Xist promoter switch mediates initiation of X inactivation. *Cell* 94, 809–817.

Johnston, C.M., Newall, A.E.T., Brockdorff, N., and Nesterova, T.B. (2002). Enox, a novel gene that maps 10 kb upstream of Xist and partially escapes X inactivation. *Genomics* 80, 236–244.

Jonkers, I., Barakat, T.S., Achame, E.M., Monkhorst, K., Kenter, A., Rentmeester, E., Grosveld, F., Grootegoed, J.A., and Gribnau, J. (2009). RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. *Cell* 139, 999–1011.

Julien, P., Brawand, D., Soumillon, M., Necsulea, A., Liechti, A., Schütz, F., Daish, T., Grützner, F., and Kaessmann, H. (2012). Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. *PLoS Biol.* 10, e1001328.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430–435.

Kalantry, S., and Magnuson, T. (2006). The Polycomb Group Protein EED Is Dispensable for the Initiation of Random X-Chromosome Inactivation. *PLoS Genet.* 2, e66.

Kalantry, S., Purushothaman, S., Bowen, R.B., Starmer, J., and Magnuson, T. (2009). Evidence of Xist RNA-independent initiation of mouse imprinted X-chromosome inactivation. *Nature*.

Kalb, R., Latwiel, S., Baymaz, H.I., Jansen, P.W.T.C., Müller, C.W., Vermeulen, M., and Müller, J. (2014). Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat. Struct. Mol. Biol.* 21, 569–571.

Kanhere, A., Viiri, K., Araújo, C.C., Rasaiyaah, J., Bouwman, R.D., Whyte, W.A., Pereira, C.F., Brookes, E., Walker, K., Bell, G.W., et al. (2010). Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Mol. Cell* 38, 675–688.

Kay, G.F., Penny, G.D., Patel, D., Ashworth, A., Brockdorff, N., and Rastan, S. (1993). Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell* 72, 171–182.

Kay, G.F., Barton, S.C., Surani, M.A., and Rastan, S. (1994). Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. *Cell* 77, 639–650.

Keniry, A., Oxley, D., Monnier, P., Kyba, M., Dandolo, L., Smits, G., and Reik, W. (2012). The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat. Cell Biol.* 14, 659–665.

Keohane, A.M., O'Neill, L.P., Belyaev, N.D., Lavender, J.S., and Turner, B.M.

- (1996a). X-Inactivation and histone H4 acetylation in embryonic stem cells. *Dev. Biol.* **180**, 618–630.
- Keohane, A.M., O'Neill, L.P., Belyaev, N.D., Lavender, J.S., and Turner, B.M. (1996b). X-Inactivation and histone H4 acetylation in embryonic stem cells. *Dev. Biol.* **180**, 618–630.
- Kobayashi, S., Totoki, Y., Soma, M., Matsumoto, K., Fujihara, Y., Toyoda, A., Sakaki, Y., Okabe, M., and Ishino, F. (2013). Identification of an Imprinted Gene Cluster in the X-Inactivation Center. *PLoS ONE* **8**, e71222.
- Krietsch, W.K., Fehla, M., Renner, P., Bücher, T., and Fundele, R. (1986). Expression of X-linked phosphoglycerate kinase in early mouse embryos homozygous at the Xce locus. *Differ. Res. Biol. Divers.* **31**, 50–54.
- Kuckenberger, P., Peitz, M., Kubaczka, C., Becker, A., Egert, A., Wardelmann, E., Zimmer, A., Brüstle, O., and Schorle, H. (2011). Lineage conversion of murine extraembryonic trophoblast stem cells to pluripotent stem cells. *Mol. Cell. Biol.* **31**, 1748–1756.
- Kumar, P.P., Bischof, O., Purbey, P.K., Notani, D., Urlaub, H., Dejean, A., and Galand, S. (2007). Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. *Nat. Cell Biol.* **9**, 45–56.
- Kunath, T., Arnaud, D., Uy, G.D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R.L., Avner, P., and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Dev. Camb. Engl.* **132**, 1649–1661.
- Kurukuti, S., Tiwari, V.K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenko, V., Reik, W., and Ohlsson, R. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10684–10689.
- Kutter, C., Watt, S., Stefflova, K., Wilson, M.D., Goncalves, A., Ponting, C.P., Odom, D.T., and Marques, A.C. (2012). Rapid turnover of long noncoding RNAs and the evolution of gene expression. *PLoS Genet.* **8**, e1002841.
- Lafrenière, R.G., and Willard, H.F. (1993). Pulsed-field map of Xq13 in the region of

the human X inactivation center. *Genomics* 17, 502–506.

Lakhotia, S.C. (2015). Divergent actions of long noncoding RNAs on X-chromosome remodelling in mammals and *Drosophila* achieve the same end result: dosage compensation. *J. Genet.* 94, 575–584.

Latham, K.E. (1996). X chromosome imprinting and inactivation in the early mammalian embryo. *Trends Genet. TIG* 12, 134–138.

Lee, J.T. (2000). Disruption of imprinted X inactivation by parent-of-origin effects at Tsix. *Cell* 103, 17–27.

Lee, J.T. (2005). Regulation of X-chromosome counting by Tsix and Xite sequences. *Science* 309, 768–771.

Lee, J.T., and Jaenisch, R. (1997). Long-range cis effects of ectopic X-inactivation centres on a mouse autosome. *Nature* 386, 275–279.

Lee, J.T., and Lu, N. (1999). Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* 99, 47–57.

Lee, J.T., Strauss, W.M., Dausman, J.A., and Jaenisch, R. (1996). A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell* 86, 83–94.

Lee, J.T., Lu, N., and Han, Y. (1999a). Genetic analysis of the mouse X inactivation center defines an 80-kb multifunction domain. *Proc. Natl. Acad. Sci. U. S. A.* 96, 3836–3841.

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

Lee, M.P., Brandenburg, S., Landes, G.M., Adams, M., Miller, G., and Feinberg, A.P. (1999c). Two novel genes in the center of the 11p15 imprinted domain escape genomic imprinting. *Hum. Mol. Genet.* 8, 683–690.

Lewin, B. (1980). Structure and evolution of interrupted genes. *Ann. N. Y. Acad. Sci.* 354, 453–465.

