# Activation of X Chromosome Inactivation

**Cheryl Mandy Maduro**

# Activation of X Chromosome Inactivation

Het activeren van X chromosoom inactivatie

Thesis

to obtain the degree of Doctor from the Erasmus University Rotterdam
by command of the rector magnificus

Prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on

Tuesday 6th December 2016 at 15:30 hours

by

**Cheryl Mandy Maduro**

born in The Hague

**Erasmus University Rotterdam**

**Doctoral Committee**

**Promotor:**                         Prof.dr. J.H. Gribnau

**Other members:**            Prof.dr. C. P. Verrijzer
Prof.dr. J.N.J Philipsen
Prof.dr. N. Geijsen

# Contents

## List of abbreviations

| | |
|---|---|
| BAC | Bacterial Artificial Chromsome |
| BF | Blocking Factor |
| bp | base pair |
| cDNA | complementary DNA |
| CF | Competence Factor |
| CT | Chromosome Territory |
| ESCs | Embryonic stem cells |
| FISH | Fluorescent in situ hybridization |
| ICM | Inner cell mass |
| iXCI | Imprinted X chromosome inactivation |
| Lif | Leukemia inhibitory factor |
| LINEs | Long interspersed nuclear elements |
| lncRNA | Long non-coding RNA |
| Mb | Megabase pairs |
| mRNA | Messenger RNA |
| MYA | Million years ago |
| PAR | Pseudo autosomal region |
| PE | Primitive endoderm |
| rXCI | Random X chromosome inactivation |
| TAD | Topologically associating domain |
| TE | Throphectoderm |
| Xa | Active X chromosome |
| XCI | X chromosome inactivation |
| XEN | eXtra embryonic endoderm |
| Xi | Inactive X chromosome |
| Xic | X inactivation center |
| Xist | X inactive specific transcript |
| Xm | Maternally derived X chromosome |
| Xp | Paternally derived X chromsome |
| Xpr | X pairing region |
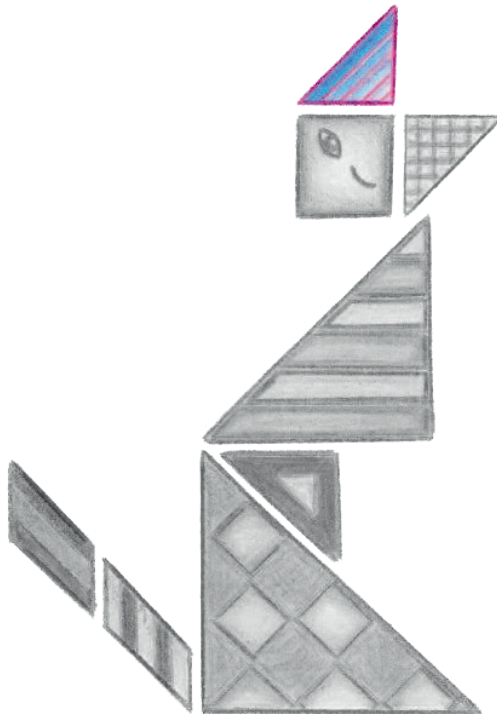
# Chapter 1

**Introduction**

**&**

**Aim and scope of this thesis**

# Chapter 1

## Introduction

Dimorphic sexes, can be found in most animal species, but determination of sex is achieved in a wide variety of ways. Some species determine sex by environmental cues [1], other species employ genetic sex determination either with or without distinguishable sex chromosomes [2]. In species with sex chromosomes, the sex chromosomes are entitled X and Y when the male is the heterogametic sex, as for example in mammals, and Z and W when the female is the heterogametic sex, as for example in birds. The mammalian sex chromosomes X and Y originated from an ancient autosomal pair about 180 MYA [3, 4]. The ancient autosomal pair evolved to the proto-X and proto-Y after acquisition of mutational changes of *Sox3* leading to the sex-determining gene *Sry* on the proto-Y. Once male specific, the proto-Y accumulated additional male beneficial genes and became progressively more diverged from the proto-X as a consequence of different rearrangements and decay of the Y chromosome, resulting in loss of recombination between the two chromosomes. Loss of recombination consequently initiated a form of selection to keep the cluster of sexually antagonistic genes as a unit on the Y chromosome [3-5]. Being unable to recombine, the X and Y chromosome evolved independently into the heterologous sex chromosomes we know today [6]. The X chromosome retained more than a thousand genes as it is still able to recombine with its homologue in females preventing its degeneration. In contrast, the Y chromosome has lost more than 95% of its original gene content as the Y chromosome degenerated over time [7]. In species with highly differentiated sex chromosomes however, dosage differences between the sexes occurs when sex chromosomes diverge and requires a form of compensation in either sex (Figure 1) [8].

*Dosage compensation*
The mammalian X and Y chromosome now differ greatly in size and in gene content. The X chromosome has been enriched in genes mainly involved in cognition, sex and reproduction [9, 10], whereas the Y chromosome mainly contains genes involved in male sex determination and spermatogenesis [11]. These differences between the sex chromosomes results in unequal gene dosage between males and females and such a dosage imbalance, equivalent to a monosomy, would affect the overall fitness of the species. For that reason, several species with sex chromosome dosage imbalances have developed different ways of coping with dosage differences between the sexes. Comparing dosage compensation mechanisms of eutherian mammals, flies and nematodes, representing three independent origins of male heterogamety, showed that the result of dosage compensation in these species was the same: equalized expression

of X-encoded genes in male and female [8]. All these species show in one sex, specific complexes composed of RNA and/or protein that target an entire chromosome for stable and inheritable changes in transcription levels through epigenetic modifications [12]. Nevertheless, the actual mechanisms for dosage compensation between eutherian mammals, flies and nematodes are very different [8]. In *Caenorhabditis elegans* for example, in which males are XO and hermaphrodites are XX, reduced expression of both X chromosomes in hermaphrodites results in an equalized dosage of X encoded gene expression between the sexes [13]. In *Drosophila melanogaster* however, with XY males and XX females, a 2 fold over expression of the male X encoded genes achieves dosage compensation [14-16]. In mammals, sex is also determined by an XX female and XY male system, but in contrast to *Drosophila melanogaster* in which expression in males is adjusted, dosage compensation in mammals is achieved by inactivating one of the two X chromosomes in females. This process is known as X-chromosome inactivation (XCI), silencing of chromosome wide transcription on one of the two X chromosomes which stays inactive throughout life.

*X chromosome inactivation: the initial hypothesis*
Random X chromosome inactivation was proposed in 1961 by Mary Lyon and was first described as the Lyon hypothesis [17, 18]. Based on the finding that the Barr body contained either one of the two X chromosomes in the female mammalian nucleus, Mary Lyon postulated that equalization of X-encoded gene dosage between males and females is achieved by transcriptional silencing of one X chromosome in females. The Lyon hypothesis consisted of four key elements with regard to XCI [17, 18]. First, each female cell contained one active X chromosome (Xa) and one inactive X chromosome (Xi). Second, inactivation of one of the two X chromosomes occurred early in development of the female embryo. Third, which X chromosome was inactivated was randomly determined; either the maternal or the paternal X chromosome was inactivated. And finally, all daughter cells would have the inactivation pattern of the parental cell, thus clonal expansion of the X inactivation pattern in the embryo would result in mosaic females. This paper, initially based solely on mouse studies, was followed by a fuller development of the hypothesis in 1962, focusing more on human X-linked conditions [19]. Mary Lyons hypothesis provided a unifying hypothesis which was brought together from various earlier work: the discovery of the Barr body [20], cytological and theoretical studies of Susumo Ohno on sex chromosome biology and Ohno's law [21, 22] and studies on sex disorders in humans [23, 24], explaining these findings for the first time. In addition, this hypothesis was also independently published by Beutler [25] based on results from a human X-linked disorder, glucose-6-phosphate dehydrogenase deficiency. Together these papers provided the scientific community with sufficient evidence to support this hypothesis and explained a longstanding problem of dosage compensation and function of sex chromosomes, now known as XCI. One of the unresolved issues regarding dosage compensation by means of XCI is described in Ohno's law [21]: "XCI itself cannot correct for the expression imbalance between the sex chromosomes and autosomes, which is the result of having only one active X chromosome in females and males, sex chromosome monosomy, as opposed to 2 active chromosomes for all the other autosomes". Therefore Ohno hypothesized that the genes on the single active X chromosome in male and female undergo a

2-fold upregulation based on the negative effects of aneuploidy. Even though there seems to be enough evidence to prove a 2-fold upregulation of X-linked genes [26-30], there also seems to be an equal number of studies showing this is actually not the case [31-35]. Whether or not Ohno's law applies to eutherian mammals is thus still controversial, mainly because of different techniques and thresholds used in the analysis of expression data and therefore still remains an open question [4, 36, 37]. Nevertheless, close examination of a wide range of studies indicates that dose sensitive genes, for instance genes encoding proteins acting in stoichiometric protein complexes, are clearly upregulated, whereas lowly expressed genes are not [32, 34].
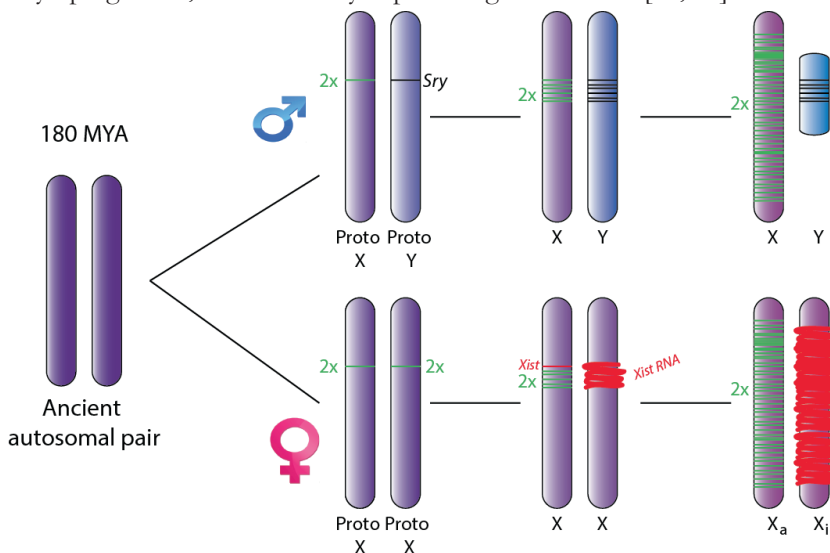


**Figure 1:** Evolution of the X and Y chromosome. The X and Y chromosome evolved over 180 MYA from an ancient autosomal pair. After acquiring the male sex-determining gene *Sry* on the proto-Y, the Y chromosome accumulated more male specific genes (black) and degenerated over time, which resulted in a need for a two-fold upregulation of genes lost from the Y chromosome (green) on the X chromosome to compensate for their loss on the Y chromosome. A two-fold upregulation of these genes leads to a 4-fold upregulation in female cells forcing the co-evolution of a second dosage compensation mechanism, XCI, which is achieved by spreading of the lncRNA Xist (red) along the X chromosome in *cis*.

**XCI in the mouse**

In the mouse, XCI occurs in two rounds (Figure 2). Around the 2- to 4-cell stage, cells in female embryos initiate imprinted XCI (iXCI), leading to exclusive inactivation of the paternal X chromosome (Xp) [38, 39]. As the embryo further develops into a blastocyst, XCI is reversed in the inner cell mass (ICM) by reactivation of the Xp [40]. Subsequent initiation of random XCI (rXCI) occurs in the embryoblast just after implantation around embryonic day 5.5 [40, 41], but iXCI persists in the extra embryonic tissues [42]. Mouse embryonic stem (ES) cells are derived from the ICM with two active X chromosomes and undergo rXCI upon differentiation, recapitulating rXCI in vitro. rXCI can be divided into three main phases: initiation, establishment and maintenance of XCI [43]. In the initiation phase, the cell ensures one active X chromosome per diploid genome, in a process regulated by activators and inhibitors of

XCI leading to mono-allelic up-regulation of the X-linked long non-coding RNA gene Xist. In the establishment phase, Xist RNA spreads and coats the entire future inactive X chromosome (Xi) in cis, resulting in loss of active histone marks and gain of inactive histone marks, contributing to the silencing process [44-47]. These modifications are catalyzed by enzymes directly and indirectly recruited by Xist through several functional domains, most of them repeats, that are also involved in Xist localization to the Xi. Once XCI is complete, the Xi is maintained and clonally propagated to all daughter cells.
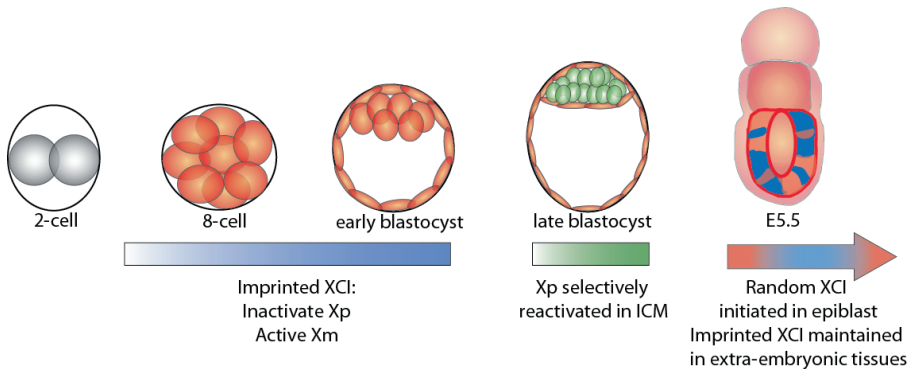


**Figure 2:** XCI in the mouse occurs in two waves. After the two-cell stage cells initiate imprinted XCI (iXCI), always inactivating the paternal X chromosome (Xp) and keep the maternal X chromosome active (shown as orange cells). As the embryo develops, cells in the inner cell mass (ICM) of the blastocyst reactivate their paternal X chromosome (green cells). After reactivation, these cells will initiate random XCI (rXCI) in the embryoblast resulting in a mosaic embryo with cells having either the Xp or Xm inactivated (orange or blue cells).

XCI counting and initiation: the models

Random XCI is initiated when the number of X chromosomes exceeds one per diploid genome. The cell must ensure that only one X chromosome is inactivated, inactivation or continued activity of both X chromosomes is lethal, underscoring the need for tight regulation of the XCI process. Over the years many models have been proposed for XCI counting and initiation including the blocking factor model, the symmetry-breaking model, the alternate states model, the X-pairing model and the stochastic model. The blocking factor model states that a limiting autosomally encoded factor blocks XCI on one of the two X chromosomes in female cells and on the single X chromosome in male cells. The cell is thus able to count the number of X chromosomes based on the excess of autosomes over X chromosomes and determine whether enough blocking factor is present to inhibit XCI [48]. However, conclusive evidence supporting this model has not been found. The symmetry breaking model is based on the blocking factor model as it also suggests blocking of XCI on one of the two X chromosomes in female cells and on the single X chromosome in male cells. However, this model predicts that it is a blocking complex, which forms on the X chromosome and inhibits XCI. This circumvents the problem of having a single factor present in the cell as proposed in the blocking factor model. There has however not been any report on complexes forming on the active X chromosome (Xa) so far [49]. Another model proposed is the alternate states model in which XCI is intrinsically determined by the chromatin state

of the X prior to XCI [50]. The different states of sister chromatid cohesin are transient and are mutually exclusively locked at the onset of XCI, ensuring random XCI. The different states may represent blocking factor binding to one chromosome, accessibility or transcriptional differences between the two X chromosomes [51]. However, further evidence for this model has yet to be found as well. The X-pairing model is based on the observation that prior to up-regulation of Xist, the regulatory regions of Xist of both X chromosomes have been found to transiently move in spatial proximity with an increased frequency [52-54]. When investigating live-cell dynamics and the outcome of pairing in differentiating ES cells, an increased X chromosome mobility was found during differentiation followed by a reduced mobility during pairing. The increased mobility during differentiation might facilitate the onset of pairing by increasing the frequency of collisions between the two loci and/or different nuclear compartments and the reduced mobility during pairing suggests some form of binding of the two regulatory regions on the X chromosomes. Pairing is very transient and the suggested outcome of pairing is monoallelic repression of XCI on one X chromosome which could lead to monoallelic promotion of XCI on the other X chromosome [55]. It is hypothesized that during these movements the future Xa and Xi are being determined. The genomic regions that were identified as required for X- pairing cover Xite and the Tsix promoter, as well as the X pairing region (Xpr) which overlaps with Slc16a2 [53, 54]. Tsix and Xite are located in close proximity to Xist and are involved in negative regulation of Xist by generating antisense transcripts. The role of the Slc16a2 region in XCI seems to be restricted to the X-pairing process itself. A recent study identified an interaction between CTCF and Tsix RNA as a key mechanism required for X-pairing [56]. The authors demonstrated that CTCF binds Tsix RNA with a higher affinity compared to DNA, and propose that during transcription, CTCF is recruited by Tsix RNA to be loaded onto its genomic locus. Both knockdown of Tsix/Xite and CTCF results in reduced pairing [56, 57], and a reduction in X-pairing leads to reduced initiation of XCI, observed as a reduction in Xist expression [56]. In contrast, in another study a reduction in Xist expression was not observed upon knockdown of CTCF, although the effects on X-pairing were not investigated [58]. Despite the effects of ES cell differentiation on spatial movement of Xist and other pairing elements, a recent study provides evidence arguing against a functional role for X- pairing in XCI [59]. In this study, XCI is unaffected on the remaining wild type X chromosome in ES cells carrying a heterozygous deletion encompassing all the regions identified as required for pairing, including Xist, Tsix, Xite and Slc16a2. Moreover, heterokaryotic ES cells containing a male and a female nucleus show no preference for XCI of an X chromosome located in the male or female nucleus [59]. X-pairing might therefore reflect changes in the transcriptional activity of XCI specific genes coincidentally recruited to nuclear locations or entities with a higher local concentration of transcription factors (such as transcription factories). Taken together X- pairing is an intriguing phenomenon, but its implications should be carefully considered with respect to a functional role for X- pairing in XCI initiation.

The above-mentioned deterministic models predict a tightly regulated XCI counting and initiation process in which a single X is inactivated in every diploid female cell. Studies with tetraploid XXXX embryos indicated that two X chromosomes are inactivated, predicting one active X chromosome per diploid genome [60]. Interestingly,

in differentiating XXXX tetraploid ES cells however, the presence of three or four inactivated X chromosomes in a significant percentage of cells suggested that XCI is not so tightly regulated as the above mentioned models imply [43, 60]. In addition, female diploid cells consistently show XCI on both X chromosomes in a small percentage of differentiating ES cells. Based on these and other observations, a stochastic model was postulated. This model states that each X chromosome has a certain independent probability to be inactivated, and that this probability is proportional to the X: ploidy ratio. This could explain all the possible outcomes of XCI in XXXX tetraploid ES cells. The outcome of the rXCI process, according to the stochastic model, is a resultant of three key driving factors: a probability for each X chromosome in a cell nucleus to be inactivated, the probability to initiate XCI being proportional to X:ploidy ratio and a feedback mechanism leading to one active X per diploid genome. Because more and more data indicate this stochastic model to be true, we opt for the stochastic model for initiation of XCI.

