UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

Functional dissection of the *Xist* lncRNA conserved sequences in X-chromosome inactivation

Cláudia Raquel Estopa Gil

Mestrado em Biologia Molecular e Genética

Dissertação orientada por:
Simão Teixeira da Rocha
Rita Zilhão

2019

Abstract

Introduction: Due to the discrepancy in the number of X chromosomes between the sexes in many animal species, dosage compensation mechanisms have evolved to equalize X-linked gene dosage between males and females. In female mammals, dosage compensation is achieved by a remarkable process known as X-chromosome inactivation (XCI), that leads to the epigenetic silencing of an entire X chromosome. XCI is master-regulated by a long non-coding RNA (lncRNA) called *X-inactive specific transcript* (*Xist* in mouse/*XIST* in human), which is exclusively expressed from the future inactive X-chromosome (Xi). *Xist* gene is peculiar, since it encodes for a 17 kb capped, spliced and polyadenylated RNA that is never translated. This lncRNA is monoallelically upregulated randomly from only one of the X chromosomes chosen for inactivation, coating the entire chromosome in *cis* and inducing transcriptional silencing and heterochromatin formation that is stably inherited through cell divisions. *Xist* lncRNA is poorly conserved at sequence level, with the exception of six tandem-repeated regions, known as the A to F repeats. These regions are believed to act as regulatory modules that exert their function through the interaction with specific RNA-binding proteins. However, the relative contribution of each of these RNA modules to XCI is not fully understood.

Methods: Using a well-described system whereby *Xist* is expressed from a tetracycline-inducible promoter in its endogenous position in J1 male mouse embryonic stem cells, we performed deletions for *Xist* D and E tandem-repeated regions by CRISPR/Cas9. Then, we validated and characterized these mutants alongside another mutant type previously generated in the lab (*Xist* ΔF). Three independent clones per mutant type were selected and their potential impact on XCI was addressed. We specifically explored several features related to XCI, namely: *Xist* coating of the X chromosome, through RNA Fluorescence *In Situ* Hybridization (FISH) for *Xist*; *Xist* capacity to silence the genes on the X chromosome, by RT-qPCR and RNA-FISH for X-linked genes; recruitment of typical heterochromatin marks of the Xi, using RNA FISH for *Xist* combined with Immunofluorescence for the histone marks H3K27me3, H2AK119ub and H4K20me1.

Results: *Xist*-TetOP mutants for D- and E-repeats were generated, and, jointly with the previous deletion for the F-repeat, successfully validated. All the characterized mutants (*Xist* ΔF, ΔD and ΔE) were able to form *Xist* domains (or clouds) around the Xi, as observed by RNA-FISH for *Xist*. *Xist* ΔF clouds visually seemed smaller, but this was not quantified. By measuring *Xist* expression levels by RT-qPCR analysis, we suspected a possible instability in *Xist* ΔD transcripts that should be further explored. The X-linked gene silencing of some candidate genes seemed not to be affected in all the *Xist* mutants, but further characterization at X chromosome global level would give a definitive answer. Finally, typical heterochromatin marks of the Xi were explored for two clones of *Xist* ΔF mutants and no major defects were observed, although the two clones (F2 and F3) exhibited some variability.

Conclusion: As a result of this work, a new series of inducible *Xist* mutants was generated, which will be a valuable set of experimental tools for further investigation in the field of XCI. Our initial characterization gave the first hints on the functional role of some unexplored conserved repeats of *Xist* during XCI. In the future, these *Xist*-inducible mutants can be used in high-throughput approaches to enquiry about gene silencing (RNA-seq), chromatin status (ChIP-seq) or to fish protein interactors of the different repeats (ChIRP-MS), to get mechanistic insights in this puzzled process of XCI.

Keywords: X-chromosome inactivation; *Xist*; lncRNA; facultative heterochromatin; CRISPR/Cas9

Resumo

Introdução: A Epigenética refere-se ao estudo de toda a informação não genómica que pode ser herdada de célula para célula ou até de indivíduo para indivíduo. Esta informação inclui modificações químicas dos nucleótidos, alterações na estrutura da cromatina e modificações pós-traducionais das histonas. Existem vários fenómenos epigenéticos que regulam a expressão dos genes nos vários organismos, sendo um deles o mecanismo de compensação da dosagem génica. Este mecanismo ocorre em várias espécies, de forma a equilibrar a discrepância de carga genética do cromossoma X entre machos e fêmeas. Nos mamíferos, este fenómeno denomina-se Inativação do Cromossoma X (XCI, na sigla inglesa), um processo complexo e extremamente regulado capaz de silenciar o cromossoma X quase na sua totalidade. A XCI ocorre durante o desenvolvimento embrionário, no blastocisto, altura em que, em cada célula, ocorre a escolha aleatória do futuro cromossoma X inativo (Xi). Esta escolha é mantida nas divisões celulares que se seguem, perpetuando-se até ao fim da vida do indivíduo. O principal regulador do processo de XCI é um RNA longo não-codificante (lncRNA) denominado *X-inactive specific transcript (Xist* em ratinho/*XIST* em humanos). A transcrição deste lncRNA é ativada aleatoriamente a partir do Xi futuro e envolve uma complexa rede de vários fatores que atuam em *cis* e *trans* na regulação desta ativação monoalélica do *Xist*. Quando a expressão do *Xist* é despoletada, os transcritos cobrem todo o cromossoma X em *cis*, ficando localizados sempre ao redor do mesmo. De seguida, o *Xist* recruta outros fatores, que causam o silenciamento e heterocromatinização do cromossoma. A cromatina do Xi sofre alterações extremas que modificam a sua estrutura 3D, transformando-a numa heterocromatina facultativa bastante estável. No processo de heterocromatinização, várias modificações ocorrem na cromatina como: a perda de marcas associadas à cromatina ativa, o recrutamento de complexos de repressão da cromatina (como as proteínas do grupo Polycomb), alterações pós-traducionais das histonas e metilação do DNA nas zonas promotoras dos genes do X. Em relação ao grande "maestro" do processo de XCI, o *Xist*, este não apresenta uma estrutura muito conservada entre espécies, à exceção de algumas regiões repetitivas denominadas de repetições de A a F. Alguns estudos indicam que estas regiões repetitivas têm funções importantes e distintas, interagindo com diferentes proteínas, de forma a desempenharem um papel específico neste processo. Atualmente, já se sabe que a região repetitiva A está envolvida no silenciamento dos genes do cromossoma X. As regiões B e C, por outro lado, participam no recrutamento dos complexos de proteínas Polycomb (PRC1 e PRC2) depositando várias marcas de histonas características, como a H3K27me3 e a H2AK119ub. O objetivo deste trabalho é estudar as restantes regiões repetitivas do *Xist* (F, D e E), pois a sua função ainda é pouco conhecida, não se sabendo o seu contributo específico para o processo de XCI.

Métodos: Usando um sistema de expressão indutível de *Xist* bastante bem descrito e conhecido no nosso laboratório, as repetições D e E do *Xist* foram removidas com recurso à técnica de edição genómica CRISPR/Cas9 em células estaminais embrionárias de ratinho macho. Neste sistema, a transcrição do *Xist* é induzida a partir do seu *locus* endógeno, com a adição de uma tetraciclina (doxiciclina), sendo possível estudar *in vitro* as várias etapas da XCI, que mimetizam a iniciação do processo tal como ocorre *in vivo*. De seguida, os mutantes gerados (*Xist* ΔD e ΔE), juntamente com outro mutante previamente desenvolvido neste sistema (*Xist* ΔF), foram validados e caracterizados, de forma a desvendar o impacto das diferentes regiões no processo de XCI. De cada um dos tipos de mutantes, foram selecionados três clones independentes para prosseguir com a análise. Foram exploradas várias funções dos mutantes, como a capacidade do lncRNA *Xist* cobrir o cromossoma X, através de RNA-FISH (Fluorescence *In Situ* Hybridization) para este lncRNA, e de silenciar os genes do cromossoma X, por RT-qPCR e RNA-FISH para genes do cromossoma X. Os mutantes *Xist* ΔF foram também avaliados em relação ao enriquecimento nas marcas repressivas de histonas

características da heterocromatina facultativa do Xi. Para tal foi utilizada a técnica de RNA-FISH, para o *Xist*, combinada com imunofluorescência, para as marcas de histonas H3K27me3, H2AK119ub e H4K20me1.

Resultados: Os mutantes gerados (*Xist* ΔD e ΔE), juntamente com outro mutante previamente criado neste sistema (*Xist* ΔF), foram validados e caracterizados com sucesso. Todos os mutantes foram capazes de formar domínios ao redor do cromossoma X inativo, tal como observado por RNA-FISH com sondas para o *Xist*, embora os domínios do clone *Xist* ΔF fossem tendencialmente mais pequenos, facto que ainda necessita de ser quantificado. Complementando a análise por RNA-FISH com a medição dos níveis de expressão do *Xist* por RT-qPCR, levantamos a suspeita de que os transcritos *Xist* ΔD possam ser mais instáveis. Esta observação prende-se com o facto de que, apesar da percentagem de células a expressar *Xist* fosse equivalente à do *Xist* não alterado, o nível global de expressão estava consideravelmente reduzido em relação ao *Xist* sem deleções. A capacidade de silenciamento dos genes no cromossoma X foi inicialmente investigada por RT-qPCR para dois genes (*Rnf12* e *Pgk1*). O silenciamento destes dois genes do cromossoma X não parece afetado em nenhuma das linhas celulares mutadas, comparando com a linha sem deleções. Complementando o RT-qPCR com RNA-FISH para o *Pgk1*, no caso dos mutantes *Xist* ΔF, confirmou-se que este silenciamento não estava afetado. Esta experiência deverá ser alargada a mais genes no cromossoma X, recorrendo, por exemplo, ao método de RNA-Seq. Através desta técnica é possível compreender como os diferentes genes se comportam nos diferentes mutantes de *Xist*. Por fim, o enriquecimento de marcas de heterocromatina facultativa foi analisado em dois clones sem a região F do *Xist* (F2 e F3) por IF/RNA FISH. Ambos os clones apresentaram enriquecimento das várias marcas de heterocromatina estudadas (H3K27me3, H2AK119ub1 e H4K20me1), embora alguma variabilidade na capacidade de recrutamento existisse entre si. De facto, um dos clones analisados (*Xist* ΔF F3) apresenta uma ligeira redução no enriquecimento destas marcas, o que tem que ser estudado futuramente. É importante referir que estes resultados são ainda preliminares e muito deles foram obtidos apenas com uma experiência.

Conclusões: Como resultado deste trabalho, foi gerada uma nova série de mutantes indutíveis para o *Xist*, que serão ferramentas muito úteis para a investigação desenvolvida na área da XCI. Foram validados e caracterizados mutantes para as regiões F, D e E, que estão pouco estudadas neste contexto. A nossa caracterização inicial é o ponto de partida para clarificar a função de cada uma destas regiões neste processo complexo. Outras experiências que podem ser efetuadas para esclarecer alguns aspetos incluem o estudo da estabilidade do *Xist* nas várias linhas celulares e o estudo do recrutamento de outras marcas de histonas, que até ao momento ainda não tenham sido investigadas. No futuro, estas novas linhas celulares podem também ser estudadas através de técnicas que avaliem aspetos a nível global, como por exemplo: o silenciamento génico, através de RNA-Seq; o estado da cromatina, por ChIP-Seq; a interação das várias regiões com proteínas especificas, através de ChIRP-MS. Até ao momento já foram identificadas algumas proteínas que interagem com o *Xist*, através do método de ChIRP-MS, no entanto não se sabe ao certo a que região se ligam. Uma destas proteínas, a proteína SAP18, que parece relacionar-se com as regiões B e C do *Xist*, já foi investigada no nosso sistema através do seu *knock-down*, método que pode ser utilizado futuramente para outros candidatos. Todos estes métodos poderão ajudar a desvendar um pouco mais sobre este processo intricado e complexo que é a inativação de um cromossoma na sua totalidade.

Palavras-chave: Inativação do cromossoma X*; Xist*; RNA não-codificante; heterocromatina facultativa; CRISPR/Cas9

Acknowledgements

I would first like to thank my thesis advisor Dr. Simão Teixeira da Rocha of the Instituto de Medicina Molecular (IMM). I felt like he was always available to hear my doubts whenever I had a question about my research or writing. He also guided me through this entire year of research and showed me the best ways to achieve my goals (and he still does). I would also like to thank to all my colleagues of IMM for their support, advices and constructive critics. A special thanks to my (mini)advisor Ana Raposo, who since the beginning taught me every technique and every trick, found me a spot right next to her and made me feel at home; and to my incredible lab friends (Catarina, Maria, Tajda and Miguel), who made my year much more fun. I would also like to thank Prof. Maria Carmo Fonseca for have given me the possibility of research in her lab.

I would like to acknowledge Prof. Rita Zilhão as the second reader of this thesis, to whom I am gratefully indebted for her valuable comments on this thesis and for all the guidance and teaching through the two master's years. I am also extremely appreciated to my master's colleagues and friends, who made my master a wonderful journey, full of shared knowledge and fellowship.

Finally, I must express my very profound gratitude to my parents, for providing me with constant support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis; to my boyfriend, who makes my life wonderful every day and pushes me to run after my dreams and aspirations; to my close friends, for always being there for me and for giving me the best moments. This accomplishment would not have been possible without them.

Thank you all.

Índex

Index of Figures

List of Abbreviations

Abbreviation	Meaning
ac	Acetylation
cDNA	Complementary DNA
ChIP-seq	Chromatin immunoprecipitation sequencing
ChIRP-MS	Comprehensive identification of RNA-binding proteins by mass spectrometry
CIZ1	Cdkn1A-interacting zinc finger protein 1
CRISPR	Clustered regularly interspaced short palindromic repeats
DAPI	4′,6-diamidino-2-phenylindole
DMEM	Dulbecco's modified Eagle's Medium
DOX	Doxycycline
ESC	Embryonic stem cell
EtOH	Ethanol
FA	Formamide
FBS	Fetal bovine serum
FISH	Fluorescence *in situ* hybridization
FL	Full-length
Fw	Forward
gRNA	Guide RNA
ICM	Inner cell mass
iDRiP	Identification of direct RNA-interacting proteins
IF	Immunofluorescence
INDELs	Insertions and/or deletions
LBR	Lamin-B receptor
LBS	LBR binding site
LIF	Leukemia inhibitory factor
lncRNA	Long non-coding RNA
MATR3	Matrin3
me	Methylation
MEFs	Mouse embryonic fibroblasts
mESC	Mouse embryonic stem cell
NHEJ	Non-homologous end joining
PBS	Phosphate-buffered saline

PCR	Polymerase chain reaction
PFA	Paraformaldehyde
ph	Phosphorylation
PRC	Polycomb repressive complex
PTBP	Polypyrimidine Tract Binding Protein
PTMs	Post-translational modifications
RAP-MS	RNA antisense purification coupled with mass spectrometry
RBP	RNA-binding proteins
RT	Room temperature
RT-qPCR	Reverse transcription quantitative PCR
Rv	Reverse
shRNA	Short hairpin RNA
SMCHD1	Structural maintenance of chromosomes hinge domain containing 1
SSC	Saline-sodium citrate
TAD	Topologically associating domain
TetOP	Tetracycline-inducible operator promoter
ub	Ubiquitylation
VRC	Vanadyl-ribonucleoside complex
WT	Wild-type
XCI	X-chromosome inactivation
Xa	Active X chromosome
Xi	Inactive X chromosome
Xic/XIC	X-inactivation center
Xist/XIST	X-inactive specific transcript
YY1	Yin-Yang 1 transcription factor

Introduction

Epigenetics: the bridge between genotype and phenotype

The term "*epigenetics*", the fusion of epigenesis (or embryonic development) and genetics, was coined by Conrad Waddington in 1942 (reprinted in (Conrad H. Waddington 2012)). Waddington defined the phenomenon as "*the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being*" and created the famous *Epigenetic Landscape* (adapted in Figure 1). During the following decades, the definition evolved as much more is known about its molecular mechanisms. Nowadays, "*epigenetics*" has a more general meaning and refers to all non-genomic information stored in cells that can be inherited, including chemical changes of nucleotides, chromatin structure and histone post-translational modifications (PTMs) (Kane and Sinclair 2019). These changes are important for the regulation of gene expression, which influences cell fate, differentiation and aging (Godini, Lafta, and Fallahi 2018).

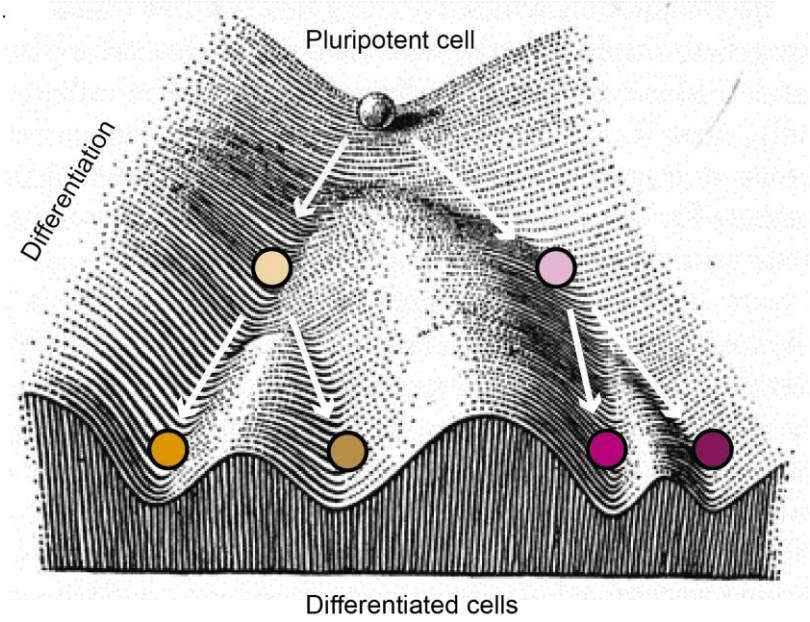


Figure 1. Epigenetic Landscape Model. The sphere represents a pluripotent cell at the verge of taking a developmental path down a hill towards more differentiated irreversible states (adapted from Conrad Waddington's Epigenetic Landscape (C. H. Waddington 1957)).

Epigenetic marks consist mostly of several chemical changes in the DNA and histone PTMs (Annex 1). Within the many DNA modifications, the covalent binding of a methyl group to the carbon 5 of a cytosine is the best studied and is frequently referred simply by DNA methylation (me). This modification is subjected to a dynamic regulation during development and is fundamental for several physiological processes, such as gene regulation, embryonic development and genomic imprinting (reviewed in (Goll and Bestor 2005)). DNA methylation correlates with transcriptional repression when occurs at gene promoters rich in CpG dinucleotides (also known as CpG islands). Mechanisms of DNA methylation, through which DNA methylation "machinery" works, involve writers (DNA methyltransferases), readers (DNA-modified interacting proteins) and erasers (DNA demethylases) (reviewed in (Torres, Kouro, and Kerr 2019)).

Histone modifications – positive and negative regulators of transcription

Gene expression can be profoundly regulated by the chromatin template, on which transcription occurs, particularly during development. Nucleosomes are the structural units of chromatin and are

constituted by an octamer of four core histones (H2A, H2B, H3 and H4) associated with DNA. The nucleosome can also interact with a linker histone (H1) to form the chromatosome for further packaging of the DNA (reviewed in (Fyodorov *et al.* 2018)). Histone proteins condense and structure the DNA of eukaryotic cells, being generally involved in repression of gene transcription, due to reduced DNA accessibility. However, they can also positively regulate gene expression, depending on nucleosome disposition and, also, on specific PTMs, that normally occur at histone N-terminal prominent tails. Based on that differences, chromatin can be classified in euchromatin (open and transcriptionally active chromatin) or heterochromatin (condensed and transcriptionally inactive chromatin). Heterochromatin, on the other hand, can be classified in constitutive (consistently formed throughout the cell cycle) or facultative (locus- and cell-type- specific) (reviewed in (Allshire and Madhani 2018)), the last playing a role in many dosage compensation processes, including the X chromosome silencing in female mammals (Bernstein and Allis 2005; J. T. Lee 2011).

Histone PTMs, like DNA chemical changes, are regulated by writers, readers and erasers. PTMs are linked to either activation or repression, depending on the modified residue, histone and type of modification. The best described PTMs are: acetylation (ac), which normally occurs on specific lysine (K) residues and plays a fundamental role in transcriptional regulation, being usually associated with euchromatin; methylation, linked to either gene expression or gene repression, most commonly happens in lysine and arginine (R) residues of histones H2A, H3 and H4; phosphorylation (ph), that affects serine (S), threonine (T) and tyrosine (Y) residues in all types of histones and is implicated in chromosome condensation during mitosis and induction of early genes; ubiquitylation (ub), that normally occurs at lysine residues, can also be linked to gene expression (e.g. ubiquitylation of histone H2B at lysine 120 - H2BK120ub) or gene repression (e.g H2AK119ub, "written" by the Polycomb repressive complex (PRC) 1) (Osley 2006).

Histone methylation has been one of the PTMs mostly studied and has very diverse ways of action. For example, methylation of histone H3 at lysine 4 (H3K4 methylation) (Byrd and Shearn 2003) is associated with gene activation and euchromatin, being "written" by the Trithorax proteins. On the other hand, the PRC2 is responsible for the histone H3 lysine 27 tri-methylation (H3K27me3) which is a repressive mark associated with facultative heterochromatin and silenced genes (Inoue *et al.* 2017). Additionally, tri-methylation of histone H3 at lysine 9 and lysine 20 (H3K9me3/H3K20me3) is characteristic of constitutive heterochromatin (reviewed in (Allshire and Madhani 2018)). Histone methylation is, therefore, an example on how the same modification can lead to completely different biological outcomes.

X-Chromosome Inactivation: an example of epigenetic gene regulation

Epigenetic regulation of genes is an imperative process to ensure the proper gene dosage, and dosage matter when we talk about X chromosomes. In 1947, Muller coined the term dosage compensation as a response triggered by the X-linked gene differences between males and females. This response indeed exists but varies between species, what is in part explained by the fact that gender is also differently determined. Male fruit fly (*Drosophila melanogaster*), with one X chromosome and one Y, has a two-fold increase in the X chromosome transcription (Belote and Lucchesi 1980) through epigenetic mechanisms that enhance initiation and elongation (Turner, Birley, and Lavender 1992; Gilfillan, Dahlsveen, and Becker 2004). The case of the model organism *Caenorhabditis elegans* is distinct, as the XX hermaphrodite nematode reduces the X chromosome transcript levels by one-half through binding of the two X chromosomes (Meyer and Casson 1986), equalizing the X-linked gene expression to the male worm (XO). In mammals, the imbalance in terms of the X-linked gene dosage is

compensated by the complete inactivation of additional X chromosomes (based on X:A (autosomes) ratio) (Starmer and Magnuson 2009).