Lifschytz, E., and Lindsley, D.L. (1972). The role of X-chromosome inactivation during spermatogenesis (*Drosophila*-allocyclic-chromosome evolution-male sterility-dosage compensation). *Proc. Natl. Acad. Sci. U. S. A.* 69, 182–186.

Lipovich, L., Johnson, R., and Lin, C.-Y. (2010). MacroRNA underdogs in a microRNA world: evolutionary, regulatory, and biomedical significance of mammalian

long non-protein-coding RNA. *Biochim. Biophys. Acta* 1799, 597–615.

Livernois, A.M., Waters, S.A., Deakin, J.E., Marshall Graves, J.A., and Waters, P.D. (2013). Independent evolution of transcriptional inactivation on sex chromosomes in birds and mammals. *PLoS Genet.* 9, e1003635.

Luikenhuis, S., Wutz, A., and Jaenisch, R. (2001). Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. *Mol. Cell. Biol.* 21, 8512–8520.

Lyle, R., Watanabe, D., te Vrugte, D., Lerchner, W., Smrzka, O.W., Wutz, A., Schageman, J., Hahner, L., Davies, C., and Barlow, D.P. (2000). The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat. Genet.* 25, 19–21.

Lyon, M.F. (1998). X-chromosome inactivation: a repeat hypothesis. *Cytogenet. Cell Genet.* 80, 133–137.

Lyon, M.F. (2006). Do LINEs have a role in X-chromosome inactivation? *J. Biomed. Biotechnol.* 2006, 59746.

Lyon, M.F., and Rastan, S. (1984). Parental source of chromosome imprinting and its relevance for X chromosome inactivation. *Differ. Res. Biol. Divers.* 26, 63–67.

Maenner, S., Blaud, M., Fouillen, L., Savoye, A., Marchand, V., Dubois, A., Sanglier-Cianféron, S., Van Dorsselaer, A., Clerc, P., Avner, P., et al. (2010). 2-D structure of the A region of Xist RNA and its implication for PRC2 association. *PLoS Biol.* 8, e1000276.

Mahmoudi, T., Katsani, K.R., and Verrijzer, C.P. (2002). GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *EMBO J.* 21, 1775–1781.

Maitra, A., Arking, D.E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassaei, K., Sui, G., Cutler, D.J., Liu, Y., Brimble, S.N., et al. (2005). Genomic alterations in cultured human embryonic stem cells. *Nat. Genet.* 37, 1099–1103.

Mak, W., Baxter, J., Silva, J., Newall, A.E., Otte, A.P., and Brockdorff, N. (2002). Mitotically stable association of polycomb group proteins eed and enx1 with the inactive x chromosome in trophoblast stem cells. *Curr. Biol. CB* 12, 1016–1020.

Mak, W., Nesterova, T.B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. (2004). Reactivation of the paternal X chromosome in early mouse

embryos. *Science* 303, 666–669.

Makhlouf, M., Ouimette, J.-F., Oldfield, A., Navarro, P., Neuillet, D., and Rougeulle, C. (2014). A prominent and conserved role for YY1 in Xist transcriptional activation. *Nat. Commun.* 5, 4878.

Mank, J.E., and Ellegren, H. (2009). All dosage compensation is local: gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. *Heredity* 102, 312–320.

Marahrens, Y. (1999). X-inactivation by chromosomal pairing events. *Genes Dev.* 13, 2624–2632.

Marahrens, Y., Panning, B., Dausman, J., Strauss, W., and Jaenisch, R. (1997). Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* 11, 156–166.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 78, 7634–7638.

Martin, G.R., Epstein, C.J., and Martin, D.W. (1978). Use of teratocarcinoma stem cells as a model system for the study of X-chromosome inactivation in vitro. *Basic Life Sci.* 12, 269–295.

Masui, S., Ohtsuka, S., Yagi, R., Takahashi, K., Ko, M.S.H., and Niwa, H. (2008). Rex1/Zfp42 is dispensable for pluripotency in mouse ES cells. *BMC Dev. Biol.* 8, 45.

Matsuura, S., Episkopou, V., Hamvas, R., and Brown, S.D. (1996). Xist expression from an Xist YAC transgene carried on the mouse Y chromosome. *Hum. Mol. Genet.* 5, 451–459.

Mattaj, I.W., Tollervey, D., and Séraphin, B. (1993). Small nuclear RNAs in messenger RNA and ribosomal RNA processing. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 7, 47–53.

Mattick, J.S., and Rinn, J.L. (2015). Discovery and annotation of long noncoding RNAs. *Nat. Struct. Mol. Biol.* 22, 5–7.

McHugh, C.A., Chen, C.-K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521, 232–236.

- McMahon, A., Fosten, M., and Monk, M. (1983). X-chromosome inactivation mosaicism in the three germ layers and the germ line of the mouse embryo. *J. Embryol. Exp. Morphol.* **74**, 207–220.
- Memili, E., Hong, Y.K., Kim, D.H., Ontiveros, S.D., and Strauss, W.M. (2001). Murine Xist RNA isoforms are different at their 3' ends: a role for differential polyadenylation. *Gene* **266**, 131–137.
- Mercer, T.R., Dinger, M.E., Sunkin, S.M., Mehler, M.F., and Mattick, J.S. (2008). Specific expression of long noncoding RNAs in the mouse brain. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 716–721.
- Mermoud, J.E., Costanzi, C., Pehrson, J.R., and Brockdorff, N. (1999). Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation. *J. Cell Biol.* **147**, 1399–1408.
- Mermoud, J.E., Popova, B., Peters, A.H.F.M., Jenuwein, T., and Brockdorff, N. (2002). Histone H3 lysine 9 methylation occurs rapidly at the onset of random X chromosome inactivation. *Curr. Biol. CB* **12**, 247–251.
- Migeon, B.R., Kazi, E., Haisley-Royster, C., Hu, J., Reeves, R., Call, L., Lawler, A., Moore, C.S., Morrison, H., and Jeppesen, P. (1999). Human X inactivation center induces random X chromosome inactivation in male transgenic mice. *Genomics* **59**, 113–121.
- Migeon, B.R., Winter, H., Kazi, E., Chowdhury, A.K., Hughes, A., Haisley-Royster, C., Morrison, H., and Jeppesen, P. (2001). Low-copy-number human transgene is recognized as an X inactivation center in mouse ES cells, but fails to induce cis-inactivation in chimeric mice. *Genomics* **71**, 156–162.
- Migeon, B.R., Lee, C.H., Chowdhury, A.K., and Carpenter, H. (2002). Species differences in TSIX/Tsix reveal the roles of these genes in X-chromosome inactivation. *Am. J. Hum. Genet.* **71**, 286–293.
- Miller, D., Summers, J., and Silber, S. (2004). Environmental versus genetic sex determination: a possible factor in dinosaur extinction? *Fertil. Steril.* **81**, 954–964.
- Milligan, M.J., and Lipovich, L. (2014). Pseudogene-derived lncRNAs: emerging regulators of gene expression. *Front. Genet.* **5**, 476.
- Minajigi, A., Froberg, J.E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., Lessing, D., Payer, B., Boukhali, M., Haas, W., et al. (2015). Chromosomes. A comprehensive

Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**.

Mirsky, A.E., and Ris, H. (1951). The desoxyribonucleic acid content of animal cells and its evolutionary significance. *J. Gen. Physiol.* **34**, 451–462.

Miska, E.A., Alvarez-Saavedra, E., Townsend, M., Yoshii, A., Sestan, N., Rakic, P., Constantine-Paton, M., and Horvitz, H.R. (2004). Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol.* **5**, R68.

Mlynarczyk-Evans, S., Royce-Tolland, M., Alexander, M.K., Andersen, A.A., Kalantry, S., Gribnau, J., and Panning, B. (2006). X chromosomes alternate between two states prior to random X-inactivation. *PLoS Biol.* **4**, e159.

Moindrot, B., Cerase, A., Coker, H., Masui, O., Grijzenhout, A., Pintacuda, G., Schermelleh, L., Nesterova, T.B., and Brockdorff, N. (2015). A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell Rep.* **12**, 562–572.

Monfort, A., Di Minin, G., Postlmayr, A., Freimann, R., Arieti, F., Thore, S., and Wutz, A. (2015). Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells. *Cell Rep.* **12**, 554–561.

Monk, M. (1978). Biochemical studies on X-chromosome activity in preimplantation mouse embryos. *Basic Life Sci.* **12**, 239–246.

Monk, M., and Harper, M.I. (1979). Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos. *Nature* **281**, 311–313.

Monkhorst, K., Jonkers, I., Rentmeester, E., Grosveld, F., and Gribnau, J. (2008). X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. *Cell* **132**, 410–421.

Monkhorst, K., de Hoon, B., Jonkers, I., Mulugeta Achame, E., Monkhorst, W., Hoogerbrugge, J., Rentmeester, E., Westerhoff, H.V., Grosveld, F., Grootegoed, J.A., et al. (2009). The probability to initiate X chromosome inactivation is determined by the X to autosomal ratio and X chromosome specific allelic properties. *PloS One* **4**, e5616.

Moreira de Mello, J.C., de Araújo, E.S.S., Stabellini, R., Fraga, A.M., de Souza, J.E.S., Sumita, D.R., Camargo, A.A., and Pereira, L.V. (2010). Random X

inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome. *PLoS One* 5, e10947.

Morey, C., Arnaud, D., Avner, P., and Clerc, P. (2001). Tsix-mediated repression of Xist accumulation is not sufficient for normal random X inactivation. *Hum. Mol. Genet.* 10, 1403–1411.

Morey, C., Navarro, P., Debrand, E., Avner, P., Rougeulle, C., and Clerc, P. (2004). The region 3' to Xist mediates X chromosome counting and H3 Lys-4 dimethylation within the Xist gene. *EMBO J.* 23, 594–604.

Murakami, K., Araki, K., Ohtsuka, S., Wakayama, T., and Niwa, H. (2011). Choice of random rather than imprinted X inactivation in female embryonic stem cell-derived extra-embryonic cells. *Dev. Camb. Engl.* 138, 197–202.

Al Nadaf, S., Waters, P.D., Koina, E., Deakin, J.E., Jordan, K.S., and Graves, J.A. (2010). Activity map of the tammar X chromosome shows that marsupial X inactivation is incomplete and escape is stochastic. *Genome Biol.* 11, R122.

Namekawa, S.H., and Lee, J.T. (2009). XY and ZW: is meiotic sex chromosome inactivation the rule in evolution? *PLoS Genet.* 5, e1000493.

Namekawa, S.H., Payer, B., Huynh, K.D., Jaenisch, R., and Lee, J.T. (2010). Two-step imprinted X inactivation: repeat versus genic silencing in the mouse. *Mol. Cell. Biol.* 30, 3187–3205.

Navarro, P., Pichard, S., Ciaudo, C., Avner, P., and Rougeulle, C. (2005). Tsix transcription across the Xist gene alters chromatin conformation without affecting Xist transcription: implications for X-chromosome inactivation. *Genes Dev.* 19, 1474–1484.

Navarro, P., Page, D.R., Avner, P., and Rougeulle, C. (2006). Tsix-mediated epigenetic switch of a CTCF-flanked region of the Xist promoter determines the Xist transcription program. *Genes Dev.* 20, 2787–2792.

Navarro, P., Chambers, I., Karwacki-Neisius, V., Chureau, C., Morey, C., Rougeulle, C., and Avner, P. (2008). Molecular coupling of Xist regulation and pluripotency. *Science* 321, 1693–1695.

Navarro, P., Chantalat, S., Foglio, M., Chureau, C., Vigneau, S., Clerc, P., Avner, P., and Rougeulle, C. (2009). A role for non-coding Tsix transcription in partitioning chromatin domains within the mouse X-inactivation centre. *Epigenetics Chromatin* 2,

8.

Navarro, P., Oldfield, A., Legoupi, J., Festuccia, N., Dubois, A., Attia, M., Schoorlemmer, J., Rougeulle, C., Chambers, I., and Avner, P. (2010). Molecular coupling of Tsix regulation and pluripotency. *Nature* 468, 457–460.

Nechanitzky, R., Dávila, A., Savarese, F., Fietze, S., and Grosschedl, R. (2012). Satb1 and Satb2 are dispensable for X chromosome inactivation in mice. *Dev. Cell* 23, 866–871.