**XCI counting and initiation: a stochastic feedback-mediated process**
During initiation of rXCI the number of X chromosomes is determined with respect to the genomic context. Initiation of XCI is a stochastic process, in which every X chromosome has an intrinsic probability to be inactivated [60]. This process is orchestrated by *trans*-acting X encoded activators and autosomally encoded inhibitors of XCI. The balance between these opposing factors will determine whether *Xist* is upregulated and if XCI will occur. In this balance, XCI-inhibitors represent the ploidy of a cell, while XCI-activators represent the number of X chromosomes present. These trans-acting regulatory factors act through a *cis*-regulatory locus controlling the expression of *Xist*.

*The cis-X inactivation center*
Genetic studies in mice and humans with X-to-autosome translocations have revealed that an X-linked control locus, the X inactivation center (Xic) is necessary for XCI to occur. Recently, it has been proposed that the Xic can be divided into a cis-Xic and a trans-Xic, containing all the *cis* and *trans* acting factors involved in XCI respectively [61]. The cis-Xic contains all *cis*- acting genes important for XCI [51]. Three key *cis*-acting genes, *Xist*, *Tsix* and *Xite* are non-coding genes from which the transcripts are spliced and polyadenylated. *Xist* is required in *cis* for silencing of the X chromosome as loss of function studies indicated that the mutated X chromosome could not be inactivated in ES cells, causing complete skewing of XCI towards the wild-type X chromosome [62, 63]. Accumulation and spreading of *Xist* RNA on the Xi in *cis* results in depletion of RNA polymerase II and other components of the transcription machinery, recruitment of the chromatin remodeling complexes, including Polycomb repressive complex 2 (PRC2), which trimethylates lysine 27 on histone H3 (H3K27me3) and several other complexes, leading to further changes of several histone modifications including histone-H3K9 methylation, histone-H4 deacetylation and histone-macroH2A accumulation which further contribute to the silencing process [64].

*Xist* RNA is retained in the nucleus and accumulation and spreading of *Xist* is restricted to the Xi and does not localize to neighboring chromosomes. The silent state is kept throughout life, but *Xist* is not required for the maintenance of XCI as

a conditional deletion of *Xist* after XCI has occurred does not reactivate silenced genes [43]. The different epigenetic modifications during XCI work synergistically in maintaining the silent state. In a study involving an inducible *Xist* expression system in ES cells, it was shown that *Xist* mediated silencing can be divided into two steps: primary inactivation and maintenance of the inactive state. The primary inactivation phase is reversible, requires *Xist* and encompasses the first 72 hours of differentiation. However, maintenance of the inactive state is irreversible, independent on *Xist* and separated from the primary inactivation phase by approximately one cell division. In ES cell differentiation there is thus a transition from reversible to irreversible XCI. *Xist* expression has to be induced early enough to allow primary inactivation to take effect before silencing becomes irreversible and independent of *Xist* expression. Thus initiation of *Xist*-mediated silencing has to occur during the critical window of ±48-72 hours of differentiation [65].

*Tsix* and X chromosome intergenic transcript element *(Xite)* are negative regulators of *Xist*. *Xite* is an enhancer of *Tsix* and was shown to be important for transiently maintaining *Tsix* expression on the Xa [66]. *Tsix* transcription is antisense to and completely overlaps with *Xist*, hence the name. *Tsix* transcripts only localize to the Xic and are expressed from both X chromosomes prior to XCI. After XCI, expression of *Tsix* only transiently continues on the future active X chromosome (Xa) [67]. The exact mechanism of *Tsix*-mediated repression of *Xist* remains unknown, however several mechanisms have been proposed. First of all, because *Tsix* is transcribed antisense to *Xist*, transcriptional interference has been proposed as a mechanism to repress *Xist* [12, 68]. In addition, the *Xist/Tsix* duplex RNA formation and processing by the RNAi pathway has been proposed to play a role by siRNA-mediated deposition of chromatin remodeling complexes [12, 69]. Recruitment of chromatin remodeling complexes by *Tsix* RNA to the *Xist* promoter was also implicated in *Tsix*-mediated repression of *Xist* [35]. Furthermore, *Tsix* was suggested to be involved in pairing of the two X chromosomes, as described above, which suggests a role for *Tsix* in initiation of XCI [54].

*The cis-X inactivation center: cis-activators*

Four non-coding *cis*-acting activators of XCI have been described: *Jpx, Ftx, Xpr* and *XistAR*. A mutation of *Ftx* in male cells results in reduced transcript levels of *Xist, Tsix* and *Jpx* [70] and a deletion of *Jpx* blocks XCI and is female lethal [71]. The *Xpr* region is suggested to be involved in X chromosome pairing and the *Xpr* could autonomously drive Xic-*trans* interactions even as an ectopic single-copy transgene. Furthermore, introduction of an *Xpr* transgene in male ES cells was strongly selected against suggesting a role in activation of *Xist* [53]. Although *Jpx, Ftx* and *Xpr* have been implicated in *trans*-regulation, a recent study indicates no trans effects and only *cis* effects upon deletion of *Jpx, Ftx* and *Xpr* [59]. Also recently, the long non-coding RNA, XistAR (Xist Activating RNA) has been described to activate Xist in cis and as required for XCI [72]. XistAR is encoded within exon 1 of the mouse *Xist* gene and is transcribed antisense to *Xist* and exclusively from the Xi. When XistAR is selectively truncated, without affecting the overlapping *Xist* gene, XCI initiation as a result of *Xist* induction is diminished by 90%.

*The trans X-inactivation center*

*Xist, Tsix* and *Xite* play an important role in *cis* inactivation, but have been shown not to be required for in *trans* communication [43], as female cells lacking one functional *Xist, Tsix* and *Xite* allele, are still able to initiate XCI on the wild type X chromosome. This suggests involvement of other factors in the counting and initiation process located outside this deleted region. The finding that these cells initiate XCI in contrast to XY males, and the observation that the XCI initiation frequency was much higher in XXXX tetraploid ES cells than XXXY ES cells suggested a model in which the probability to initiate XCI is proportional to the X chromosome:ploidy ratio. These observations also provided evidence for the presence of an X-linked gene(s) encoding a *trans*-acting factor(s) that is involved in promoting XCI. Autosomally encoded inhibitors set the threshold for XCI to occur whereas the X-linked activators help to overcome this threshold. Female cells have two X chromosomes and thus double the amount of X-linked activators as compared to male cells with only one X chromosome. The two copies of X-linked activator(s) in female cells can overcome the threshold to allow XCI to occur. Inactivation of one X chromosome in females, results in *cis* inactivation of the activator gene, thereby lowering the concentration of X-linked activator in *cis* below the threshold required for XCI to occur and similar to male cells. Because of the stochasticity of the process a female cell can also initiate XCI on both X chromosomes. Accumulation of *Xist* on both X chromosomes leads to reactivation of the X chromosomes because a specific level of XCI-activator activity may be required for sustained expression of *Xist* and maintenance of the Xi [59]. In addition, in a cell with two active X chromosomes are blocked in differentiation, extending the time frame where XCI can be initiated. The balance between activators and inhibitors of XCI determines the probability to initiate XCI and the feedback through inactivation of XCI activator genes in *cis* prevents inactivation of both X chromosomes in most female cells.

*The trans X-inactivation center: XCI inhibitors*

The pluripotency factors *Nanog, Rex1, Klf4, Oct4, c-Myc* and *Sox2*, germline factor *Prdm14* [73-75] as well as factors like *Ctcf* and *YY1* [76, 77] have been reported to inhibit the XCI process directly by repressing *Xist*, activating Tsix, or indirectly by repressing activators of XCI by binding of different combinations of these factors at different loci. NANOG, REX1, OCT4, SOX2 and PRDM14 have been implicated in *Xist* repression and OCT4, SOX2, REX1, c-MYC, KLF4 as well as CTCT and YY1 in *Tsix* activation. *Rex1* and other XCI inhibitors, including *Nanog, Sox2, Oct4*, are central factors in the pluripotency network, which link XCI to loss of pluripotency (Figure 3A). Recently, the MOF-associated complexes [78], involved in dosage-compensation in *Drosophila*, and chromatin organizers Satb1 and Satb2 [79] have also been associated with a possible role in inhibiting XCI, but this involvement needs to be further elucidated.

*The trans X-inactivation center: XCI activators*

Autosomally encoded XCI inhibitors or autosomally encoded Xist-activators such as YY1 [80] will be equally expressed in male and female cells and can therefore not be the determining factors in the counting process in initiation of XCI (Figure 3B). In contrast, X-encoded XCI activators will be differentially expressed between male and female cells. Activators of XCI promote *Xist* expression, either directly or indirectly

by repression of *Tsix*. Recently, *Rnf12, Xpr, Jpx* and *Ftx,* have been reported as XCI activators (Figure 3A), and interestingly they are all located in close proximity to *Xist*, which may facilitate a rapid feedback mechanism. Introduction of additional copies of mouse or human *Rnf12* in ES cells increased the probability to initiate XCI, which resulted in XCI on the single X chromosome in male cells and in a high percentage of female cells on both X chromosomes. These studies supported a dose-dependent role for RNF12 in the activation of XCI. *Rnf12* is essential for XCI, as *Rnf12*[+/-] and *Rnf12*[-/-] differentiated ESCs showed reduced XCI, with the latter having almost no XCI initiation at all, although the genetic background may influence the phenotype, as another study showed a less severe phenotype in vivo [81]. Unlike *Rnf12* transgenic male ES cell lines, male cell lines with *Ftx, Jpx* and *Xpr* transgenes did not show XCI induction. Female cells lacking the region containing *Jpx, Ftx,* and *Rnf12* showed further affected XCI as compared to a deletion of *Rnf12* alone, whereas a deletion of *Jpx* or *Ftx* alone did not appear to have affected XCI [59]. This indicates that *Jpx* and *Ftx* might only work alongside *Rnf12* in activating *Xist* expression. Moreover, it appears that the function of *Jpx* and *Ftx* is mostly if not exclusively in *cis*, as only when introducing *Rnf12* back into these cells as a randomly integrated transgene, the phenotype was rescued and a combination of *Rnf12, Jpx,* and *Ftx* transgenes did not have additional effects on XCI [59]. These findings argue against a role in trans activation of *Xist,* and could be explained by a different role for these genes in initiation of XCI in *cis*. For instance through co-activation or another unidentified mechanism. Thus, only *Rnf12* has been described as an activator of XCI shown to activate XCI in *trans* [82, 83]. The presence of female *Rnf12*[+/-] cells that did initiate XCI also indicated that one or more additional X-encoded XCI-activators are involved in initiation of XCI [51] (Figure 3B).
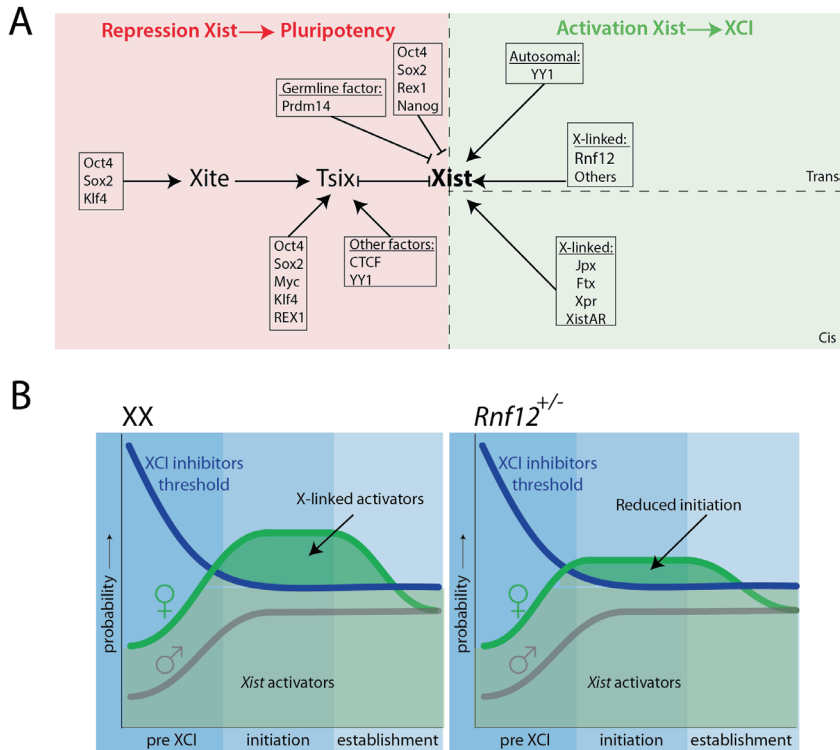
**Figure 3:** A) Molecular coupling of XCI to loss of pluripotency. Mainly pluripotency factors are inhibiting *Xist* expression and thus XCI initiation. Loss of pluripotency, thus loss of inhibiting factors repressing *Xist* in addition to activation of *Xist* through *Rnf12, Jpx, Ftx, Xpr, XistAR, Yy1* and other yet to be identified activator(s) would lead to initiation of XCI. B) Factors affecting threshold for XCI initiation (adapted from Barakat et al, 2014 [59]). Xist activators (light green shaded area) and XCI inhibitors (blue line) set the threshold for XCI to occur, whereas X-linked activators (dark green shaded area), present in double the amount in females (green line), allow females with two X chromosomes to overcome the threshold. Males (grey line) with one X chromosome will not overcome the threshold and will not initiate XCI. When one X chromosome, along with the X-linked activators is silenced in female cells, they will decrease to levels similar to males underneath the threshold and will not initiate XCI on the other X chromosome. Normal situation (left) in which females initiate XCI, because females have double the amount of X-linked activators to overcome the threshold for XCI to occur, whereas males do not. In *Rnf12*[+/-] females (right), a reduced initiation of XCI is observed. Since XCI is not completely abolished, this would point to the presence of other remaining, yet unidentified XCI activator(s).

### *Rnf12* and *Rex1:* a balance controlling XCI

*Rnf12* has been identified as a potent XCI-activator [83]. *Rnf12* encodes an E3 ubiquitin-ligase that activates *Xist* indirectly by targeting its repressor REX1 for proteasomal degradation [84]. REX1 is a transcription factor and pluripotency factor maintaining the pluripotent state in ES cells by suppression of genes involved in differentiation, which again links loss of pluripotency to XCI [85, 86]. REX1 mediated repression of XCI as mentioned before is two-way as it activates *Tsix* and represses *Xist*, the latter most likely by competing for binding sites with its ortholog and *Xist* activator *YY1* [80]. In male ES cells addition of one copy of *Rnf12* or removal of one copy of *Rex1* results in ectopic XCI [84]. REX1 binding sites were found in *Xist* and *Tsix* regulatory

regions and over-expression of REX1 led to inhibition of *Xist* expression. Genetic studies and reporter studies indicate that the prime regulatory target of REX1 is *Xist,* although previous studies indicated a role for REX1 in transcription elongation of *Tsix* [87]. The mechanism proposed is therefore that REX1 represses *Xist* possibly partially mediated through *Tsix*, and sets the threshold for XCI to occur along with other XCI inhibitors. Male cells have only one copy of *Rnf12*, and degradation of REX1 by RNF12 will not be sufficient to allow the threshold to drop below the level required for initiation of XCI to occur. Female cells on the other hand have two copies of *Rnf12* and upon differentiation *Rnf12* will be upregulated and RNF12 will target REX1 for degradation relieving the repression of *Xist* and promoting XCI by activating *Xist*. Auto-ubiquitination of RNF12 facilitates its high turnover, which together with its location in close proximity to *Xist* ensures that, in most cells, only one X chromosome will be inactivated as the RNF12 dose will decrease below the threshold. Inactivation of the XCI-activators lowers their nuclear concentration, which lowers the probability of the remaining X chromosome to initiate XCI, nevertheless by chance some cells will upregulate *Xist* on both X chromosomes [83, 84]. A recent study in an XX ES cell line, analyzing the effect of heterozygous deletion of both *Xist* and *Rnf12* on the same chromosome, indicated a severe XCI defect [59]. Loss of XCI was attributed to exclusive initiation of XCI on the wild type X chromosome, leading to *Rnf12* null cells with upregulated REX1, which blocks continued expression of *Xist*, underscoring the robustness of the feedback mechanisms involved in XCI. The role of RNF12 and REX1 in XCI is a nice example of the regulation of XCI by interplay between its inhibitors and activators.

**Establishment of XCI: locking and spreading of Xist along the Xi**

Initiation of XCI leads to *cis* spreading of *Xist* coating the 150Mb mouse X chromosome in order to silence it. How exactly *Xist* RNA spreads and coats the entire X chromosome is not completely understood. From previous studies we know that as *Xist* RNA spreads along the X chromosome, the basic transcriptional machinery becomes depleted and euchromatic marks such as histone 3 lysine 4 di-and tri- methylation (H3K4me2/me3) and histone H3 and histone H4 acetylation  (H3/H4 Ac) are lost [45, 46, 88]. After the loss of euchromatic marks, repressive marks accumulate such as histone 3 lysine 9 di-methylation (H3K9me2) [46, 47] and Polycomb complex 1 and 2 (PRC1 and PRC2) are recruited [89-92] catalyzing the monoubiquitylation of lysine 119 of histone H2A (H2A119ub) [91] and histone 3 lysine 27 tri-methylation (H3K27me3) [47], macroH2A is incorporated [93, 94] and finally CpG island methylation accumulates [95, 96], altogether contributing to the silencing process.

*Loading of Xist RNA onto the X chromosome in cis*

For *Xist* RNA to carry out its function, it needs to be loaded onto the X chromosome it is transcribed from to allow further spreading along the X in *cis*. How exactly this is achieved in such a tight manner, especially considering it is not affecting the other genetically identical X chromosome in *trans*, has been a long standing question in the field. It was proposed that *Xist* RNA might act trough a nucleation center, with high affinity binding sites for *Xist* RNA in close proximity to the Xic, facilitating spread

along the Xi in *cis* to other binding sites along the X chromosome [97, 98]. Recently, has been suggested that *Xist* RNA is tethered to the nucleation center on the Xi by YY1 to subsequently allow its further spreading in *cis*. YY1 is a bivalent protein, capable of binding both DNA and RNA. In XCI, YY1 binds *Xist* exon 1 via a trio of binding sites near conserved repeat F of *Xist* exon 1 and in turn repeat C and potentially repeat B of *Xist* RNA makes direct contact with YY1 [99]. Conversely, an earlier study generated a mutated *Xist* cDNA transgene, which happened to lack these YY1 binding sites, and found that spreading and silencing was not affected [100]. This study did indicate that regions mediating *Xist* RNA localization were scattered throughout *Xist* and were found to be functionally redundant. Therefore, it cannot be excluded that the presence of other low-affinity binding sites were bound by YY1 serving as a new nucleation center near the transgene to allow YY1 to tether *Xist* RNA to the X chromosome. Along with being tethered to the nucleation center to allow spreading in *cis*, the RNA and DNA binding matrix protein heterogenous nuclear protein U (hnRNPU; also known as SAF-A) might also be an important factor locking *Xist* RNA on the Xi. hnRNPU was found enriched on the Xi [101, 102] and depletion of hnRNPU results in detachment of *Xist* RNA from the Xi, diffusion of *Xist* RNA away from the Xi, and ES cells lacking hnRNPU fail to form an Xi [103]. HnRNPU recruitment is *Xist* dependent, but independent on gene silencing, as hnRNPU was still recruited to the Xi by an *Xist* lacking the A-repeat [104], indicating its involvement in locking *Xist* to the Xi, but not in silencing of the Xi.