The first insights on X-chromosome inactivation (XCI) were made in 1949, when a Canadian researcher observed a highly stained nuclear body in neurons of female cats (Barr and Bertram 1949). This Barr body was also later observed in female cells of other animals, including in humans (Klinger and Schwarzacher 1960). Ohno and his colleagues speculated that the Barr body was a condensed X chromosome (Ohno, Kaplan, and Kinosita 1959), through the observation of heteropyknotic (differently stained) bodies only in female cells, a discovery later confirmed by *in situ* hybridization with specific X chromosome probes (Walker et al. 1991). In 1961, Mary F. Lyon formulated, for the first time, the XCI hypothesis or "Lyonisation". Experimenting in the mouse, the "Grande dame of mouse genetics" made remarkable advances in this field. Lyon hypothesized that one of the female X chromosomes, paternal or maternal, was genetically inactivated in different cells of the same animal, due to its random inactivation early in embryonic development. Once established, the XCI would be stably maintained through cell divisions (Lyon 1961). Few years later, Ohno also developed a complementary hypothesis, stating that the Y chromosome degradation would lead to an upregulation of the X chromosome in males and females, which would be afterwards compensated by the XCI in females (Ohno 1967). This hypothesis was later refuted using gene expression studies comparing sex chromosomes with autosomes (Xiong et al. 2010) and studies of gene expression evolution of the mammalian X chromosome (Lin et al. 2012).

In the female mouse, two waves of XCI occur during embryonic development. First, the paternal X chromosome is inactivated from the 4-cell stage onwards, due to a maternal imprint preventing the maternal X from being inactivated (Mak *et al.* 2004; Okamoto *et al.* 2005; Inoue *et al.* 2017). Imprinted XCI persists during pre-implantation embryogenesis up to the blastocyst stage, where the paternal X chromosome is reactivated in the inner cell mass (ICM) but remains inactive in cells that give rise to extra-embryonic tissues (Patrat *et al.* 2009). In the epiblast, the ICM-derived lineage that gives rise to the embryo, loss of pluripotency is accompanied by a second wave of XCI, where either the paternal or the maternal X chromosome can become inactivated randomly in each cell. The inactive X chromosome (Xi) choice is then stably maintained through cell divisions, resulting in female adult individuals with several degrees of mosaicism for X-linked gene expression, depending on the choice of each precursor cell. Mouse embryonic stem cells (mESCs) derived from the ICM are used as a powerful model system to study XCI *in vitro*, since they still have two active X chromosomes (Xa) that undergo random XCI upon differentiation (Figure 2).

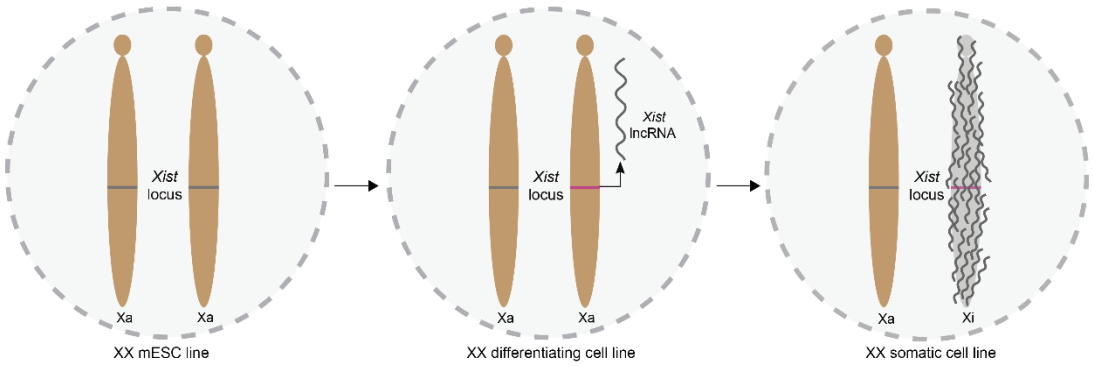


Figure 2. X-chromosome inactivation in mouse female cells. Initially, female mESCs have two Xa; upon differentiation, *Xist* transcription is upregulated from the future Xi and the transcripts coat the entire chromosome in *cis*. Afterwards, X chromosome is silenced and becomes heterochromatic. mESC, mouse embryonic stem cell; *Xist*, *X-inactive specific transcript*; lncRNA, long non-coding RNA; Xa, active X chromosome; Xi, inactive X chromosome.

X-inactivation center: the core of XCI regulation

XCI in eutherians is primarily controlled by a long non-coding RNA (lncRNA)-rich region, located on the X chromosome, known as the *X-inactivation center* (*Xic* in mouse and *XIC* in human) (Brown, Lafreniere, *et al.* 1991). This region contains the master regulator of XCI – the *Xist* (*X-inactive specific transcript*) gene (*Xist* in mouse and *XIST* in human) - that remains active upon X chromosome silencing, contrasting with the majority of other *loci* of the same chromosome. A main repressor element of *Xist* is its antisense-non-coding transcript *Tsix* (Lee *et al.*, 1999). Within *Xic*, *Xist* and *Tsix* promoters are located in adjacent topologically associating domains (TADs), which divide the *Xic* in two separate compartments that adopt opposite transcriptional fates during mESCs differentiation (Nora *et al.* 2012). TADs are 200 kb to 1 Mb regions where dynamic interactions between enhancers and their target promoters preferentially occur, being critical to ensure the proper developmental timing of XCI interacting domains (Galupa and Heard 2015; van Bemmel *et al.* 2019). The *Xist* TAD seems to contain *Xist* activators (e.g. *Rnf12*, *Ftx* and *Jpx*), while *Tsix* TAD seems to contain *Tsix* activators (and, consequently, *Xist* repressors) (e.g. *Tsx* and *Linx*) (Figure 3).

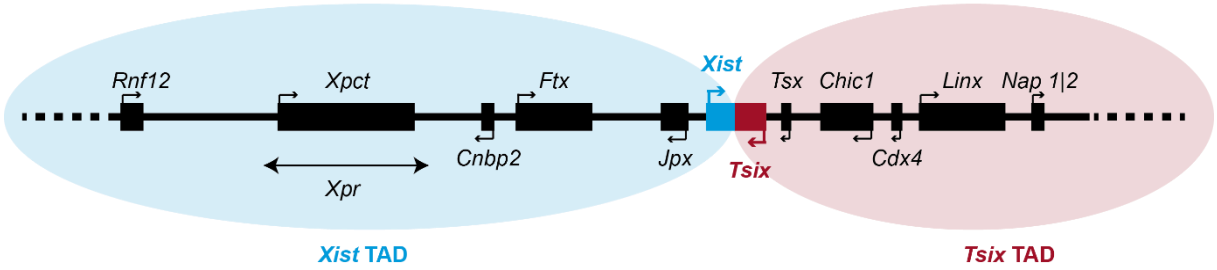


Figure 3. The mouse *X-inactivation center* bipartite organization. *Xic* is a region rich in lncRNAs (e.g. *Xist, Tsix, Jpx, Ftx and Linx)*, being also the location of some protein-coding genes (e.g. *Rnf12*, *Chic1*, *Nap1l2* and *Cdx4*), and is located at a 100- to 500-kb domain in the XqD band (genomic coordinates (GRCm38/mm10) chrX:103,128,040-104,007,846). *Xic* harbors the master regulator of XCI – the *Xist* lncRNA. *Xist* and *Tsix*, despite their overlapping transcriptional units, adopt opposite transcriptional fates during mESCs differentiation, due to their promoters' localization in adjacent TADs. *Xist, X-inactive specific transcript*; TAD, topologically associated domains.

The *Xist* lncRNA is monoallelically upregulated during both imprinted and random XCI, coating the future Xi in *cis* (Figure 2). Its expression is modulated by multiple *cis*- and *trans*-acting factors, however, the main factor responsible for its activation remains unknown. Dose-dependent degradation of the pluripotency factor REX1 by the X-linked activator RNF12 (protein-coding gene on the *Xic*), during differentiation of female mESCs, has been proposed to act as a crucial mechanism in this process (Gontan *et al.* 2012). Decrease of REX1 levels leads to a downregulation of *Tsix*, increasing *Xist* expression and ensuring *Xist* induction on one of the X chromosomes. In turn, *Xist* upregulation inhibits immediately *Rnf12* expression in *cis*, lowering the dose of RNF12 protein to insufficient levels to degrade REX1, preventing the inactivation of the second X chromosome (J. T. Lee, Davidow, and Warshawsky 1999; Gayen *et al.* 2015). *Jpx* and *Ftx* may have some influence on *Xist* activation as well, but how they act is not completely understood yet (Carmona et al. 2018; Sun et al. 2013; Furlan et al. 2018).

Xist lncRNA: the master regulator of XCI

Human *XIST* lncRNA was discovered in the early 90's, when Carolyn Brown and her colleagues were searching for complementary DNA (cDNA) of the human steroid sulfatase locus. They found a cDNA with a characteristic feature: it was only expressed from the Xi in female cells (Brown, Ballabio, *et al.* 1991). This cDNA, was mapped in the region Xq13, where the human *XIC* lies (Brown, Lafreniere, *et al.* 1991). Independently, a mouse *Xist* cDNA was isolated and mapped in the same year by Neil Brockdorff and his colleagues, using *XIST* cDNA probes (Brockdorff *et al.* 1991). Few years later, it

was suggested a role for *Xist* in the initiation of XCI, due to the onset of *Xist* expression in mouse development preceding the XCI (Kay *et al.* 1993).

The *Xist* gene encodes a capped, spliced and polyadenylated, and yet untranslated RNA (Brockdorff *et al.* 1992; Brown, Ballabio, *et al.* 1991). It is a remarkably long RNA with approximately 17 kb in size and poorly conserved at the sequence level, with the exception of six tandem-repeated regions, named A to F (Figure 4). These regions have similarities at sequence level between different mammals (Nesterova *et al.* 2001; Yen *et al.* 2007) and are believed to interact with specific RNA-binding proteins (RBPs), performing distinct functions to ensure a correct XCI process (Chu *et al.* 2015; Bousard *et al.* 2019).

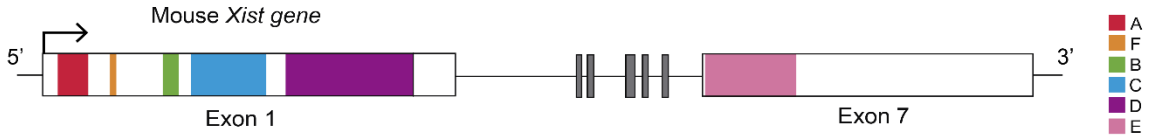


Figure 4. Mouse *Xist* gene structure with tandem-repeated regions. Representation of *Xist*-201 transcript (ENSMUST00000127786.3). Repeats are represented with different colors: repeats A-D + F are localized on exon 1; E-repeat localizes on exon 7. *Xist, X-inactive specific transcript.*

To study the different repeats of *Xist* in the context of XCI, several systems have been used, each one with its advantages and disadvantages. Female mESCs have been the most extensively used model, because they carry two Xa and faithfully mimic the pluripotent environment of the ICM. Moreover, mESCs *in vitro* differentiation reproduces what happens *in vivo* in terms of XCI (*Xist* upregulation, transcriptional inactivation of X-linked genes and histone marks accumulation on the Xi). The disadvantage of using female mESCs as a model is that XCI occurs only in a subset of differentiated cells in a non-synchronized manner. This is because they are subjected to an intricate *cis*- and *trans*-regulation, which results in inefficient *Xist* upregulation in most of the cells. To overtake that problems, *Xist* inducible systems have been generated. In this type of systems, *Xist* expression is induced by the addition of an external substance, normally of the tetracycline class, like doxycycline (DOX). Many *Xist* inducible systems have been generated in the context of autosomes or the X chromosome in both male and female mESCs (Wutz, Rasmussen, and Jaenisch 2002; Tang et al. 2010; Engreitz et al. 2013) and also in human somatic cells (Chow et al. 2007). *Xist* cDNA transgenes on autosomes are quite easy to engineer, however they provide the less efficient system to study XCI. This is due to position-effects that can undermine *Xist* transgenic expression or by the general reduction capability of *Xist* to silence autosomal genes (Loda et al. 2017). Male mESC lines expressing an inducible *Xist* from the endogenous *locus* allow an easy manipulation of *Xist* DNA, due to the existence of a single X chromosome per cell. That fact provides an easy read-out for determining XCI efficiency, since correct inactivation leads to cell death due to X chromosome nullisomy (Wutz, Rasmussen, and Jaenisch 2002; Engreitz et al. 2013; Simão Teixeira da Rocha et al. 2014; Bousard et al. 2019). Recently, it has also been generated several female *Xist* inducible cell lines (Schulz *et al.* 2014; Nesterova *et al.* 2019), that mimic *in vivo* XCI in a synchronized manner in the majority of cells, without causing cell death. A possible, and very contested, disadvantage of inducible systems is their *Xist* super-physiological levels (Engreitz *et al.* 2013; Bousard *et al.* 2019), which might influence the obtained results and might not necessarily simulate what normal *Xist* levels do.

In any case, *Xist*-inducible transgenes have been very useful to understand the dynamics of XCI. Indeed, *Xist*-inducible transgenes on autosomes of male mESCs showed that XCI is mediated by *Xist* and can be described in a three-stage manner (initiation, establishment, and maintenance) during mESCs differentiation (Wutz and Jaenisch 2000) (Figure 5). Initiation phase occurs in the first 1.5 days of

mESCs differentiation, when *Xist* is monoallelically upregulated from the X chromosome chosen for inactivation. During the initiation phase, *Xist* spreads and coats the entire chromosome in *cis* and gene silencing is initiated, at least for some genes. Establishment phase occurs during the following brief 24-hour period and can be seen as a period in which the silenced state is established. Until this phase, *Xist* needs to be attached to the Xi to avoid transcriptional reactivation. The maintenance phase is an irreversible phase during which Xi is stably silenced by factors brought by *Xist* RNA. At this point, expression of *Xist* RNA is no longer required to maintain the inactive state and the Xi can be stably propagated through cell divisions.

Following *Xist* accumulation on the X chromosome, chromatin modifications also occur sequentially during the differentiation of female mESCs or in the embryo (reviewed in (Martin Escamilla-Del-Arenal, da Rocha, and Heard 2011)). Upon *Xist* coating, occurs a rapid loss of histone modifications linked with gene activity, such as mono- and tri-methylation of histone H3 at lysine 4 and acetylation of histone H3 at lysine 9 and lysine 27 (J Chaumeil et al. 2002; Jeppesen and Turner 1993), and depletion of RNA polymerase II machinery (J Chaumeil et al. 2002), as seen by RNA fluorescence i*n situ* hybridization (FISH) combined with immunofluorescence (IF) experiments. More recently, native chromatin immunoprecipitation sequencing (ChIP-seq) experiments showed that H3K27 deacetylation is not only one of the earliest chromatin alterations during XCI but it is also tightly linked to transcriptional silencing of X-linked genes (Żylicz *et al.* 2019).

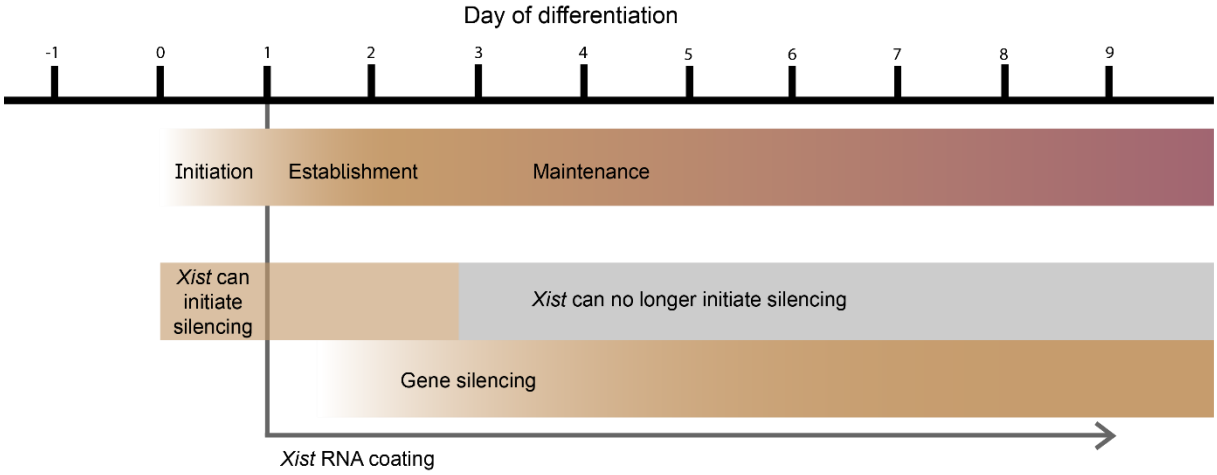


Figure 5. The three steps of the *Xist* RNA-mediated silencing process: initiation, establishment, and maintenance. Initiation phase occurs within the first 1.5 days of mESCs differentiation and is when induction of *Xist* RNA can initiate gene silencing. The, following and brief, establishment phase is a period in which the silenced state is locked. The maintenance phase is distinguished by being irreversible and independent of *Xist* RNA. *Xist, X-inactive specific transcript.*

After histone deacetylation and initiation of gene silencing, *Xist* recruits several chromatin repressive complexes. The best described ones are the Polycomb repressive complexes (PRC1 and PRC2), which are large multi-component machineries that modify the histone tails with repressive marks, functioning through chromatin-based and epigenetic mechanisms. PRC1 is responsible for the mono-ubiquitylation of histone H2A at lysine 119 (Shao *et al.* 1999; H. Wang *et al.* 2004), while PRC2 is responsible for the mono-, di-, and tri-methylation of histone H3 at lysine 27 (Cao *et al.* 2002). PRC1, in fact, can be subdivided into two distinct types of complexes, both of which are recruited to the Xi as discovered by Tavares *et al.* (Tavares et al. 2012): canonical PRC1 complexes, that bind to H3K27me3 to occupy chromatin modified by PRC2, however with a small contribution for H2AK119ub deposition (Almeida et al. 2017); variant PRC1 complexes, the most active H2AK119 ubiquitin ligases *in vitro*, being the driving force of H2AK119ub deposition on the X chromosome (Almeida et al. 2017). H2AK119ub1 accumulation occurs rapidly, showing evidence for the initial recruitment of the PRC1 by *Xist*, followed by PRC2 and other non-canonical PRC1 complexes (Almeida *et al.* 2017; Pintacuda

et al. 2017). A previous study with female mESCs showed that PRC2 recruitment occurs, after initial H2AK119ub1 accumulation, via JARID2 or AEBP2, leading to H3K27me3 accumulation slightly later (Simão Teixeira da Rocha *et al.* 2014; Cooper *et al.* 2016).

Apart from the Polycomb marks, there are other classically defined marks of the facultative heterochromatin of the Xi, such as H3K9me2 and H4K20me1. H3K9me2 seems to mediate the Cdyl interaction with the Xi, a reader for this mark (M. Escamilla-Del-Arenal et al. 2013). However, the histone methyltransferase responsible for its enrichment on Xi remains unclear, being G9a and ESET possible candidates. H4K20me1 mark, catalyzed by PR-Set7, is also associated with *Xist* expression and initiation of XCI (Kohlmaier *et al.* 2004; Calabrese *et al.* 2012), but how it gets recruited and the actual importance of this mark for XCI remains unknown.

Finally, during maintenance phase, Xi becomes late replicating and incorporates a specific histone variant, macroH2A, which becomes enriched on the whole chromosome and is believed to be important for maintenance of silencing (Costanzi and Pehrson 1998; Hernández-Muñoz *et al.* 2005). Around the same time, promoter sequences of X-linked genes undergo CpG methylation to lock the inactive state (Blewitt *et al.* 2008; Gendrel *et al.* 2012). The final outcome is a stable silencing status of most Xi genes, with the exception of a 3-7% genes that escape XCI in mouse and are known as escapees (Yang et al. 2010; Berletch et al. 2015).

During XCI, the Xi not only changes at chromatin level but also undertakes a massive 3D re-folding. The Xi acquires two large interacting domains separated by a tandem repeat locus (Dxz4), with a global attenuation of TADs (Minajigi et al. 2015; Deng et al. 2015; Giorgetti et al. 2016). The Xi "compartment-less" bipartite structure contrasts with the typical A/B compartmentalization of mammalian chromosomes and it seems necessary during *de novo* XCI to allow the spreading of *Xist* RNA (Giorgetti et al. 2016; C.-Y. Wang et al. 2019). The transition between A/B compartments and the "compartment-less" structure might occur through S1/S2 compartments, that are merged by the structural maintenance of chromosomes hinge domain containing 1 (SMCHD1) protein (Blewitt et al. 2008; Gendrel et al. 2012; C.-Y. Wang et al. 2019).

Xist as a multi-tasking molecule

Xist is a lncRNA that can structurally and functionally transform an active chromosome into a silenced and heterochromatic chromosome organized in a unique 3D conformation ((C.-Y. Wang *et al.* 2019); reviewed in (Simão T. Da Rocha and Heard 2017)). The molecular mechanisms through which *Xist* coordinates these structural and functional alterations are still being unveiled. A great leap forward was the discovery of the *Xist* RNA-protein interactome, as revealed by three independent proteomic studies in 2015: a comprehensive identification of RBPs by mass spectrometry (ChIRP-MS) (Chu *et al.* 2015) revealed a list of 81 direct or indirect *Xist* interacting proteins, like HNRNPU, HNRNPK, SPEN, WTAP, RNF20, MATR3, SAP18 and Polypyrimidine Tract Binding Protein (PTBP)1; a RNA antisense purification coupled with mass spectrometry (RAP-MS) (McHugh *et al.* 2015) identified 10 *Xist* binding proteins, including SPEN, HNRNPU, Lamin-B receptor (LBR) and PTBP1; an identification of direct RNA-interacting proteins (iDRiP) (Minajigi *et al.* 2015) identified SPEN and other proteins involved in maintenance phase, such as cohesins, condesins and chromatin factors. All three methods include a cross-linking step to preserve *in vivo Xist*-protein interactions, which were then captured with antisense oligonucleotides complementary to *Xist* and identified by mass spectrometry. Also in 2015, forward genetic studies performed in mESCs also identified proteins that have an impact on *Xist*-mediated gene silencing (Moindrot et al. 2015; Monfort et al. 2015), some of which were common to the candidate *Xist* interactors found in the proteomic approaches (Chu et al. 2015; McHugh et al. 2015; Minajigi et al. 2015). Moindrot and colleagues with a dual pooled lentiviral short hairpin RNA (shRNA) approach

identified SPEN, RBM15 and WTAP among the main proteins involved in *Xist*-mediated gene silencing (Moindrot *et al.* 2015). Monfort *et al.* through a gene-trap screen in haploid cells identified 6 proteins involved in *Xist*-mediated gene-silencing, being SPEN on the top of the list (Monfort *et al.* 2015) (reviewed in (Simão T. Da Rocha and Heard 2017)) (see Annex 2 for details). Some of the isolated proteins were then reported to bind to specific *Xist* repeated regions, which are believed to perform specific functions (Wutz, Rasmussen, and Jaenisch 2002; Chu et al. 2015; Sunwoo et al. 2017; Ridings-Figueroa et al. 2017; Bousard et al. 2019).