Necsulea, A., Soumillon, M., Warnefors, M., Liechti, A., Daish, T., Zeller, U., Baker, J.C., Grützner, F., and Kaessmann, H. (2014). The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* 505, 635–640.

Nesterova, T.B., Slobodyanyuk, S.Y., Elisaphenko, E.A., Shevchenko, A.I., Johnston, C., Pavlova, M.E., Rogozin, I.B., Kolesnikov, N.N., Brockdorff, N., and Zakian, S.M. (2001). Characterization of the genomic Xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. *Genome Res.* 11, 833–849.

Nesterova, T.B., Popova, B.C., Cobb, B.S., Norton, S., Senner, C.E., Tang, Y.A., Spruce, T., Rodriguez, T.A., Sado, T., Merckenschlager, M., et al. (2008). Dicer regulates Xist promoter methylation in ES cells indirectly through transcriptional control of Dnmt3a. *Epigenetics Chromatin* 1, 2.

Nesterova, T.B., Senner, C.E., Schneider, J., Alcayna-Stevens, T., Tattermusch, A., Hemberger, M., and Brockdorff, N. (2011). Pluripotency factor binding and Tsix expression act synergistically to repress Xist in undifferentiated embryonic stem cells. *Epigenetics Chromatin* 4, 17.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* 4, 487–492.

Nichols, J., Evans, E.P., and Smith, A.G. (1990). Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. *Dev. Camb. Engl.* 110, 1341–1348.

Nitsche, A., Rose, D., Fasold, M., Reiche, K., and Stadler, P.F. (2015). Comparison of splice sites reveals that long noncoding RNAs are evolutionarily well conserved. *RNA* 21, 801–812.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and

- Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* 123, 917–929.
- Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385.
- Nora, E.P., Dekker, J., and Heard, E. (2013). Segmental folding of chromosomes: a basis for structural and regulatory chromosomal neighborhoods? *BioEssays News Rev. Mol. Cell. Dev. Biol.* 35, 818–828.
- Norris, D.P., Brockdorff, N., and Rastan, S. (1991). Methylation status of CpG-rich islands on active and inactive mouse X chromosomes. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* 1, 78–83.
- Ogawa, Y., Sun, B.K., and Lee, J.T. (2008). Intersection of the RNA interference and X-inactivation pathways. *Science* 320, 1336–1341.
- Ohno, S., Kaplan, W.D., and Kinoshita, R. (1959). Formation of the sex chromatin by a single X-chromosome in liver cells of *Rattus norvegicus*. *Exp. Cell Res.* 18, 415–418.
- Okada, Y., Shimazaki, T., Sobue, G., and Okano, H. (2004). Retinoic-acid-concentration-dependent acquisition of neural cell identity during in vitro differentiation of mouse embryonic stem cells. *Dev. Biol.* 275, 124–142.
- Okamoto, I., Tan, S., and Takagi, N. (2000). X-chromosome inactivation in XX androgenetic mouse embryos surviving implantation. *Dev. Camb. Engl.* 127, 4137–4145.
- Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* 303, 644–649.
- Okamoto, I., Arnaud, D., Le Baccon, P., Otte, A.P., Disteche, C.M., Avner, P., and Heard, E. (2005). Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* 438, 369–373.
- Okamoto, I., Patrat, C., Thépot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J.-P., Renard, J.-P., Duranthon, V., et al. (2011). Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* 472, 370–374.
- Ouyang, Y., Salstrom, J., Diaz-Perez, S., Nahas, S., Matsuno, Y., Dawson, D.,

Teitell, M.A., Horvath, S., Riggs, A.D., Gatti, R.A., et al. (2005). Inhibition of Atm and/or Atr disrupts gene silencing on the inactive X chromosome. *Biochem. Biophys. Res. Commun.* 337, 875–880.

Panning, B., and Jaenisch, R. (1996). DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev.* 10, 1991–2002.

Panning, B., Dausman, J., and Jaenisch, R. (1997). X chromosome inactivation is mediated by Xist RNA stabilization. *Cell* 90, 907–916.

Pask, A., and Graves, J.A. (1999). Sex chromosomes and sex-determining genes: insights from marsupials and monotremes. *Cell. Mol. Life Sci. CMLS* 55, 864–875.

Pask, A., and Graves, J.A. (2001). Sex chromosomes and sex-determining genes: insights from marsupials and monotremes. *EXS* 71–95.

Pask, A., Toder, R., Wilcox, S.A., Camerino, G., and Graves, J.A. (1997). The candidate sex-reversing DAX1 gene is autosomal in marsupials: implications for the evolution of sex determination in mammals. *Genomics* 41, 422–426.

Pask, A., Renfree, M.B., and Marshall Graves, J.A. (2000). The human sex-reversing ATRX gene has a homologue on the marsupial Y chromosome, ATRY: implications for the evolution of mammalian sex determination. *Proc. Natl. Acad. Sci. U. S. A.* 97, 13198–13202.

Paterno, G.D., and McBurney, M.W. (1985). X chromosome inactivation during induced differentiation of a female mouse embryonal carcinoma cell line. *J. Cell Sci.* 75, 149–163.

Patrat, C., Okamoto, I., Diabangouaya, P., Vialon, V., Le Baccon, P., Chow, J., and Heard, E. (2009). Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. *Proc. Natl. Acad. Sci. U. S. A.* 106, 5198–5203.

Pedersen, J.F. (1980). Ultrasound evidence of sexual difference in fetal size in first trimester. *Br. Med. J.* 281, 1253.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. *Nature* 379, 131–137.

Pergament, E., Fiddler, M., Cho, N., Johnson, D., and Holmgren, W.J. (1994). Sexual differentiation and preimplantation cell growth. *Hum. Reprod. Oxf. Engl.* 9, 1730–1732.