*Xist RNA functioning exclusively in cis*
*Xist* RNA was allegedly non-diffusible and also unstable when not bound to the Xi and could therefore only act in *cis* [100, 105-107]. Interestingly, in the study by Jeon and Lee [99] and Hasegawa [108], *Xist* RNA appeared to diffuse. In the former study, transgenic *Xist* clouds also contained *Xist* particles from the cells own Xi clouds resulting in squelching of the endogenous clouds. *Xist* RNA might therefore not be strictly cis-acting as was long believed and does remain stable when not bound to chromatin. In contrast, the active X chromosome (Xa) nor the male X chromosome have ever been shown to be enriched for *Xist* RNA from the Xi or from transgenic *Xist* particles [65, 100, 105-107]. Also, previous work in which a tetraploid XXXX cell line was used with tagged *Xist,* showed exclusive *Xist* RNA spreading in *cis* in cells with two Xi's [105]. For this experiments ES cells were differentiated to analyze the effects after XCI, whereas in the study describing diffusion of *Xist* in *trans,* an inducible transgene was introduced *de novo* into post-XCI cells. In this way the transgene and autosome are not exposed to the usual developmental programming, which might explain the reported phenomenon of trans diffusion [99]. In *Drosophila melanogaster,* dosage compensation on the X chromosome in males involves spreading of *Rox1/2* RNA, which is dependent on a balance between expression of the *roX* RNA and the abundance of MSL protein components [109-111]. One could imagine that over-expression of an *Xist* transgene might deplete the pool of chromatin factors and hence facilitate spreading in *trans*. Taken together these observations postulate that *Xist* RNA is actually diffusible and stable, but this would never be observed *in vivo* as it is 'locked' into place by proteins including hnRNPU.

*Xist spreading and LINE-1 repeats*

Now that we have an idea on how *Xist* RNA is loaded onto the Xi and only attaches to the X chromosome it is transcribed from, we look further into how it actually spreads and completely coats a 150Mb chromosome with only a few escaping genes as exceptions. The long-standing hypothesis was that *Xist* RNA is able to spread through way stations on the X chromosome [112-114]. LINE-1 repeats were suggested to serve as way-stations for *Xist* RNA spreading considering they are enriched on the X chromosome over autosomes [113], their density was correlated with the low efficiency of XCI spreading on autosomes [106, 115, 116] and escaping genes were found to be LINE-1 poor in human [117]. Recently, evidence is accumulating against this hypothesis. Inefficient spreading and silencing of an autosome might be the result of a lethal monosomy if *Xist* RNA would silence an autosome. This is evident from efficient *XIST* RNA spreading and silencing of one of the trisomy 21 chromosomes of a Down syndrome hiPS cell [118]. In addition three recent studies show that LINE-1 repeats were not enriched over input in a *Xist* CHART-seq [119], no significant relationship between *Xist* localization and LINE1 repeats was found using an RNA antisense purification (RAP) technology [120] and finally, when looking at a spatial relationship, LINE-1 dense regions were more frequently found exterior to H3K27me3 and *Xist* RNA domains [121]. Taken together, these observations point to a mechanism other than LINE-1 elements as way stations for spreading of *Xist* RNA .

Even though LINE-1 repeats might not be involved in the direct spreading of *Xist* RNA, these repeats could still have other important roles in XCI [122]. Two classes of LINEs do seem to function in XCI. The first are LINEs that are silenced upon *Xist* RNA coating, prior to silencing of other X-linked genes. These LINEs are thought to facilitate assembly of a heterochromatic nuclear compartment into which genes can be recruited and silenced. The other class of LINEs are expressed on both X chromosomes prior to XCI. After differentiation, these LINEs are exclusively expressed from the Xi and silenced on the Xa, and thought to facilitate the spread of silencing into regions on the X chromosome that are difficult to silence through the regular spreading mechanism, such as regions in close proximity to escaping genes.

*Models for spreading of Xist RNA along the X chromosome*

If *Xist* RNA does not spread by means of LINE-1 way stations or other specific binding sites which could act as way stations, how does it spread along an entire X chromosome? High-resolution maps of *Xist* binding on the X chromosome generated using allele-specific CHART-seq, suggested a two-step model for spreading of *Xist* [119]. In cells undergoing XCI, *Xist* RNA targets gene-rich domains first, and the remaining gene-poor domains thereafter [119, 120]. However it was found that gene density alone did not explain early *Xist* RNA localization patterns, as chromosome wide correlation between *Xist* localization and gene density was relatively modest and therefore a proximity transfer model was hypothesized instead [120]. In this model the spatial organization of the X chromosome is instructive in *Xist* RNA targeting gene-rich regions in close proximity to the *Xist* locus first, and subsequently spreading further into gene-poor regions afterwards. This proximity hypothesis was validated in a male cell line with an inducible *Xist* transgene integrated into the *Hprt* locus (±50Mb proximal to *Xist*), where

*Xist* RNA localization strongly correlated with proximity contacts at the *Hprt* integration site but not with the endogenous *Xist* localization sites. This study also observed that most genes targeted by *Xist* RNA are not active in ES cells and targeting initially occurs independent of the conserved A repeat of *Xist*. After spreading to the inactive gene-rich regions, *Xist*, now dependent on a functional A repeat, is able to spread and silence the active genes as well. Additionally, Marks et al [123] used high-resolution allele specific RNA-seq to study the dynamics of gene silencing on the Xi. Genes proximal to the *Xist* were found to be silenced prior to distal genes and lowly expressed genes also showed faster XCI dynamics than highly expressed genes [123]. Furthermore, it was hypothesized that X-linked TADs might function as modular domain structures initially targeted by *Xist* RNA spreading, and that secondary spread of XCI might occur within these TADs. These studies point to a stepwise model for the spreading of *Xist* RNA and suggest LINE-1 repeats might be less important in *Xist* RNA spreading than anticipated. *Xist* RNA spreads to (inactive) gene-rich regions first, and spreading along the rest of the chromosome follows from thereon. Whether it is independent of sequence specificity and most likely due to proximity transfer [120] or that TADs might be involved in *Xist* RNA spreading remain intriguing questions to be explored.

*Silencing the X, novel Xist interactors*

After *Xist* has spread along the X chromosome it needs to ensure silencing of the genes coated by *Xist* RNA. One of the most studied relationship between *Xist* and recruitment of silencing complexes to silence chromosome-wide are *Xist* and PRC2. PRC2 is composed of SUZ12, EED, RBBP4/7 and EZH2 of which EZH2 convers histone methyltransferase activity, catalyzing tri-methylation of histone H3 at lysine 27 [124, 125]. Recruitment of PRC2 to the Xi is dependent on *Xist* RNA [92] and to some extent on repeat A of *Xist* [89, 126, 127]. *Xist* and PRC2 seem to be functionally tethered when observing *Xist* clouds at 20-nm resolution using STORM [128], and "identification of Direct RNA interacting Proteins" (iDRiP) [129] identified PRC2 as an *Xist* interacting partner, suggesting direct recruitment of PRC2 by *Xist*. However, significant spatial separation and no co-localization between PRC2 proteins and *Xist* RNA respectively was observed with 3D-SIM [130, 131], and PRC2 binding to *Xist* [132] was not confirmed in two other studies applying alternative technologies (RNA Antisense Purification-Mass Spectrometry (RAP-MS) and Comprehensive Identification of RNA-binding Proteins by Mass Spectrometry (ChIRP-MS)) to identify *Xist* interacting proteins [133, 134]. Although this discrepancy might have resulted from application of different techniques, averaging for whole populations as compared to single cells and using male versus female cells, these findings may indicate that PRC2 recruitment is indirect. This could involve PRC2 mediated recruitment by JARID2 [135] and ATRX [136]. Loss of *Jarid2*, normally interacting with *Xist* repeat B and F, results in inefficient recruitment of PRC2 and loss of H3K27me3 enrichment on the Xi, despite proper *Xist* RNA spreading [135]. Loss of ATRX, which is enriched on the Xi, also leads to an XCI defect [136, 137]. ATRX promotes binding of PRC2 to repeat A of *Xist* RNA by forming a ternary complex of ATRX, PRC2 and *Xist* RNA [136]. Recent work also indicates that PRC1 mediated accumulation of H2AK119ub is sufficient for PRC2 recruitment [138], and interestingly, components of the PRC1 complex were detected in the two *Xist*-interactor screens [132, 139]. These findings indicate that

PRC2 recruitment might be secondary following initial recruitment of PRC1 by *Xist*, but also suggests that multiple mechanisms are involved tethering PRC2 to the X.

The RAP-MS, ChIRP and iDRiP identified many other interesting *Xist* RNA binding proteins [129, 132, 133]. Spen, was among the 9 proteins identified in all three studies, and its involvement in XCI was further confirmed in two independent studies applying forward genetic [140] and pooled shRNA [141] screens. *Spen* (also known as *Sharp*) is a novel silencing factor required for and recruited by *Xist* RNA to silence the Xi [132, 133, 140, 141]. Spen was initially identified as an RNA binding protein involved in the recruitment of the SMRT/NcoR co-repressor complex [142]. Several Spen co-factors were also identified in different screens including, HDAC3, SMRT and RBM15, and knockdown or knockout of all of these factors resulted in XCI phenotypes. *Spen* interacts with *Xist* via the A repeat [132, 140] in ESC but this interaction is intensified in differentiated cells [132]. The exact mechanism through which Spen acts in establishment of the Xi is not clear, but most likely involves deposition of histone deacetylase activity through recruitment of HDAC3. Several studies indicate that active histone marks antagonize polycomb group complex targeting (reviewed in [143]), and therefore loss of histone acytelation might pave the way for PRC1 and PRC2 to be able to catalyze H2A119ub and H3K27me3, further installing the silent state. Besides Spen and its interactors several more interesting *Xist* interacting proteins were identified with these screens. These include the m6A methyltransferase WTAP and HnrnpK, which are both required for XCI, and it will be worth looking into more detail at possible roles for these proteins in XCI. A summary of Xist binding proteins and the regulatory network at the Xic is shown in Figure 4.
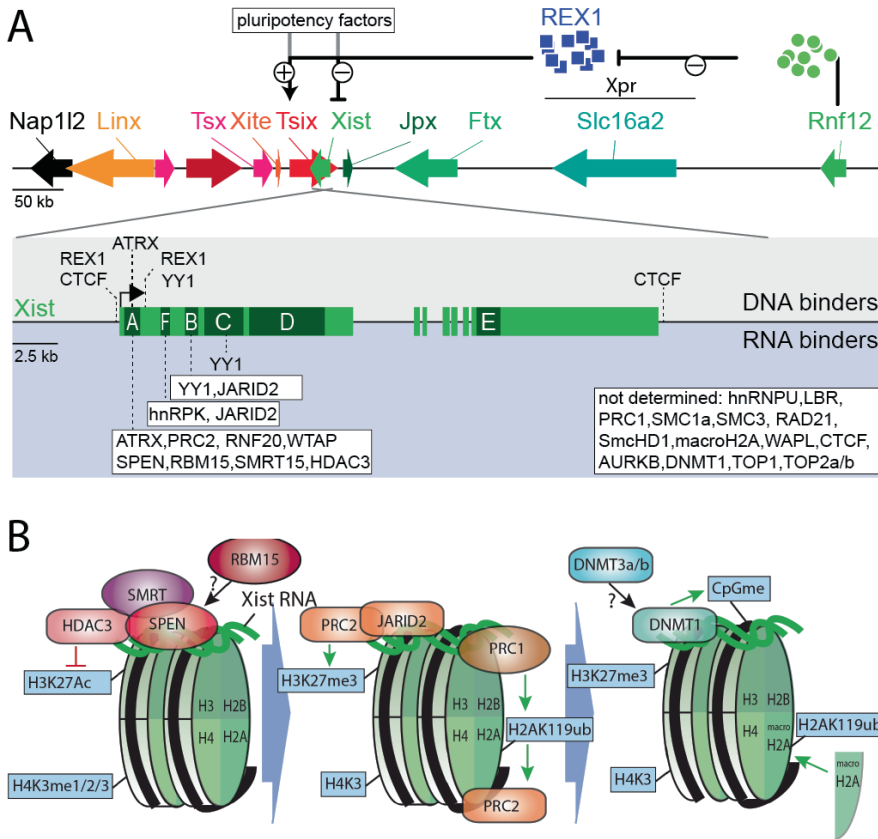
**Figure 4:** A)The Xic and *Xist* binding proteins. Shown are different genes located within the Xic and the action sites of the trans acting factors (top). Also indicated are the binding sites of *Xist* regulatory and nucleation factors (DNA binders), and *Xist* RNA-binding proteins involved in establishment of the Xi (RNA binders). B) Model of the sequence of events leading to silencing of the X chromosome. Xist RNA interacts with SPEN and RBM15. SPEN recruits SMRT, resulting in recruitment and/or activation of HDAC3 and deacetylation of H3K27. This allows for PRC2 and PRC1 recruitment which respectively methylate H3K27 and ubiquitylate H2AK119. Eventually,macroH2A is recruited to the X, together with the DNA methylation machinery. Figure adapted from [144] and [145].

*Barr body formation*

The final stage in establishment of the Xi is the formation of the Barr body [20], but how and when exactly during XCI this chromatin compaction towards a Barr body occurs was still unknown because of limiting resolution with the standard microscopy techniques. Now using 3D-SIM and STORM with better resolution, the structural and spatial arrangement of the Barr body in the nucleus can be observed in greater detail. *Xist* RNA is organized as distinct foci distributed throughout the Barr body with less than 100 foci, each consisting of about 3 (up to 10) transcripts *Xist* RNA [128, 131]. The Barr body maintains principle ultra-structural features of a general chromosome territory (CT) architecture, only in the Barr body the active nuclear compartment (ANC) appears to be collapsed and preferential localization of *Xist* RNA foci was observed

within and at boundaries of these collapsed ANC channels [131]. The nuclear matrix protein hnRNP U/SAF-A [101, 108] was enriched within the Barr body and found in close spatial proximity to *Xist* RNA [131]. Multimer formation of hnRNP U/SAF-A, triggered by *Xist* RNA, could create a scaffold that helps maintain the integrity of the Barr body structure [131]. Interestingly, hnRNP U/SAF-A was identified in all three RAP-MS, ChIRP and iDRiP screens as interaction partner of *Xist*, in support of a crucial role for this protein in XCI [129, 132, 133]. In addition, RAP-MS, and iDRiP also identified Lamin-B receptor [129, 133], which might play a role in tethering the Barr body near the nuclear periphery.

Time resolved 3D-SIM on differentiating ESC revealed initial spreading of *Xist* RNA foci around day 3 of differentiation into a de-condensed chromatin environment, by day 4 focal *Xist* RNA spreading and RNA PolII exclusion was observed, but no evidence for chromatin compaction. Distinct global compaction of a *Xist* RNA defined Xi territory towards a Barr body was evident at day 5. Exclusion of RNA PolII and removal of active marks could result in silenced genes quickly undergoing chromatin condensation at this stage. Distinctive marking of the Barr body was observed in most cells after day 7 and a higher level of chromatin compaction in the Barr body was observed at day 9 and might represent a stable stage of chromatin arrangement in the Xi mediated by late stage repressive signatures such as DNA methylation. One interesting observation worth mentioning is that autosomal Barr body formation, obtained by inducible *Xist* transgenes [65], clearly differ from the female Xi and resemble more active chromatin up to day 10 of differentiation [131] potentially explaining discrepancies in different studies.

Taken together, studies on the establishment of the Xi are rapidly unravelling how *Xist* is loaded onto the X chromosome, how it spreads along the Xi in cis and how it manages to silence chromosome-wide and finally form a highly condensed Barr body. Using the proteins found in the unbiased proteomic screens in functional analysis can help identify the missing players in the establishment of the Xi.

**Maintenance of XCI**

Once the Xi has been established and a Barr body is formed, the Xi is maintained in an inactive state and clonally propagated throughout cell divisions. Even though maintenance of the Xi seems to be very stringent and maintained throughout cell divisions, escaping genes seems to vary between tissues [146] which could indicate tissue specific maintenance. Genes escaping XCI include genes in the pseudo-autosomal regions (PAR) and individual genes escaping XCI called escapers, with or without a Y homolog. Genes in the PAR, the region homologous to the Y chromosome, escape XCI but are not differentially expressed between the sexes as males also have two copies of genes located in this region. Escapers that have a Y homolog are also present in two copies in both males and females. It is only the escapers without a functional Y homolog which are differentially expressed between the sexes, thus absolute equalization of gene dosage between the sexes is not the case [7].

The maintenance phase is the least studied phase of the rXCI process and therefore a lot still remains to be explored. In a study using MEFs harboring a GFP transgene on the X chromosome and exposing these cells to a mouse shRNA library, factors involved in maintenance of the Xi could be identified as a result of GFP reactivation [147]. 32 proteins involved in maintenance of the Xi were identified in this study, of which only DNMT1 was also identified in one of the three *Xist* interactor screens [129]. One of these studies involving the identification of Xist-interactors also tested the effects of knockdown of specific genes, revealing a role for cohesion factors and topoisomerase in maintenance of the Xi. These findings might indicate that different factors act in establishment and maintenance of the Xi, but might also reflect differences in the technical setup. It will clearly be worth looking into more detail at to unravel the unknowns of this last phase of rXCI.

## XCI in different species: choosing the right model

XCI occurs in all therian mammals, thus marsupials and eutherians. Even though the end result of XCI, dosage compensation by inactivating one X chromosome in female cells, is the same, two major differences between these two clades are evident: marsupial XCI involves genomic imprinting, always inactivating the paternal X chromosome (Xp), whereas XCI in eutherians is mostly random with respect to parental origin of the X chromosome, only showing iXCI in extraembryonic tissues in mouse [148, 149], rat [150], common voles [151] and cow [152]. In addition, marsupial XCI involves a lncRNA *Rsx* [153], whereas eutherian XCI involves the lncRNA *Xist. Rsx* and *Xist* are non-homologous and have a different origin but share several common features. Both lncRNAs are large, repeat-rich, spliced and polyadenylated non-coding RNAs only expressed in female cells. Also, both contain a high GC content and are highly enriched in tandem repeats which are biased towards the 5'end [153]. Furthermore, *Xist* and Rsx are exclusively transcribed from and coat the Xi resulting in formation of distinct RNA clouds which can be visualized by RNA-FISH. Marsupial XCI and eutherian XCI have evolved separately, which explains the differences between the two forms of XCI, even though both forms achieve inactivation of one X chromosome in female cells. Nevertheless, it seems that there are differences in XCI, especially regarding timing and escaping genes across eutherians as well.

Most eutherian studies show rXCI in the embryo and adult or somatic tissues and *Xist* is conserved between them, yet some differences in initiation, form of XCI in extraembryonic tissues and the actual process of XCI have been reported for different eutherian species. Therefore, one of the major debates in the field is whether the mouse is a good model to study XCI as it appears that not everything can be extrapolated to other species as well. So far, as mentioned before, iXCI has only been reported in mouse, rat, common voles, cow and some contradicting reports in human as well [154-158]. Other species such as rabbit [159], horse [160], mule [160], monkey [161] show rXCI in their extra embryonic tissues. In addition, it seems that the mouse is the only species which has an *Xist* gene with a completely overlapping *Tsix* gene, whereas it is truncated before reaching the Xist promoter in other species [162]. Other major differences are the timing of XCI initiation and silencing. Some of these differences in timing are a result of differences in development of the embryo of different species. In mouse female embryos *Xist* is upregulated and spreads to inactivate the X chromosome

directly, whereas in human and monkey female embryos, *XIST* can be detected around the blastocyst stage on both X chromosome without signs of gene silencing. In culture there are even more differences and hES and hIPS can be found in different states with respect to XCI [163]. Most of these and other discrepancies between mouse/human XCI and other species seems to lie in the ability to obtain pluripotent cells and maintain them in the naïve state when analyzing them in culture.