The best studied and most conserved *Xist* repeated region is the A-repeat, which lies on *Xist*'s 5'end and is crucial for the X-linked gene silencing of Xi (Wutz, Rasmussen, and Jaenisch 2002). Deletion of the A-repeat results in total absence of gene silencing of practically all the genes in mESCs and female embryos (Wutz, Rasmussen, and Jaenisch 2002; Sakata *et al.* 2017). Several studies showed that this region can interact with specific proteins important for its function, such as: SPEN (an RBP involved in transcriptional repression), RNF20 (an E3 ubiquitin ligase involved in H2BK120 ubiquitylation), WTAP (a component of N6-adenosine–m6A RNA-methylation machinery) and RBM15 (member of the SPEN family) (Chu *et al.* 2015; McHugh *et al.* 2015; Moindrot *et al.* 2015; Patil *et al.* 2016).

The B-repeat, alone or in conjunction with C-repeat, was shown to be involved in PRC1/2 recruitment, impacting primarily on the heterochromatinization process and, secondarily, on *Xist*-mediated gene silencing (Simão Teixeira da Rocha *et al.* 2014; Zhao *et al.* 2008; Bousard *et al.* 2019; Colognori *et al.* 2019). It was identified that this region binds to HNRNPK protein, recruiting the non-canonical PRC1 (Cirillo et al. 2016; Pintacuda et al. 2017; Bousard et al. 2019)

Not much is known about the short repeat right after A-repeat, the F-repeat. Deletions containing this region caused loss or significant weakening of *Xist* clouds and a reduction in *Xist* RNA level (Colognori *et al.* 2019; Jeon and Lee 2011). In humans, this region binds to HNRNPM protein (Nostrand *et al.* 2016), a negative regulator of transcription, that was also described as being a mouse *Xist*-binding protein (Chu et al. 2015; McHugh et al. 2015; Minajigi et al. 2015). Interestingly, in the vicinity of the F-repeat region, there are a few elements also with some degree of conservation which are worth mentioning. For example, several binding sites for the Yin-Yang 1 transcription factor (YY1) lie a few base pairs 5'end to the F-repeat. Binding of YY1 to this region seems to be important for *Xist* expression in both mouse and human cells (Makhlouf et al. 2014; Chapman et al. 2014) and perhaps for *Xist* localization, acting as a putative docking site for *Xist* spreading (Jeon and Lee 2011). Additionally, F-repeat is part in one of the several LBR binding sites (LBS). LBR protein was found in some proteomic studies (McHugh *et al.* 2015; Minajigi *et al.* 2015) and Chen *et al.*, using a male *Xist*-tetracycline-inducible operator promoter (TetOP) mESC line, showed that it recruits the future Xi to the nuclear lamina. The deletion of the LBS region which includes *Xist* F-repeat (ΔLBS) in mESCs was shown to abolish the X-linked gene silencing (Chen *et al*. 2016). More recently, it was shown that *Xist* ΔLBS has a mild effect on X-linked gene silencing but a stronger effect than the LBR KO, which could indicate additional functions for the LBS region (Nesterova *et al.* 2019).

The D-repeat is the longest repeat of *Xist* and, still, has an unknown function. In human cells (HEK and Hela), the deletion of *XIST* D-repeat significantly decreased the expression of *XIST* RNA and led to an upregulation of at least six X-linked genes, comparing with control cells (without D-repeat KO) (Lv *et al.* 2016). In humans, by enhanced Cross-Linking and ImmunoPrecipitation, this region was identified as the HNRNPK protein binding site (Nostrand *et al.* 2016), a protein that in mouse *Xist* binds to B- and C-repeats (Bousard et al. 2019). This interesting finding suggests the possibility that, although

the similarities at sequence level between mouse and human D-repeat, there might be differential RBPs binding to it and, as a consequence, a functional specification of the repeated region in different species.

The only tandem-repeated region of the second longest exon of *Xist* (exon 7) – the E-repeat – is a highly repetitive AT-rich region. This repeat was previously shown to be important for *Xist cis* localization, as severe disruption and aberrant morphologies of *Xist* clouds were seen in different types of endogenous and inducible *Xist* ΔE mutants in mouse cells (Ridings-Figueroa *et al.* 2017; Sunwoo *et al.* 2017; Yue *et al.* 2017). Several RBPs were shown to bind to this region, such as: CIZ1 (Cdkn1A-interacting zinc finger protein 1) (Ridings-Figueroa *et al.* 2017); PTBP1 and PTBP2 (known as differentiation-induced *Xist* splicing regulators) (Chen et al. 2016; Vuong, Black, and Zheng 2016) and MATR3 (Matrin3 protein; normally binds directly to intronic pyrimidine-rich sequences and controls alternative splicing) (The ENCODE Project Consortium 2012). MATR3 and PTBP1 have a putative role in *Xist*-mediated gene silencing (Moindrot *et al.* 2015), but not much is known about their function on the XCI context. *Xist* ΔE mutants in female mouse embryonic fibroblasts (MEFs) showed near-complete loss of H3K27me3 from the Xi (Sunwoo *et al.* 2017), but this could simply be a consequence of poor chromosome coating. A study in human K562 cells showed that *XIST* E-repeat seems to be important for both *XIST* cloud formation and XCI maintenance (H. J. Lee *et al.* 2019).

The fact that *Xist* is able to silence genes through the A-repeat and recruit Polycomb proteins through the B- and C-repeat highlights for the fact that *Xist* has several different RNA modules and works as a multi-tasking molecule to ensure a chromosome-wide stable Xi. However, the available data often comes from diverse cellular models, masking or confounding the precise contributions of each one of the repeated regions within *Xist* for the initiation of XCI. Previous work from the host lab dissected the role of *Xist* A-, B- and C-repeats in the context of an inducible *Xist* allele expressed from its endogenous location in XY mESCs (Wutz, Rasmussen, and Jaenisch 2002; Bousard *et al.* 2019). To gain an integrative view of the role of all the repeats within *Xist* gene, we decided to address the role of the remaining repeats (F, D and E) in the same cellular model. We choose to use this system because it allows a synchronized expression of *Xist* in a large fraction of cells, is easy to manipulate and *Xist* mutant allele expression is not skewed toward the wild-type (WT) allele (because it is an inducible system). The knowledge of each repeat function will be crucial to dissect the molecular mechanisms through which this unusual and multi-tasking lncRNA acts to ensure a stable inactive state of a full chromosome.

Aims of the Study

In this study, our main goal is understanding the multi-tasking nature of the *Xist* lncRNA at regulating XCI. More precisely, we aim to reveal the functions of the F-, D- and E-repeats of *Xist* in this process. For this purpose, we intend to use a mutagenesis approach whereby these different repeats will be deleted using the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas9 genome editing technique. Our specific objectives are the following:

(1) Generation of novel *Xist*-inducible mutants by CRISPR/Cas9, involving D- and E-repeats (F-repeat deletion was previously performed in the host lab, but not studied);
(2) Screening and molecular characterization of the novel *Xist* mutants for F-, D- and E-repeats using molecular biology techniques, such as Polymerase Chain Reaction (PCR), Reverse Transcription Quantitative PCR (RT-qPCR) and Sanger sequencing;
(3) Functional dissection of the novel *Xist*-inducible mutants in terms of ability to coat the X-chromosome (by RNA-FISH), to induce facultative heterochromatin (by IF/RNA-FISH) and to silence X-linked genes (by nascent-transcript RNA-FISH and RT-qPCR).

Materials and Methods

All buffer solutions were prepared using autoclaved distilled water and stored at room temperature (RT) unless stated otherwise. All reagents are stored on ice when in use.

Cell Culture

The previously published *Xist*-TetOP XY mESCs (Wutz, Rasmussen, and Jaenisch 2002) and *Xist*-TetOP XY ΔF mESCs (Raposo, unpublished results) were maintained in classic embryonic stem cells (ESCs) medium - Dulbecco's modified Eagle's Medium 1x with high glucose, L-glutamine, and sodium pyruvate (DMEM) (Gibco) containing 15% ESC Fetal Bovine Serum (FBS), 1% Glutamine, 10^3 U/ml Leukaemia Inhibitor Factor (LIF) (Millipore), 10^{-4} mM 2-mercaptoethanol, 50 U/ml penicillin and 50 µg/ml of streptomycin. *Xist*-TetOP XY mESC line was used to generate the *Xist* mutants ΔD and ΔE (see Generation of *Xist*-TetOP mutants by CRISPR/Cas9 genome editing).

All ESCs were grown at 37°C in 5% CO_2 and medium was changed daily. Inducible expression of *Xist* driven by a TetOP promoter was achieved by adding DOX (1.5 µg/ml) while differentiating the ESCs in LIF withdrawal medium - DMEM media containing 10% FBS, 1% Glutamine, 10^{-4} mM 2-mercaptoethanol, 50 U/ml penicillin and 50 µg/ml of streptomycin, for 2 days.

Generation of *Xist*-TetOP mutants by CRISPR/Cas9 genome editing

All guide RNAs (gRNAs) were designed using a combination of CRISPR design tools available online: Benchling (https://benchling.com/crispr), CHOPCHOP (http://chopchop.cbu.uib.no), CRISPOR (http://crispor.tefor.net) and IDT (https://eu.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM) and selected based on a conjugation of the best rated sequences in terms of on- and off-targets effects (sequences of the gRNAs used for each *Xist* mutant are displayed in Annex 3). Each synthesized oligonucleotide (after annealing) has overhangs 5' CACC on the RNA antisense strand and 5' AAAC on the sense strand that allow for cloning into pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene, plasmid #62988) digested with BpiI/BbsI restriction enzyme. gRNAs were ligated to PX459 plasmids using T4 DNA ligase (Thermo Fisher) and transformed in DH5α competent cells. Correct ligation of the gRNA oligonucleotides was confirmed by Sanger sequencing (see Sanger sequencing) with plasmid specific sequencing primer (Annex 3). Upon sequencing validation, plasmids were further amplified in DH5α competent cells and then extracted with Maxiprep kit (NZYTech) according to the manufacturer's instructions.

To generate *Xist* ΔD and ΔE mutants, around $4x10^5$ cells were co-electroporated with 2.5 µg of DNA of two PX459 plasmids (Addgene) expressing the Cas9 endonuclease and chimeric gRNAs flanking the region to delete. Delivery of plasmids into mESCs was performed by electroporation using the Neon™ Transfection System (Thermo Fisher) at 1400 Volts, 10 ms, 3 pulses setting. After delivery of gRNA/Cas9 plasmids, transfected cells were separated by sequential dilutions and cultured without selection drug for around 8 days. When single clones were visible and properly separated, they were picked under a microscope to a 96-well plate. 96 clones were picked per mutant type. When suitable, cells in the 96-well plate were split in two new 96-well plates (one for DNA extraction and the other for freezing). The whole plate was screened by genomic PCR using primer sets (Annex 3) for each specific deleted region. Positive clones of each *Xist* mutant type were expanded and further validated for the mutation and absence of *Xist* full-length (FL) band. Amplicons from the PCR for the *Xist* deletions were Gel-purified (ΔF mutants) and PCR-cleaned (ΔD and ΔE mutants) using NZYGelpure kit (NZYTech) and sequenced by the Sanger method (see Sanger sequencing).

Genomic DNA Preparation from Cultured Cells

mESCs were cultured until reaching 70-90% confluency and collected by detachment induced by trypsin treatment, resuspended with normal ESCs medium and centrifuged at 1,000 rpm for 3 min. After centrifugation, supernatant was removed and the pellet resuspended in 1x phosphate-buffered saline (PBS). Cells were centrifuged again at 1,000 rpm for 3 min and the supernatant discarded. The cell pellet was then lysed overnight at 56ºC, with 500 µl of Lysis buffer (100 mM NaCl, 10 mM Tris pH 8.0, 25 mM EDTA pH 8.0 and 0.5% SDS) containing proteinase K (0.2 µg/µl).

DNA was purified with 55 µl of 3 M Sodium Acetate pH 5.2 and 500 µl of phenol:chloroform:isoamyl alcohol, mixed vigorously and centrifuged at max speed (14,000 rpm) for 10 min at RT. The clear top aqueous phase was carefully transferred to a fresh Eppendorf tube and 500 µl of Chloroform was added, mixed vigorously and centrifuge at max speed for 10 min at RT. The clear top aqueous phase was carefully transferred again to a fresh Eppendorf tube and 450 µl of Isopropanol was added, mixed by inverting several times and incubated at -20ºC (for at least 1 h). Then the mixture was centrifuged at max speed for 30 min at RT and the supernatant removed carefully. After the addition of 700 µL of freshly made 70% ethanol (EtOH), the sample was centrifuged again at max speed for 5 min at RT and the supernatant removed carefully. Pellet was air dried with open Eppendorf (15-30 min) and resuspended in 200 µl of DNAse/RNAse free dH_2O by gently mixing up and down. Incubation occurred for a few hours at 37ºC for the pellet to be well resuspended and the concentration measured using Nanodrop 1000/2000 (Thermo Fisher).

Polymerase Chain Reaction

PCRs were performed with the Xpert Fast Hotstart Mastermix (2x) (Grisp, ref #GE35.5001), following the conditions in Annex 4. Forward (Fw) and reverse (Rv) primers were designed to flank the regions of interest (sequences in Annex 3). After PCR, DNA samples ran in a 1%-2% agarose electrophoresis gel with Xpert Green (Grisp) in 1x TAE (2 M Trizma Base; 0.5 M EDTA pH 8; 5.7% acetic acid) for 30-45 min at 85-100 V. Product sizes were compared with a 1kb plus DNA ladder (Invitrogen) and visualized with UV light in the Imaging System Chemidoc (Bio-Rad).

Sanger sequencing

CRISPR/Cas9 plasmids (PX459 + gRNAs) and DNA sequence of *Xist*-TetOP mutants mESCs were sequenced by the Sanger method (StabVida) using the Fw primer (Annex 3). The required DNA amounts were: 10 µl of DNA (100 ng/µl) + 3 µl of Fw primer (10 µM).

RNA Isolation, cDNA synthesis and RT-qPCR expression analysis

Total RNA was isolated from the different *Xist*-TetOP mutant mESCs at D2 (from both DOX and noDOX conditions). After growth media removal from culture dish, 1 ml of NZYol Reagent was added directly to the cells (on a 6-well-plate) and incubated for 5 min at RT to allow complete dissociation of the nucleoprotein complex. After pipetting up and down, to destroy cell aggregates, the samples were transferred to an eppendorff tube. 0.2 ml of chloroform was added per 1 ml of NZYol used for homogenization and the tube shaken vigorously by hand for 15 sec. After incubation for 2-3 min at RT, samples were centrifuged at 12000 g for 15 min at 4ºC. Aqueous phase was removed carefully and transferred into a new tube. To precipitate the RNA, 1 µl of glycogen blue (15 µg/µl) and 0.5 ml of 100% isopropanol were added to the aqueous phase. After incubate for 10 min at RT, samples were centrifuged at 12000 g for 10 min at 4ºC. The supernatant was discarded and the RNA pellet washed with 1 ml of 75% ethanol. The sample was vortexed briefly and centrifuged at 7500 g for 5 min at 4ºC. The wash was discarded and the RNA pellet air dried for 10 min. RNA pellet was resuspended

in 30 µl RNase-free water by passing the solution up and down several times through a pipette tip and further quantified in Nanodrop.

Extracted RNA was DNAse I treated (Roche) for 30 min at 37°C, to remove contaminating genomic DNA, accordingly to manufacturer's protocol. Subsequently, the RNA template was reverse transcribed using the Transcriptor High Fidelity cDNA Synthesis Kit (Roche), according to the manufacturer's instructions. For quantitative PCR, gene expression levels were quantified using 2x SYBR Green PCR Master Mix (Applied Biosystems) in a ViiA 7 384-well (Applied Biosystems) using gene specific primers (Applied Biosystems) (Annex 5). Expression levels were normalized to GAPDH using the Δ^{-CT} method (2^{-dCT} formula).

Stellaris RNA-FISH

Stellaris RNA-FISH was performed to analyze *Xist* coating in *Xist* FL, ΔF, ΔD and ΔE mESCs differentiated for 2 days in DOX conditions on gelatin-coated 22 × 22 mm coverslips. The set of Stellaris RNA-FISH probes was designed using the Stellaris® Probe Designer software (Biosearch Technologies) within *Xist* exon 7. The set comprised 48 singly labeled oligonucleotides labeled with Quasar® 570 dye, 46 of which mapped outside the *Xist* ΔE deletion (Annex 6). Hybridization conditions for RNA-FISH were followed according to Stellaris® guidelines using a final concentration of 125 nM of probe set per coverslip. Briefly, cells were washed with PBS and fixed with 3.7% paraformaldehyde (PFA) in PBS for 10 min at RT. After rinsing with PBS and washed one time with 70% EtOH, samples were incubated with 70% EtOH for 1h at RT. Then, samples were washed in washing buffer (10% formamide (FA), 2x saline-sodium citrate (SSC)) for 5–10 min at RT before probe hybridization. The coverslips containing the samples were then removed from the washing buffer and transferred to parafilm containing 25 µl of hybridization buffer (10% dextran sulfate, 10% FA, 2x SSC) with 125 nM of each probe set per coverslip and incubated overnight at 37°C in a moist chamber. The following day, cells were washed twice with washing buffer (30 min at 37°C), followed by a single wash with 2x SSC (5 min at RT). After, nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI) (Sigma-Aldrich), diluted 1:10000 in 2x SCC for 5 min at RT, followed by two washes in 2x SSC (5 min at RT), before being mounted with mounting media (0.1% phenylenediamine (Sigma-Aldrich), 10% PBS 10x and 90% Sterile Glicerol (Sigma-Aldrich)). Z-stack images (30 slices at 0.4 µm) of each sample were acquired in a wide-field fluorescence microscope Zeiss Axio Observer (Carl Zeiss MicroImaging) equipped with an Axiocam 506 mono CCD camera using a 63x/1.4 oil objective (Plan-Apochromat) and filter sets FS43HE (HE DsRed) for Quasar 570, FS38HE (HE GFP) and FS49 for DAPI. The acquired z-stacks were deconvolved using the Huygens Remote Manager software (Scientific Volume Imaging, The Netherlands, http://svi.nl), using the CMLE algorithm, with SNR:50 and 100 iterations. Deconvolved z-stacks were then processed and analyzed in FIJI ((Schindelin et al. 2009), https://fiji.sc/).

Combined RNA-FISH for *Xist* and the nascent-transcript of *Pgk1*

RNA-FISH probes for *Xist* (a 19 kb genomic λ clone 510) (Julie Chaumeil et al. 2008) and *Pgk1* (a 15-16 kb genomic sequence starting 1.6 kb upstream of *Pgk1* gene up to its intron 6) (kind gift from T. Nesterova, Univ. of Oxford) (Moindrot et al. 2015) were prepared using the Nick Translation Kit (Abbot) with red and/or green dUTPs (Enzo Life Sciences).

RNA-FISH was done in *Xist*-TetOP mutant differentiating mESCs accordingly to established protocols (Julie Chaumeil et al. 2008), with minor modifications. Briefly, cells were dissociated with trypsin (Gibco) and adsorbed onto Poly-L-Lysine(Sigma-Aldrich)-coated 22x22 mm coverslips for 10 min. Cells were then fixed in 3% PFA in PBS for 10 min at RT, washed with PBS and permeabilized with 0.5% Triton X-100 diluted in PBS with 2 mM vanadyl-ribonucleoside complex (VRC) (New

England Biolabs) for 4 min on ice. Coverslips were then washed three times with PBS and incubated 3 min in increasing concentrations of EtOH (70%, 80%, 95% and 100%) with agitation. Coverslips were air-dried quickly in filter paper before hybridization with the fluorescent labelled probes.

Probes (3 µl/coverslip) were EtOH precipitated with sonicated salmon sperm DNA (1 mg/ml), 3 M NaAc (1:10 dilution) and mouse *Cot1* DNA (1 µg/µl) (for *Pgk1* probes), diluted in FA and denatured at 75ºC for 7 min (in the case of *Pgk1* probes, they were let incubating at 37ºC after denaturation). *Xist* (in green) and *Pgk1* (in red) probes were co-hybridized in RNA-FISH Hybridization Buffer (40% dextran sulfate, 20% 20x SSC, 20% bovine serum albumin (10 mg/µl) and 1% VRC (10mM)) just before the preparation of the slides. Cells in coverslips were then incubated overnight with the probe solution, at 37ºC in a FA/SSC humid chamber.

Next day, washes were carried out with FA/SSC solution (50% FA and 1% 20x SSC) three times for 7 min at 42ºC and then only with 2x SSC, three times for 5 min at 42ºC. After that, nuclei were stained with DAPI (Sigma-Aldrich) diluted 1:10000 in 2x SSC, for 5 min at RT, and mounted with mounting media (see above).

Z-stack images (30 slices at 0.4 µm) of each sample were acquired in a wide-field fluorescence microscope Zeiss Axio Observer (Carl Zeiss MicroImaging) equipped with an Axiocam 506 mono CCD camera using a 63x/1.4 oil objective (Plan-Apochromat) using the filter sets FS43HE (HE DsRed), FS38HE (HE GFP) and FS49 (DAPI). Digital images were analyzed with the FIJI platform. At least 400 cells per single experiment were counted to determine the number of cells with an *Xist* cloud signal, corresponding to an *Xist*-coated X-chromosome, or to determine the expression of the X-linked gene studied.