- Peschansky, V.J., and Wahlestedt, C. (2014). Non-coding RNAs as direct and indirect modulators of epigenetic regulation. *Epigenetics* 9, 3–12.
- Pessia, E., Makino, T., Bailly-Bechet, M., McLysaght, A., and Marais, G.A.B. (2012). Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. *Proc. Natl. Acad. Sci.* 109, 5346–5351.
- Peters, A.H.F.M., Mermoud, J.E., O’Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N., and Jenuwein, T. (2002). Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nat. Genet.* 30, 77–80.
- Petrascheck, M., Escher, D., Mahmoudi, T., Verrijzer, C.P., Schaffner, W., and Barberis, A. (2005). DNA looping induced by a transcriptional enhancer in vivo. *Nucleic Acids Res.* 33, 3743–3750.
- Philippidou, P., and Dasen, J.S. (2013). Hox genes: choreographers in neural development, architects of circuit organization. *Neuron* 80, 12–34.
- Pierpont, M.E., and Yunis, J.J. (1977). Localization of chromosomal RNA in human G-banded metaphase chromosomes. *Exp. Cell Res.* 106, 303–308.
- Pinter, S.F., Sadreyev, R.I., Yildirim, E., Jeon, Y., Ohsumi, T.K., Borowsky, M., and Lee, J.T. (2012). Spreading of X chromosome inactivation via a hierarchy of defined Polycomb stations. *Genome Res.* 22, 1864–1876.
- Piper, A.A., Bennett, A.M., Noyce, L., Swanton, M.K., and Cooper, D.W. (1993). Isolation of a clone partially encoding hill kangaroo X-linked hypoxanthine phosphoribosyltransferase: sex differences in methylation in the body of the gene. *Somat. Cell Mol. Genet.* 19, 141–159.
- Plath, K., Mlynarczyk-Evans, S., Nusinow, D.A., and Panning, B. (2002). *Xist* RNA and the Mechanism of X Chromosome Inactivation. *Annu. Rev. Genet.* 36, 233–278.
- Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., Cecile, C., Otte, A.P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. *Science* 300, 131–135.
- Plenge, R.M., Percec, I., Nadeau, J.H., and Willard, H.F. (2000). Expression-based assay of an X-linked gene to examine effects of the X-controlling element (Xce) locus. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* 11, 405–408.
- Politz, J.C.R., Zhang, F., and Pederson, T. (2006). MicroRNA-206 colocalizes with

ribosome-rich regions in both the nucleolus and cytoplasm of rat myogenic cells. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 18957–18962.

Politz, J.C.R., Hogan, E.M., and Pederson, T. (2009). MicroRNAs with a nucleolar location. *RNA N. Y. N* *15*, 1705–1715.

Rastan, S. (1982). Timing of X-chromosome inactivation in postimplantation mouse embryos. *J. Embryol. Exp. Morphol.* *71*, 11–24.

Rastan, S. (1983). Non-random X-chromosome inactivation in mouse X-autosome translocation embryos—location of the inactivation centre. *J. Embryol. Exp. Morphol.* *78*, 1–22.

Rastan, S., and Robertson, E.J. (1985). X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *J. Embryol. Exp. Morphol.* *90*, 379–388.

Ray, P.F., Winston, R.M., and Handyside, A.H. (1997). XIST expression from the maternal X chromosome in human male preimplantation embryos at the blastocyst stage. *Hum. Mol. Genet.* *6*, 1323–1327.

Rebuzzini, P., Neri, T., Zuccotti, M., Redi, C.A., and Garagna, S. (2008). Chromosome number variation in three mouse embryonic stem cell lines during culture. *Cytotechnology* *58*, 17–23.

Reik, W., Murrell, A., Lewis, A., Mitsuya, K., Umlauf, D., Dean, W., Higgins, M., and Feil, R. (2004). Chromosome loops, insulators, and histone methylation: new insights into regulation of imprinting in clusters. *Cold Spring Harb. Symp. Quant. Biol.* *69*, 29–37.

Rens, W., Grützner, F., O'Brien, P.C.M., Fairclough, H., Graves, J.A.M., and Ferguson-Smith, M.A. (2004). Resolution and evolution of the duck-billed platypus karyotype with an X1Y1X2Y2X3Y3X4Y4X5Y5 male sex chromosome constitution. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 16257–16261.

Rens, W., O'Brien, P.C.M., Grützner, F., Clarke, O., Graphodatskaya, D., Tsend-Ayush, E., Trifonov, V.A., Skelton, H., Wallis, M.C., Johnston, S., et al. (2007). The multiple sex chromosomes of platypus and echidna are not completely identical and several share homology with the avian Z. *Genome Biol.* *8*, R243.

Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., and Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro.

Nat. Biotechnol. 18, 399–404.

Da Rocha, S.T., Edwards, C.A., Ito, M., Ogata, T., and Ferguson-Smith, A.C. (2008). Genomic imprinting at the mammalian Dlk1-Dio3 domain. Trends Genet. TIG 24, 306–316.

Da Rocha, S.T., Boeva, V., Escamilla-Del-Arenal, M., Ancelin, K., Granier, C., Matias, N.R., Sanulli, S., Chow, J., Schulz, E., Picard, C., et al. (2014). Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. Mol. Cell 53, 301–316.

Rougeulle, C., and Avner, P. (1996). Cloning and characterization of a murine brain specific gene Bpx and its human homologue lying within the Xic candidate region. Hum. Mol. Genet. 5, 41–49.

Rougeulle, C., and Avner, P. (2004). The role of antisense transcription in the regulation of X-inactivation. Curr. Top. Dev. Biol. 63, 61–89.

Rougeulle, C., Chaumeil, J., Sarma, K., Allis, C.D., Reinberg, D., Avner, P., and Heard, E. (2004). Differential Histone H3 Lys-9 and Lys-27 Methylation Profiles on the X Chromosome. Mol. Cell. Biol. 24, 5475–5484.

Royce-Tolland, M.E., Andersen, A.A., Koyfman, H.R., Talbot, D.J., Wutz, A., Tonks, I.D., Kay, G.F., and Panning, B. (2010). The A-repeat links ASF/SF2-dependent Xist RNA processing with random choice during X inactivation. Nat. Struct. Mol. Biol. 17, 948–954.

Sado, T., Fenner, M.H., Tan, S.S., Tam, P., Shioda, T., and Li, E. (2000). X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. Dev. Biol. 225, 294–303.

Sado, T., Wang, Z., Sasaki, H., and Li, E. (2001). Regulation of imprinted X-chromosome inactivation in mice by Tsix. Development 128, 1275–1286.

Sado, T., Okano, M., Li, E., and Sasaki, H. (2004). De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. Dev. Camb. Engl. 131, 975–982.