Mouse embryonic stem (ES) cells are the most frequent used ex vivo model system for XCI studies as these pluripotent stem cells are derived from the ICM and initiate random XCI upon differentiation in female cells but not in male cells [18]. Ultimately, we need to consider: are mice really recapitulating human XCI? So far, studies of XCI in mouse have been very important in translation to X-linked diseases. In addition ES cells can be genetically manipulated and subsequently used to generate genetically manipulated mice. Mice are therefore the main model organism used for XCI, as XCI can be studied both in *vitro* and in *vivo* [19]. In contrast, XCI in early human development has not been studied as intensively because of ethical and other restrictions in culture. Thus while mouse ES cells can recapitulate XCI regulation in vitro, XCI in hES and hIPS cells shows variations and switches from one state to the other in culture, making it much more difficult to study the XCI process, requiring more knowledge about human XCI in *vivo*.

In summary, mouse XCI seems to be more complete, having least escaping genes and a very tightly regulated XCI in place as compared to other eutherian species. Understanding the main differences in XCI between species and the different degrees of silencing resulting from these differences can help us understand silencing and skewing which can be used to target specific human diseases in the future.

**References**

1. Vickaryous, M.K. and McLean, K.E. (2011) Reptile embryology. Methods Mol Biol 770, 439-55.

2. Barske, L.A. and Capel, B. (2008) Blurring the edges in vertebrate sex determination. Curr Opin Genet Dev 18 (6), 499-505.

3. Veitia, R.A. and Potier, M.C. (2015) Gene dosage imbalances: action, reaction, and models. Trends Biochem Sci 40 (6), 309-17.

4. Veitia, R.A. et al. (2015) X chromosome inactivation and active X upregulation in therian mammals: facts, questions, and hypotheses. J Mol Cell Biol 7 (1), 2-11.

5. Charlesworth, B. (1991) The evolution of sex chromosomes. Science 251 (4997), 1030-3.

6. Prothero, K.E. et al. (2009) Dosage compensation and gene expression on the mammalian X chromosome: one plus one does not always equal two. Chromosome Res 17 (5), 637-48.

7. Skaletsky, H. et al. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423 (6942), 825-37.

8. Mank, J.E. (2009) The W, X, Y and Z of sex-chromosome dosage compensation. Trends Genet 25 (5), 226-33.

9. Graves, J.A. et al. (2002) Evolution of the human X--a smart and sexy chromosome that controls speciation and development. Cytogenet Genome Res 99 (1-4), 141-5.

10. Graves, J.A. (2002) Evolution of the testis-determining gene--the rise and fall of SRY. Novartis Found Symp 244, 86-97; discussion 97-101, 203-6, 253-7.

11. Lahn, B.T. and Page, D.C. (1997) Functional coherence of the human Y chromosome. Science 278 (5338), 675-80.

12. Pontier, D.B. and Gribnau, J. (2011) Xist regulation and function explored. Hum Genet 130 (2), 223-36.

13. Meyer, B.J. et al. (2004) Sex and X-chromosome-wide repression in Caenorhabditis elegans. Cold Spring Harb Symp Quant Biol 69, 71-9.

14. Straub, T. and Becker, P.B. (2011) Transcription modulation chromosome-wide: universal features and principles of dosage compensation in worms and flies. Curr Opin Genet Dev 21 (2), 147-53.

15. Gelbart, M.E. and Kuroda, M.I. (2009) Drosophila dosage compensation: a complex voyage to the X chromosome. Development 136 (9), 1399-410.

16. Gelbart, M.E. et al. (2009) Drosophila MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. Nat Struct Mol Biol 16 (8), 825-32.

17. Lyon, M.F. (1961) Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature 190, 372-3.

18. Harper, P.S. (2011) Mary Lyon and the hypothesis of random X chromosome inactivation. Hum Genet 130 (2), 169-74.

19. LYON, M.F. (1962) Sex chromatin and gene action in the mammalian X-chromosome. Am J Hum Genet 14, 135-48.

20. BARR, M.L. and BERTRAM, E.G. (1949) A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. Nature 163 (4148), 676.

21. Ohno, S. (1967) Sex Chromosomes and Sex-Linked Genes, Springer.

22. OHNO, S. et al. (1959) Formation of the sex chromatin by a single X-chromosome in liver cells of Rattus norvegicus. Exp Cell Res 18, 415-8.

23. POLANI, P.E. et al. (1954) Chromosomal sex in Turner's syndrome with coarctation of the aorta. Lancet 267 (6829), 120-1.

24. POLANI, P.E. et al. (1956) Colour-blindness in ovarian agenesis (gonadal dysplasia). Lancet 271 (6934), 118-20.

25. BEUTLER, E. et al. (1962) The normal human female as a mosaic of X-chromosome activity: studies using the gene for C-6-PD-deficiency as a marker. Proc Natl Acad Sci U S A 48, 9-16.

26. Nguyen, D.K. and Disteche, C.M. (2006) Dosage compensation of the active X chromosome in mammals. Nat Genet 38 (1), 47-53.

27. Yildirim, E. et al. (2012) X-chromosome hyperactivation in mammals via nonlinear relationships between chromatin states and transcription. Nat Struct Mol Biol 19 (1), 56-61.

28. Adler, D.A. et al. (1997) Evidence of evolutionary up-regulation of the single active X chromosome in mammals based on Clc4 expression levels in Mus spretus and Mus musculus. Proc Natl Acad Sci U S A 94 (17), 9244-8.

29. Lin, H. et al. (2011) Relative overexpression of X-linked genes in mouse embryonic stem cells is consistent with Ohno's hypothesis. Nat Genet 43 (12), 1169-70; author reply 1171-2.

30. Marks, H. et al. (2015) Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. Genome Biol 16, 149.

31. Xiong, Y. et al. (2010) RNA sequencing shows no dosage compensation of the active X-chromosome. Nat Genet 42 (12), 1043-7.

32. Lin, F. et al. (2012) Expression reduction in mammalian X chromosome evolution refutes Ohno's hypothesis of dosage compensation. Proc Natl Acad Sci U S A 109 (29), 11752-7.

33. Julien, P. et al. (2012) Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. PLoS Biol 10 (5), e1001328.

34. Pessia, E. et al. (2012) Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. Proc Natl Acad Sci U S A 109 (14), 5346-51.

35. Chen, X. and Zhang, J. (2015) No X-chromosome dosage compensation in human proteomes. Mol Biol Evol 32 (6), 1456-60.

36. Jue, N.K. et al. (2013) Determination of dosage compensation of the mammalian X chromosome by RNA-seq is dependent on analytical approach. BMC Genomics 14, 150.

37. Castagné, R. et al. (2011) The choice of the filtering method in microarrays affects the inference regarding dosage compensation of the active X-chromosome. PLoS One 6 (9), e23956.

38. Huynh, K.D. and Lee, J.T. (2003) Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. Nature 426 (6968), 857-862.

39. Patrat, C. and Corbel, C. (2009) Regulation of the inactivation of the X chromosome during early development. Biofutur (304), 24-27.

40. Mak, W. et al. (2004) Reactivation of the paternal X chromosome in early mouse embryos. Science 303 (5658), 666-669.

41. Okamoto, I. et al. (2004) Epigenetic dynamics of imprinted X inactivation during early mouse development. Science 303 (5658), 644-649.

42. Harper, M.I. et al. (1982) Preferential Paternal X-Inactivation in Extra-Embryonic Tissues of Early Mouse Embryos. Journal of Embryology and Experimental Morphology 67 (Feb), 127-135.

43. Barakat, T.S. et al. (2010) X-changing information on X inactivation. Experimental Cell Research 316 (5), 679-687.

44. Plath, K. et al. (2003) Role of histone H3 lysine 27 methylation in X inactivation. Science 300 (5616), 131-135.

45. O'Neill, L.P. et al. (2008) Differential loss of histone H3 isoforms mono-, di- and

tri-methylated at lysine 4 during X-inactivation in female embryonic stem cells. Biol Chem 389 (4), 365-70.

46. Heard, E. et al. (2001) Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. Cell 107 (6), 727-38.

47. Rougeulle, C. et al. (2004) Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. Mol Cell Biol 24 (12), 5475-84.

48. Rastan, S. (1983) Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. J Embryol Exp Morphol 78, 1-22.

49. Nicodemi, M. and Prisco, A. (2007) Symmetry-breaking model for X-chromosome inactivation. Phys Rev Lett 98 (10), 108104.

50. Mlynarczyk-Evans, S. et al. (2006) X chromosomes alternate between two states prior to random X-inactivation. PLoS Biol 4 (6), e159.

51. Wutz, A. and Gribnau, J. (2007) X inactivation Xplained. Curr Opin Genet Dev 17 (5), 387-93.

52. Bacher, C.P. et al. (2006) Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat Cell Biol 8 (3), 293-9.

53. Augui, S. et al. (2007) Sensing X chromosome pairs before X inactivation via a novel X-pairing region of the Xic. Science 318 (5856), 1632-6.

54. Xu, N. et al. (2006) Transient homologous chromosome pairing marks the onset of X inactivation. Science 311 (5764), 1149-52.

55. Masui, O. et al. (2011) Live-cell chromosome dynamics and outcome of X chromosome pairing events during ES cell differentiation. Cell 145 (3), 447-58.

56. Kung, J.T. et al. (2015) Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. Mol Cell 57 (2), 361-75.

57. Xu, N. et al. (2007) Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. Nat Genet 39 (11), 1390-6.

58. Yang, F. et al. (2015) The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. Genome Biol 16, 52.

59. Barakat, T.S. et al. (2014) The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. Mol Cell 53 (6), 965-78.

60. Monkhorst, K. et al. (2008) X inactivation counting and choice is a stochastic

process: evidence for involvement of an X-linked activator. Cell 132 (3), 410-21.

61. Gribnau, J. and Grootegoed, J.A. (2012) Origin and evolution of X chromosome inactivation. Curr Opin Cell Biol 24 (3), 397-404.

62. Plath, K. et al. (2002) Xist RNA and the mechanism of X chromosome inactivation. Annu Rev Genet 36, 233-78.

63. Brown, C.J. et al. (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature 349 (6304), 38-44.

64. Spatz, A. et al. (2004) X-chromosome genetics and human cancer. Nat Rev Cancer 4 (8), 617-29.

65. Wutz, A. and Jaenisch, R. (2000) A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol Cell 5 (4), 695-705.

66. Sun, S. et al. (2010) Characterization of Xpr (Xpct) reveals instability but no effects on X-chromosome pairing or Xist expression. Transcription 1 (1), 46-56.

67. Lee, J.T. et al. (1999) Tsix, a gene antisense to Xist at the X-inactivation centre. Nat Genet 21 (4), 400-4.

68. Luikenhuis, S. et al. (2001) Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. Mol Cell Biol 21 (24), 8512-20.

69. Ogawa, Y. et al. (2008) Intersection of the RNA interference and X-inactivation pathways. Science 320 (5881), 1336-41.

70. Chureau, C. et al. (2011) Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. Hum Mol Genet 20 (4), 705-18.

71. Tian, D. et al. (2010) The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. Cell 143 (3), 390-403.

72. Sarkar, M.K. et al. (2015) An Xist-activating antisense RNA required for X-chromosome inactivation. Nat Commun 6, 8564.

73. Navarro, P. et al. (2008) Molecular coupling of Xist regulation and pluripotency. Science 321 (5896), 1693-5.

74. Navarro, P. and Avner, P. (2009) When X-inactivation meets pluripotency: an intimate rendezvous. FEBS Lett 583 (11), 1721-7.

75. Payer, B. et al. (2013) Tsix RNA and the germline factor, PRDM14, link X reactivation and stem cell reprogramming. Mol Cell 52 (6), 805-18.

76. Navarro, P. et al. (2010) Molecular coupling of Tsix regulation and pluripotency. Nature 468 (7322), 457-60.

77. Donohoe, M.E. et al. (2007) Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. Mol Cell 25 (1), 43-56.

78. Chelmicki, T. et al. (2014) MOF-associated complexes ensure stem cell identity and Xist repression. Elife 3, e02024.

79. Nechanitzky, R. et al. (2012) Satb1 and Satb2 are dispensable for X chromosome inactivation in mice. Dev Cell 23 (4), 866-71.

80. Makhlouf, M. et al. (2014) A prominent and conserved role for YY1 in Xist transcriptional activation. Nat Commun 5, 4878.

81. Shin, J. et al. (2010) Maternal Rnf12/RLIM is required for imprinted X-chromosome inactivation in mice. Nature 467 (7318), 977-81.

82. Barakat, T.S. et al. (2011) RNF12 activates Xist and is essential for X chromosome inactivation. PLoS Genet 7 (1), e1002001.

83. Jonkers, I. et al. (2009) RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. Cell 139 (5), 999-1011.

84. Gontan, C. et al. (2012) RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature 485 (7398), 386-90.

85. Hosler, B.A. et al. (1989) Expression of REX-1, a gene containing zinc finger motifs, is rapidly reduced by retinoic acid in F9 teratocarcinoma cells. Mol Cell Biol 9 (12), 5623-9.

86. Scotland, K.B. et al. (2009) Analysis of Rex1 (zfp42) function in embryonic stem cell differentiation. Dev Dyn 238 (8), 1863-77.

87. Navarro, P. and Avner, P. (2010) An embryonic story: analysis of the gene regulative network controlling Xist expression in mouse embryonic stem cells. Bioessays 32 (7), 581-8.

88. Goto, Y. et al. (2002) Differential patterns of histone methylation and acetylation distinguish active and repressed alleles at X-linked genes. Cytogenet Genome Res 99 (1-4), 66-74.

89. Zhao, J. et al. (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 322 (5902), 750-6.

90. Plath, K. et al. (2004) Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. J Cell Biol 167 (6), 1025-35.

91. de Napoles, M. et al. (2004) Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev Cell 7 (5), 663-76.

92. Schoeftner, S. et al. (2006) Recruitment of PRC1 function at the initiation of X

inactivation independent of PRC2 and silencing. EMBO J 25 (13), 3110-22.

93. Costanzi, C. et al. (2000) Histone macroH2A1 is concentrated in the inactive X chromosome of female preimplantation mouse embryos. Development 127 (11), 2283-9.

94. Mietton, F. et al. (2009) Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome. Mol Cell Biol 29 (1), 150-6.

95. Grant, M. et al. (1992) Methylation of CpG sites of two X-linked genes coincides with X-inactivation in the female mouse embryo but not in the germ line. Nat Genet 2 (2), 161-6.

96. Kaslow, D.C. and Migeon, B.R. (1987) DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation. Proc Natl Acad Sci U S A 84 (17), 6210-4.

97. Riggs, A.D. et al. (1985) Methylation of the PGK promoter region and an enhancer way-station model for X-chromosome inactivation. Prog Clin Biol Res 198, 211-22.

98. McBurney, M.W. (1988) X chromosome inactivation: a hypothesis. Bioessays 9 (2-3), 85-8.

99. Jeon, Y. and Lee, J.T. (2011) YY1 tethers Xist RNA to the inactive X nucleation center. Cell 146 (1), 119-33.

100. Wutz, A. et al. (2002) Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat Genet 30 (2), 167-74.

101. Helbig, R. and Fackelmayer, F.O. (2003) Scaffold attachment factor A (SAF-A) is concentrated in inactive X chromosome territories through its RGG domain. Chromosoma 112 (4), 173-82.

102. Fackelmayer, F.O. (2005) A stable proteinaceous structure in the territory of inactive X chromosomes. J Biol Chem 280 (3), 1720-3.

103. Hasegawa, Y. et al. (2010) The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. Dev Cell 19 (3), 469-76.

104. Pullirsch, D. et al. (2010) The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. Development 137 (6), 935-43.

105. Jonkers, I. et al. (2008) Xist RNA is confined to the nuclear territory of the silenced X chromosome throughout the cell cycle. Molecular and Cellular Biology 28 (18), 5583-94.

106. Duthie, S.M. et al. (1999) Xist RNA exhibits a banded localization on the inactive

X chromosome and is excluded from autosomal material in cis. Hum Mol Genet 8 (2), 195-204.

107. Lee, J.T. et al. (1999) Genetic analysis of the mouse X inactivation center defines an 80-kb multifunction domain. Proc Natl Acad Sci U S A 96 (7), 3836-41.

108. Hasegawa, Y. et al. (2010) The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. Dev Cell 19 (3), 469-76.

109. Oh, H. et al. (2003) Local spreading of MSL complexes from roX genes on the Drosophila X chromosome. Genes Dev 17 (11), 1334-9.

110. Kelley, R.L. et al. (2008) Transcription rate of noncoding roX1 RNA controls local spreading of the Drosophila MSL chromatin remodeling complex. Mech Dev 125 (11-12), 1009-19.

111. Park, Y. et al. (2002) Extent of chromatin spreading determined by roX RNA recruitment of MSL proteins. Science 298 (5598), 1620-3.

112. Gartler, S.M. and Riggs, A.D. (1983) Mammalian X-chromosome inactivation. Annu Rev Genet 17, 155-90.

113. Boyle, A.L. et al. (1990) Differential distribution of long and short interspersed element sequences in the mouse genome: chromosome karyotyping by fluorescence in situ hybridization. Proc Natl Acad Sci U S A 87 (19), 7757-61.

114. Bala Tannan, N. et al. (2014) DNA methylation profiling in X;autosome translocations supports a role for L1 repeats in the spread of X chromosome inactivation. Hum Mol Genet 23 (5), 1224-36.

115. Popova, B.C. et al. (2006) Attenuated spread of X-inactivation in an X;autosome translocation. Proc Natl Acad Sci U S A 103 (20), 7706-11.

116. Keohane, A.M. et al. (1999) H4 acetylation, XIST RNA and replication timing are coincident and define x;autosome boundaries in two abnormal X chromosomes. Hum Mol Genet 8 (2), 377-83.

117. Bailey, J.A. et al. (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc Natl Acad Sci U S A 97 (12), 6634-9.

118. Jiang, J. et al. (2013) Translating dosage compensation to trisomy 21. Nature 500 (7462), 296-300.

119. Simon, M.D. et al. (2013) High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. Nature 504 (7480), 465-9.

120. Engreitz, J.M. et al. (2013) The Xist lncRNA exploits three-dimensional genome

architecture to spread across the X chromosome. Science 341 (6147), 1237973.

121. Calabrese, J.M. et al. (2012) Site-specific silencing of regulatory elements as a mechanism of X inactivation. Cell 151 (5), 951-63.

122. Chow, J.C. et al. (2010) LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. Cell 141 (6), 956-69.

123. Marks, H. et al. (2015) Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. Genome Biol 16 (1), 149.

124. Muller, J. et al. (2002) Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell 111 (2), 197-208.

125. Cao, R. et al. (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298 (5595), 1039-43.

126. Kohlmaier, A. et al. (2004) A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. PLoS Biol 2 (7), E171.

127. Maenner, S. et al. (2010) 2-D structure of the A region of Xist RNA and its implication for PRC2 association. PLoS Biol 8 (1), e1000276.

128. Sunwoo, H. et al. (2015) The Xist RNA-PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. Proc Natl Acad Sci U S A 112 (31), E4216-25.

129. Minajigi, A. et al. (2015) Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. Science 349 (6245).

130. Cerase, A. et al. (2014) Spatial separation of Xist RNA and polycomb proteins revealed by superresolution microscopy. Proc Natl Acad Sci U S A 111 (6), 2235-40.