RNA-FISH combined with IF

To perform RNA-FISH combined with IF (IF/RNA-FISH), cells were previously cultured on gelatin-coated coverslips for at least 24 hours. After permeabilization washes, cells were blocked for at least 15 min at RT with Blocking solution (5% Gelatin in PBS). Then, they were incubated with the primary antibodies for H3K27me3 and H2AK1119ub marks diluted in Blocking Solution with Ribonuclease Inhibitor (11.5 µl/ml) (Euromedex) at the desired concentration (Annex 7) for 45 min at RT in a H_2O humid chamber (and protected from light for the rest of the protocol). Coverslips were putted back into the wells (with cells facing up) and washed three times with PBS for 5 min at RT with agitation. Then, they were incubated with the secondary antibody for 45 min at RT (Annex 7). After that, cells were washed three times with 1x PBS for 5 min at RT with agitation, post-fixed with 3% PFA in PBS for 10min at RT and rinsed three times in PBS and twice in 2x SSC. Excess of 2x SSC was removed and cells were hybridized with a *Xist* p510 probe labelled with Alexa green or red dUTPs (prepared and hybridized as mentioned in the RNA-FISH protocol above). After that, nuclei were stained with DAPI in 2x SSC for 5min at RT. The slides were prepared in a similar way as in RNA-FISH.

Statistical analysis

Statistical significance between the different groups was assessed by unpaired student *t*-test. Statistical significance was set at $P < 0.05$.

Results

Generation of novel *Xist*-inducible mutants by CRISPR/Cas9

To study the functional relevance of the different repeats within *Xist* gene, we took advantage of a well-described system whereby *Xist* is expressed from a TetOP in its endogenous position upon DOX addition in J1 male mESCs (Figure 6A). The main hallmarks of XCI are recapitulated in this system, including coating of the X chromosome, heterochromatin formation and transcriptional silencing of the whole chromosome (Wutz, Rasmussen, and Jaenisch 2002). Previously, this system was used in the lab to successfully delete several regulatory regions such as the B- and C-repeats (Bousard *et al.* 2019) and also the F-repeat, which had not yet been characterized (Raposo, unpublished results).

Building from the previous work in the lab, we decided to also delete the D- and E-repeats, which were the two other repeats with putative functional importance not yet manipulated in this system. CRISPR/Cas9 genome editing technique (Ran *et al.* 2013) was used to perform these deletions. This method uses a gRNA complementary to the target DNA sequence, coupled with a Cas9 endonuclease. After the gRNA-DNA interaction, the Cas9 induces a double strand break in the DNA. Then, the two broken ends are linked to each other and repaired by non-homologous end joining (NHEJ), in the absence of an additional homologous sequence, what can originate insertions and/or deletions (INDELs) (Figure 6B). To create new mutants within *Xist* exons 1 and 7 (*Xist* ΔD and ΔE) we designed two different gRNAs flanking each one of these regions. Two paired DNA oligonucleotides corresponding to each gRNA were cloned individually downstream of a U6 promoter in the PX459 plasmid containing the sequence encoding for Cas9 endonuclease (see Materials and Methods for details). Then, these pairs of plasmids were electroporated into *Xist*-TetOP XY mESCs and 96 individual clones were picked for further molecular analysis (see Materials and Methods for details).

Screening and molecular characterization of the new *Xist*-inducible mutants

Screening for the desired ΔD and ΔE deletions were performed by PCR in 96 clones per mutation with specific primers across the mutated region (Figure 7A; primers sequences in Annex 3). Clones that showed the presence of only one band with the expected mutated size (419 bp to *Xist* ΔD clones and 206 bp to *Xist* ΔE clones) marked with an asterisk in Figure 7A corresponded to positive clones, some of which were selected to pursue the analysis. The ones with a band corresponding to *Xist* FL (without the deletion) were immediately excluded (marked with a purple arrow in Figure 7A). We selected three positive clones for each deletion that were further grown from the P96-well into T25 flasks. After, we repeated the PCRs with specific primers (Annex 3) for the deletions, and mapped them by Sanger sequencing (Figure 7B; see Annex 8 for *Xist* mutants' DNA sequences). Additionally, we also rule out for the presence of the wild-type band, confirming the purity of the clones (data not shown). The screening was performed not only in the newly generated clones (*Xist* ΔD and ΔE), but also on the previously generated, but not yet characterized, ΔF clones (Raposo, unpublished results). The three chosen clones differ on the extent of deleted region, and consequently on the size of DNA gel band (*Xist* ΔF clones: F1 198 bp, F2 458 bp, F3 168 bp; *Xist* ΔD clones: D1 380 bp, D2 385 bp, D3 277 bp; *Xist* ΔE clones: E1 174 bp, E2 160 bp, E3 393 bp) (Figure 7B), but they all do not contain the regions we wanted to delete (namely F-, D- and E-repeats) (exception for the *Xist* ΔE clone E3, where a small part of the E-repeat was not deleted; see below). For each mutant type, generated clones have a variety of deleted sizes, due to the fact that upon Cas9 cleavage NHEJ repaired the DNA in an inconsistent way, a phenomenon that is frequently noticed upon CRISPR/Cas9 induced double strand breaks (Ran et al. 2013). This occurrence led to clones with differences on the number of nucleotides, comparing several clones with deletion of the same repeated region. In the case of *Xist* ΔF mutation, one of the generated clones (clone F2) has an inverted insertion of a portion of the complementary chain, that was reinserted

after Cas9 cleavage (Figure 7B). The chosen *Xist* ΔD clones are more similar in size, differing only in a few nucleotides. One of the chosen *Xist* ΔE clones has a bigger PCR size (meaning a smaller deletion), because the Cas9 double cut occurred 3' end further in the sequence (approx 200 bp). This could be explained by the fact that one of the gRNAs, mapped at the beginning of the highly repetitive sequence E-repeat, might have bound to a similar region in the sequence further downstream, leading to a smaller deletion and therefore only truncating part of the E-repeat (Figure 7B**;** Annex 8).

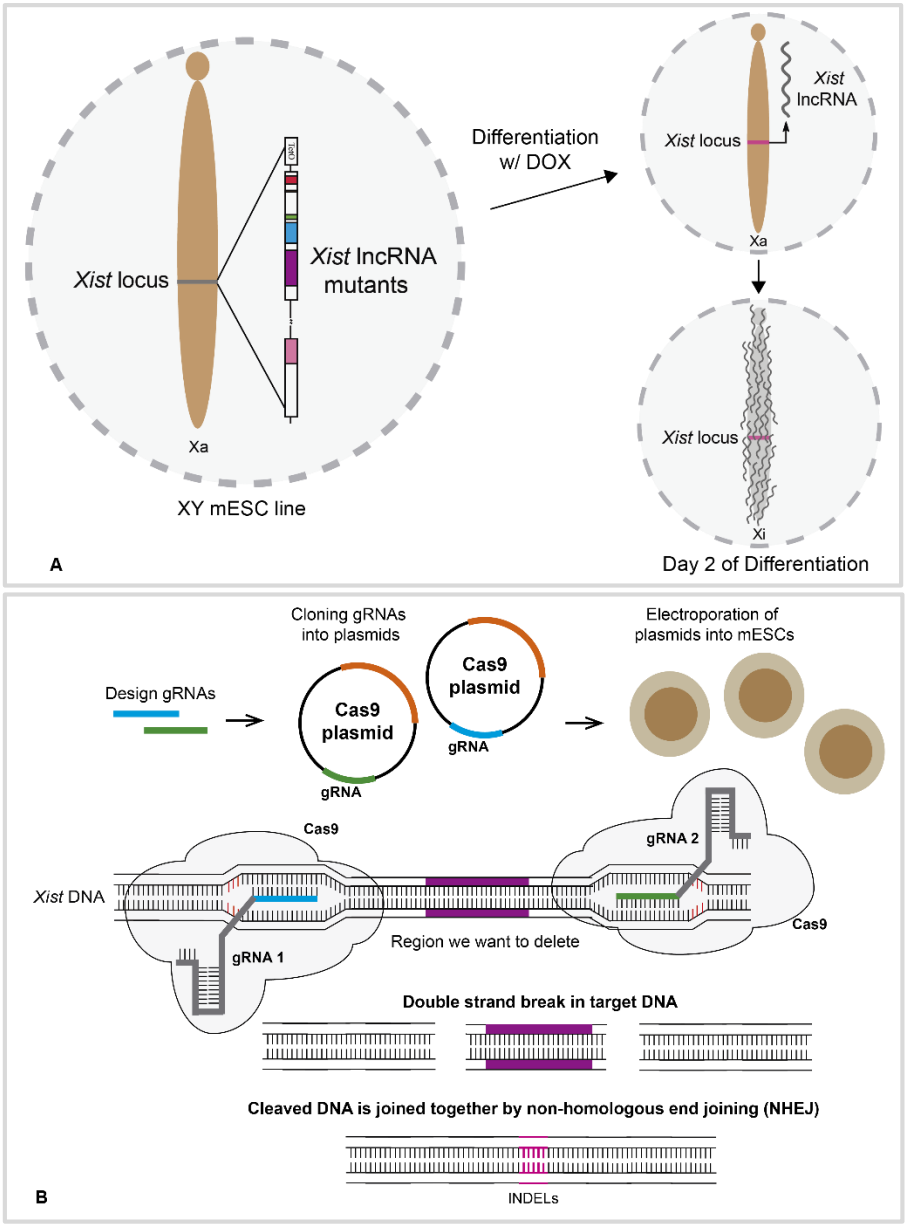


Figure 6. Schematic flow of the generation of novel *Xist*-TetOP mutants by CRISPR/Cas9. A – Representation of the *Xist*-TetOP J1 ESC system used to perform *Xist* deletions by CRISPR/Cas9. Upon DOX induction, *Xist* is expressed from its endogenous location to coat the X chromosome in *cis*, leading to silencing and heterochromatinization of the entire chromosome after 2 days under differentiation conditions; different *Xist* repeats are highlighted in color boxes; *Xist*, *X-inactive specific transcript*; DOX, doxycycline; Xa, active X chromosome; Xi, inactive X chromosome; TetOP, tetracycline-inducible promoter. B – CRISPR/Cas9 approach to generate novel *Xist* inducible mutants in mESCs. gRNAs (marked in blue and green) are designed to flank the region we want to delete and are inserted into plasmids also encoding for a Cas9 endonuclease (marked in orange); plasmids are then electroporated into mESCs; gRNAs bind to the specific region of DNA and Cas9 cuts the DNA in a double strand manner; in the absence of an additional sequence, the two ends bind to each other by NHEJ. Purple area corresponds to the region of *Xist* we wanted to delete; gRNA, guide RNA; mESCs, mouse embryonic stem cells; INDELs, insertions and deletions.

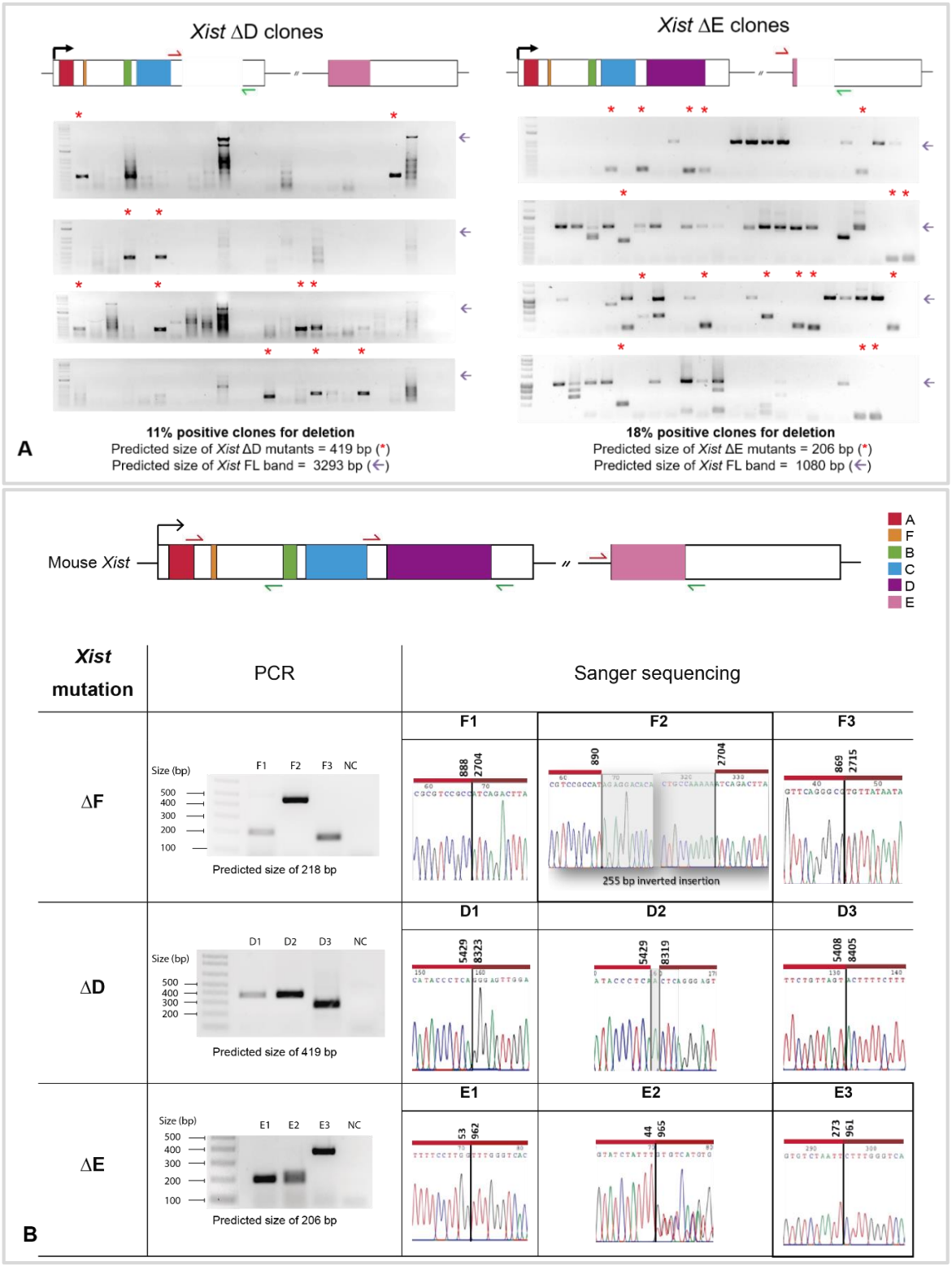
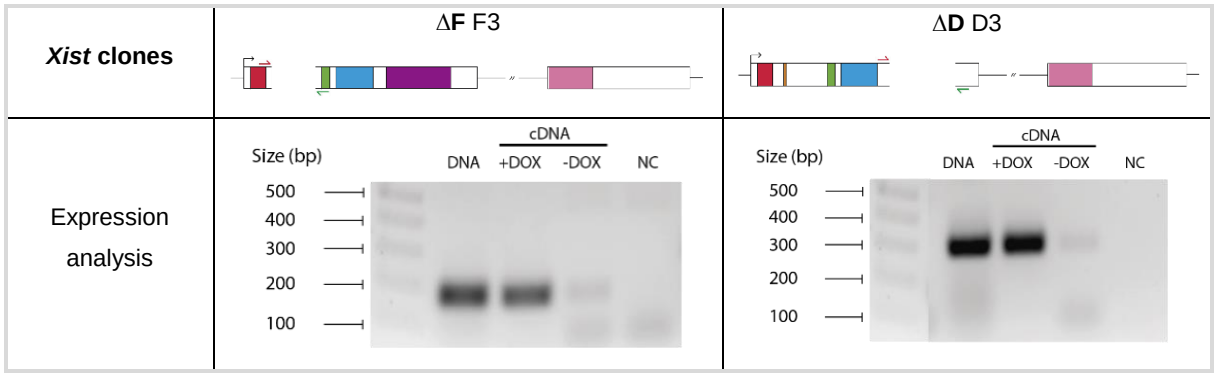
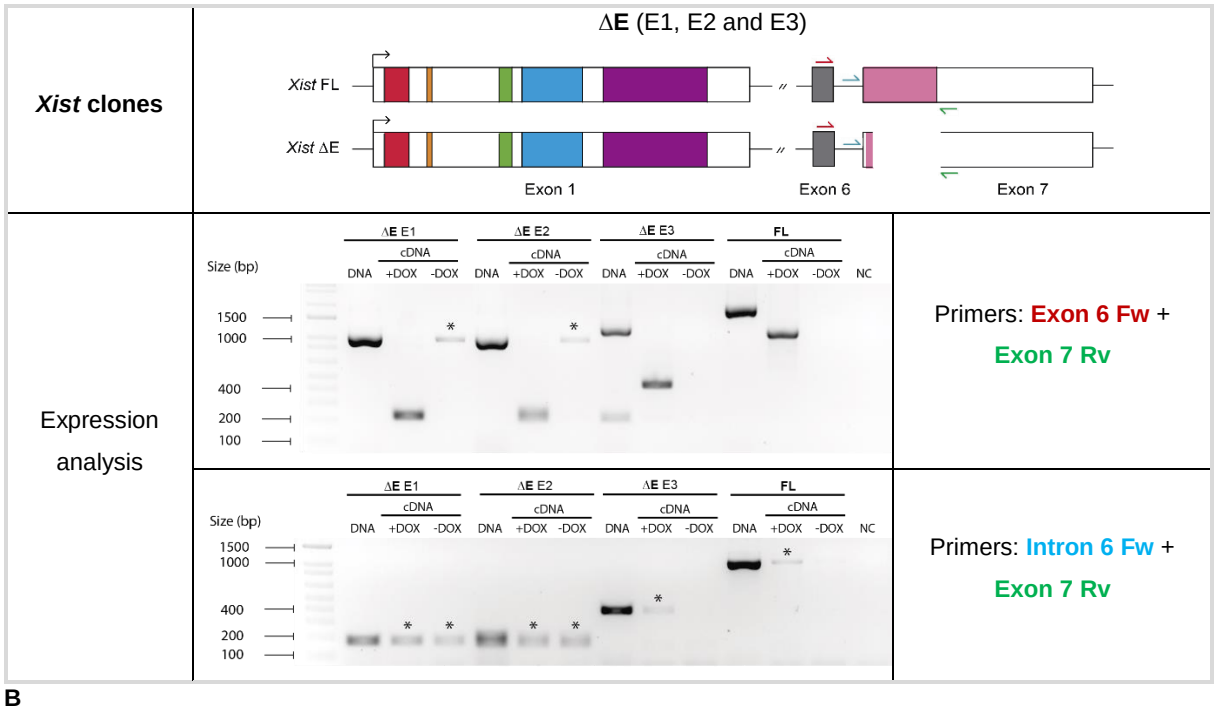


Figure 7. Screening and validation of new *Xist*-TetOP mutants. A – Screening of the newly generated *Xist* mutants (ΔD and ΔE) performed by PCR with specific primers (represented in *Xist* DNA structure above). From the 96 clones analyzed per deletion, 11% were positive for the ΔD mutants and 18% positive for the ΔE mutants (positive clones are represented with red asterisks). *Xist* FL bands are point out by purple arrows; FL, full-length. B – Characterization of *Xist* ΔF, ΔD and ΔE mutants by PCR and Sanger sequencing with specific primers (represented in *Xist* DNA structure above). Sequenced *Xist* mutants' DNA are represented for each clone (see Annex 8 for details). Red arrows represent Fw primers; green arrows represent Rv primers; different colors correspond to the several tandem-repeated regions; *Xist*, *X-inactive specific transcript*; bp, base-pairs; NC, negative control; shaded section on the chromatogram from the ΔF F2 mutant represents an insertion in the opposite direction of a portion of the deleted region with 255 bp; shaded section on the chromatogram from the ΔD D2 mutant represents an added nucleotide.

Characterization of the selected clones was also made at RNA level, through the analysis of mutants' cDNA. In the case of *Xist* ΔF and ΔD mutants, for the deletion PCR, we expected to obtain bands of the same size for DNA and cDNA upon DOX addition, because these repeats are in the middle of the long exon 1 (Figure 8A). This analysis was only performed for one clone of each *Xist* mutant (ΔF F3 and ΔD D3), that were the samples later selected for RT-qPCR analysis to determine expression levels of *Xist* and other X-linked genes (see below). The expected result was obtained for both clones, with the cDNA band size matching the one from DNA upon DOX addition (*Xist* ΔF F3: 168 bp; *Xist* ΔD D3: 277 bp). In non-inducible (noDOX) conditions, we only obtained weak bands for the cDNA, because in these conditions *Xist* RNA might be weakly expressed (Bousard et al. 2019). We cannot also completely rule out a mild DNA contamination, since the band size for DNA will not be distinguishable from cDNA.



A



B

Figure 8. Characterization of the novel *Xist*-TetOP mutants at RNA level. A – Expression analysis across deleted region in the novel *Xist* ΔF and ΔD mutants was made by PCR with specific primers (represented in *Xist* DNA structure for each mutant) for both DNA and cDNA samples. B – Expression analysis across deleted region in the novel *Xist* ΔE mutants was made by PCR with specific primers (represented in *Xist* DNA structure above) with mutants' DNA and cDNA samples. Red arrows indicate Fw primers in exons; blue arrows indicate Fw primers in introns; green arrows represent Rv primers; different colors correspond to the several tandem-repeated regions; asterisks represent samples with possible contamination with genomic DNA; cDNA, complementary DNA; FL, full-length; DOX, doxycycline; bp, base-pairs; NC, negative control.

For the *Xist* ΔE mutants, we performed the expression analysis with the three selected clones (E1, E2 and E3). In this case the situation was different, due to the fact that E-repeat localized right at the beginning of exon 7. To detect the RNA expression, we used a Fw primer on exon 6 and the Rv primer on exon 7 after the deleted region (Figure 8B). In this case, we would not expect DNA and RNA sizes to match, since at the RNA level, the intron will be skipped. We were also concerned with the fact that CRISPR/Cas9 double strand break at the beginning of exon 7 would interfere with splicing DNA elements between exon 6 and 7 and result in abnormal splicing pattern. For this reason, we also designed a primer within intron 6 and we used it with the same Rv primer after E-repeat deletion on exon 7. In this situation, we would expect to see only a band at DNA level, but not at the RNA level. Using this double PCR approach, we could confirm that splicing occurred properly. The exon 6-exon 7 pair of primers showed a difference in size which matched the sizes predicted for DNA and cDNA for each mutant (*Xist* ΔE E1: DNA 1008 bp, cDNA 227 bp; *Xist* ΔE E2: DNA 994 bp, cDNA 213 bp; *Xist* ΔE E3: DNA 1227 bp, cDNA 446 bp; *Xist* FL: DNA 1960 bp, cDNA 1133 bp). Concerning the intron 6-exon 7 pair of primers, the correct band was observed for the DNA samples (*Xist* ΔE E1: 174 bp; *Xist* ΔE E2: 160 bp; *Xist* ΔE E3: 393 bp; *Xist* FL: 1080 bp). Interestingly, a weaker band of the same size was seen in the majority of cDNA samples, in both DOX and noDOX conditions. This could be due to the presence of small levels of unspliced transcript. Alternatively, this could also be caused by residual DNA contamination in our cDNA samples. Since the band appears in both DOX (*Xist* is expressed) and noDOX (*Xist* is not or only weakly expressed) conditions, we favored this last hypothesis. This could also explain why a weak band is seen in noDOX samples (but not in DOX samples) with the same size of the DNA band when exon 6-exon 7 pair was used (Figure 8B). A negative control for the reverse transcriptase reaction (commonly known as –RT) should have been used to rule out DNA contamination in our cDNA samples. Overall, we concluded that correct pattern of splicing occurred and no sign of altered splicing was noted. In some instances, we could still observe the unspliced transcripts, but DNA contamination needs to be rule out.