Sado, T., Hoki, Y., and Sasaki, H. (2005). Tsix silences Xist through modification of chromatin structure. Dev. Cell 9, 159–165.

Sarkar, M.K., Gayen, S., Kumar, S., Maclary, E., Buttigieg, E., Hinten, M., Kumari, A., Harris, C., Sado, T., and Kalantry, S. (2015). An Xist-activating antisense RNA

required for X-chromosome inactivation. *Nat. Commun.* **6**, 8564.

Sarma, K., Levasseur, P., Aristarkhov, A., and Lee, J.T. (2010). Locked nucleic acids (LNAs) reveal sequence requirements and kinetics of Xist RNA localization to the X chromosome. *Proc. Natl. Acad. Sci.* **107**, 22196–22201.

Sarma, K., Cifuentes-Rojas, C., Ergun, A., del Rosario, A., Jeon, Y., White, F., Sadreyev, R., and Lee, J.T. (2014). ATRX Directs Binding of PRC2 to Xist RNA and Polycomb Targets. *Cell* **159**, 869–883.

Savarese, F., Flahndorfer, K., Jaenisch, R., Busslinger, M., and Wutz, A. (2006). Hematopoietic precursor cells transiently reestablish permissiveness for X inactivation. *Mol. Cell. Biol.* **26**, 7167–7177.

Sayegh, C.E., Sayegh, C., Jhunjhunwala, S., Riblet, R., and Murre, C. (2005). Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. *Genes Dev.* **19**, 322–327.

Schmidt, J.V., Matteson, P.G., Jones, B.K., Guan, X.J., and Tilghman, S.M. (2000). The *Dlk1* and *Gtl2* genes are linked and reciprocally imprinted. *Genes Dev.* **14**, 1997–2002.

Schulz, E.G., Meisig, J., Nakamura, T., Okamoto, I., Sieber, A., Picard, C., Borensztein, M., Saitou, M., Blüthgen, N., and Heard, E. (2014). The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell* **14**, 203–216.

Scott, W.J., and Holson, J.F. (1977). Weight differences in rat embryos prior to sexual differentiation. *J. Embryol. Exp. Morphol.* **40**, 259–263.

Seller, M.J., and Perkins-Cole, K.J. (1987). Sex difference in mouse embryonic development at neurulation. *J. Reprod. Fertil.* **79**, 159–161.

Shao, C., and Takagi, N. (1990). An extra maternally derived X chromosome is deleterious to early mouse development. *Dev. Camb. Engl.* **110**, 969–975.

Sharman, G.B. (1971). Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature* **230**, 231–232.

Sheardown, S.A., Duthie, S.M., Johnston, C.M., Newall, A.E., Formstone, E.J., Arkell, R.M., Nesterova, T.B., Alghisi, G.-C., Rastan, S., and Brockdorff, N. (1997). Stabilization of Xist RNA mediates initiation of X chromosome inactivation. *Cell* **91**, 99–107.

- Shen, Y., Matsuno, Y., Fouse, S.D., Rao, N., Root, S., Xu, R., Pellegrini, M., Riggs, A.D., and Fan, G. (2008). X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 4709–4714.
- Shevchenko, A.I., Zakharova, I.S., Elisaphenko, E.A., Kolesnikov, N.N., Whitehead, S., Bird, C., Ross, M., Weidman, J.R., Jirtle, R.L., Karamysheva, T.V., et al. (2007). Genes flanking Xist in mouse and human are separated on the X chromosome in American marsupials. *Chromosome Res.* *15*, 127–136.
- Shibata, S., and Lee, J.T. (2003). Characterization and quantitation of differential Tsix transcripts: implications for Tsix function. *Hum. Mol. Genet.* *12*, 125–136.
- Shimosato, D., Shiki, M., and Niwa, H. (2007). Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. *BMC Dev. Biol.* *7*, 80.
- Shin, J., Bossenz, M., Chung, Y., Ma, H., Byron, M., Taniguchi-Ishigaki, N., Zhu, X., Jiao, B., Hall, L.L., Green, M.R., et al. (2010). Maternal Rnf12/RLIM is required for imprinted X-chromosome inactivation in mice. *Nature* *467*, 977–981.
- Shin, J., Wallingford, M.C., Gallant, J., Marcho, C., Jiao, B., Byron, M., Bossenz, M., Lawrence, J.B., Jones, S.N., Mager, J., et al. (2014). RLIM is dispensable for X-chromosome inactivation in the mouse embryonic epiblast. *Nature* *511*, 86–89.
- Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T.B., Webster, Z., Peters, A.H.F.M., Jenuwein, T., Otte, A.P., and Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev. Cell* *4*, 481–495.
- Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. *Cell* *138*, 722–737.
- Silva, S.S., Rowntree, R.K., Mekhoubad, S., and Lee, J.T. (2008). X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 4820–4825.
- Simmler, M.C., Cattanaach, B.M., Rasberry, C., Rougeulle, C., and Avner, P. (1993). Mapping the murine Xce locus with (CA)_n repeats. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* *4*, 523–530.

- Simmler, M.C., Heard, E., Rougeulle, C., Cruaud, C., Weissenbach, J., and Avner, P. (1997). Localization and expression analysis of a novel conserved brain expressed transcript, Brx/BRX, lying within the Xic/XIC candidate region. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* **8**, 760–766.
- Simon, M.D., Wang, C.I., Kharchenko, P.V., West, J.A., Chapman, B.A., Alekseyenko, A.A., Borowsky, M.L., Kuroda, M.I., and Kingston, R.E. (2011). The genomic binding sites of a noncoding RNA. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20497–20502.
- Simon, M.D., Pinter, S.F., Fang, R., Sarma, K., Rutenberg-Schoenberg, M., Bowman, S.K., Kesner, B.A., Maier, V.K., Kingston, R.E., and Lee, J.T. (2013). High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504**, 465–469.
- Smeets, D., Markaki, Y., Schmid, V.J., Kraus, F., Tattermusch, A., Cerase, A., Sterr, M., Fiedler, S., Demmerle, J., Popken, J., et al. (2014). Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. *Epigenetics Chromatin* **7**, 8.
- Smith, K.P., Byron, M., Clemson, C.M., and Lawrence, J.B. (2004). Ubiquitinated proteins including uH2A on the human and mouse inactive X chromosome: enrichment in gene rich bands. *Chromosoma* **113**, 324–335.
- Soma, M., Fujihara, Y., Okabe, M., Ishino, F., and Kobayashi, S. (2014). Ftx is dispensable for imprinted X-chromosome inactivation in preimplantation mouse embryos. *Sci. Rep.* **4**.
- Song, S.-H., Hou, C., and Dean, A. (2007). A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. *Mol. Cell* **28**, 810–822.
- Spencer, R.J., del Rosario, B.C., Pinter, S.F., Lessing, D., Sadreyev, R.I., and Lee, J.T. (2011). A boundary element between Tsix and Xist binds the chromatin insulator Ctf and contributes to initiation of X-chromosome inactivation. *Genetics* **189**, 441–454.
- Spilianakis, C.G., and Flavell, R.A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat. Immunol.* **5**, 1017–1027.
- Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008). Defining

molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2, 230–240.