131. Smeets, D. et al. (2014) Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. Epigenetics Chromatin 7, 8.

132. Chu, C. et al. (2015) Systematic discovery of Xist RNA binding proteins. Cell 161 (2), 404-16.

133. McHugh, C.A. et al. (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature 521 (7551), 232-6.

134. Chu, C. et al. (2015) Systematic Discovery of Xist RNA Binding Proteins. Cell 161 (2), 404-416.

135. da Rocha, S.T. et al. (2014) Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. Mol Cell 53 (2), 301-16.

136. Sarma, K. et al. (2014) ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. Cell 159 (4), 869-83.

137. Baumann, C. and De La Fuente, R. (2009) ATRX marks the inactive X chromosome (Xi) in somatic cells and during imprinted X chromosome inactivation in trophoblast stem cells. Chromosoma 118 (2), 209-22.

138. Cooper, S. et al. (2014) Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment. Cell Rep 7 (5), 1456-70.

139. Blackledge, N.P. et al. (2014) Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. Cell 157 (6), 1445-59.

140. Monfort, A. et al. (2015) Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells. Cell Rep 12 (4), 554-61.

141. Moindrot, B. et al. (2015) A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. Cell Rep 12 (4), 562-72.

142. Mikami, S. et al. (2014) Structural insights into the recruitment of SMRT by the corepressor SHARP under phosphorylative regulation. Structure 22 (1), 35-46.

143. van Kruijsbergen, I. et al. (2015) Recruiting polycomb to chromatin. Int J Biochem Cell Biol 67, 177-87.

144. Maduro, C. et al. (2016) Fitting the Puzzle Pieces: the Bigger Picture of XCI. Trends Biochem Sci 41 (2), 138-47.

145. Mira-Bontenbal, H. and Gribnau, J. (2016) New Xist-Interacting Proteins in X-Chromosome Inactivation. Curr Biol 26 (8), R338-42.

146. Peeters, S.B. et al. (2014) Variable escape from X-chromosome inactivation: identifying factors that tip the scales towards expression. Bioessays 36 (8), 746-56.

147. Chan, K.M. et al. (2011) Diverse factors are involved in maintaining X chromosome inactivation. Proc Natl Acad Sci U S A 108 (40), 16699-704.

148. Huynh, K.D. and Lee, J.T. (2003) Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. Nature 426 (6968), 857-62.

149. Patrat, C. et al. (2009) Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. Proceedings of the National Academy of Sciences of the United States of America 106 (13), 5198-5203.

150. Wake, N. et al. (1976) Non-random inactivation of X chromosome in the rat yolk sac. Nature 262 (5569), 580-1.

151. Shevchenko, A.I. et al. (2009) Mosaic heterochromatin of the inactive X chromosome in vole Microtus rossiaemeridionalis. Mamm Genome 20 (9-10), 644-53.

152. Xue, F. et al. (2002) Aberrant patterns of X chromosome inactivation in bovine clones. Nat Genet 31 (2), 216-20.

153. Grant, J. et al. (2012) Rsx is a metatherian RNA with Xist-like properties in X-chromosome inactivation. Nature 487 (7406), 254-8.

154. Harrison, K.B. and Warburton, D. (1986) Preferential X-chromosome activity in human female placental tissues. Cytogenet Cell Genet 41 (3), 163-8.

155. Harrison, K.B. (1989) X-chromosome inactivation in the human cytotrophoblast. Cytogenet Cell Genet 52 (1-2), 37-41.

156. Goto, T. et al. (1997) Paternal X-chromosome inactivation in human trophoblastic cells. Mol Hum Reprod 3 (1), 77-80.

157. Vasques, L.R. et al. (2002) X chromosome inactivation: how human are mice? Cytogenet Genome Res 99 (1-4), 30-5.

158. Moreira de Mello, J.C. et al. (2010) Random X inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome. PLoS One 5 (6), e10947.

159. Okamoto, I. et al. (2011) Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature 472 (7343), 370-4.

160. Wang, X. et al. (2012) Random X inactivation in the mule and horse placenta. Genome Res 22 (10), 1855-63.

161. Tachibana, M. et al. (2012) X-chromosome inactivation in monkey embryos and pluripotent stem cells. Dev Biol 371 (2), 146-55.

162. Sado, T. and Sakaguchi, T. (2013) Species-specific differences in X chromosome inactivation in mammals. Reproduction 146 (4), R131-9.

163. Fan, G. and Tran, J. (2011) X chromosome inactivation in human and mouse pluripotent stem cells. Hum Genet 130 (2), 217-22.

## Aim and scope of this thesis

In mammals, males are the heterogametic sex having an X chromosome and a Y chromosome whereas females have two X chromosomes. Despite originating from an ancient homologous autosomal pair, the X and Y chromosome now differ greatly in size and gene content after ~180 MY of evolution. The X chromosome retained over 1000 genes, whereas the Y chromosome degenerated over time and only contains about a hundred genes mainly involved in male spermatogenesis. Females have two X chromosomes and thus double the amount of X-encoded genes as compared to males which creates an imbalance of X-encoded genes between males and females. Mammals achieve dosage compensation of this imbalance of X-encoded genes by inactivating one of the two X chromosomes in female cells by a process called X chromosome inactivation (XCI). XCI is a stochastic process in which each X chromosome has an equal probability to be inactivated. Even though XCI is stochastic, it is a tightly regulated process to obtain one active X chromosome per diploid genome. Regulation of XCI is achieved by the X inactivation center (Xic), located on the X chromosome, which harbors all the necessary elements for XCI to occur. These elements are regulated by XCI activators and inhibitors.

The aim of this thesis is to gain more insight into the molecular mechanism of XCI. Specifically, how XCI is regulated by activators and inhibitors of XCI, which regions of the X chromosome are involved in this process and whether imprinted XCI is different from random XCI with respect to the inactivated X (Xi).

XCI is regulated by X-encoded activators and autosomally encoded inhibitors. The only described trans-activator so far is the E3 ubiquitin ligase *Rnf12*. RNF12 was shown to target the XCI inhibitor REX1 for proteasomal degradation. RNF12 mediated degradation of REX1 highlights the interplay between XCI activators and XCI inhibitors to determine whether *Xist* is upregulated and if XCI is initiated or not. In **Chapter 2**, different aspects of the RNF12 mediated ubiquitin pathway involved in XCI are explored. These aspects include involvement of the X-encoded E1 activating enzyme UBA1, whether specific lysines are indispensable for the proteasomal degradation of REX1 as well as determining whether a close homologue of RNF12, RNF6, might be involved in XCI. In **Chapter 3,** we focus on the dynamics of the key players in XCI, *Xist* and *Tsix*. *Xist* and *Tsix* are two overlapping antisense transcribed non coding RNAs. *Xist* RNA coats and spreads along the entire X chromosome recruiting chromatin remodeling complexes to silence the X chromosome in *cis* whereas *Tsix* transcription and/or the produced *Tsix* RNA represses *Xist*. Using reporter cell lines we studied the dynamics of *Xist* and *Tsix* independently. In **Chapter 4**, we set out to identify the region(s) involved in activation of XCI, containing novel X-encoded XCI activator(s). In this chapter a large mega base deletion was generated to narrow down the region(s) of interest. In **Chapter 5**, we explored the Xi epigenetic landscape, examining Xi specific chromatin marks in cell types representing different embryonic lineages and developmental stages of the mouse embryo. We compared different cell types displaying imprinted and random XCI to identify lineage specific markers that could explain the difference between these two forms of XCI. In **Chapter 6,** we discuss the main findings of the described chapters taking into consideration related work in the XCI field and possible future directions to continue this work. In **Addendum A**, an overview is given on candidate genes explored to search for the remaining XCI activators which were found not to be involved in activation XCI.
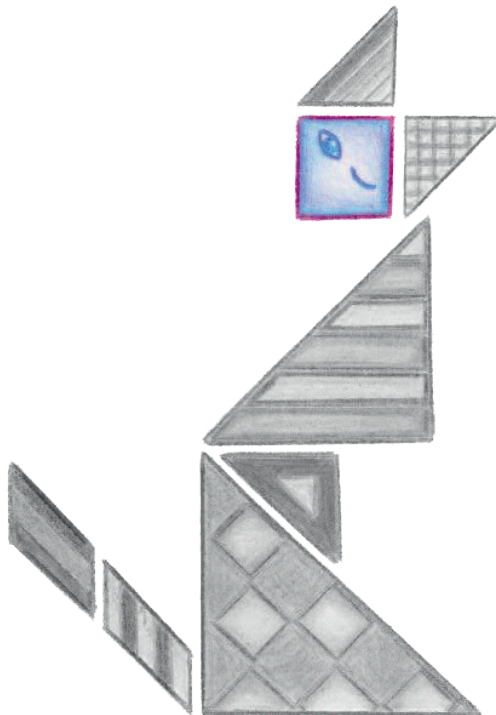
# Chapter 2

*Activation of XCI: a role for ubiquitination*

# Chapter 2

*Maduro C.M., Rentmeester E., Gontan C., Mulugeta E., Sleddens-Linkels E. and Gribnau J.*

*Maduro C.1, Rentmeester E.1, Gontan C.1, Mulugeta E.2, Sleddens-Linkels1 E. and Gribnau J.1\**

*1 Department of Developmental Biology Erasmus MC, University Medical Center, Rotterdam, The Netherlands*
*2Curie Institute, CNRS UMR3215, INSERM U934, 26 rue d'Ulm, Paris 75248, France*
*\*Correspondence: j.gribnau@erasmusmc.nl*

## Abstract

X chromosome inactivation (XCI) in female mammals is a dosage compensation mechanism to equalize X encoded gene dosage of XX females to XY males. Initiation of XCI is a tightly regulated process directed by autosomally encoded inhibitors and X encoded activators of XCI. The XCI activator and E3 ubiquitin ligase, RNF12, activates XCI through proteasomal degradation of one of the XCI inhibitors, REX1. Here we have studied the ubiquitin pathway in context of XCI in more detail. We show that two copies of the X encoded E1 activating enzyme, UBA1, are necessary for proper initiation of XCI and is a novel XCI activator. In addition, we show that the RNF12 mediated downregulation of REX1 is robust, involving multiple ubiquitin targets and XCI still proceeds in the absence of these targets. Finally, we show that the E3 ubiquitin ligase *Rnf6,* an autosomal close homologue of *Rnf12* that is equally expressed in male and female cells, acts as an *Xist* activator most likely through degradation of REX1.

## Introduction

X chromosome inactivation (XCI) is required for proper dosage compensation of X encoded genes between mammalian XX females and XY males, leading to near complete inactivation of one of the two X chromosomes in female cells. XCI can be divided into three main phases: initiation, establishment and maintenance of the inactive X chromosome (Xi) [1]. Initiation of XCI is the crucial phase, regulated by activators and inhibitors of XCI, determining the number of X chromosomes in a nucleus and directing initiation of XCI. Inhibitors of XCI are autosomally encoded, acting as denominators, setting up the threshold for XCI to occur. In contrast, XCI activators are X encoded, acting as numerators, expressed at twice the level in female as compared to male cells and therefore trigger female specific initiation of XCI. Several XCI inhibitors have been identified, mainly consisting of pluripotency factors, clearly linking XCI initiation to loss of pluripotency [2-5]. The E3 ubiquitin ligase *Rnf12* has been identified as a potent *trans* activator of XCI [6]. RNF12 is important for XCI as *Rnf12*$^{-/-}$ cells fail to initiate XCI, but represents not the sole XCI activator, as *Rnf12*$^{+/-}$ cells, though delayed, still undergo XCI which is never observed in male cells.

In XCI, RNF12 targets the XCI inhibitor and pluripotency factor REX1 for proteasomal degradation [7, 8]. This pathway starts with ATP-dependent activation of ubiquitin by the E1 activating enzyme and transfer to an E2 enzyme to be recruited by the ring finger of an E3 ubiquitin ligase. Bridging of the active ubiquitin charged E2 and the substrate protein results in ubiquitination of the target protein, and repetitive recruitment of the E2 results in poly-ubiquitination and subsequent degradation of substrates [9-11]. The E1 activating enzyme is indispensable for proteasomal degradation and deletion of *Uba1* in *C. elegans* is lethal [12-14]. In vertebrates, two E1s exist, *Uba1* and *Uba6,* however *Uba1* appears to be the main activator of most E2s [9, 15-17]. Interestingly, *Uba1* is located on the X chromosome and dose dependent degradation of target proteins might therefore be affected by different expression levels of *Uba1* in male and female ES cells. In mice, *Uba1* is inactivated upon XCI [18], and might be involved as an upstream activator in a ubiquitin mediated cascade to up-regulate *Xist*, through dose dependent RNF12 mediated downregulation of REX1, or other ubiquitin related pathways. *Rnf12* is a member of the RING finger family, consisting of over 100 members. Sequence comparison by principal component analysis (Figure 1A) of members of the RNF family showed highest homology of *Rnf12* to *Rnf6,* clustering together yet secluded from the clustering of the rest of the members of this gene family. The sequence homology of different functional domains of these proteins is very high ranging in the order of 70% [19], and although *Rnf6* is autosomally encoded and equally expressed in male and female cells, this E3 ubiquitin ligase might attribute to the turnover of REX1, lowering the threshold for initiation of XCI. To address this and explore the role for of the ubiquitin pathway in activation of XCI, we first studied the specificity of RNF12 mediated targeting of REX1 by examining the effects of modifying the substrate on ubiquitination and degradation. In subsequent genetic ablation experiments we studied the dose dependent role of *Uba1* and *Rnf6* in XCI. Our studies underscore the impact of the ubiquitin pathway on the regulation of XCI.

**Experimental procedures**

*BACs, plasmids and antibodies*

The BAC covering *Rnf6* (CHORI) used in this study was RP24-304C5. Recombination between the BAC and an ori-less recombination vector to introduce a kanamycin/neomycin resistance cassette, allowing selection in ES cells, into the BAC backbone was performed according to Barakat et al (2015) [20]. Plasmids used in the REX1 and RNF12 ubiquitination experiments were REX1-mCherry, Rex1-FLAG-V5, RNF12-EGFP and RNF12-FLAG [8]. *Rnf6* cDNA was cloned into the pEGFP-N1 vector (Clonetech) to obtain an RNF6-EGFP expression vector. The pSpCas9(BB)-2A-Puro (PX459) (Addgene plasmid # 48139) was used to insert the CRISPR guide RNAs [21]. Antibodies used were recognizing REX1 (Abcam), RNF12 (Abnova), β-Actin (Sigma), V5 (Invitrogen) and Ub (Enzo). The secondary antibodies were α-Mouse IgG-peroxidase (Sigma) for RNF12 and Ub, and α-Rabbit IgG-peroxidase (Sigma) for REX1.

*Cell lines*

Cell lines, culture media and culture conditions for ESC culture and differentiation were described previously [7, 22]. BACs (linearized) and CRISPR vectors were introduced in hybrid wildtype F1 129sv/CAST [23] male or female ES cells by electroporation in a 0.2cm electroporation cuvette (BioRad) at 118kV, 1200μF and ∞Ω in a Gene Pulser Xcell Electroporation System (BioRad). After 24 hours of recovery, ES cells were grown on ES medium containing neomycin

(270µg/ml) or puromycin (1ug/ml) for seven days or 24 hours for BACs or CRISPR vectors respectively.

*Xist RNA Fluorescent in situ hybridization*
Procedure, probe labeling and antibodies have been described previously [7, 22].

*RNA isolation from ES cells and cDNA synthesis*
TRIzol® (Invitrogen) was added to the cells and RNA was isolated according to manufacturer's instructions. RNA samples were diluted to contain 2µg RNA in 10µl. Before cDNA synthesis, 1µl DNase (Invitrogen) was added to the RNA samples, for 30 minutes at 37°C to remove any contaminating genomic DNA. The DNase was inactivated at 65°C for 10 minutes. 1µl Random Hexamers (Applied BioSystems) and 1µl dNTP mix (Invitrogen) was added and incubated at 65°C for 5 minutes and subsequently cooled on ice for 1 minute. 4µl 5x First Strand Buffer, 2µl 0.1M DTT and 1µl RNaseOUT (Invitrogen) were then added to the samples and incubated at 25°C for 2 minutes before adding 1µl SuperscriptII Reverse transcriptase (Invitrogen). The samples were then incubated at 25°C for 10 minutes followed by 50 minutes at 42°C and finally, 15 minutes at 70°C.

*Rnf6-EGFP fusion protein*
Primers were designed flanking *Rnf6* cDNA with *XhoI* and *KpnI* sites, leaving the ATG (start site) intact but removing the Stop-codon: Fw 5' CTCGAGATGGATCCGTCTAGATCTAGATCA 3' and Rv 5' CGGTACCCCACTGCTTGTGGCTCCGAATTC 3'. This was then cloned into TOPO by Zero Blunt® TOPO® PCR Cloning (Invitrogen) and sequenced with MF-20i and MR-Invitrogen primers. Using the *XhoI* and *KpnI* sites *Rnf6* was cloned into the pEGFP-N1 vector with T4 DNA ligase (Roche). *Rnf12-EGFP* and *Rex1-mCherry* were described previously [8].

*Transient transfection of HEK293 cells*
HEK293 cells were co-transfected with Rnf6 (-EGFP or FLAG) and Rex1 (-mCherry or FLAG-V5) [8] to determine the degree of ubiquitination and degradation of REX1 by RNF6. As a control, RNF12 (-EGFP or FLAG) [8] was used. *Rnf6/12* and *Rex1* were transfected in the ratio 3:1 when using EGFP/mCherry and 1:1 when using the pCAG-FLAG-V5 vectors respectively. Polyethylenimine (Polysciences) was used to transfect the HEK293 cells. For the EGFP/mCherry transfection, cells were grown on round cover slips treated with poly D-lysine in a 12-well plate until ~80% confluent. After 24 hours, the cells were fixed with 4% PFA for 20 minutes at room temperature, washed 3 times in PBS, incubated for 20 minutes in 0.1% triton and washed 3 times in PBS before mounting with VECTASHIELD® Mounting Medium containing 4', 6-diamidino-2-phenylindole (DAPI). Images were acquired using a fluorescence microscope (Axioplan2; Carl Zeiss). The transfection with Flag-Rnf6/Rnf12 and Flag-V5-Rex1 was done in a similar way as described above, but in a 10 cm dish per transfection. After 48 hours, nuclear proteins were isolated.

*Nuclear protein isolation*

For nuclear protein isolation a confluent 10 cm dish of cells was used. The cells were scraped into 1 ml cold PBS containing protease inhibitor (Complete, EDTA-free, Protease Inhibitor Cocktail Tablets (Roche)) and kept on ice. The cells were spun down for 10 seconds at 4ºC and resuspended in 400µl cold buffer A (10mM Hepes 7.6, 1.5mM MgCl2, 10mM KCl, 1xPI, 0.5mM DTT and 15µM MG132) and subsequently incubated on ice for 10 minutes. The cells were vortexed for 10 seconds and briefly spun down again at 4ºC. The supernatant fraction containing the cytoplasmic proteins was discarded and the pellet was resuspended in 150µl Buffer C (20mM Hepes 7.6, 1.5mM MgCl2, 25% glycerol, 420mM NaCl, 0.2mM EDTA, 1xPI, 0.5mM DTT and 15µM MG132) and incubated on ice for 20 minutes. After centrifugation at 13200 rpm for 2 minutes at 4ºC, the DNA binding proteins were obtained in the supernatant fraction.