Functional analysis of *Xist*-TetOP mutants

The newly generated *Xist*-TetOP mutants (ΔF, ΔD and ΔE) were tested for *Xist* capability to coat the X-chromosome, to induce chromosome-wide gene silencing and to form facultative heterochromatin. For that, the different *Xist*-TetOP mutant mESCs were differentiated in DOX or noDOX conditions for 2 days. Induction of differentiation was performed by withdrawal of the self-renewal factor LIF. The reason for that is the fact that we would like to mimic the physiological conditions in which female ESCs spontaneously express *Xist* and start XCI *in vitro*.

Xist coating and expression analysis

Xist expression and coating of the X chromosome were analyzed after 2 days of differentiation in DOX and noDOX conditions, by RT-qPCR and RNA-FISH (Figure 9A). First, RNA-FISH with Stellaris probes for *Xist* was performed, to understand whether the different mutants were able to coat the X chromosome. For this experiment, we decided to use a set of 48 Stellaris customized oligo-probes, that we designed on exon 7 in a region that is not deleted in any of our novel mutants (with the exception of 2 out of 48 probes that lay within ΔE region). This way, we could compare our different mutants' phenotype with *Xist* FL and appreciate their degree of *Xist* coating, without the influence of the probe specificity. To perform this analysis, we used one clone for ΔF (F3) and two clones for ΔD (D1 and D3) and ΔE (E1 and E3), and the *Xist* FL cell line (as a control). *Xist* Stellaris RNA-FISH showed that all *Xist* mutants studied were able to form a *Xist* domain upon DOX induction **(**Figure 9B**)**. Overall, *Xist* mutants seemed to have similar clouds, with the exception of *Xist* ΔF clouds that looked smaller (Figure 9B), but this was not quantified. For the ΔE deletions in particular, the *Xist* coating capability was not affected in neither ΔE clone with a partial deletion (E3) nor a full deletion (E1). Therefore, in this

inducible system, it seems that E-repeat does not influence Xi coating. Interestingly, the proportion of cells with a *Xist*-coated X chromosome varied somewhat between mutated clones (*Xist* FL: 44%; *Xist* ΔF F3: 19%; *Xist* ΔD D1: 44%; *Xist* ΔD D3: 49%; *Xist* ΔE E1: 32%; *Xist* ΔE E3: 27%) (Annex 9). The observed variability is often seen in this system, due to a fraction of cells that never upregulate *Xist* and that seems to differ from cell line to cell line, independently of the mutant type (Bousard et al. 2019).

After, the results obtained by RNA-FISH were related with the *Xist* expression levels measured by RT-qPCR. Using this technique, *Xist* expression in all the studied mutated clones (*Xist* ΔF F3, *Xist* ΔD D3 and *Xist* ΔE E3) was found to be significantly lower comparing with *Xist* FL clone (Figure 9B). Interestingly, in *Xist* FL, ΔF F3 and ΔE E3 cell lines, *Xist* expression levels mirror the percentage of cells that showed *Xist* coating, so, the difference in *Xist* overall levels of expression might only indicate differences in the proportion of cells that are able to induce *Xist*. Remarkably, *Xist* ΔD mutants showed the same proportion of cells with *Xist* domains when compared to *Xist* FL (*Xist* FL: 44%; *Xist* ΔD D1: 44%; *Xist* ΔD D3: 49%) by Stellaris RNA-FISH. Although, by RT-qPCR, they exhibited a tendency to have a reduced *Xist* expression level (Figure 9B). This result might suggest that *Xist* ΔD-expressing cells express less *Xist* than *Xist* FL-expressing cells. On the other hand, this result might also suggest a decreased RNA stability of *Xist* ΔD transcripts. *Xist* transcript stability (or half-life) should be explored in the future by the RNA quantification at different time-points, after RNA polymerase transcriptional inhibition with actinomycin D. In any case, it is important to stress that all these results are preliminary, because each experiment was just performed once.

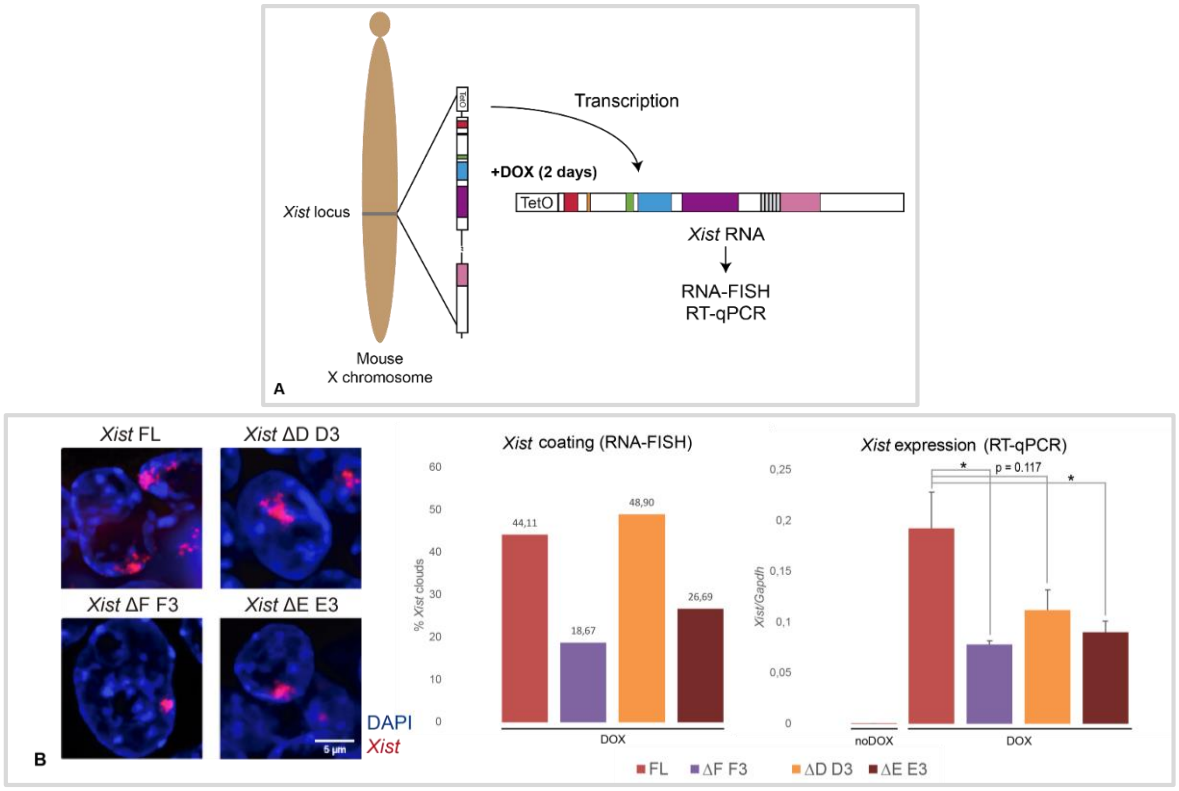


Figure 9. *Xist* coating and expression analysis of the novel *Xist* mutants by RNA-FISH and RT-qPCR. A – Model system with representation of the *Xist* gene *locus* on the mouse X chromosome (genomic coordinated (GRCm38/mm10): *Xist* chrX:103,460,366-103,483,254). *Xist*'s transcription is induced with the addition of DOX for 2 days in differentiating conditions. RNA-FISH and RT-qPCR are performed after that period**.** TetO, tetracycline-inducible promoter. B – Analysis of *Xist* coating and expression by *Stellaris* RNA-FISH and RT-qPCR. Representative RNA-FISH images for *Xist* (red) in *Xist*-TetOP lines at day 2 of differentiation in the presence of DOX; nuclei are stained in blue (DAPI); different colors represent different *Xist* cell lines; a minimum of 200 cells were counted per sample; *Xist*, *X-inactive specific transcript*; DOX, doxycycline; FL, full-length. Scale bar is 5 µm. Only P-values corresponding to significant differences comparing mutants to *Xist* FL DOX are indicated as * ($P < 0.05$), unpaired Student's *t*-test. Images of Stellaris RNA-FISH were deconvolved using the Huygens Remote Manager software (see Materials and Methods for details).

Transcriptional silencing of X-linked genes

The ability of the newly generated *Xist*-TetOP mutants to induce X-linked gene silencing was assessed for two X-linked genes (*Rnf12* and *Pgk1*) after 2 days of differentiation with or without DOX. *Rnf12* expression level was evaluated only by RT-qPCR, while *Pgk1* expression was measured through RT-qPCR complemented with RNA-FISH for nascent transcripts (Figure 10A). First, by RT-qPCR, the expression of *Rnf12* and *Pgk1* was analyzed in one clone of each *Xist* mutant type (*Xist* ΔF F3, *Xist* ΔD D3 and *Xist* ΔE E3), with *Xist* FL cell line as a control. Unfortunately, for the *Xist* ΔE mutants, we only used the clone with a partial deletion of the E-repeat, since we had technical problems and we did not get results for the other two clones with a complete ΔE deletion. Three biological replicates were used for each sample. This analysis did not show noticeable differences between the mutants and the *Xist* FL cells, with both *Pgk1* and *Rnf12* becoming silencing in all the *Xist* mutants at comparable levels to *Xist* FL. These data suggest that absence of these portions, in each one of these mutants, does not impact majorly the X-linked gene silencing. Interestingly, *Pgk1* expression was slightly reduced in *Xist* ΔF F3 clone, when compared to *Xist* FL (Figure 10B).

To obtain more detailed results for the potential enhancement in *Pgk1* silencing of the *Xist* ΔF F3 clone, we went to study the *Pgk1* silencing by an alternative method: nascent-transcript RNA-FISH. For this, we used two *Xist* ΔF clones (clone F2 and clone F3) and one *Xist* FL clone as a control. RNA-FISH analysis confirmed that *Pgk1* is silenced in both *Xist* ΔF mutants, arguing that loss of the F-repeat does not compromise X-linked gene silencing in an inducible context. Interestingly, the *Xist* ΔF F3 clone showed a slightly reduced *Pgk1* silencing, comparing with *Xist* ΔF F2 and *Xist* FL clones (Figure 10C). This result differed from the one obtained by RT-qPCR, which suggested a more efficient *Pgk1* silencing in *Xist* ΔF F3 clone. The discrepancies in the results obtained with alternative techniques might suggest that these differences are likely to be small. Moreover, it points up to the need to repeat these experiments to draw proper conclusions.

Facultative heterochromatin formation

Finally, we went to evaluate the chromatin-associated phenotypes (enrichment of histone marks) by IF/RNA-FISH. The Xi is characterized by the presence of several facultative heterochromatin marks, such as H3K27me3, H2AK119ub and H4K20me1. It is already known that H3K27me3 mark is recruited to the Xi through PRC2, while PRC1 is responsible for H2AK119ub deposition, and the B- and C-repeats are involved (Bousard et al. 2019). About the mechanisms behind the H4K20me1 enrichment not much is yet known and the involvement of the different repeats has not yet been investigated. Here, we evaluated the enrichment of these three histone modifications in the previously generated *Xist* ΔF mutated clones (F2 and F3) and we compared the results with the accumulation that occurs in *Xist* FL cell line. Both selected *Xist* ΔF clones showed accumulation of the three histone marks on the Xi. Interestingly, *Xist* ΔF F3 clone showed a slightly decreased enrichment, comparing with *Xist* FL clone (Figure 11). Additional experiments are needed to take any conclusions, because these are still preliminary results. The analysis of the *Xist* ΔD and ΔE mutants would also be interesting, in order to investigate the importance of those repeats in the formation of facultative heterochromatin.

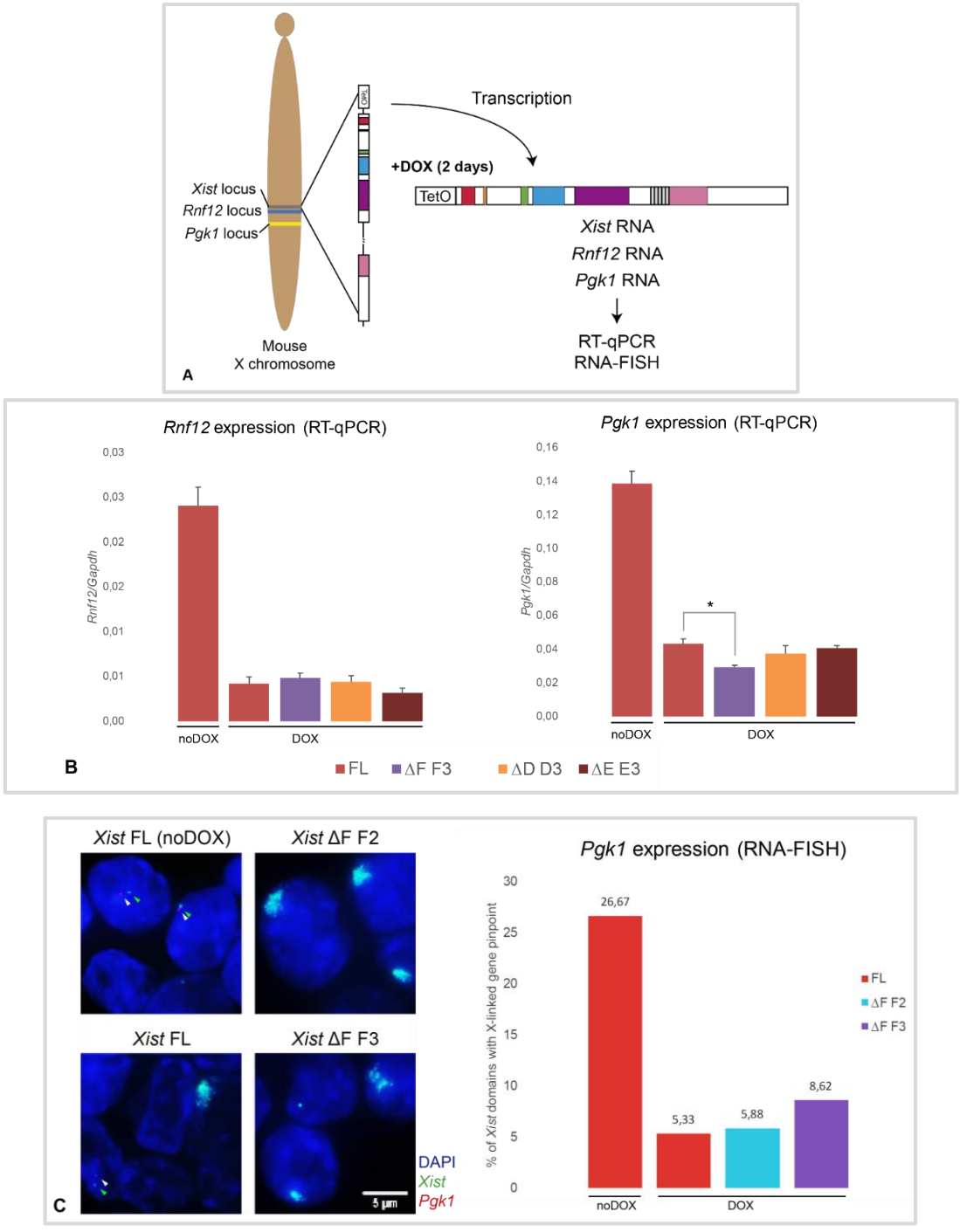


Figure 10. X-linked gene expression analysis of the novel *Xist* mutants by RT-qPCR and RNA-FISH. A – Model system with representation of the X-linked genes *loci* in the mouse X chromosome (genomic coordinated (GRCm38/mm10): *Xist* chrX:103,460,366-103,483,254; *Rlim* (*Rnf12*) chrX:103,957,163-103,981,284; *Pgk1* chrX:106,187,100-106,203,699). *Xist* FL and *Xist*-TetOP mutants were differentiated for two days in DOX and noDOX conditions and X-linked gene silencing was analyzed by RT-qPCR (for *Pgk1* and *Rlim/Rnf12*) and RNA-FISH for the *Pgk1* gene (only in *Xist* FL and *Xist* ΔF mutants). TetOP, tetracycline-inducible promoter. B – Analysis of X-linked gene expression (*Rnf12* and *Pgk1*) in *Xist*-TetOP mutants (ΔF, ΔD and ΔE) by RT-qPCR**.** C – Analysis of X-linked gene expression (*Pgk1*) in *Xist* ΔF mutants by nascent-transcript RNA-FISH. Representative RNA-FISH images for *Xist* (green) and nascent transcript of *Pgk1* (red); nascent transcript of *Tsix* pinpoint (green) is represented with green arrows (the p510 DNA probe used for *Xist* recognize also *Tsix* in noDOX conditions because both sense and antisense transcripts are detected); nascent transcript of *Pgk1* pinpoint is represented with white arrows; nuclei are stained in blue (DAPI). In the graph, numbers represent % of Xist-coated X chromosomes with active *Pgk1* gene (except for *Xist* FL noDOX, where numbers represent % of cells with *Pgk1* active gene); a minimum of 70 *Xist*-coated X chromosomes were counted per sample (615 cells were counted in *Xist* FL noDOX sample); different colors represent different *Xist* cell lines; DOX, doxycycline; FL, full-length. Scale bar is 5 µm. Only P-values corresponding to significant differences comparing mutants to *Xist* FL DOX are indicated as * ($P < 0.05$), unpaired Student's *t*-test.

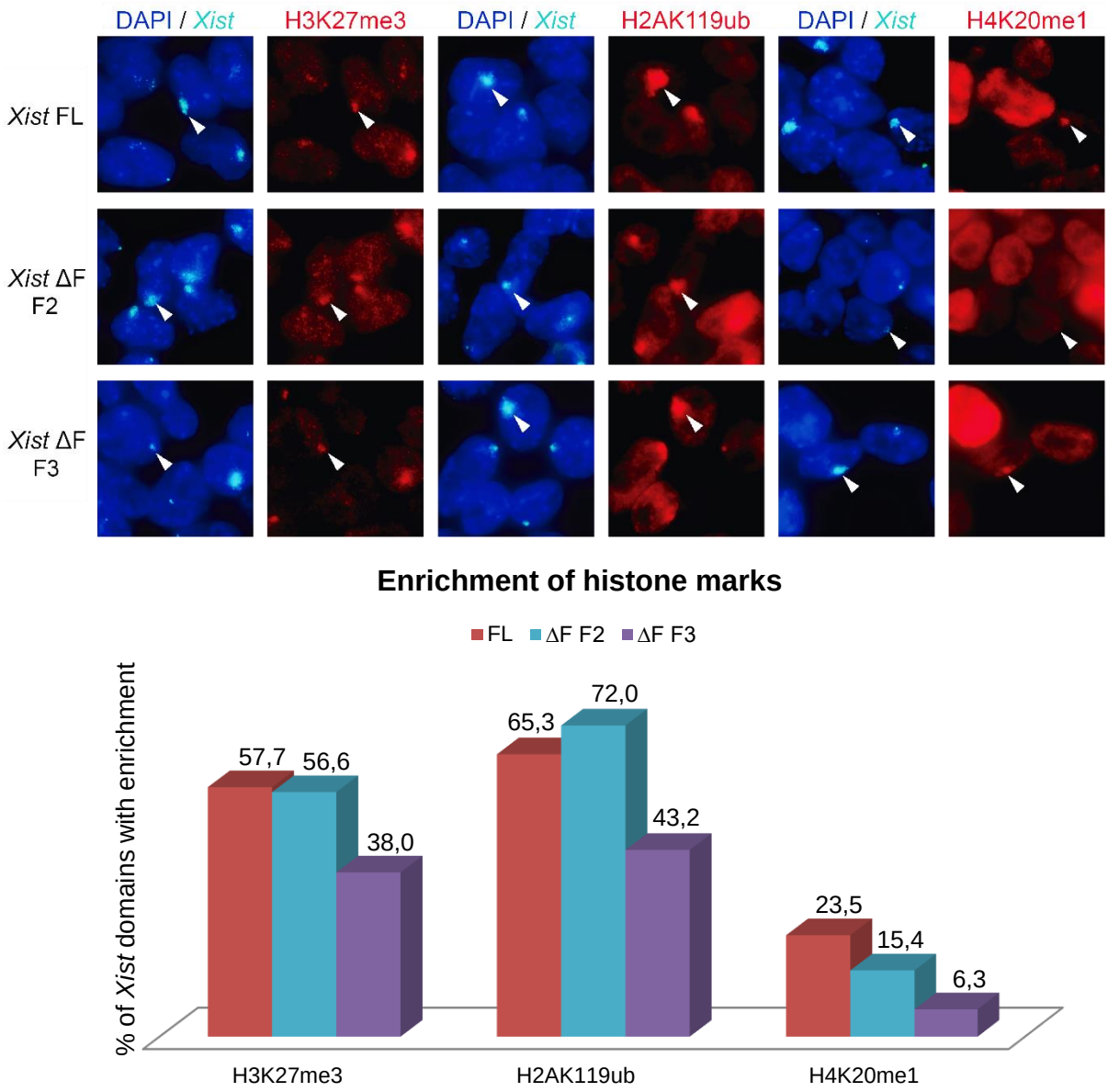


Figure 11. Characterization of the accumulation of histone marks on *Xist* ΔF mutants by IF/RNA-FISH. A – Representative images of combined IF for H3K27me3, H2AK119ub1 or H4K20me1 (red) with RNA-FISH for *Xist* (green) in *Xist*-TetOP ΔF lines (clones F2 and F3) and *Xist* FL at day 2 of differentiation in DOX conditions. *Xist*'s clouds and accumulation of histone marks are represented with white arrows; nuclei are stained in blue (DAPI). B – Graph represents the % of *Xist*-coated X chromosomes enriched for H3K27me3, H2AK119ub1 or H4K20me1 in the different clones of *Xist*-TetOP ΔF mutant and *Xist* FL; a minimum of 100 *Xist*-coated X chromosomes were counted per sample; different colors represent different *Xist* cell lines; FL, full-length.