Stavropoulos, N., Lu, N., and Lee, J.T. (2001). A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice. *Proc. Natl. Acad. Sci. U. S. A.* 98, 10232–10237.

Strome, S., Kelly, W.G., Ercan, S., and Lieb, J.D. (2014). Regulation of the X chromosomes in *Caenorhabditis elegans*. *Cold Spring Harb. Perspect. Biol.* 6.

Sugawara, A., Goto, K., Sotomaru, Y., Sofuni, T., and Ito, T. (2006). Current status of chromosomal abnormalities in mouse embryonic stem cell lines used in Japan. *Comp. Med.* 56, 31–34.

Suh, M.-R., Lee, Y., Kim, J.Y., Kim, S.-K., Moon, S.-H., Lee, J.Y., Cha, K.-Y., Chung, H.M., Yoon, H.S., Moon, S.Y., et al. (2004). Human embryonic stem cells express a unique set of microRNAs. *Dev. Biol.* 270, 488–498.

Sun, B.K., Deaton, A.M., and Lee, J.T. (2006a). A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol. Cell* 21, 617–628.

Sun, B.K., Deaton, A.M., and Lee, J.T. (2006b). A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol. Cell* 21, 617–628.

Sun, S., Del Rosario, B.C., Szanto, A., Ogawa, Y., Jeon, Y., and Lee, J.T. (2013). Jpx RNA Activates Xist by Evicting CTCF. *Cell* 153, 1537–1551.

Sun, S., Payer, B., Namekawa, S., An, J.Y., Press, W., Catalan-Dibene, J., Sunwoo, H., and Lee, J.T. (2015). Xist imprinting is promoted by the hemizygous (unpaired) state in the male germ line. *Proc. Natl. Acad. Sci. U. S. A.* 112, 14415–14422.

Sutton, E., Hughes, J., White, S., Sekido, R., Tan, J., Arboleda, V., Rogers, N., Knowler, K., Rowley, L., Eyre, H., et al. (2011). Identification of SOX3 as an XX male sex reversal gene in mice and humans. *J. Clin. Invest.* 121, 328–341.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr. Biol. CB* 11, 1553–1558.

Tada, T., Takagi, N., and Adler, I.D. (1993a). Parental imprinting on the mouse X chromosome: effects on the early development of X0, XXY and XXX embryos. *Genet. Res.* 62, 139–148.

Tada, T., Tada, M., and Takagi, N. (1993b). X chromosome retains the memory of its

parental origin in murine embryonic stem cells. *Dev. Camb. Engl.* 119, 813–821.

Tada, T., Obata, Y., Tada, M., Goto, Y., Nakatsuji, N., Tan, S., Kono, T., and Takagi, N. (2000). Imprint switching for non-random X-chromosome inactivation during mouse oocyte growth. *Dev. Camb. Engl.* 127, 3101–3105.

Taft, R.J., Pheasant, M., and Mattick, J.S. (2007). The relationship between non-protein-coding DNA and eukaryotic complexity. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 29, 288–299.

Takagi, N. (1980). Primary and secondary nonrandom X chromosome inactivation in early female mouse embryos carrying Searle's translocation T(X; 16)16H. *Chromosoma* 81, 439–459.

Takagi, N., and Abe, K. (1990). Detrimental effects of two active X chromosomes on early mouse development. *Dev. Camb. Engl.* 109, 189–201.

Takagi, N., Sugawara, O., and Sasaki, M. (1982). Regional and temporal changes in the pattern of X-chromosome replication during the early post-implantation development of the female mouse. *Chromosoma* 85, 275–286.

Takagi, N., Yoshida, M.A., Sugawara, O., and Sasaki, M. (1983). Reversal of X-inactivation in female mouse somatic cells hybridized with murine teratocarcinoma stem cells in vitro. *Cell* 34, 1053–1062.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.

Tam, P.P., Zhou, S.X., and Tan, S.S. (1994). X-chromosome activity of the mouse primordial germ cells revealed by the expression of an X-linked lacZ transgene. *Dev. Camb. Engl.* 120, 2925–2932.

Tattermusch, A., and Brockdorff, N. (2011). A scaffold for X chromosome inactivation. *Hum. Genet.* 130, 247–253.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448, 196–199.

Thomas, C.A. (1971). The genetic organization of chromosomes. *Annu. Rev. Genet.* 5, 237–256.

Thomson, J.A., Kalishman, J., Golos, T.G., Durning, M., Harris, C.P., Becker, R.A., and Hearn, J.P. (1995). Isolation of a primate embryonic stem cell line. *Proc. Natl.*

Acad. Sci. U. S. A. 92, 7844–7848.

Thorvaldsen, J.L., Krapp, C., Willard, H.F., and Bartolomei, M.S. (2012). Nonrandom X chromosome inactivation is influenced by multiple regions on the murine X chromosome. *Genetics* 192, 1095–1107.

Tian, D., Sun, S., and Lee, J.T. (2010). The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* 143, 390–403.

Tsai, C.-L., Rowntree, R.K., Cohen, D.E., and Lee, J.T. (2008). Higher order chromatin structure at the X-inactivation center via looping DNA. *Dev. Biol.* 319, 416–425.

Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* 147, 1537–1550.

Valdivia, R.P., Kunieda, T., Azuma, S., and Toyoda, Y. (1993). PCR sexing and developmental rate differences in preimplantation mouse embryos fertilized and cultured in vitro. *Mol. Reprod. Dev.* 35, 121–126.