*Western blot*

Samples for western blots were diluted with 2xLaemmli Buffer to obtain 180µg protein in 60µl. The samples were boiled at 99ºC for 5 minutes. The SDS-PAGE gel consisted of a 5% stacking gel and a 10% running gel and was run in electrophoresis buffer (25mM Tris (pH8.3), 192mM glycine, 0.1%SDS from BioRad) in a Biorad cell. After electrophoresis, the proteins were transferred onto a nitrocellulose membrane by blotting for 45 minutes at 200mA in blot buffer (25mM Tris (pH8.3), 192mM glycine (BioRad) and 20% Methanol) in a Trans-Blot SD semi-dry electrophoretic transfer cell (BioRad). Blocking of non-specific sites was done for 30 minutes in 3% Milk (or 1% BSA for Ubiquitin)/1xTBS/0.1%Tween. After blocking, the primary antibody was added overnight at 4ºC. Primary antibodies used in this study: α-Rnf12 (1:3000), α-Rex1 (1:3000), α-Actin-HRP (1:50000), α-Ubiquitin (1000). Secondary HRP labeled antibodies (all diluted 1:5000) used in this study: α-Mouse (Rnf12, Ubiquitin) and α-Rabbit (Rex1). The membrane was exposed to ECL Western Blotting detecting reagents in a ratio of 1:1 (GE Healthcare) for 1 minute and exposed to light-sensitive film (Kodak or GE Healthcare) and developed.

*Site-directed mutagenesis*

The QuickChange site-directed mutagenesis kit (Invitrogen) was used to introduce single point mutations. *Rex1* cDNA [8] was used as template and two overlapping primers (Table 1) containing the desired mutations were used to generate a mutated plasmid containing staggered nicks. The PCR product was treated with *DpnI*, which is specific for methylated and hemi methylated DNA and therefore digests only the template DNA. The desired mutated plasmid was then transformed into XL1-Blue super competent cells by a heat shock for 45 seconds at 42ºC. When correct mutations were confirmed, the *Rex1* mutant sequences were cloned into a pCAG-Flag-V5 vector [8].

Table 1: Primers for site-directed mutagenesis, creating point mutations changing lysine into arginines in REX1. Point mutations are underlined in the sequence.

| Fw: 5'-3'and Rv: 3'-5' | Mutation |
| --- | --- |
| CGACAGACTGACCCTAAGGCAAGACGAGGCAAGGC | K33 → A |
| GGAGTACATGACAAGGGGGACGAGGCAAGAGAGGAGAGAGAGGTC | K126, K129, K132 → A |
| CACGGAGAGAGCTCGAGACTAAGGCGACATTTTCTGGTG | K213, K215 → A |

*V5 Immunoprecipitation*
The nuclear proteins were diluted 1:1 with buffer C-0 (20mM Hepes (pH 7.6), 20% glycerol, 1.5mM MgCl2, 0.2mM EDTA, 1xPI, 0.5mM DTT, 15µM MG132). Precipitates caused by dilution were removed by centrifugation at maximum rpm for 5 minutes after which the supernatant was transferred to non-stick tubes containing 15µl washed α-V5 agarose beads (Sigma). The beads with nuclear proteins were rotated at 4°C for 90 minutes and then washed four times with buffer C-150 (20mM Hepes (pH 7.6), 10% glycerol, 150mM KCl, 1.5mM Mgcl2, 0.2mM EDTA, 0.02% NP40, 1xPI, 0.5mM DTT, 15µM MG132). 2x Laemmli buffer was added to the beads, mixed and boiled for 4 minutes at 99°C. The samples were then centrifuged for 1 minute at 3000 rpm and transferred to a new tube and checked by immunoblotting on an 8% polyacrylamide gel.

*Deletions with the CRISPR/Cas9 system*
CRISPR guides were designed by using the CRISPR design tool (http://crispr.mit.edu/). The designed CRISPR guide oligos with 5'- CACC and 3'- CAAA overhangs (Table 2) were cloned into the pX459 CRISPR vector (Addgene) containing a U6 RNA polymerase II promoter. The CRISPR guides were inserted into the pSpCas9 (BB)-2A-Puro (PX459) plasmid (Addgene) [21] by a simultaneous digestion-ligation reaction [24]. First, the pX459 vector was digested with *BbsI* (NEB) to allow the replacement of the restriction sites with direct insertion of the annealed oligo guides. This was then directly used to transform heat competent bacteria. A combination of two guides (beginning and end) was used in each targeting and each targeting was first screened as a pool to ensure the deletion is present, followed by single clone screens (Table 3).

Table 2: Oligo's for CRISPR guides for deletion of Rnf12, Rnf6 and Uba1. 5'- CACC and 3'- CAAA overhangs are indicated in bold and the preferred G as a starting nucleotide is underlined.

| Gene | Location | Guide |
|------|----------|-------|
| **Rnf12** | Intron 2 | Fw: 5' **CACC**<u>G</u>AAAGCGCTGTACAAAAAGTT 3' |
| | | Rv: 5'**AAAC**AACTTTTTGTACAGCGCTTT<u>C</u>3' |
| | | |
| | 3'UTR | Fw: 5' **CACC**<u>G</u>GAACAAGTACTCTAAACTA 3' |
| | | Rv: 5' **AAAC**TAGTTTAGAGTACTTGTTC<u>C</u> 3' |
| | | |
| **Rnf6** | Exon1 | Fw: 5'**CACC**<u>G</u> TGCAAATAGAACCCGATCTA 3' |
| | | Rv: 5'**AAAC** TAGATCGGGTTCTATTTGCA <u>C</u>3' |
| | | |
| | Exon5 | Fw: 5'**CACC**<u>G</u> AGTCGGGGGCCTCGATTAAC 3' |
| | | Rv: 5'**AAAC** GTTAATCGAGGCCCCCGACT <u>C</u>3' |
| | | |
| **Uba1** | Exon1 | Fw: 5'**CACC**<u>G</u> TTGCGCGGAGCTCGGAAGCG 3' |
| | | Rv: 5'**AAAC** CGCTTCCGAGCTCCGCGCAA <u>C</u>3' |
| | | |
| | Exon27 | Fw: 5'**CACC**<u>G</u> TATAGCTGGAGTAGCCCATT 3' |
| | | Rv: 5'**AAAC** AATGGGCTACTCCAGCTATA <u>C</u>3' |

*Table 3: Primers for detection wildtype allele, deletion and allele check of Rnf12, Rnf6 and Uba1*

| | Wildtype allele | PCR over deletion | Allele check |
|---|---|---|---|
| Rnf12 | CTGCAGTGAATCCTCCTTGA | CTGCAGTGAATCCTCCTTGA | GCTCACAACCATCCATAACG |
| | AGCAAACTTGGTTGGGAAAC | TGAACAAATTGTGTGCATGG | TCTATCCCCTCTGCCATTGTT |
| | | | Bstx1 digest on 129 allele |
| Rnf6 | GGCTGCATCTCCATTACTCA | GGCCATTGCTTTGATTCTCT | CTTCCCAGAGTTGCAAATGA |
| | GGCCATTGCTTTGATTCTCT | GGAAGGGTCATTTGGTGAGT | CTGAGGGGGTAGAGCAGAAG |
| | | | EcoRV digest on CAS allele |
| Uba1 | GAGGCGCAGAAGTAGAGAGAA | CCCCTCTATTCAGGCGTCTA | TCTCCTTCACAGCTCCTTGA |
| | AAAATTCTGCCCTCCAACAG | CCCTTCAGACCCACCATTAC | TGATTAAGAGCACCGACTGC |
| | | | EcoNI digest on 129 allele |

## Results

*RNF12 mediated degradation of REX1: Redundancy in preferentially ubiquitinated lysines*

To obtain more insight into the mechanism of RNF12-mediated degradation of REX1, mass spectrometric analysis was performed to identify the ubiquitinated sites within REX1 in female ES cells. An *in vitro* ubiquitination study identified five peptides derived from REX1 with the diglycine signature corresponding to putative lysine acceptor sites for ubiquitin linkage. These five peptides corresponded to ubiquitination of K33, K129, K132, K213 and K215 of REX1 [25]. To determine whether these sites are indispensable for the degradation of REX1, the five lysines identified, including one additional lysine (K126) in close proximity to two ubiquitinated lysines, were mutated into arginines by site-directed mutagenesis. We replaced lysines into arginines as these amino acids are most similar; ionizable basic amino acids with a net positive charge in the pH ranges found in the cell. Different *Rex1* mutant expression vectors were generated with point mutations at different positions within REX1 (Figure 1B). The generated *Rex1* mutant vectors were subsequently transiently co-transfected with RNF12 in HEK293 cells. Nuclear proteins were analyzed by western blot for REX1 protein levels in the absence (no ubiquitination) and presence of RNF12 (ubiquitination). Despite equal amounts of expression vectors transfected in HEK cells, all mutated REX1 proteins showed reduced expression levels in the absence of RNF12, possibly related to protein miss-folding and destabilization (Figure 1C and 1D lower panel). Co-transfection of *Rnf12* with *Rex1* single, double and triple mutations still let to degradation of REX1, highlighting redundancy in the degradation pathway. We therefore generated REX1 expression vectors with five or six mutated lysines. These REX1 mutants 4 and 5 displayed near similar REX1 protein levels in the presence and absence of RNF12, suggesting that these mutants are less prone to degradation by RNF12. To determine the degree of ubiquitination of REX1 mutant 4 and 5, IP followed by immuno-detection of ubiquitinated REX1 was performed. This analysis indicated the presence of an intense smear corresponding to poly-ubiquitinated mutant REX1, indicating activity of RNF12 towards mutant REX1 (Figure 3C). However, levels of REX1 poly-ubiquitination were reduced for mutants 4 and 5 compared to wild type in the presence of RNF12, suggesting that the new targets for poly-ubiquitination were less efficiently recognized by the proteasome.
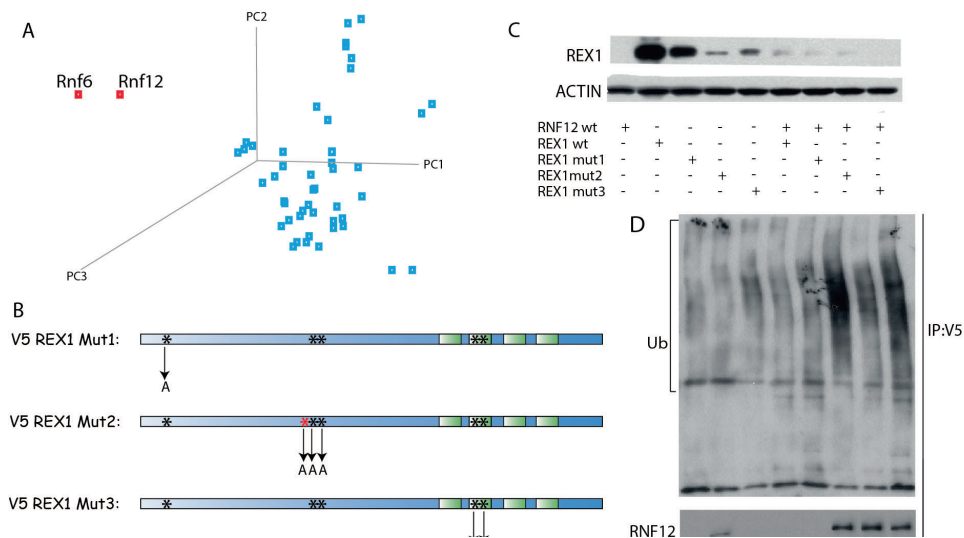
**Figure 1**: A) *Rnf12* and *Rnf6* cluster together and away from the rest of the RNF family in a Principal component analysis. A spatial representation of the similarities of the sequences of the RNF gene family. The sequences are displayed as points in 'similarity space' and similar sequences tend to lie near each other in this space. B) Five different *Rex1* mutants (V5 REX1 mut1-5) were created by site directed mutagenesis. Green shaded areas represent the zinc finger domains. The five identified lysines are indicated with a * and the additional lysine mutated in their close proximity is shown in red. C) Western blot with nuclear proteins of transient co-transfections of V5 REX1 mut1-3. Levels of REX1 and RNF12 were detected in the different transfected samples and actin was used as a loading control. D) Western blot (lower panels) and IP (upper panel) with nuclear proteins of transient co-transfections of combinations of RNF12 and V5 REX1 mut4 and mut5 expression vectors. Levels of REX1 and RNF12 were detected in the different transfected samples and actin was used as a loading control in the western blot (lower panel). The IP was performed with an anti-V5 antibody followed by immuno-detection with an anti-ubiquitin antibody to determine the degree of ubiquitination of the different REX1 mutants by RNF12 as compared to wild type REX1 (upper panel).

## *Loss of Uba1 in Rnf12⁺/⁻ cells affects XCI*

*Rnf12* expression levels play an important role in XCI upon ES cell differentiation. Nevertheless, the presence of female ES cells inducing XCI in a *Rnf12*$^{+/-}$ heterozygous background indicates that more XCI activators are involved. The main E1 activating enzyme in mammals, *Uba1,* is X-encoded and might represent a putative XCI activator increasing the rate of ubiquitination by boosting the levels of activated E1 in female cells. To test whether *Uba1* acts as an activator of XCI, we aimed to generate *Rnf12*$^{+/-}$:*Uba1*$^{+/-}$ heterozygous knockout female ES cells, using the CRISPR/Cas9 technology. First, *Rnf12* was removed from one allele in female F1 Cast:129/Sv ES cells. These hybrid ES cells have 1 SNP in every 100bp facilitating identification of properly targeted clones and allele specific follow up of phenotypes. *Rnf12* was targeted to the 129/Sv allele using two guides recognizing intron 2 and the 3'UTR of *Rnf12*, generating a deletion of approximately 18.5kb, removing the complete open reading frame of *Rnf12* (Supplementary Figure 1A-C). This new *Rnf12*$^{+/-}$ ES cell line recapitulated the reported XCI phenotype observed
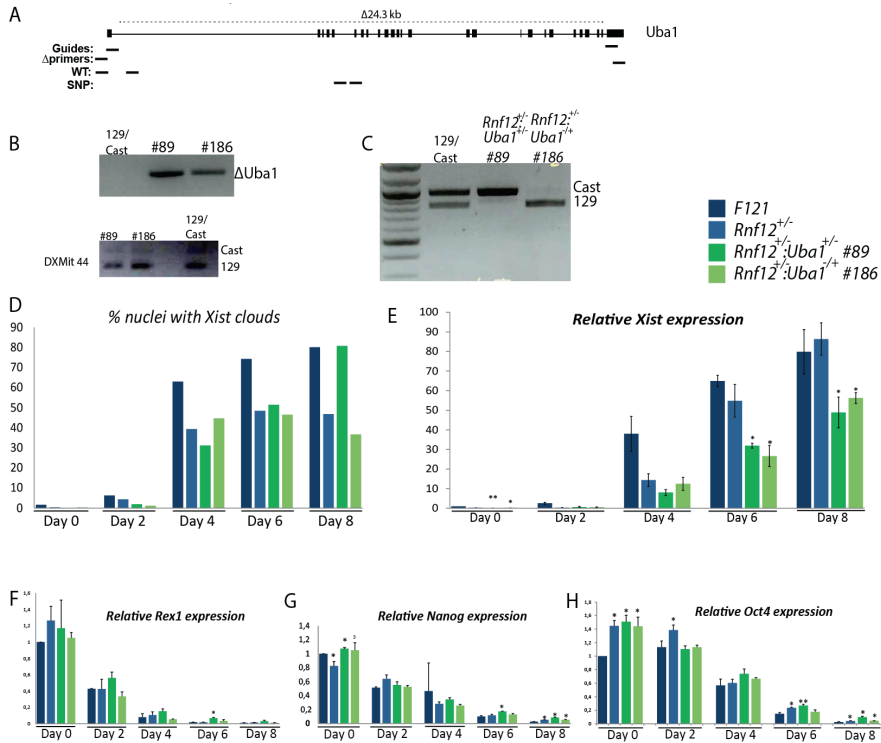
**Figure 2:** A) Representation of *Uba1* deletion guides and primers. B)PCR spanning the deleted region of *Uba1*. Two clones (89 and 186) show the deletion not detected in control 129/Sv:Cast cells (top). PCR spanning DXMit44 length polymorphism marker. Both clones show the presence of both Cast and 129 allele as seen in the wildtype (129/Cast) cells representing the presence of two X chromosomes (bottom). C) Allele specific PCR of *Uba1* allele identifying targeted allele; 129/Sv for clone 89 and Cast for clone 186 . D) Percentage of cells with *Xist* clouds at day 0,2,4,6 and 8 of differentiation in F121 (wildtype; n=298 d0; n= 256 d2; n=241 d4; n=206 d6; n= 272 d8) *Rnf12+/-* cells (n= 223 d0; n= 251 d2; n=241 d4; n=233 d6; n=218 d8) and two *Rnf12+/- :Uba1+/-* clones (clone 89: n= 226 d0; n= 232 d2; n=250 d4; n= 206 d6 n=260 d8; clone 186: n= 266 d0; n= 264 d2; n=208 d4; n=202 d6; n= 259 d8). E) Relative *Xist* expression normalized to *β-actin* at day 0,2,4,6 and 8 of differentiation of F121 (wildtype) *Rnf12+/-* cells and two *Rnf12+/- :Uba1+/-* clones. Error bars represent standard deviation (Comparison of *Rnf12+/- :Uba1+/- vs Rnf12+/-* t test, * p<0.05, ** p<0.001). F) Relative *Rex1* expression normalized to *β-actin* at day 0,2,4,6 and 8 of differentiation of F121 (wildtype) *Rnf12+/-* cells and two *Rnf12+/- :Uba1+/-* clones. Error bars represent standard deviation (Comparison of *Rnf12+/- :Uba1+/- vs Rnf12+/-* t test, * p<0.05, ** p<0.001). G)As in F but now for *Nanog* (G) and *Oct4* (H).

*RNF6 can target REX1 for proteasomal degradation*

The prominent role of *Rnf12* and *Uba1* in XCI, highlights the importance of the ubiquitin pathway in initiation of XCI. Interestingly, *Rnf6* is a close homologue of *Rnf12*, and is expressed in ES cells. Although RNF6 is autosomally encoded, it might still contribute to the turnover of REX1 or other RNF12 target proteins in the activation of XCI. To investigate this, we generated

a *Rnf6*-EGFP fusion expression construct and tested whether REX1 could act as a substrate of RNF6. Expression vectors encoding RNF6-EGFP and RNF12-EGFP were co-transfected with REX1-mCherry fusion proteins in HEK293 cells that do not endogenously express any of these proteins. REX1-mCherry alone transfections revealed bright red cells, whereas co-transfection of RNF12-EGFP with REX1-mCherry, only resulted in EGFP positive cells (Figure 3A). In contrast co-transfection of RNF6-EGFP with REX1-mCherry resulted in EGFP positive cells that also showed dimmed expression of mCherry (Figure 3A). This finding suggested that RNF6 could also target REX1 for degradation, but also indicated that RNF12 was more efficient. To further analyze and support this observation, transfections were performed in HEK293 cells using combinations of a tagged expression vector encoding V5-REX1 and expression vectors encoding untagged RNF12/RNF6. Nuclear extracts were isolated from cells that were either treated or not treated with the proteasome inhibitor MG132, followed by immuno-precipitation of REX1 with anti-V5 agarose beads, and detection with anti-V5 and anti-ubiquitin to visualize poly-ubiquitinated REX1. We found that in the absence of MG132 REX1 levels are reduced in the presence of RNF12 and RNF6, although the effect is less pronounced for RNF6 (Figure 3B, left bottom panels). Immunoprecipitation of REX1 and detection with an anti-ubiquitin antibody also indicated poly-ubiquitinated REX1 in the presence of RNF12 and RNF6, but confirmed that the activity of RNF6 was less pronounced compared to RNF12 (Figure 3B, right panels). These results indicate that *Rnf6* might be involved in XCI by lowering the threshold set by REX1. To test this hypothesis, we introduced extra copies of *Rnf6* in male 1.3 ES cells using BACs. Our previous studies have shown that introduction of additional copies of *Rnf12* resulted in ectopic XCI in transgenic male and female cells [6]. Also in our *Rnf6* transgenic male cells ectopic *Xist* clouds were observed, which were never observed in male control cells (Figure 3C). Nevertheless, the percentage of cells with *Xist* clouds was very low, with most clones showing between 1% and only one clone 2.4% of cells with *Xist* clouds (Figure 3D). This percentage of *Xist* clouds did not correlate with the copy number, and *Xist* qPCR analysis indicated no significant increase in *Xist* expression, reflecting that ectopic XCI was limited to a small number of cells.