Discussion

Brief summary

In this study, we report on the generation of novel *Xist*-inducible mutants for the F-, D- and E-repeats. Several clones of *Xist* ΔD and ΔE mutants were generated and isolated, while previously generated clones for *Xist* ΔF mutants were successfully characterized. The selected clones were validated and characterized at DNA and RNA levels and their deletion properly mapped. We also performed an exploratory functional characterization of the different *Xist*-inducible mutants, namely, on *Xist* coating and expression, on *Xist* capacity to induce X-linked transcriptional silencing and on the *Xist* ability to recruit the typical facultative heterochromatin marks of the Xi. No major defects in these parameters were gathered from our preliminary analysis in the three recently generated mutants, but more complete analysis should be performed in the future. The generation of the *Xist* ΔF, ΔD and ΔE mutants increases the current repertoire of mutants for the tandem repeated regions in this inducible system of *Xist* expression in J1 XY male mESCs (Wutz, Rasmussen, and Jaenisch 2002; Bousard et al. 2019) and are important research tools for further dissection of *Xist* function in the context of XCI.

Inducible system for *Xist* expression

To study the different *Xist* repeats' function in the XCI process, we used an inducible system where *Xist* expression is upregulated from its endogenous *locus* in male mESCs (Wutz, Rasmussen, and Jaenisch 2002). Female mESCs, that initially have two active X chromosomes, are an alternative model for studying XCI initiation and have been used for decades now (reviewed in Escamilla-Del-Arenal et al., 2011). However, they might not be ideal for performing *Xist* deletions with the aim of study the downstream functions of *Xist* lncRNA at the initiation stages of XCI. Indeed, when we perform mutations in one of the *Xist* alleles, there is a tendency for XCI skewing toward the wild-type allele (Hoki et al. 2009; Senner et al. 2011; Lv et al. 2016). As a consequence, the mutated *Xist* gene never gets expressed and we cannot study its respective phenotype. Alternatively, if we mutate both alleles, we might not achieve *Xist* expression from any of the two alleles. Indeed, within *Xist* gene there are several important genomic regions for correct activation of the endogenous *Xist* promoter such as the A-repeat region (Sakata et al. 2017) and the YY1 sites around the F-repeat (Makhlouf et al. 2014). To overcome these problems related to the endogenous activation of *Xist*, *Xist* inducible systems have long been implemented to study *Xist* functions (Wutz and Jaenisch 2000; Wutz, Rasmussen, and Jaenisch 2002).

A clear advantage of inducible systems is that *Xist* expression is independent of the intricate regulation of endogenous *Xist* expression, allowing the expression of *Xist* mutants without genomic regions necessary for normal *Xist* regulation. Furthermore, *Xist* induction results in a synchronized *Xist* expression which is very beneficial for experimental purposes. Different inducible systems have been used to express *Xist* either from its endogenous locus or from another autosomal or X chromosome locations (through the insertion of transgenes) (Wutz and Jaenisch 2000; Wutz, Rasmussen, and Jaenisch 2002; Engreitz et al. 2013; Loda et al. 2017; Pintacuda et al. 2017). Nevertheless, concerning these multiple available *Xist* inducible systems, the best one is the one where inducible expression occurs at the *Xist* endogenous *locus*, because of two important reasons: (1) *Xist* capability to silence genes is higher, due to position-effects that affect *Xist* expression/coating (Loda et al. 2017); (2) this system phenocopies better the natural environment from which *Xist* is normally expressed and spreads, which affects the timing of transcriptional silencing of each X-linked gene (Engreitz et al. 2013).

An alternative method to our system would be the use of a female mESC line carrying an inducible promotor in *Xist* endogenous *locus* in one of the X chromosomes. This inducible *locus* could

be used to induce the mutated *Xist*. However, this system was only developed recently and it was not available in our lab (Schulz et al. 2014; Nesterova et al. 2019). Besides the clear advantages, this system has also some drawbacks, such as the propensity of female cells to become XO in culture. Therefore, we decided to use the male cell line due to its availability in the lab and, also, due to the previous work performed with this system in house (Bousard et al. 2019). An additional reason for this choice was the fact that it is an easy system to work with, due to the presence of only one X chromosome per cell, allowing efficient genetic editing and molecular characterization of the deletions of the *Xist* locus.

Despite the clear advantages of inducible systems, they have also some drawbacks. One of the disadvantages of these systems is the super-physiological levels of *Xist*, that increase the number of *Xist* RNA molecules surrounding the X chromosome (Engreitz et al. 2013; Bousard et al. 2019) and could mask or attenuate some phenotypes. This has been controlled in this study by the use of the WT version (*Xist* FL) in the same system. Another potential problem of this system is the fact that it is developed in a male cell line. In this case, inactivation of the single X chromosome will eventually lead to cell lethality and prevent any meaningful analysis after day 5 of DOX induction. Therefore, this system does not allow the study of the maintenance stage of XCI. Regardless of *Xist* expression being inducible, another aspect of this system that we should be aware is the variability of cells capable of expressing *Xist* in the presence of DOX, in the different *Xist* mutants and even within clones from the same mutant type (Figure 9B; (Wutz, Rasmussen, and Jaenisch 2002; Bousard et al. 2019)). This occurs due to the fact that one fraction of cells, which varies between cell lines and even between experiments, never get to express *Xist*. Aware of this feature, we always complemented any population-averaged method used (e.g. RT-qPCR) with single-cell analyses (for example by RNA-FISH, alone or combined with IF) to compare and interpret our results.

Generation of *Xist* mutants by CRISPR/Cas9

To perform deletions on *Xist* endogenous locus, we decided to use the CRISPR/Cas9 genome editing technique. This method is suitable for this purpose because it is easily applied in this system and it allowed, previously, the successful generation of several *Xist*-TetOP mutants (e.g. *Xist* ΔB+C) (Bousard et al. 2019). Using this approach, we obtained a high number of clones with the desired deletion. In the case of *Xist* ΔD mutants, 11% out of the 96 isolated clones were positive for the deletion, while the number of *Xist* ΔE positive clones was even higher (18%), confirming the success of the chosen genome editing approach. Despite its advantages, this method has also some possible drawbacks, such as the off-targets effects that can occur genome-wide or the introduction of INDELs in the target region after the Cas9 cleavage (Ran et al. 2013). To deal with these possible problems, we selected three independent clones for each mutant type (including for the previously generated ΔF mutants). This way, any observed phenotypes in one specific clone per mutant type could be confirmed by the other two.

Xist coating and expression analysis

All of the generated *Xist* mutants (ΔF, ΔD and ΔE) were able to create *Xist* clouds in this system. Indeed, no major differences were seen when compared to *Xist* FL, maybe with the exception of ΔF clone. In a *Xist* ΔF clone, *Xist* domains around the Xi looked, on average, smaller. This feature needs to be properly quantified in the future through measurements of the area and total intensity of each *Xist* foci. Still, this observation would be consistent with previous studies that show a loss or significant weakening in *Xist* ΔF mutants' clouds in a DOX-inducible *Xist* transgene system (Jeon and Lee 2011) or in a system with deletions in two alleles of *Xist* endogenous *locus* (Colognori et al. 2019), both in female transformed tetraploid MEFs. Jeon and Lee also identified a region in the vicinity of the F-repeat, known as a binding site for the YY1, as a putative docking site for initiation of *Xist* spreading (Jeon and Lee 2011). Since our *Xist* ΔF mutant is able to coat to some extent, our results argue against the need of

this docking site for initiation of *Xist* spreading. Interestingly, Chen and colleagues also performed a deletion of the portion of *Xist* that binds to LBR, that encompasses the F-repeat, in a male *Xist*-TetOP mESC line. Interestingly, they observed a strong depletion of *Xist* RNA localization across regions of actively transcribed genes as mapped by RNA antisense purification (RAP), a method to map *Xist* RNA interactions with chromatin (Chen et al. 2016). This might be reminiscent of the putative decreased in *Xist* ΔF clouds seen by Stellaris RNA-FISH.

Xist E-repeat was previously reported to be important for *cis* localization, due to the *Xist* clouds' phenotypes observed in *Xist* ΔE mutants in female ESCs and MEFs. In these cell lines, *Xist* ΔE mutants showed *Xist* clouds with aberrant morphologies and widespread dispersal of the RNA throughout the nucleus (Yamada et al. 2015; Ridings-Figueroa et al. 2017; Sunwoo et al. 2017; Yue et al. 2017; Colognori et al. 2019). This observation was attributed to a failure on the recruitment of the nuclear matrix protein CIZ1 (Ridings-Figueroa et al. 2017; Sunwoo et al. 2017). However, in our system, no bigger differences were seen between *Xist* ΔE and *Xist* FL clouds in order to say that this repeat is involved in *Xist* localization. Although, this kind of inducible systems have higher transcript levels than normal cells and, therefore, the absence of the E-repeat might have been compensated by other *Xist* regions, which might also have a role in *Xist* coating.

Our *Xist* ΔD mutants, lacking the longest repeat of *Xist*, did not show significant differences in terms of *Xist* coating, comparing with *Xist* FL, having both an identical number of cells with *Xist* clouds. Although, *Xist* expression in *Xist* ΔD clone (D3) tended to be lower than in *Xist* FL clone. This is concordant with data from human cells (HEK and Hela) with a knock-out of this repeat, showing a significant decreased of *XIST* RNA expression (Lv *et al.* 2016) measured by RT-qPCR. This observation could mean that: (1) *Xist* ΔD transcripts are unstable and have a lower half-life; (2) that *Xist* FL and *Xist* ΔD have the same number of cells expressing *Xist*, but *Xist* ΔD mutants express less transcripts per cell. RNA polymerase II inhibition experiments (e.g. with actinomycin treatment) could help to distinguish between these two options.

Transcriptional silencing of X-linked genes

We also studied the initiation of the X-linked gene silencing by RT-qPCR and RNA-FISH. By RT-qPCR, we used one clone of each *Xist* mutant type (ΔF F3, ΔD D3 and ΔE E3) and we targeted two X-linked genes (*Rnf12* and *Pgk1*). We observed that all types of mutants were able to initiate gene silencing of these two genes. *Xist* ΔF F3 clone, specifically, was able to silence the *Pgk1* gene even better, comparing with *Xist* FL. However, using single-cell RNA-FISH analysis on two *Xist* ΔF clones (F2 and F3), we did not see this improvement on *Pgk1* silencing of ΔF F3 clone and there was actually a small relaxation of silencing comparing with *Xist* FL and with the other *Xist* ΔF clone. Interestingly, a failure in silencing was previously reported in a mutant for the LBS region (within the F-repeat region that we deleted) on an inducible system in male mESCs identical to ours (Chen *et al*. 2016). The differences observed between our results and Chen *et al*. results can be due to the used methods and also to the studied X-linked genes. Chen and colleagues analyzed the X-linked gene expression of 5 genes (*Atrx, Gpc4, Rbmx, Smc1a* and *Mecp2*) in male mESCs treated with DOX for 16 hours, using single-molecule RNA-FISH. In our case, we just analyzed the expression of two genes (*Rnf12* and *Pgk1*), but we treated the cells with DOX for 48 hours and we measured the expression by RT-qPCR complemented with RNA-FISH. Concluding, these results need further characterization to understand the role of *Xist* F-repeat in X-linked gene silencing.

The importance of D-repeat for the X-linked gene silencing was not yet stablished. Although, a study with human cells (HEK and Hela) showed that the absence of this long repetitive region of *XIST* leads to a relaxation of X-linked gene silencing. The authors attributed this to a reduce stability or

decreased expression of *XIST* ΔD RNA (Lv et al. 2016). In our case, our analysis by RT-qPCR shows that the *Xist* ΔD clone (D3) is able to silence the X-linked genes *Rnf12* and *Pgk1*. Nevertheless, our *Xist* ΔD expression analysis give us also indication for a possible transcripts' instability (see above).

Previous studies using female mESC lines showed that *Xist* ΔE mutants can induce X-linked gene silencing, but cannot sustain transcriptional repression on the Xi (Yamada et al. 2015) and have alterations in escapee genes' expression (Yue et al. 2017). This results disparity comparing to our results might be due to differences in the systems itself or because we did not specifically study these features, namely the maintenance of X-linked gene silencing and the escapee genes' expression.

Taking all into account, it is necessary to perform a high-throughput analysis in order to measure the expression of other genes, beyond *Rnf12* and *Pgk1*. These two X-linked genes are among the early inactivated genes (Patrat et al. 2009) and a more complete analysis, including later inactivated genes, would be more informative. In the future, could be interesting to do this kind of analysis in all the newly generated mutants to understand the specific contribution of each one of them for transcriptional silencing of the X chromosome.

Facultative heterochromatin formation

During the XCI process, the Xi becomes enriched in histone marks characteristic of the facultative heterochromatin, such as H3K27me3, H2AK119ub and H4K20me1. It was previously shown in this system that *Xist* B- and C-repeats recruit PRC1/PRC2 proteins and other proteins to stabilize the repressive state of at least some X-linked genes. Whether other repeats are also involved in this process is not known. Concerning the H4K20me1 enrichment of the Xi, no mechanisms have been proposed to its recruitment and we do not know which segments of *Xist* RNA could be involved. Here, we specifically evaluate whether *Xist* ΔF mutants, lacking the repeat immediately upstream of the B- and C-repeats, exhibited these typical heterochromatin marks (H3K27me3, H2AK119ub and H4K20me1) over the *Xist*-coated X chromosome. By IF/RNA-FISH, we observed an enrichment in all these marks in *Xist* ΔF-coated X chromosomes. However, *Xist* ΔF F3 clone tends to be less enriched in the three analyzed marks. This reduction fits with the decreased H2AK119ub and H3K27me3 accumulation on the already established Xi observed in a study using female transformed tetraploid MEFs (Colognori *et al.* 2019). However, reduction of Polycomb marks might be secondary to the reduced *Xist* levels observed in the absence of the F-repeat in that cell line. In the future, we intend to study the enrichment on these histone marks on the rest of the generated mutants (ΔD and ΔE) and also explore other putative histone modifications.

Conclusions and future perspectives

As a result of this work, a new series of inducible *Xist* mutants was generated, which will be a valuable set of experimental tools for further investigation in the field of XCI. Our initial characterization gave the first hints on the functional role of some unexplored conserved repeats of *Xist* during XCI. In the future, these *Xist*-inducible mutants can be used in high-throughput approaches to enquiry about gene silencing (RNA-seq), chromatin status (ChIP-seq) or to fish the protein interactors of the different repeats (ChIRP-MS), which will increase our knowledge on how this non-coding RNA molecule master-regulates XCI. Some of the *Xist*-protein interactors have already been mapped to specific repeats, such as MATR3 and PTBP1/2 to the E-repeat, and complementary experiments have already been planned to knock-down some of these factors. Until now, was performed a knock-down of the SAP18 protein, which seems to interact with B- and C-repeats. This analysis was recently published in a paper to which I also contributed (Bousard et al. 2019).

Nowadays, some discrepancies remain between the results obtained using different systems and cell lines. A way to control part of these disparities would be a proper evaluation and quantification of *Xist* features. In order to achieve that, an efficient system comprising the study of all the repeats would be required. The generation of our novel mutants in a well-established system in house, can give us new insights into the XCI process. With this work, I complete a full set of *Xist* mutants where all the single tandem repeated regions are deleted, building up from previous studies (Wutz, Rasmussen, and Jaenisch 2002; Bousard et al. 2019). This now allows us to easily compare the results and notice changed features between the mutants. We hope this will open new chapters to understand how this multi-tasking lncRNA, even without being translated, exerts its chromosome-wide functions.

References

Allshire, Robin C, and Hiten D Madhani. 2018. “Ten Principles of Heterochromatin Formation and Function.” *Nature Reviews Molecular Cell Biology* 19: 229–244. https://doi.org/10.1038/nrm.2017.119.

Almeida, Mafalda, Greta Pintacuda, Osamu Masui, Yoko Koseki, Michal Gdula, Andrea Cerase, David Brown, et al. 2017. “PCGF3/5 – PRC1 Initiates Polycomb Recruitment in X Chromosome Inactivation.” *Science* 356: 1081–84. https://doi.org/10.1126/science.aal2512.

Barr, Murray L., and Ewart G. Bertram. 1949. “A Morphological Distinction between Neurones of the Male and Female, and the Behaviour of the Nucleolar Satellite during Accelerated Nucleoprotein Synthesis.” *Nature* 163 (4148): 676–77. https://doi.org/10.1038/163676a0.

Belote, John M., and John C. Lucchesi. 1980. “Control of X Chromosome Transcription by the Maleless Gene in Drosophila.” *Nature* 285: 573–75. https://doi.org/10.1038/285573a0.

Bemmel, Joke G. van, Rafael Galupa, Chris Gard, Nicolas Servant, Christel Picard, James Davies, Anthony James Szempruch, et al. 2019. “The Bipartite TAD Organization of the X-Inactivation Center Ensures Opposing Developmental Regulation of Tsix and Xist.” *Nature Genetics* 51 (6): 1024–34. https://doi.org/10.1038/s41588-019-0412-0.

Berletch, Joel B., Wenxiu Ma, Fan Yang, Jay Shendure, William S. Noble, Christine M. Disteche, and Xinxian Deng. 2015. “Escape from X Inactivation Varies in Mouse Tissues.” *PLoS Genetics* 11 (3): 1–26. https://doi.org/10.1371/journal.pgen.1005079.

Bernstein, Emily, and C. David Allis. 2005. “RNA Meets Chromatin.” *Genes & Development* 19: 1635–55. https://doi.org/10.1101/gad.1324305.GENES.

Blewitt, Marnie E., Anne Valerie Gendrel, Zhenyi Pang, Duncan B. Sparrow, Nadia Whitelaw, Jeffrey M. Craig, Anwyn Apedaile, et al. 2008. “SmcHD1, Containing a Structural-Maintenance-of-Chromosomes Hinge Domain, Has a Critical Role in X Inactivation.” *Nature Genetics* 40 (5): 663–69. https://doi.org/10.1038/ng.142.

Bousard, Aurélie, Ana Cláudia Raposo, Jan Jakub Żylicz, Christel Picard, Vanessa Borges Pires, Yanyan Qi, Cláudia Gil, et al. 2019. “The Role of Xist-mediated Polycomb Recruitment in the Initiation of X-chromosome Inactivation.” *EMBO Reports*, 1–18. https://doi.org/10.15252/embr.201948019.

Brockdorff, Neil, Alan Ashworth, Graham F. Kay, Veronica M. McCabe, Dominic P. Norris, Penny J. Cooper, Sally Swift, and Sohaila Rastan. 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: 515–26. https://doi.org/10.1016/0092-8674(92)90519-i.

Brockdorff, Neil, Alan Ashworth, Graham F Kay, Penny Cooper, Sandy Smith, Veronica M. McCabe, Dominic P. Norris, Graeme D Penny, Dipika Patel, and Sohaila Rastan. 1991. “Conservation of Position and Exclusive Expression of Mouse Xist from the Inactive X Chromosome.” *Nature* 351: 329–31. https://doi.org/10.1038/351329a0.

Brown, Carolyn J., Andrea Ballabio, James L. Rupert, Ronald G. Lafreniere, Markus Grompe, Rossana Tonlorenzi, and Huntington F. Willard. 1991. “A Gene from the Region of the Human X Inactivation Centre Is Expressed Exclusively from the Inactive X Chromosome.” *Nature* 349: 38–44. https://doi.org/10.1038/349038a0.

Brown, Carolyn J., Ronald G. Lafreniere, Vicki E. Powers, Gianfranco Sebastio, Andrea Ballabio, Anjana L. Pettigrew, David H. Ledbetter, and Huntington F. Willard. 1991. “Localization of the X Inactivation Centre on the Human X Chromosome in Xq13.” *Nature* 349: 82–84. https://doi.org/10.1038/349082a0.

Byrd, Kristin Nastase, and Allen Shearn. 2003. “ASH1, a Drosophila Trithorax Group Protein , Is Required for Methylation of Lysine 4 Residues on Histone H3.” *PNAS* 100 (20): 11535–40. https://doi.org/10.1073/pnas.1933593100.

Calabrese, J. Mauro, Wei Sun, Lingyun Song, Joshua W. Mugford, Lucy Williams, Della Yee, Joshua Starmer, Piotr Mieczkowski, Gregory E. Crawford, and Terry Magnuson. 2012. “Site-Specific Silencing of Regulatory Elements as a Mechanism of X Inactivation.” *Cell* 151 (5): 951–63. https://doi.org/10.1016/j.cell.2012.10.037.

Cao, Ru, Liangjun Wang, Hengbin Wang, Li Xia, Hediye Erdjument-Bromage, Paul Tempst, Richard S. Jones, and Yi Zhang. 2002. “Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing.” *Science* 298 (November): 1039–44. https://doi.org/10.1126/science.1076997.

Carmona, Sarah, Benjamin Lin, Tristan Chou, Katti Arroyo, and Sha Sun. 2018. “LncRNA Jpx Induces Xist Expression in Mice Using Both Trans and Cis Mechanisms.” *PLoS Genetics* 14 (5): 1–21. https://doi.org/10.1371/journal.pgen.1007378.

Chapman, Andrew G., Allison M. Cotton, Angela D. Kelsey, and Carolyn J. Brown. 2014. “Differentially Methylated CpG Island within Human XIST Mediates Alternative P2 Transcription and YY1 Binding.” *BMC Genetics* 15 (89): 1–11. https://doi.org/10.1186/s12863-014-0089-4.

Chaumeil, J, I Okamoto, M Guggiari, and E Heard. 2002. “Integrated Kinetics of X Chromosome Inactivation in Differentiating Embryonic Stem Cells.” *Cytogenet Genome Res* 99: 75–84. https://doi.org/10.1159/000071577.

Chaumeil, Julie, Sandrine Augui, Jennifer C. Chow, and Edith Heard. 2008. “Combined Immunofluorescence, RNA Fluorescent in Situ Hybridization, and DNA Fluorescent in Situ Hybridization to Study Chromatin Changes, Transcriptional Activity, Nuclear Organization, and X-Chromosome Inactivation.” *Methods in Molecular Biology* 1 (18): 297–308. https://doi.org/10.1007/978-1-59745-406-3-18.