Vallot, C., Huret, C., Lesecque, Y., Resch, A., Oudrhiri, N., Bennaceur-Griscelli, A., Duret, L., and Rougeulle, C. (2013). XACT, a long noncoding transcript coating the active X chromosome in human pluripotent cells. *Nat. Genet.* 45, 239–241.

Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* 26, 2041–2051.

Veyrunes, F., Waters, P.D., Miethke, P., Rens, W., McMillan, D., Alsop, A.E., Grützner, F., Deakin, J.E., Whittington, C.M., Schatzkamer, K., et al. (2008). Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome Res.* 18, 965–973.

Veyrunes, F., Chevret, P., Catalan, J., Castiglia, R., Watson, J., Dobigny, G., Robinson, T.J., and Britton-Davidian, J. (2010). A novel sex determination system in a close relative of the house mouse. *Proc. Biol. Sci.* 277, 1049–1056.

Vigneau, S., Augui, S., Navarro, P., Avner, P., and Clerc, P. (2006). An essential role for the DXPas34 tandem repeat and Tsix transcription in the counting process of X chromosome inactivation. *Proc. Natl. Acad. Sci.* 103, 7390–7395.

Wake, N., Takagi, N., and Sasaki, M. (1976). Non-random inactivation of X

chromosome in the rat yolk sac. *Nature* 262, 580–581.

Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A., and Magnuson, T. (2001). Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat. Genet.* 28, 371–375.

Wang, X., Douglas, K.C., VandeBerg, J.L., Clark, A.G., and Samollow, P.B. (2014). Chromosome-wide profiling of X-chromosome inactivation and epigenetic states in fetal brain and placenta of the opossum, *Monodelphis domestica*. *Genome Res.* 24, 70–83.

Warshawsky, D., Stavropoulos, N., and Lee, J.T. (1999). Further examination of the Xist promoter-switch hypothesis in X inactivation: evidence against the existence and function of a P(0) promoter. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14424–14429.

Wen, W.-S., Yuan, Z.-M., Ma, S.-J., Xu, J., and Yuan, D.-T. (2016). CRISPR-Cas9 systems: versatile cancer modelling platforms and promising therapeutic strategies. *Int. J. Cancer* 138, 1328–1336.

West, J.D., Frels, W.I., Papaioannou, V.E., Karr, J.P., and Chapman, V.M. (1977). Development of interspecific hybrids of *Mus*. *J. Embryol. Exp. Morphol.* 41, 233–243.

Wichterle, H., Lieberam, I., Porter, J.A., and Jessell, T.M. (2002). Directed differentiation of embryonic stem cells into motor neurons. *Cell* 110, 385–397.

Williams, L.H., Kalantry, S., Starmer, J., and Magnuson, T. (2011a). Transcription precedes loss of Xist coating and depletion of H3K27me3 during X-chromosome reprogramming in the mouse inner cell mass. *Dev. Camb. Engl.* 138, 2049–2057.

Williams, L.H., Kalantry, S., Starmer, J., and Magnuson, T. (2011b). Transcription precedes loss of Xist coating and depletion of H3K27me3 during X-chromosome reprogramming in the mouse inner cell mass. *Dev. Camb. Engl.* 138, 2049–2057.

Williamson, C.M., Turner, M.D., Ball, S.T., Nottingham, W.T., Glenister, P., Fray, M., Tymowska-Lalanne, Z., Plagge, A., Powles-Glover, N., Kelsey, G., et al. (2006). Identification of an imprinting control region affecting the expression of all transcripts in the *Gnas* cluster. *Nat. Genet.* 38, 350–355.

Würtele, H., and Chartrand, P. (2006). Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 14, 477–495.

- Wutz, A., and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell* 5, 695–705.
- Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* 30, 167–174.
- Xu, K.P., Yadav, B.R., King, W.A., and Betteridge, K.J. (1992). Sex-related differences in developmental rates of bovine embryos produced and cultured in vitro. *Mol. Reprod. Dev.* 31, 249–252.
- Xu, N., Tsai, C.-L., and Lee, J.T. (2006). Transient homologous chromosome pairing marks the onset of X inactivation. *Science* 311, 1149–1152.
- Xu, N., Donohoe, M.E., Silva, S.S., and Lee, J.T. (2007). Evidence that homologous X-chromosome pairing requires transcription and Ctf protein. *Nat. Genet.* 39, 1390–1396.
- Xue, F., Tian, X.C., Du, F., Kubota, C., Taneja, M., Dinnyes, A., Dai, Y., Levine, H., Pereira, L.V., and Yang, X. (2002). Aberrant patterns of X chromosome inactivation in bovine clones. *Nat. Genet.* 31, 216–220.
- Yadav, B.R., King, W.A., and Betteridge, K.J. (1993). Relationships between the completion of first cleavage and the chromosomal complement, sex, and developmental rates of bovine embryos generated in vitro. *Mol. Reprod. Dev.* 36, 434–439.
- Yamada, N., Hasegawa, Y., Yue, M., Hamada, T., Nakagawa, S., and Ogawa, Y. (2015). Xist Exon 7 Contributes to the Stable Localization of Xist RNA on the Inactive X-Chromosome. *PLOS Genet.* 11, e1005430.
- Yang, F., Babak, T., Shendure, J., and Disteche, C.M. (2010). Global survey of escape from X inactivation by RNA-sequencing in mouse. *Genome Res.* 20, 614–622.
- Yang, F., Deng, X., Ma, W., Berletch, J.B., Rabaia, N., Wei, G., Moore, J.M., Filippova, G.N., Xu, J., Liu, Y., et al. (2015). The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. *Genome Biol.* 16.
- Zhang, L.-F., Huynh, K.D., and Lee, J.T. (2007). Perinucleolar Targeting of the Inactive X during S Phase: Evidence for a Role in the Maintenance of Silencing. *Cell* 129, 693–706.

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.-J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322, 750–756.

Zwingman, T., Erickson, R.P., Boyer, T., and Ao, A. (1993). Transcription of the sex-determining region genes *Sry* and *Zfy* in the mouse preimplantation embryo. *Proc. Natl. Acad. Sci. U. S. A.* 90, 814–817.

ANNEXES