*Loss of Rnf6 in female Rnf12$^{+/-}$ cells leads to further affected XCI*

Although the level of *Xist* induction was very low in male *Rnf6* transgenic ES cells, the presence of a small number of ectopic *Xist* clouds and *in vitro* ubiquitination of XCI inhibitor REX1 by RNF6, suggested a role for *Rnf6* in XCI initiation. To test whether *Rnf6* is indeed involved in XCI, we deleted *Rnf6* from in *Rnf12$^{+/-}$* cells using the CRISPR/Cas9 technology. *Rnf6*, located on chromosome 5, was deleted using two guides recognizing exon1 and exon5, resulting in an approximate deletion of 10kb (Figure 3E). We obtained one *Rnf12$^{+/-}$:Rnf6$^{+/-}$* clone with a heterozygous deletion of *Rnf6* and one *Rnf12$^{+/-}$:Rnf6$^{-/-}$* clone with a homozygous deletion of *Rnf6* (Figure 3 F, G). To test the effect of these deletions on XCI, wildtype, *Rnf12$^{+/-}$, Rnf12$^{+/-}$:Rnf6$^{+/-}$* and *Rnf12$^{+/-}$:Rnf6$^{-/-}$* ES cells were differentiated up to 6 days, and analyzed by *Xist*-RNA FISH to determine the percentage of cells with *Xist* clouds. We foundreduced *Xist* cloud formation after 4 and 6 days of differentiation (Figure 3H) and *Xist* qPCR analysis indicated that cells with a deletion of *Rnf6* displayed affected XCI. In addition, lower levels of *Xist* expression did not recover after day 6 of differentiation, as found for *Rnf12$^{+/-}$* cells (Figure 3I). These results indicate that *Rnf6* facilitates in activation of *Xist*, likely through the degradation of REX1 and possibly by affecting other targets in XCI.
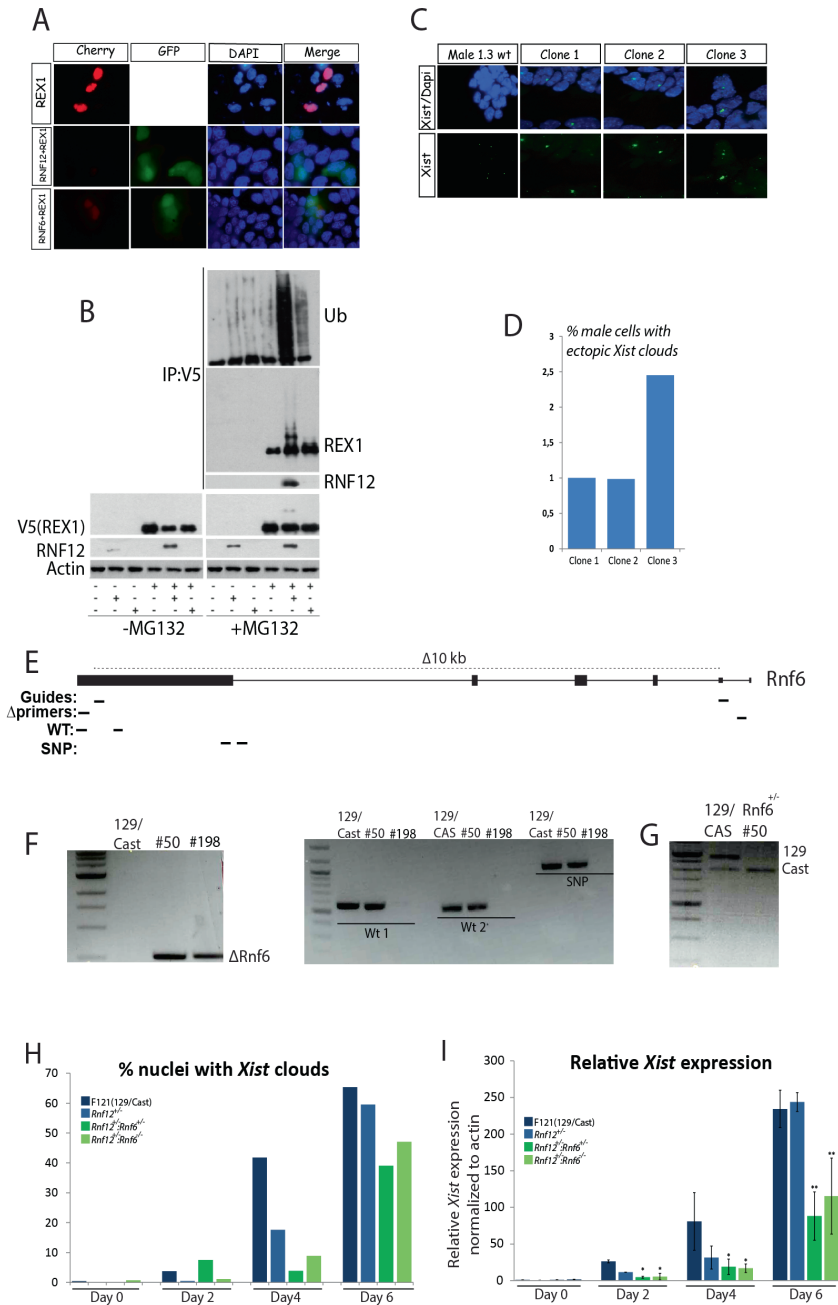
**Figure 3:** A) Transient transfections of (from top to bottom) REX1-mCherry, co-transfection of RNF12-EGFP with REX1-mCherry and co-transfection of RNF6-EGFP with REX1-mCherry. From left to right, mCherry, EGFP, DAPI and a merge of the different transfected cells. B) Immuno-detection (lower panel)

and immunoprecipitation with nuclear protein extracts of transfection of combinations of expression vectors V5-REX1, RNF12 and RNF6 (RNF12 antibody weakly detects RNF6 (74kD), RNF12 (66kD)). Cells were either not treated (left) or treated with MG132 (right). C) *Xist* RNA FISH on male 1.3 wildtype cells and three male clones with overexpression of *Rnf6*. Wildtype cell show only pinpoints, whereas the three clones show ectopic *Xist* cloud formation in a small percentage of cells. D) The three *Rnf6* transgenic clones of (C) with corresponding percentage of cells showing ectopic *Xist* cloud formation E) Representation of *Rnf6* deletion: guides and primers. F) PCR spanning the deleted region of *Rnf6* (left) and the wildtype *Rnf6* (right) revealing a heterozygous clone (#50) and a knockout clone (#198) which is negative on both alleles for all PCR amplifications in the deleted region. G) Digestion of SNP PCR in figure 3F with *EcoRV* indicates the deletion of clone #50 on the 129/Sv allele of chromosome 5. H) Percentage of cells showing Xist clouds. I) Relative *Xist* expression normalized to *β-actin* at day 0, day 2, day 4, day 6 and day 8 of differentiation of F121 (wildtype) $Rnf12^{+/-}$, $Rnf12^{+/-}:Rnf6^{+/-}$, $Rnf12^{+/-}:Rnf6^{-/-}$ cells. Error bars represent standard deviation (Comparison of clones *vs* $Rnf12^{+/-}$ t test, * $p<0.05$, ** $p<0.001$).

## Discussion and conclusion

In XCI, RNF12 targets REX1 for proteasomal degradation [8]. RNF12 is X-encoded, facilitating female specific *Xist* activation by dose dependent down regulation of REX1, which acts as a repressor of *Xist*. Mass spectrometry analysis identified five of the 26 lysines within REX1 as targets for ubiquitination: K33, K129, K132, K213 and K215 respectively [8]. To determine whether these specific lysines were indispensable for degradation of REX1, combinations of lysines were mutated into arginines. In a study applying a similar approach, the specific lysines important for degradation of p53 were determined [28]. For p53 it was found that C-terminal lysine residues were the main sites of ubiquitin ligation and crucial for targeting p53 for proteasomal degradation. Mutating these sites to arginines resulted in a mutant p53 which was resistant to Mdm2 induced degradation and refractory to Mdm2-mediated ubiquitination [28]. This requirement of specific lysines for proteasomal degradation was also apparent for NF-κB and IκBα [28]. For RNF12 we found that only when most lysines were mutated this resulted in a reduction of REX1 ubiquitination, nevertheless ubiquitination still occurred, suggesting redundancy in targeted lysines in the degradation of REX1. Redundant action in ubiquitination has also been reported before for other proteins [29-32] and would make the RNF12/REX1 pathway more robust.

To further investigate the role of the ubiquitin pathway in XCI we studied the effects of mutating the main ubiquitin activation enzyme Uba1 on XCI. *Uba1* is located on the X chromosome and dose dependent break down of REX1 might be further augmented by increased levels of UBA1 in female as compared to male cells. Our studies indeed showed a further reduction in XCI in differentiating $Rnf12^{+/-}:Uba1^{+/-}$ ES cells when compared to $Rnf12^{+/-}$ cells. This effect was more prominent at later stages of differentiation not affecting the percentage of *Xist* clouds but impacting at the *Xist* expression level. This finding suggests that *Uba1* facilitates robust expression of *Xist* in a dose dependent manner, and demonstrates that *Uba1* is an activator of XCI. Whether dose dependent levels of UBA1 are important for the action of RNF12 or other E3 ubiquitin ligases in XCI, requires further investigation. Similar to *Rnf12*, *Uba1* is silenced upon XCI in mouse and the mouse Y homolog, *UbaY*, only has male-specific functions [27], providing an important feedback mechanism to block XCI of all X chromosomes present [18].

Both *Rnf12* and its close homologue *Rnf6* are ubiquitously expressed during mouse embryogenesis [19]. To test whether these proteins had overlapping functions in XCI we performed co-transfection studies in HEK293 cells. These experiments indicated that REX1 is a much better

target for RNF12 than RNF6 in proteasomal degradation, but also indicate that RNF6 is able to target REX1 for ubiquitination. Nevertheless, the impact of RNF6 on the degradation of REX1 appears less prominent compared to RNF12, which might be related to differences in expression levels, target specificity and nuclear localization. RNF6 is mainly present in the cytoplasm whereas RNF12 is located to the nucleus. This seems to be determined by differences in NLS, showing most of the sequence convergence between the two E3 ubiquitin ligases, showing only 20% homology in this region [19]. Replacement of these NLSs even resulted in nuclear RNF6 and cytoplasmic RNF12 [19]. Although addition of *Rnf6* transgenes to male cells did show little effect with respect to *Xist* expression and XCI, knockout of one or both copies of *Rnf6* in a heterozygous *Rnf12* background had a clear impact on XCI. In contrast to RNF12, RNF6 is not X-encoded, and is equally expressed in male and female cells, and hence cannot be considered an XCI activator. Interestingly, *Rnf6* dosage appears not to affect XCI as we observed a similar XCI phenotype in heterozygous and knockout *Rnf6* cells. Nevertheless, *Rnf6* could stimulate the XCI process by lowering the threshold for XCI (in male and female cells) through degradation of REX1 and potential other targets. A similar mode of action has been proposed for YY1, which competes for binding sites with REX1 to lower the threshold for *Xist* activation [33].

Taken together, these results identify UBA1 as a novel XCI activator, and RNF6 as an *Xist* activator. We propose a model of an activation cascade in which the E1 activating enzyme UBA1 together with RNF12 and RNF6 target REX1 and potential other targets for proteasomal degradation (Figure 4). Increased expression of X-encoded *Uba1*, and *Rnf12* in female cells facilitates female exclusive initiation of XCI, and proper feedback once the inactive X chromosome is established (Figure 4). As XCI was not completely abolished in $Rnf12^{+/-}$ :$Uba1^{+/-}$ cells our results also indicate that additional putative X-encoded XCI activators are present, most likely acting through otherpathways.
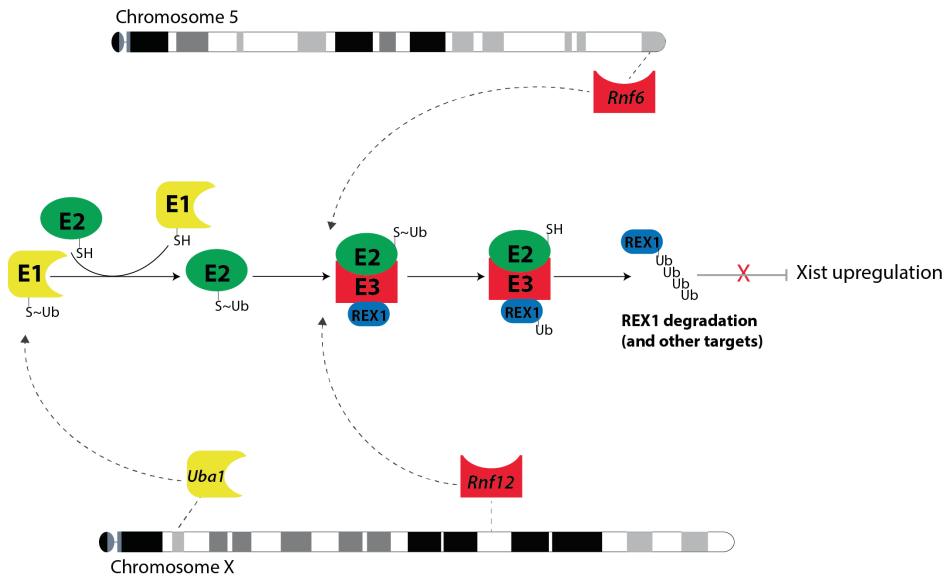


**Figure 4:** Model of the activation cascade involving the E1 activating enzyme UBA1, which allows ubiquitination of the target protein, REX1 and other potential targets, by the E3 ubiquitin ligases RNF12 and RNF6 resulting in proteasomal degradation of REX1. Reduced levels of REX1 relieve repression of *Xist* activating upregulation of *Xist* and initiation of XCI.
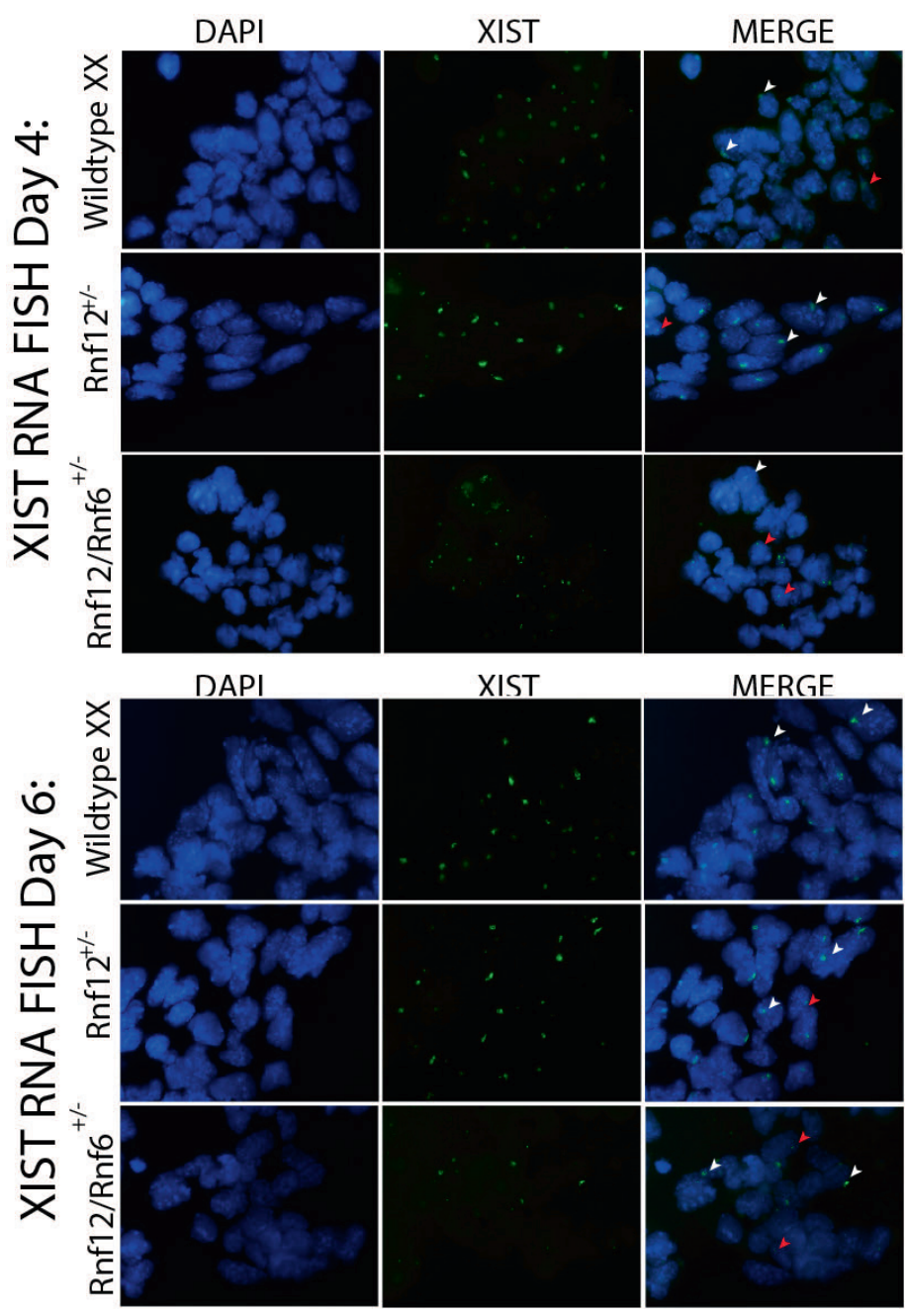
## Supplementary Figures
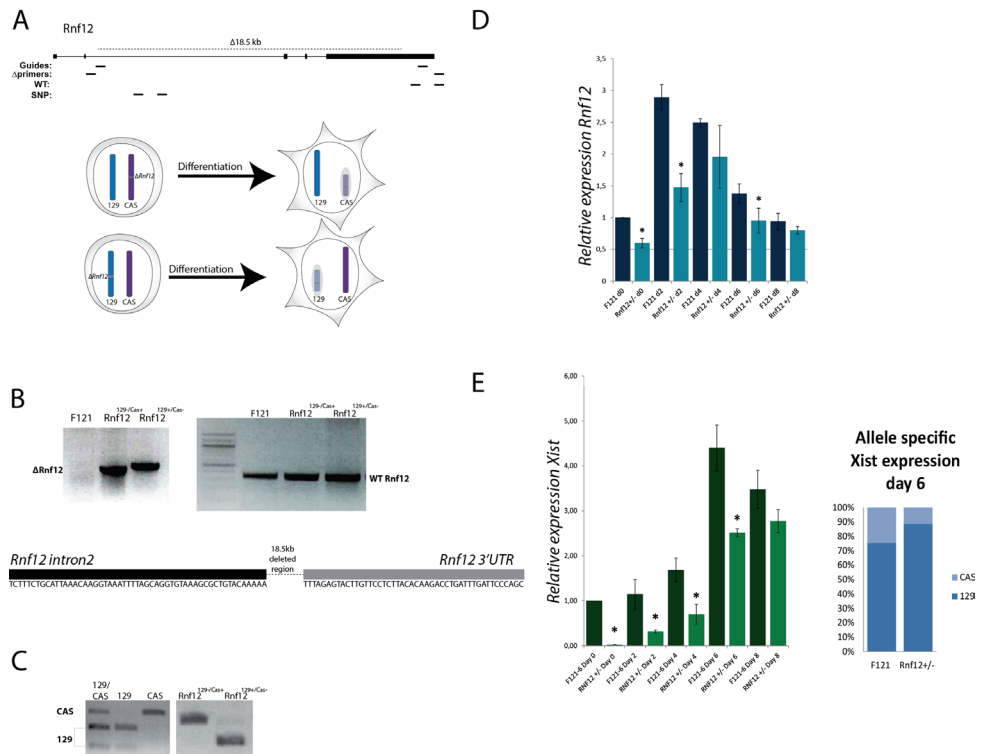