Chen, Chun-Kan, Mario Blanco, Constanza Jackson, Erik Aznauryan, Noah Ollikainen, and Mitchell Guttman. 2016. “Xist Recruits the X Chromosome to the Nuclear Lamina to Enable Chromosome-Wide Silencing.” *Science* 354 (6311): 468–73. https://doi.org/10.1126/science.aae0047.

Chow, Jennifer C, Lisa L Hall, Sarah E L Baldry, Nancy P Thorogood, Jeanne B Lawrence, and Carolyn J Brown. 2007. “Inducible XIST-Dependent X-Chromosome Inactivation in Human Somatic Cells Is Reversible.” *PNAS* 104 (24): 10104–9. https://doi.org/10.1073/pnas.0610946104.

Chu, Ci, Qiangfeng Cliff Zhang, Simão Teixeira da Rocha, Ryan A. Flynn, Maheetha Bharadwaj, J. Mauro Calabrese, Terry Magnuson, Edith Heard, and Howard Y. Chang. 2015. “Systematic Discovery of Xist RNA Binding Proteins.” *Cell* 161 (2): 404–16. https://doi.org/10.1002/cncr.31084.Talking.

Cirillo, Davide, Mario Blanco, Alexandros Armaos, Andreas Buness, Philip Avner, Mitchell Guttman, Andrea Cerase, and Gian Gaetano Tartaglia. 2016. “Quantitative Predictions of Protein Interactions with Long Noncoding RNAs.” *Nature Methods* 14 (1): 5–6. https://doi.org/10.1038/nmeth.4100.

Colognori, David, Hongjae Sunwoo, Andrea J. Kriz, Chen Yu Wang, and Jeannie T. Lee. 2019. “Xist Deletional Analysis Reveals an Interdependency between Xist RNA and Polycomb Complexes for Spreading along the Inactive X.” *Molecular Cell* 74 (1): 1–17. https://doi.org/10.1016/j.molcel.2019.01.015.

Cooper, Sarah, Anne Grijzenhout, Elizabeth Underwood, Katia Ancelin, Tianyi Zhang, Tatyana B Nesterova, Burcu Anil-kirmizitas, et al. 2016. “Jarid2 Binds Mono-Ubiquitylated H2A Lysine 119 to Mediate Crosstalk between Polycomb Complexes PRC1 and PRC2.” *Nature Communications* 7 (13661): 1–8. https://doi.org/10.1038/ncomms13661.

Costanzi, Carl, and John R Pehrson. 1998. “Histone MacroH2A1 Is Concentrated in the Inactive X Chromosome of Female Mammals.” *Nature* 393: 599–601. https://doi.org/10.1038/31275.

Deng, Xinxian, Wenxiu Ma, Vijay Ramani, Andrew Hill, Fan Yang, Ferhat Ay, Joel B. Berletch, et al. 2015. “Bipartite Structure of the Inactive Mouse X Chromosome.” *Genome Biology* 16 (1): 1–21. https://doi.org/10.1186/s13059-015-0728-8.

Engreitz, Jesse M., Amy Pandya-Jones, Patrick McDonel, Alexander Shishkin, Klara Sirokman, Christine Surka, Sabah Kadri, et al. 2013. “The Xist LncRNA Exploits Three-Dimensional Genome Architecture to Spread across the X Chromosome.” *Science* 341 (6147). https://doi.org/10.1126/science.1237973.

Escamilla-Del-Arenal, M., Simão Teixeira da Rocha, C. G. Spruijt, O. Masui, O. Renaud, Arne H. Smits, R. Margueron, M. Vermeulen, and Edith Heard. 2013. “Cdyl, a New Partner of the Inactive X Chromosome and Potential Reader of H3K27me3 and H3K9me2.” *Molecular and Cellular Biology* 33 (24): 5005–20. https://doi.org/10.1128/MCB.00866-13.

Escamilla-Del-Arenal, Martin, Simão Teixeira da Rocha, and Edith Heard. 2011. “Evolutionary Diversity and Developmental Regulation of X-Chromosome Inactivation.” *Human Genetics* 130: 307–27. https://doi.org/10.1007/s00439-011-1029-2.

Furlan, Giulia, Nancy Gutierrez Hernandez, Christophe Huret, Edith Heard, Rafael Galupa, Joke Gerarda Van Bemmel, and Claire Rougeulle. 2018. “The Ftx Noncoding Locus Controls X Chromosome Inactivation Independently of Its RNA Products Article The Ftx Noncoding Locus Controls X Chromosome Inactivation Independently of Its RNA Products.” *Molecular Cell* 70: 462–72. https://doi.org/10.1016/j.molcel.2018.03.024.

Fyodorov, Dmitry V., Bing Rui Zhou, Arthur I. Skoultchi, and Yawen Bai. 2018. “Emerging Roles of Linker Histones in Regulating Chromatin Structure and Function.” *Nature Reviews Molecular Cell Biology* 19 (3): 192–206. https://doi.org/10.1038/nrm.2017.94.

Galupa, Rafael, and Edith Heard. 2015. “X-Chromosome Inactivation: New Insights into Cis and Trans Regulation.” *Current Opinion in Genetics and Development* 31: 57–66. https://doi.org/10.1016/j.gde.2015.04.002.

Gayen, Srimonta, Emily Maclary, Michael Hinten, Sundeep Kalantry, Srimonta Gayen, Emily Maclary, Emily Buttigieg, Michael Hinten, and Sundeep Kalantry. 2015. “A Primary Role for the Tsix LncRNA in Maintaining Random X-Chromosome Inactivation.” *Cell Reports* 11 (8): 1251–65. https://doi.org/10.1016/j.celrep.2015.04.039.

Gendrel, Anne Valerie, Anwyn Apedaile, Heather Coker, Ausma Termanis, Ilona Zvetkova, Jonathan Godwin, Y. Amy Tang, et al. 2012. “Smchd1-Dependent and -Independent Pathways Determine Developmental Dynamics of CpG Island Methylation on the Inactive X Chromosome.” *Developmental Cell* 23 (2): 265–79. https://doi.org/10.1016/j.devcel.2012.06.011.

Gilfillan, Gregor D., Ina K. Dahlsveen, and Peter B. Becker. 2004. “Lifting a Chromosome: Dosage Compensation in Drosophila Melanogaster.” *FEBS Letters* 567 (1): 8–14. https://doi.org/10.1016/j.febslet.2004.03.110.

Giorgetti, Luca, Bryan R. Lajoie, Ava C. Carter, Mikael Attia, Ye Zhan, Jin Xu, Chong Jian Chen, et al. 2016. “Structural Organization of the Inactive X Chromosome in the Mouse.” *Nature* 535 (7613): 575–79. https://doi.org/10.1038/nature18589.

Godini, Rasoul, Haider Yabr Lafta, and Hossein Fallahi. 2018. “Epigenetic Modifications in the Embryonic and Induced Pluripotent Stem Cells.” *Gene Expression Patterns* 29: 1–9. https://doi.org/10.1016/j.gep.2018.04.001.

Goll, Mary Grace, and Timothy H. Bestor. 2005. “Eukaryotic Cytosine Methyltransferases.” *Annual Review of Biochemistry* 74 (1): 481–514. https://doi.org/10.1146/annurev.biochem.74.010904.153721.

Gontan, Cristina, Eskeatnaf Mulugeta Achame, Jeroen Demmers, Tahsin Stefan Barakat, Eveline Rentmeester, Wilfred Van Ijcken, J. Anton Grootegoed, and Joost Gribnau. 2012. “RNF12 Initiates X-Chromosome Inactivation by Targeting REX1 for Degradation.” *Nature* 485 (7398): 386–90. https://doi.org/10.1038/nature11070.

Häfner, Sophia J., and Anders H. Lund. 2016. “Great Expectations – Epigenetics and the Meandering Path from Bench to Bedside.” *Biomedical Journal* 39 (3): 166–76. https://doi.org/10.1016/j.bj.2016.01.008.

Hernández-Muñoz, Inmaculada, Anders H. Lund, Petra Van Der Stoop, Erwin Boutsma, Inhua Muijrers, Els Verhoeven, Dmitri A. Nusinow, Barbara Panning, York Marahrens, and Maarten Van Lohuizen. 2005. “Stable X Chromosome Inactivation Involves the PRC1 Polycomb Complex and Requires Histone MACROH2A1 and the CULLIN3/SPOP Ubiquitin E3 Ligase.” *Proceedings of the National Academy of Sciences of the United States of America* 102 (21): 7635–40. https://doi.org/10.1073/pnas.0408918102.

Hoki, Yuko, Naomi Kimura, Minako Kanbayashi, Yuko Amakawa, Tatsuya Ohhata, Hiroyuki Sasaki, and Takashi Sado. 2009. “A Proximal Conserved Repeat in the Xist Gene Is Essential as a Genomic Element for X-Inactivation in Mouse.” *Development* 136 (1): 139–46. https://doi.org/10.1242/dev.026427.

Inoue, Azusa, Lan Jiang, Falong Lu, and Yi Zhang. 2017. “Genomic Imprinting of Xist by Maternal H3K27me3.” *Genes and Development* 31 (19): 1927–32. https://doi.org/10.1101/gad.304113.117.

Jeon, Yesu, and Jeannie T. Lee. 2011. “YY1 Tethers Xist RNA to the Inactive X Nucleation Center.” *Cell* 146 (1): 119–33. https://doi.org/10.1016/j.cell.2011.06.026.

Jeppesen, Peter, and Bryan M Turner. 1993. “The Inactive X Chromosome in Female Mammais Is Distinguished by a Lack of Histone H4 Acetylation , a Cytogenetic Marker for Gene Expression.” *Cell* 74 (2): 281–89. https://doi.org/10.1016/0092-8674(93)90419-Q.

Kane, Alice E., and David A. Sinclair. 2019. “Epigenetic Changes during Aging and Their Reprogramming Potential.” *Critical Reviews in Biochemistry and Molecular Biology* 54 (1): 1–23. https://doi.org/10.1080/10409238.2019.1570075.

Kay, Graham F., Graeme D. Penny, Dipika Patel, Alan Ashworth, Neil Brockdorff, and Sohaila Rastan. 1993. “Expression of Xist during Mouse Development Suggests a Role in the Initiation of X Chromosome Inactivation.” *Cell* 72: 171–82. https://doi.org/10.1016/0092-8674(93)90658-d.

Klinger, Harold P., and Hans G. Schwarzacher. 1960. “The Sex Chromatin and Heterochromatic Bodies in Human Diploid and Polyploid Nuclei.” *The Journal of Biophysical and Biochemical Cytology* 8: 345–64. https://doi.org/10.1083/jcb.8.2.345.

Kohlmaier, Alexander, Fabio Savarese, Monika Lachner, Joost Martens, Thomas Jenuwein, and Anton Wutz. 2004. “A Chromosomal Memory Triggered by Xist Regulates Histone Methylation in X Inactivation.” *PLoS Biology* 2 (7): 0991–1003. https://doi.org/10.1371/journal.pbio.0020171.

Lee, Hyeon J, Ramu Gopalappa, Hongjae Sunwoo, Seo-Won Choi, Suresh Ramakrishna, Jeannie T Lee, Hyongbum H Kim, and Jin-Wu Nam. 2019. “En Bloc and Segmental Deletions of Human XIST Reveal X Chromosome Inactivation-Involving RNA Elements.” *Nucleic Acids Research*, 1–13. https://doi.org/10.1093/nar/gkz109.

Lee, Jeannie T. 2011. “Gracefully Ageing at 50, X-Chromosome Inactivation Becomes a Paradigm for RNA and Chromatin Control.” *Nature Reviews Molecular Cell Biology* 12 (12): 815–26. https://doi.org/10.1038/nrm3231.

Lee, Jeannie T, Lance S Davidow, and David Warshawsky. 1999. “Tsix, a Gene Antisense to Xist at the X-Inactivation Centre.” *Nature Genetics* 21 (april): 400–404. https://doi.org/10.1038/7734.

Lin, Fangqin, Ke Xing, Jianzhi Zhang, and Xionglei He. 2012. “Expression Reduction in Mammalian X Chromosome Evolution Refutes Ohno’s Hypothesis of Dosage Compensation.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (29): 11752–57. https://doi.org/10.1073/pnas.1201816109.

Loda, Agnese, Johannes H. Brandsma, Ivaylo Vassilev, Nicolas Servant, Friedemann Loos, Azadeh Amirnasr, Erik Splinter, et al. 2017. “Genetic and Epigenetic Features Direct Differential Efficiency of Xist-Mediated Silencing at X-Chromosomal and Autosomal Locations.” *Nature Communications* 8 (690): 1–16. https://doi.org/10.1038/s41467-017-00528-1.

Lv, Qingyan, Lin Yuan, Yuning Song, Tingting Sui, Zhanjun Li, and Liangxue Lai. 2016. “D-Repeat in the XIST Gene Is Required for X Chromosome Inactivation.” *RNA Biology* 13 (2): 172–76. https://doi.org/10.1080/15476286.2015.1137420.

Lyon, Mary F. 1961. “Gene Action in the X-Chromosome of the Mouse (Mus Musculus L.).” *Nature* 190: 372–73. https://doi.org/10.1038/190372a0.

Mak, Winifred, Tatyana B Nesterova, Mariana De Napoles, and Neil Brockdorff. 2004. “Reactivation of the Paternal X Chromosome in Early Mouse Embryos.” *Science* 303: 666–69. https://doi.org/10.1126/science.1092674.

Makhlouf, Mélanie, Jean François Ouimette, Andrew Oldfield, Pablo Navarro, Damien Neuillet, and Claire Rougeulle. 2014. “A Prominent and Conserved Role for YY1 in Xist Transcriptional Activation.” *Nature Communications* 5 (4878): 1–12. https://doi.org/10.1038/ncomms5878.

McHugh, Colleen A., Chun-Kan Chen, Amy Chow, Christine F. Surka, Christina Tran, Patrick McDonel, Amy Pandya-Jones, et al. 2015. “The Xist LncRNA Directly Interacts with SHARP to Silence Transcription through HDAC3.” *Nature* 521 (7551): 232–36. https://doi.org/10.1038/nature14443.The.

Meyer, Barbara J., and Lawrence P. Casson. 1986. “Caenorhabditis Elegans Compensates for the Difference in X Chromosome Dosage between the Sexes by Regulating Transcript Levels.” *Cell* 47 (6): 871–81. https://doi.org/10.1016/0092-8674(86)90802-0.

Minajigi, Anand, John Froberg, Chunyao Wei, Hongjae Sunwoo, Barry Kesner, David Colognori, Derek Lessing, et al. 2015. “A Comprehensive Xist Interactome Reveals Cohesin Repulsion and an RNA-Directed Chromosome Conformation.” *Science* 349 (6245): 289–313. https://doi.org/110.1016/j.bbi.2017.04.008.

Mira-Bontenbal, Hegias, and Joost Gribnau. 2016. “New Xist-Interacting Proteins in X-Chromosome Inactivation.” *Current Biology* 26 (8): R338–42. https://doi.org/10.1016/j.cub.2016.03.022.

Moindrot, Benoit, Andrea Cerase, Heather Coker, Osamu Masui, Anne Grijzenhout, Greta Pintacuda, Lothar Schermelleh, Tatyana B. Nesterova, and Neil Brockdorff. 2015. “A Pooled ShRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing.” *Cell Reports* 12 (4): 562–72. https://doi.org/10.1016/j.celrep.2015.06.053.

Monfort, Asun, Giulio Di Minin, Andreas Postlmayr, Remo Freimann, Fabiana Arieti, Stéphane Thore, and Anton Wutz. 2015. “Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells.” *Cell Reports* 12 (4): 554–61. https://doi.org/10.1016/j.celrep.2015.06.067.

Nesterova, Tatyana B., Sergey Ya Slobodyanyuk, Eugene A. Elisaphenko, Alexander I. Shevchenko, Colette Johnston, Marina E. Pavlova, Igor B. Rogozin, Nikolay N. Kolesnikov, Neil Brockdorff, and Suren M. Zakian. 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 Research* 11 (5): 833–49. https://doi.org/10.1101/gr.174901.

Nesterova, Tatyana B, Guifeng Wei, Heather Coker, Greta Pintacuda, Joseph S Bowness, Tianyi Zhang, Mafalda Almeida, et al. 2019. “Systematic Allelic Analysis Defines the Interplay of Key Pathways in X Chromosome Inactivation.” *Nature Communications* 10 (3129): 1–15. https://doi.org/10.1038/s41467-019-11171-3.

Nora, Elphège P., Bryan R. Lajoie, Edda G. Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Piolot, et al. 2012. “Spatial Partitioning of the Regulatory Landscape of the X-Inactivation Centre.” *Nature* 485 (7398): 381–85. https://doi.org/10.1038/nature11049.

Nostrand, Eric L Van, Gabriel a Pratt, Alexander a Shishkin, Chelsea Gelboin-, Mark Y Fang, Balaji Sundararaman, Steven M Blue, et al. 2016. “Robust Transcriptome-Wide Discovery of RNA Binding Protein Binding Sites with Enhanced CLIP (ECLIP).” *Nature Methods* 13 (6): 508–14. https://doi.org/10.1038/nmeth.3810.Robust.

Ohno, S. 1967. *Sex Chromosomes and Sex-Linked Genes*. *Springer-Verlag*.

Ohno, S., W. D. Kaplan, and R. Kinosita. 1959. “Formation of the Sex Chromatin by a Single X-Chromosome in Liver Cells of Rattus Norvegicus.” *Experimental Cell Research*, 415–18. https://doi.org/10.1016/0014-4827(59)90031-X.

Okamoto, Ikuhiro, Danielle Arnaud, Patricia Le Baccon, Arie P. Otte, Christine M. Disteche, Philip Avner, and Edith Heard. 2005. “Evidence for de Novo Imprinted X-Chromosome Inactivation Independent of Meiotic Inactivation in Mice.” *Nature* 438 (7066): 369–73. https://doi.org/10.1038/nature04155.

Osley, Mary Ann. 2006. “Regulation of Histone H2A and H2B Ubiquitylation.” *Briefings in Functional Genomics and Proteomics* 5 (3): 179–89. https://doi.org/10.1093/bfgp/ell022.

Patil, Deepak P., Chun Kan Chen, Brian F. Pickering, Amy Chow, Constanza Jackson, Mitchell Guttman, and Samie R. Jaffrey. 2016. “M6 A RNA Methylation Promotes XIST-Mediated Transcriptional Repression.” *Nature* 537 (7620): 369–73. https://doi.org/10.1038/nature19342.

Patrat, Catherine, Ikuhiro Okamoto, Patricia Diabangouaya, Vivian Vialon, Patricia Le Baccon, Jennifer Chow, and Edith Heard. 2009. “Dynamic Changes in Paternal X-Chromosome Activity during Imprinted X-Chromosome Inactivation in Mice.” *PNAS* 106 (13): 5198–5203. https://doi.org/10.1073/pnas.0810683106.

Pintacuda, Greta, Guifeng Wei, Chloë Roustan, Burcu Anil Kirmizitas, Nicolae Solcan, Andrea Cerase, Alfredo Castello, et al. 2017. “HnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing.” *Molecular Cell* 68 (5): 955-969.e10. https://doi.org/10.1016/j.molcel.2017.11.013.

Ran, F Ann, Patrick D Hsu, Jason Wright, Vineeta Agarwala, David A Scott, and Feng Zhang. 2013. “Genome Engineering Using the CRISPR-Cas9 System.” *Nature Protocols* 8 (11): 2281–2308. https://doi.org/10.1038/nprot.2013.143.

Ridings-Figueroa, Rebeca, Emma R. Stewart, Tatyana B. Nesterova, Heather Coker, Greta Pintacuda, Jonathan Godwin, Rose Wilson, et al. 2017. “The Nuclear Matrix Protein CIZ1 Facilitates Localization of Xist RNA to the Inactive X-Chromosome Territory.” *Genes and Development* 31 (9): 876–88.

https://doi.org/10.1101/gad.295907.117.

Rocha, Simão T. Da, and Edith Heard. 2017. "Novel Players in X Inactivation: Insights into Xist-Mediated Gene Silencing and Chromosome Conformation." *Nature Structural and Molecular Biology* 24 (3): 197–204. https://doi.org/10.1038/nsmb.3370.

Rocha, Simão Teixeira da, Valentina Boeva, Martin Escamilla-del-arenal, Katia Ancelin, Camille Granier, Neuza Reis Matias, Danny Reinberg, A Francis Stewart, Anton Wutz, and Edith Heard. 2014. "Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome." *Molecular Cell* 53: 301–16. https://doi.org/10.1016/j.molcel.2014.01.002.

Sakata, Yuka, Koji Nagao, Yuko Hoki, Hiroyuki Sasaki, Chikashi Obuse, and Takashi Sado. 2017. "Defects in Dosage Compensation Impact Global Gene Regulation in the Mouse Trophoblast." *Development* 144: 2784–97. https://doi.org/10.1242/dev.149138.

Schindelin, Johannes, Ignacio Arganda-Carrera, Erwin Frise, Kaynig Verena, Longair Mark, Pietzsch Tobias, Preibisch Stephan, et al. 2009. "Fiji - an Open Platform for Biological Image Analysis." *Nature Methods* 9 (7): 1–15. https://doi.org/10.1038/nmeth.2019.Fiji.

Schulz, Edda G, Johannes Meisig, Tomonori Nakamura, Ikuhiro Okamoto, Anja Sieber, Christel Picard, Maud Borensztein, Mitinori Saitou, Nils Blüthgen, and Edith Heard. 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 (2): 203–16. https://doi.org/10.1016/j.stem.2013.11.022.

Senner, Claire E., Tatyana B. Nesterova, Sara Norton, Hamlata Dewchand, Jonathan Godwin, Winifred Mak, and Neil Brockdorff. 2011. "Disruption of a Conserved Region of Xist Exon 1 Impairs Xist RNA Localisation and X-Linked Gene Silencing during Random and Imprinted X Chromosome Inactivation." *Development* 138 (8): 1541–50. https://doi.org/10.1242/dev.056812.

Shao, Zhaohui, Florian Raible, Ramin Mollaaghababa, Jeffrey R Guyon, Chao-ting Wu, Welcome Bender, and Robert E Kingston. 1999. "Stabilization of Chromatin Structure by PRC1, a Polycomb Complex." *Cell* 98 (7): 37–46. https://doi.org/10.1016/S0092-8674(00)80604-2.

Starmer, Joshua, and Terry Magnuson. 2009. "A New Model for Random X Chromosome Inactivation." *Development* 136: 1–10. https://doi.org/10.1242/dev.025908.

Sun, Sha, Brian C Del Rosario, Attila Szanto, Yuya Ogawa, Yesu Jeon, and Jeannie T Lee. 2013. "Jpx RNA Activates Xist by Evicting CTCF." *Cell* 153 (7): 1537–51. https://doi.org/10.1016/j.cell.2013.05.028.

Sunwoo, Hongjae, David Colognori, John E. Froberg, Yesu Jeon, and Jeannie T. Lee. 2017. "Repeat E Anchors Xist RNA to the Inactive X Chromosomal Compartment through CDKN1A-Interacting Protein (CIZ1)." *Proceedings of the National Academy of Sciences* 114 (40): 10654–59. https://doi.org/10.1073/pnas.1711206114.

Tang, Y Amy, Derek Huntley, Giovanni Montana, Andrea Cerase, Tatyana B Nesterova, and Neil Brockdorff. 2010. "Efficiency of Xist-Mediated Silencing on Autosomes Is Linked to Chromosomal Domain Organisation." *Epigenetics and Chromatin* 3 (10): 1–11. https://doi.org/10.1186/1756-8935-3-10.

Tavares, Lígia, Emilia Dimitrova, David Oxley, Judith Webster, Raymond Poot, Jeroen Demmers, Karel Bezstarosti, et al. 2012. "RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3." *Cell* 148 (4): 664–78. https://doi.org/10.1016/j.cell.2011.12.029.

The ENCODE Project Consortium. 2012. "An Integrated Encyclopedia of DNA Elements in the Human Genome." *Nature* 489 (7414): 57–74. https://doi.org/10.1038/nature11247.

Torres, Rodrigo F., Ricardo Kouro, and Bredford Kerr. 2019. "Writers and Readers of DNA Methylation/Hydroxymethylation in Physiological Aging and Its Impact on Cognitive Function." *Neural Plasticity* 2019: 1–8. https://doi.org/10.1155/2019/5982625.

Turner, Bryan M., Andrew J. Birley, and Jayne Lavender. 1992. "Histone H4 Isoforms Acetylated at Specific Lysine Residues Define Individual Chromosomes and Chromatin Domains in Drosophila Polytene Nuclei." *Cell* 69: 375–84. https://doi.org/10.1016/0092-8674(92)90417-b.

Vuong, Celine K, Douglas L Black, and Sika Zheng. 2016. "The Neurogenetics of Alternative Splicing." *Nat Rev Neurosci* 17 (5): 265–81. https://doi.org/10.1038/nrn.2016.27.

Waddington, C. H. 1957. “The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology.” *Allen & Unwin, London*.

Waddington, Conrad H. 2012. “The Epigenotype.” *International Journal of Epidemiology* 41 (1): 10–13. https://doi.org/10.1093/ije/dyr184.

Walker, Cheri L., Colyn B. Cargile, Kimberly M. Floy, Michael Delannoy, and Barbara R. Migeon. 1991. “The Barr Body Is a Looped X Chromosome Formed by Telomere Association.” *Proceedings of the National Academy of Sciences of the United States of America* 88 (14): 6191–95. https://doi.org/10.1073/pnas.88.14.6191.

Wang, Chen-Yu, David Colognori, Hongjae Sunwoo, Danni Wang, and Jeannie T. Lee. 2019. “PRC1 Collaborates with SMCHD1 to Fold the X-Chromosome and Spread Xist RNA between Chromosome Compartments.” *Nature Communications* 10 (1): 2950. https://doi.org/10.1038/s41467-019-10755-3.

Wang, Hengbin, Liangjun Wang, Hediye Erdjument-Bromage, Miguel Vidal, Paul Tempst, Richard S. Jones, and Yi Zhang. 2004. “Role of Histone H2A Ubiquitination in Polycomb Silencing.” *Nature* 431 (October): 873–78. https://doi.org/10.1038/nature02926.

Wutz, Anton, and Rudolf Jaenisch. 2000. “A Shift from Reversible to Irreversible X Inactivation Is Triggered during ES Cell Differentiation.” *Molecular Cell* 5 (4): 695–705. https://doi.org/10.1016/S1097-2765(00)80248-8.

Wutz, Anton, Theodore P. Rasmussen, and Rudolf Jaenisch. 2002. “Chromosomal Silencing and Localization Are Mediated by Different Domains of Xist RNA.” *Nature Genetics* 30 (2): 167–74. https://doi.org/10.1038/ng820.

Xiong, Yuanyan, Xiaoshu Chen, Zhidong Chen, Xunzhang Wang, Suhua Shi, Xueqin Wang, Jianzhi Zhang, and Xionglei He. 2010. “RNA Sequencing Shows No Dosage Compensation of the Active X-Chromosome.” *Nature Genetics* 42 (12): 1043–47. https://doi.org/10.1038/ng.711.

Yamada, Norishige, Yuko Hasegawa, Minghui Yue, Tomofumi Hamada, Shinichi Nakagawa, and Yuya Ogawa. 2015. “Xist Exon 7 Contributes to the Stable Localization of Xist RNA on the Inactive X-Chromosome.” *PLoS Genetics* 11 (8): 1–21. https://doi.org/10.1371/journal.pgen.1005430.

Yang, Fan, Tomas Babak, Jay Shendure, and Christine M. Disteche. 2010. “Global Survey of Escape from X Inactivation by RNA-Sequencing in Mouse.” *Genome Research* 20 (5): 614–22. https://doi.org/10.1101/gr.103200.109.

Yen, Ziny C., Irmtraud M. Meyer, Sanja Karalic, and Carolyn J. Brown. 2007. “A Cross-Species Comparison of X-Chromosome Inactivation in Eutheria.” *Genomics* 90 (4): 453–63. https://doi.org/10.1016/j.ygeno.2007.07.002.

Yue, Minghui, Akiyo Ogawa, Norishige Yamada, John Lalith Charles Richard, Artem Barski, and Yuya Ogawa. 2017. “Xist RNA Repeat E Is Essential for ASH2L Recruitment to the Inactive X and Regulates Histone Modifications and Escape Gene Expression.” *PLoS Genetics* 13 (7): 1–23. https://doi.org/10.1371/journal.pgen.1006890.

Zhao, Jing, Bryan K. Sun, Jennifer A. Erwin, Ji Joon Song, and Jeannie T. Lee. 2008. “Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome.” *Science* 322 (5902): 750–56. https://doi.org/10.1126/science.1163045.

Żylicz, Jan Jakub, Aurélie Bousard, Kristina Žumer, Francois Dossin, Eusra Mohammad, Simão Teixeira da Rocha, Björn Schwalb, et al. 2019. “The Implication of Early Chromatin Changes in X Chromosome Inactivation.” *Cell* 176: 1–16. https://doi.org/10.1016/j.cell.2018.11.041.

Annexes

Annex 1 - Classical “epigenetic” chromatin marks

(adapted from (Häfner and Lund 2016))

DNA modifications

1. Methylation (5 meC)
    a. Methyl group added on cytosine or adenine bases
    b. Mainly studied in CpG context, might be frequent in other sequences
    c. Leads to gene silencing via DNA compaction, recruitment of transcriptional repressors and exclusion of transcriptional activators
    d. Initially deposited by de novo methyltransferases guided by sequence, DNA binding proteins, long noncoding RNAs or RNA interference
    e. Mitotically inherited in a semi-conservative manner via maintenance methyltransferases
    f. Enzymatic activities responsible for demethylation in mammals are still controversial
2. Hydroxymethylation (5 hmeC)
    a. Found in many mammalian tissues
    b. Effect on gene expression still relatively unknown

Histone modifications

1. Post-translational modifications of specific serine, lysine and arginine residues of the histone amino-terminal tail or histone variants
2. Histone-modifying enzymes are recruited by specific DNA-sequences or guided intermediate protein and/or RNA complexes or RNA interference
3. The replication and inheritance of histone modifications during mitosis is unclear
4. Types:
    i. Acetylation: associated with transcriptional activation
    ii. Methylation: mono-, di- or tri-methylation; either repression- or activation-associated, depending on the targeted residue
    iii. Phosphorylation: associated with transcriptional activation
    iv. Ubiquitylation
    v. Sumoylation: associated with transcriptional repression
    vi. Histone variants: macroH2A is associated with inactive chromatin; H3.3 might be mitotically heritable and accumulates in active chromatin

Annex 2 – *Xist* interactors identified by different studies

(adapted from (Mira-Bontenbal and Gribnau 2016; Simão T. Da Rocha and Heard 2017))

Study		Cell lines used	Techniques (crosslinking)	Proteins identified
Xist interactors	Chu *et al.* (Chu *et al.* 2015)	XY mESCs and differentiating mESCs with a *Xist*-TetOP transgene on chromosome 11 Mouse XX EpiSCs and TSCs	**ChIRP-MS** Biotinylated probes to pulldown *Xist* followed by mass spectrometry (PFA)	**SPEN/SHARP** **RBM15** WTAP HNRNPU MATR3 SAP18 PTBP1 HNRNPC HNRNPK HNRNPL HNRNPAO RNF20 *Xist* ΔA lacks SPEN, WTAP, RBM15
	Bousard *et al.* (Bousard *et al.* 2019)	XY mESCs and differentiating mESCs with a *Xist*-TetOP transgene on endogenous locus		**SPEN** HNRNPUL2 RBMXL1 **RBM15** TRIM71 ZFR RNF20 SAFB *Xist* ΔB+C lacks RNF2, PCGF5, SAP18, RYBP, HNRNPK
	McHugh *et al.* (McHugh *et al.* 2015)	Differentiating XY mESCs with a *Xist*-TetOP at the endogenous locus	**RAP-MS** Biotinylated probes to pulldown Xist from SILAC-labelled cells followed by mass spectrometry (UV)	**SPEN/SHARP** **RBM15** MYEF2 CELF1 HNRNPC LBR SAF-A/HNRNPU RALY HNRNPM PTBP1
	Minajigi *et al.* (Minajigi *et al.* 2015)	Mouse XX fibroblasts	**iDRiP** Biotinylated probes to pulldown Xist followed by mass spectrometry (UV)	**SPEN/SHARP** **RBM15** HNRNPU LBR PTBP1 HNRNPK HNRNPL HNRNPAO
Xist interactors involved in silencing	Moindrot *et al.* (Moindrot *et al.* 2015)	XY mESCs with a *Xist*-TetOP transgene on chromosome 17; GFP reporter in cis with transgene	**shRNA screen** shRNA library genes with functions in ubiquitylation, and SUMOylation and genes encoding nuclear proteins	**SPEN/SHARP** **RBM15** WTAP MATR3 SAP18 PTBP1 HNRNPC HNRNPL HNRNPAO
	Monfort *et al.* (Monfort *et al.* 2015)	Haploid mESCs with a DOX-inducible promoter within the *Xist* locus	**Gene-trap screen** Viral gene trap mutagenesis strategy	**SPEN/SHARP**

Annex 3 – List of gRNAs and primers sequences used for CRISPR/Cas9 editing and validation of the different *Xist*-TetOP mutants

Xist mutant	5'end sgRNAs	3'end sgRNAs	Primer to confirm oligo integration on PX459 plasmid	Primers to confirm deletion (F/R)	Primers to confirm loss of FL sequence (F/R)
ΔF	GGTAATA CCGAAGA CGCT	GATAGT CTGAAT GACAAT AT	GAGGGCCT ATTTCCCAT GATT	CTGCTGATCGTTTGGTG CTG/TCTAAGAGGAGTT CTGAGTC (**Tm = 61ºC**)	TACTGACCTACAGTCCT CAT/TCTAAGAGGAGTTC TGAGTC (**Tm = 60ºC**)
ΔD	TTGTCAC CCATTAG GGTATG	ATATTGC AGGCAA GTAGCT C		CAGCCAAAAACCTCTGC ACT/CTTTTCTGGCAGTT GGTCCT (**Tm = 60ºC**)	CAGCCAAAAACCTCTGC ACT/TTCTCTCAAACCAC CACACG (**Tm = 60ºC**)
ΔE	AGAATTA GACACAC AGACCA	GCCTTTC GTGACT CCCCTTT		GCCCTTCCCTGTTTCTTG TT/AAGAACAAGTGGGG TGAGCA (**Tm = 66ºC**)	CCTTCCTTTGCATGTCTC CT/ACAGAGAAAGTGGC CCAAGA (**Tm = 60ºC**)

Annex 4 – PCR conditions

Reagents	1x (µl)		Conditions
Xpert Fast Hotstart Mastermix (2x)	12.5		95ºC 3 min
Fw primer (10 µM)	1.25		
Rw primer (10 µM)	1.25		(95ºC 30 sec → TmºC 1 min → 72ºC 1 min) x 35
H_2O RNase/DNase free	9		
DNA (20-100 ng/µl)	1		72ºC 10 min
Total	25		

Annex 5 - RT-qPCR primer-probe sequences

Gene	Fw primer	Rv primer
Xist	GCTGGTTCGTCTATCTTGTGGG	CAGAGTAGCGAGGACTTGAAGAG
Pgk1	CTCCGCTTTCATGTAGAGGAAG	GACATCTCCTAGTTTGGACAGTG
Rnf12	CCCCAGGTGAAAGTACTGAGG	CTCTCCAGCTCTATTTTCATCG
GapdH	AACTTTGGCATTGTGGAAGG	ACACATTGGGGGTAGGAACA

Annex 6 - Sequences of the set of Stellaris RNA-FISH oligo-probes

Exon 7 Set	Oligo-probes	Exon 7 Set	Oligo-probes
* 1	acaggcaccagagaaagtga	25	gtggtgttaggtgataagca
2	gagccatagctagtgaagac	26	tctatgtgttcgctcaacac
3	agattatttctcagggcagt	27	tttggtctagatttcactcc
4	gaccggacggaatgatgtat	28	taacatttagcacactgcct
5	acttcagagccacttgaatc	29	ggaagtatcctagacagcat
6	tcacaggtgtcctgtagaaa	30	tttactggcaaggtgtttgt
7	tatgcttggacttagctcag	31	tcactctggccctataaaac
8	ctttatgggcaatggcaaca	32	aggacctaggttagcatatt
9	agacccttgcacaatactat	33	aggaagtctccagtttatgt
10	caaatggctaagacccagtt	34	ctgagccaggcaatgaacaa
11	gatttagttccgtctcaagt	35	ctgtagtctcaaggtgtgac
12	tatgcatgcttattcctagg	36	tgtttttcccagctaaagtg
13	aggatggcagtatgcatatc	37	ggaggggagcaacttttat
14	tttctctagatagcctgaca	38	agtagcatcttctgcaatgt
15	tagaaaaccataccctgctc	39	ctgtaagtcacactgagtgc
16	actactatgagcagggagtt	40	ttcaacctctgaggcaaact
17	tcaggtacttagtctgatgc	41	tccttgtagaacaggcgaac
18	tcagggcttgatcaagtaca	42	agggcttctgtaactttagg
19	ttcatacgtttcctggttac	43	cttaatctgtttcattccca
20	ggaagttacagtaggcttca	44	agctctaaggcaacttggat
21	gcatgttattcctagagcaa	45	gtaacacttttgatgcccaa
22	cctattttataggcagcttt	46	gaaatgtaagcccattccta
23	gtgcatgcctgggataaaag	47	gatgatggtaggatgtgctt
24	acaaagctccacacagatct	48	cagttggtgggaagatgact

* Oligo-probes that not bind to *Xist* RNA in the mutants without the E-repeat.

Annex 7 – Antibodies dilutions used for IF/RNA-FISH

Experiment	Antibody	Against	Reference	Dilution
IF/RNA-FISH	Primary (Rabbit)	H3K27me3	H3K27me3 pAb Rabbit (Active Motif #39155)	1/200
		H2AK119ub	Ubiquityl-Histone H2A (Lys119) (D27C4) XP® Rabbit mAb (Cell Signaling Technology, #8240)	1/200
		H4K20me1	Anti-Histone H4 (mono methyl K20) antibody - ChIP Grade (abcam, #ab9051)	1/200
	Secondary (Goat)	Rabbit (Green)	Alexa Fluor 488-conjugated AffiniPure Goat Anti-Rabbit IgG (H+L) (Jackson ImmunoResearch, #111545144)	1/100
		Rabbit (Cy5 Red)	Goat Anti-Rabbit IgG H&L (Cy5 ®) preadsorbed (Abcam, ab97077)	

Annex 8 – DNA sequences of the generated *Xist* mutants (Sanger method)

Xist mutant	Clone	Fw Primer	Sequence
ΔF	F1	GACGACT AGCAAAC CACGAC	CCATGGCTTAAAACGACTTTGCTCTTAACTGAGTGGGTGTTCAGGGCG TGGAGAGCCCGCGTCCGCCATCAGACTTACTGTTATAATAGTAATATT AAGGCCTACATTTCAACTTTCTGTGTGTTCTTGCCTTTATGGCATCTAG ATTCTCCTCAAGACTCAGATCA
	F2		AGGGCTTAAAACGACTTTGCTCTTAACTGAGTGGGTGTTCAGGGCGT GGAGAGCCCGCGTCCGCCATAGAGGACACAGAGCATTACAATTCAAG GCTCTTGCTGACAAGTAAAACCATTAGTCAATATAGCATCAACAATA GCAACAAAAGCAAAGCCTGAACTGGTATAGAACCTCAAATACTCCTG ACATCCAGTCATAATGTTCACTTTACATTAACAGTATCTGTGTTATAA GACTTAAAAATTTGAAGAGTAGCTCGGTGGATGAGTTTGAAAGAAAG TACATTTAAGTGCCAAGTAATTTGAGTATCATCTGCCAAAAAATCAG ACTTACTGTTATAATAGTAATATTAAGGCCTACATTTCAACTTTCTGT GTGTTCTTGCCTTTATGGCATCTAGATTCTCCTCAAAGACTCAGAAG
	F3		GCTAAAAAACGACTTTGCTCTTAACTGAGTGGGTGTTCAGGGCGTGTT ATAATAGTAATATTAAGGCCTACATTTCAACTTTCTGTGTGTTCTTGC CTTTATGGCATCTAGATTCTCCTCAAGACTCAGAAA
ΔD	D1	AGCTAGT CCAGCCC TGTGTG	GGGAATCCTGACTAGGAATTGAGTCTTTTCTCAAGGTCCTAATACTAC CCTTGCTTTATGTAAAGAGGGTGCTGATTACTTAATGCCTCTTACACA ATTGTGCAAAATTGCAGTTGTTCAAGTCCCCTTCTGTTAGTAACCAAG ATCCCATACCCTCAGGGAGTTGGATTGCCACCTTTTACTTGGGGCTTT CCTTTACAGTATGAACTGAAAATTGTCTTCCTGAGAAGGAAGCTTAGC ACTTTTCTTTCCATTCTTCCTCCAGGAAGGAGCCAACTGTCTGCTTAA GAAACTTTAAGCCCGATTTTGTATATTGCTACTGTACAGGACCAACTG CCAGAAAAGAAG
	D2		GGGGAATCCTGACTAGGAATTGAGTCTTTTCTCAAGGTCCTAATACTA CCCTTGCTTTATGTAAAGAGGGTGCTGATTACTTAATGCCTCTTACAC AATTGTGCAAAATTGCAGTTGTTCAAGTCCCCTTCTGTTAGTAACCAA GATCCCATACCCTCAACTCAGGGAGTTGGATTGCCACCTTTTACTTGG GGCTTTCCTTTACAGTATGAACTGAAAATTGTCTTCCTGAGAAGGAAG CTTAGCACTTTTCTTTCCATTCTTCCTCCAGGAAGGAGCCAACTGTCT GCTTAAGAAACTTTAAGCCCGATTTTGTATATTGCTACTGTACAGGAC CAACTGCCAGAAAAGAA
	D3		TCCTGACTAGGATTGAGTCTTTTCTCAAGGTCCTAATACTACCCTTGC TTTATGTAAAGAGGGTGCTGATTACTTAATGCCTCTTACACAATTGTG CAAAATTGCAGTTGTTCAAGTCCCCTTCTGTTAGTACTTTTCTTTCCAT TCTTCCTCCAGGAAGGAGCCAACTGTCTGCTTAAGAAACTTTAAGCCC GATTTTGTATATTGCTACTGTACAGGACCAACTGCCAGAAAAGAAC
ΔE	E1	GCCCTTC CCTGTTT CTTGTT	**TGCTTCTATCATCATTAG**TGTGTATTGTGGGTGTGTCTATTTCTTGTT TTATGTATCTATTTTTTCCTTGGTTTGGGTCACATGTTGCATGCATCCC TCTCTTTTTCTTGTGCTCACCCCACTTGTTCTTAAAAC
	E2		**GGGGTTTGCTTACTATCATCATTACG**TGTGTATTGTGGGTGTGTCTA TTTCTTGTTTTATGTATCTATTTGTGTCATGTGTTGCATGCATACGTCT CATGCTCTTGTGCTCATCTCTCTCACTGTGTTCTTACCACTTGTTCTTA A
	E3		**GGTTTGCTTCTATCATCATTAG**TGTGTATTGTGGGTGTGTCTATTTC TTGTTTTATGTATCTATTTTTTCTGTGTGTCTAATTCTTTGTTACATCTA TTTCTTCCTTGCTTTGTGTGTCTATTTCTTCCTTGCTTTGTGTGTCTATT TCTTCCTTGCATTATGTCTAATTCTTTGTTATATCTATTTCTTCCTTGCT TTGTGTCTATTTCTTCCTTGCAGTTGTGTCTAATTCTTTGTTACATCTAT TTCTTCCTTGCTTTGTGTGTCTATTTCTTCCTTGCATTGTGTCTAATTCT TTGGGTCACATGTTGCATGCATCCCTCTCTTTTTCTTGTGCTCACCCCA CATTGTTCTTAAT

Note: Introns are represented in bold; green represents the sequence upstream of the deletion; yellow represents the sequence downstream of the deletion.

Annex 9 – *Xist*-TetOP mutants' characterization

Xist mutant	Clone number	Size of deletion	Deleted region	% of cells with *Xist* domains	*Xist* coating	Nº of experiments	Total cell counted
FL	n.a.	n.a.	n.a.	44 %	+++++	1	492
ΔF	F1	1,815 bp	Δ888-2,704				
	F2	1,814 bp	Δ890-(255bp)-2,704				
	F3	1,846 bp	Δ869-2,715	19%	+++	1	300
ΔD	D1	2,994 bp	Δ5,429-8,323	44%	+++++	1	290
	D2	2,890 bp	Δ5,429-8,319				
	D3	2,997	Δ5,408-8,405	49%	+++++	1	362
ΔE	E1	909 bp	Δ53-962	32%	+++++	1	225
	E2	921 bp	Δ44-965				
	E3	688 bp	Δ273-961	27%	+++++	1	236