**Supplementary Figure 1**: A) The *Rnf12* gene showing guide RNAs and PCR primers used. A heterozygous deletion of *Rnf12* results in skewing of XCI towards the mutated allele. B) Deletion of *Rnf12* in F121 (129/Cas) female ES cells. Top left: PCR spanning the deletion (ΔRnf12). Top right: wildtype *Rnf12* allele (Wt *Rnf12*). Bottom: Sequencing showing the deletion of 18.5 kb by CRISPR/Cas9 fusion of *Rnf12* exon1 and *Rnf12* intron 3 by NHEJ. C) PCR of a region containing a SNP to determine targeted allele for the two clones; first clone deletion on the 129/Sv allele, second clone deletion on the Cast allele. D) Q-PCR analyzing *Rnf12* expression in F121 (dark blue) and *Rnf12*[+/-] cells (deletion on the 129/Sv allele) (light blue) differentiated for 0, 2, 4, 6 and 8 days. E) qPCR analyzing *Xist* expression in F121 (dark green) and *Rnf12*[+/-] cells (deletion on the 129/Sv allele) (light green) differentiated for 0, 2, 4, 6 and 8 days (left) and an allele specific quantification of *Xist* from the 129 (dark blue) and CAS (light blue) allele in F121 and *Rnf12*[+/-] cells (deletion on the 129/Sv allele).
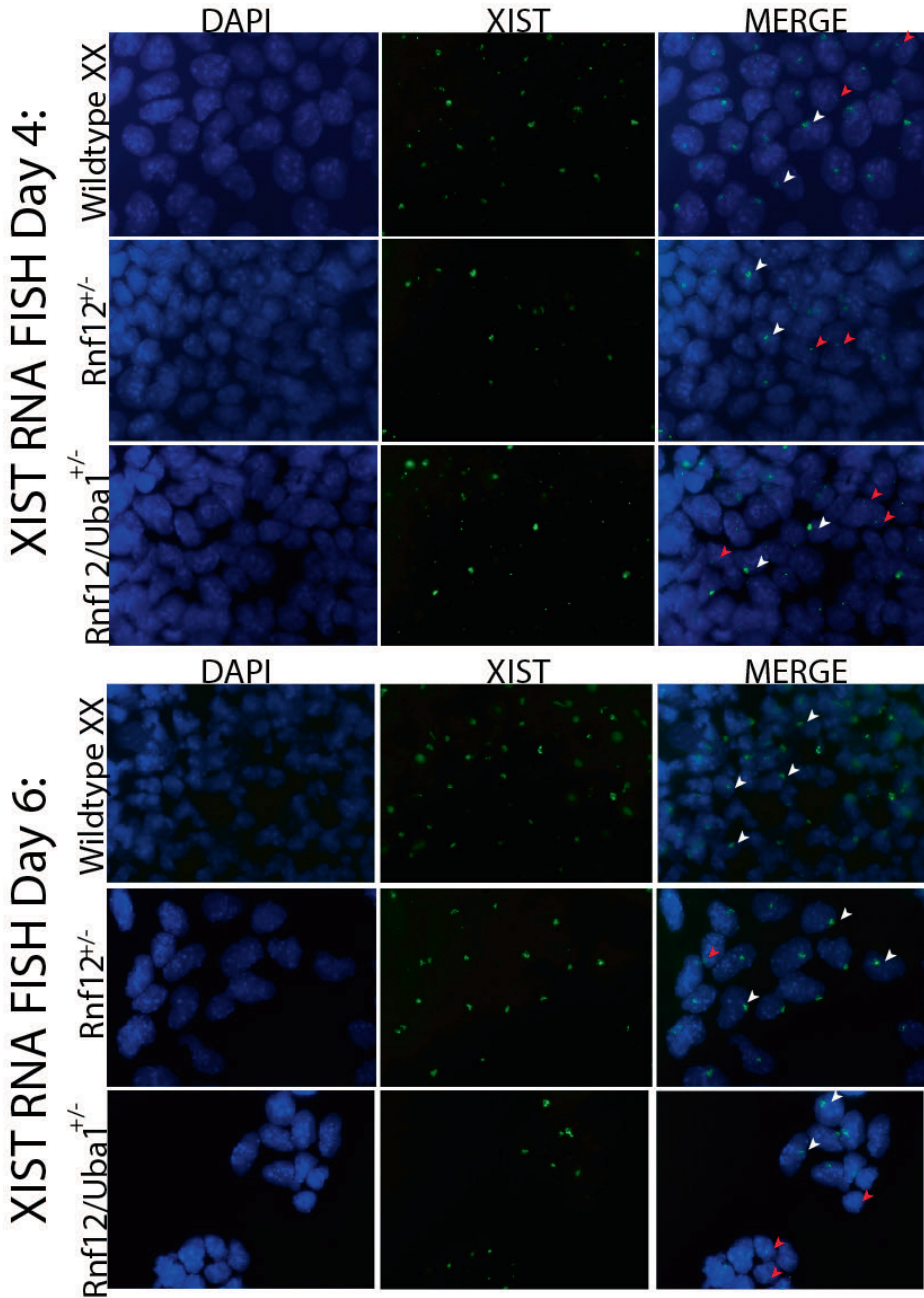
**Supplementary Figure 2**: *Xist* RNA FISH on day 4 and 6 differentiated wildtype, *Rnf12*[+/-] and *Rnf12*[+/-] :*Uba1*[+/-] cells. The right panel merge shows nuclei with *Xist* clouds (white arrows) and *Xist*/*Tsix* pinpoints (red arrows).

**Supplementary Figure 3:** Xist RNA FISH day 4 and 6 differentiated wildtype, *Rnf12*[+/-] and *Rnf12*[+/-] :*Rnf6*[+/-] cells. The right panel merge shows nuclei with Xist clouds (white arrows) and Xist/Tsix pinpoints (red arrows).

**References**

1. Barakat, T.S. et al. (2010) X-changing information on X inactivation. Experimental Cell Research 316 (5), 679-687.

2. Navarro, P. et al. (2008) Molecular coupling of Xist regulation and pluripotency. Science 321 (5896), 1693-5.

3. Navarro, P. and Avner, P. (2009) When X-inactivation meets pluripotency: an intimate rendezvous. FEBS Lett 583 (11), 1721-7.

4. Navarro, P. et al. (2010) Molecular coupling of Tsix regulation and pluripotency. Nature 468 (7322), 457-60.

5. Donohoe, M.E. et al. (2009) The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. Nature 460 (7251), 128-32.

6. Jonkers, I. et al. (2009) RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. Cell 139 (5), 999-1011.

7. Barakat, T.S. et al. (2011) RNF12 activates Xist and is essential for X chromosome inactivation. PLoS Genet 7 (1), e1002001.

8. Gontan, C. et al. (2012) RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature 485 (7398), 386-90.

9. Schulman, B.A. and Harper, J.W. (2009) Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. Nat Rev Mol Cell Biol 10 (5), 319-31.

10. Hershko, A. et al. (2000) Basic Medical Research Award. The ubiquitin system. Nat Med 6 (10), 1073-81.

11. Polge, C. et al. (2013) Deciphering the ubiquitin proteome: Limits and advantages of high throughput global affinity purification-mass spectrometry approaches. Int J Biochem Cell Biol 45 (10), 2136-46.

12. Jones, D. et al. (2002) Functional and phylogenetic analysis of the ubiquitylation system in Caenorhabditis elegans: ubiquitin-conjugating enzymes, ubiquitin-activating enzymes, and ubiquitin-like proteins. Genome Biol 3 (1), RESEARCH0002.

13. Kamath, R.S. et al. (2003) Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature 421 (6920), 231-7.

14. Sönnichsen, B. et al. (2005) Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. Nature 434 (7032), 462-9.

15. Jin, J. et al. (2007) Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. Nature 447 (7148), 1135-8.

16. Pelzer, C. et al. (2007) UBE1L2, a novel E1 enzyme specific for ubiquitin. J Biol Chem 282 (32), 23010-4.

17. Chiu, Y.H. et al. (2007) E1-L2 activates both ubiquitin and FAT10. Mol Cell 27 (6), 1014-23.

18. Murakami, K. et al. (2009) Identification of the chromatin regions coated by non-coding Xist RNA. Cytogenet Genome Res 125 (1), 19-25.

19. Jiao, B. et al. (2013) Functional activity of RLIM/Rnf12 is regulated by phosphorylation-dependent nucleocytoplasmic shuttling. Mol Biol Cell 24 (19), 3085-96.

20. Barakat, T.S. and Gribnau, J. (2015) Generation of knockout alleles by RFLP based BAC targeting of polymorphic embryonic stem cells. Methods Mol Biol 1227, 143-80.

21. Ran, F.A. et al. (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8 (11), 2281-308.

22. Jonkers, I. et al. (2008) Xist RNA is confined to the nuclear territory of the silenced X

chromosome throughout the cell cycle. Molecular and Cellular Biology 28 (18), 5583-94.

23. Monkhorst, K. et al. (2008) X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. Cell 132 (3), 410-21.

24. Cong, L. et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339 (6121), 819-23.

25. (!!! INVALID CITATION !!! [8]).

26. Barakat, T.S. et al. (2014) The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. Mol Cell 53 (6), 965-78.

27. Lévy, N. et al. (2000) The ubiquitin-activating enzyme E1 homologous genes on the mouse Y chromosome (Ube1y) represent one functional gene and six partial pseudogenes. Mamm Genome 11 (2), 164-8.

28. Rodriguez, M.S. et al. (2000) Multiple C-terminal lysine residues target p53 for ubiquitin-proteasome-mediated degradation. Mol Cell Biol 20 (22), 8458-67.

29. Yang, X. et al. (2013) Drosophila Vps36 regulates Smo trafficking in Hedgehog signaling. J Cell Sci 126 (Pt 18), 4230-8.

30. Barrera, S.P. et al. (2015) PKC-Dependent GlyT1 Ubiquitination Occurs Independent of Phosphorylation: Inespecificity in Lysine Selection for Ubiquitination. PLoS One 10 (9), e0138897.

31. Sheldon, A.L. et al. (2008) Ubiquitination-mediated internalization and degradation of the astroglial glutamate transporter, GLT-1. Neurochem Int 53 (6-8), 296-308.

32. Yamada, H. et al. (2012) The high-affinity choline transporter CHT1 is regulated by the ubiquitin ligase Nedd4-2. Biomed Res 33 (1), 1-8.

33. Makhlouf, M. et al. (2014) A prominent and conserved role for YY1 in Xist transcriptional activation. Nat Commun 5, 4878.
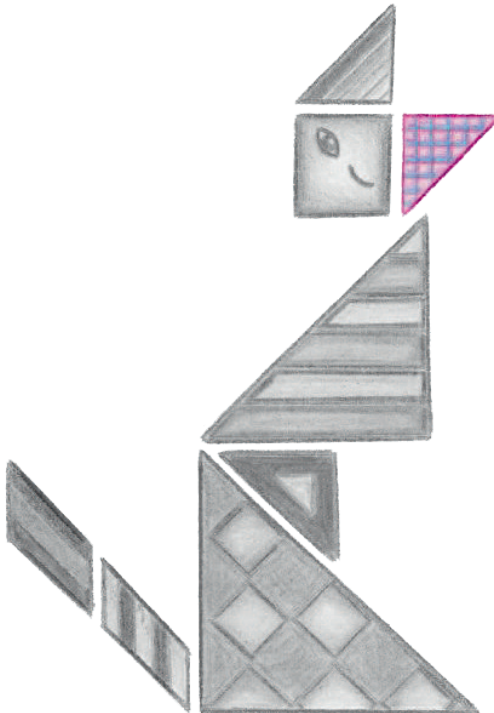
# *Chapter 3*

**Xist and Tsix transcription dynamics is regulated by the X-to-autosome ratio and semi-stable transcriptional states**

# Chapter 3

*Friedemann Loos[1], Cheryl Maduro[1#], Agnese Loda[1#], Johannes Lehmann[2], Gert-Jan Kremers[3] , Derk ten Berge[2], J. Anton Grootegoed[1] and Joost Gribnau1[\*]*

*[1]Department of Developmental Biology, Erasmus MC, University Medical Center, Rotterdam, The Netherlands.*
*[2]Erasmus MC Stem Cell Institute, Erasmus MC, University Medical Center, Rotterdam, The Netherlands.*
*[3] Optical Imaging Center, Erasmus MC, University Medical Center, Rotterdam, The Netherlands.*
*[#] Equal contribution.*
*[\*]To whom correspondence should be addressed. Email: j.gribnau@erasmusmc.nl*

**Abstract**

In female mammals, X chromosome inactivation (XCI) is a key process in the control of gene dosage compensation between X-linked genes and autosomes. *Xist* and *Tsix*, two overlapping antisense transcribed noncoding genes, are central elements of the X inactivation center (*Xic*) regulating XCI. *Xist* up-regulation results in coating of the entire X chromosome by Xist RNA *in cis*, whereas *Tsix* transcription acts as a negative regulator of *Xist*. Here, we generated *Xist* and *Tsix* reporter mouse embryonic stem (ES) cell lines, to study the genetic and dynamic regulation of these genes upon differentiation. Our results revealed mutually antagonistic roles for *Tsix* on *Xist* and vice versa, and indicate the presence of semi-stable transcriptional states of the *Xic* predicting the outcome of XCI. These transcriptional states are instructed by the X to autosome ratio, directed by regulators of XCI, and can be modulated by tissue culture conditions.

**Introduction**

Early during mammalian development one of the two X chromosomes in female cells is transcriptionally inactivated. This X chromosome inactivation (XCI) process is initiated early during development, and is then clonally propagated through a near infinite number of cell divisions. Two X-linked non-coding genes, *Xist* and *Tsix* play a key role in the regulation of XCI in mouse. *Xist* expression is up-regulated on the future inactive X chromosome (Xi) (1, 2), and *cis*-spreading of Xist leads to recruitment of chromatin remodelling complexes that render the X inactive (3, 4). *Tsix* is transcribed anti-sense to *Xist* and fully overlaps with *Xist* (5). *Tsix* transcription and/or the produced Tsix RNA are involved in repression of *Xist* which includes Tsix mediated chromatin changes at the *Xist* promoter (6-9).

       *Xist* and *Tsix* are key components of the *Xic*, the master switch locus that is regulated by XCI-activators and inhibitors of XCI. XCI-activators either activate *Xist* and/or repress *Tsix*, whereas XCI-inhibitors are involved in repression of *Xist* and/or the activation of *Tsix*. In recent years several XCI-inhibitors have been described, including the pluripotency factors NANOG, SOX2, OCT4, REX1, and PRDM14, which provide a direct link between cell differentiation and initiation of XCI (10-13). These factors, and other ubiquitously expressed

XCI-inhibitors including CTCF (14, 15), repress initiation of XCI through binding to multiple gene regulatory elements of *Xist* and *Tsix*. Genetic studies indicate that several of these elements might fulfil redundant roles in the regulation of XCI (16-18).

The X-linked gene *Rnf12* encodes a potent XCI-activator, as overexpression of *Rnf12* results in ectopic initiation of XCI in differentiating transgenic embryonic stem cells (ESCs) (19). The encoded protein RNF12 is an E3 ubiquitin ligase, which targets the XCI-inhibitor REX1 for degradation (20). Degradation of REX1 by RNF12 is dose dependent and two-fold expression of RNF12 in female cells prior to XCI is important for female specific initiation of this process. ChIP-seq studies indicated REX1 binding in both *Xist* and *Tsix* regulatory regions. REX1 mediated repression of *Xist* involves indirect mechanisms including activation of *Tsix*, as well as direct regulation of *Xist* by a competition mechanism, where REX1 and YY1 compete for shared binding sites in the F repeat region in *Xist* exon 1 (21).

*Rnf12* knockout studies revealed a reduction of XCI in differentiating female *Rnf12*[+/-] ES cells, and a near loss in XCI initiation in *Rnf12*[-/-] ES cells (16). However, remained initiation of XCI in a subpopulation of *Rnf12*[+/-] cells also indicates the presence of additional XCI activators, as XCI is not initiated in male cells. This is supported by *in vivo* studies revealing that mice with a conditional deletion of *Rnf12* in the developing epiblast are born alive (22). *Jpx* and *Ftx* have been described as putative XCI-activators (15, 23, 24). Both genes are located in a region 10-100kb distal to *Xist*, and knockout studies indicated that both genes are involved in *Xist* activation. Although transgene studies implicated *Jpx* as a *trans*-activator of *Xist*, recent studies involving a knockout of a region from *Xite* up to the *Xpr* region did not reveal a trans effect, suggesting that the predominant function of *Ftx* and *Jpx* in XCI is the *cis*-activation of *Xist* (25).

Interestingly, examination of the higher order chromatin structure revealed *Xist* and *Tsix* to be located in two distinct neighboring topological associated domains (TADs) (26, 27). Positive regulators of *Xist*, including *Jpx* and *Ftx* are located in the same TAD. Similarly, the *Tsix* positive regulators *Xite*, *Tsx* and *Linx* are located in the *Tsix* TAD, suggesting that these two TADs represent the minimal X inactivation center covering all *cis*-regulatory elements, which are regulated by *trans*-acting XCI-activators and –inhibitors (27-29). During development or ES cell differentiation the XCI-activator concentration in female cells will be two fold higher compared to male cells, which is sufficient to direct female exclusive initiation of XCI. Stochastic initiation of XCI and rapid feedback mechanisms, including the shutdown of *Tsix*, *Rnf12* and other XCI-activators *in cis*, direct a highly efficient XCI process, facilitated by the requirement of loss of pluripotency for initiation of XCI (30).

The overlapping gene bodies of *Xist* and *Tsix* and the mutually antagonistic roles of these two genes hamper clear insights in the regulatory mechanisms that govern *Xist* and *Tsix* transcription. To be able to study the independent pathways directing *Xist* and *Tsix* transcription we have generated *Xist* and *Tsix* reporter alleles, with fluorescent reporters replacing the first exon of *Xist* and/or *Tsix*. Our studies indicate antagonistic roles for both *Xist* and *Tsix*, and show that RNF12 and REX1 regulate XCI through both repression of *Tsix* and activation of *Xist*. Live cell imaging confirms a reciprocal correlation of *Xist* and *Tsix* transcription, but also reveals that their regulation is not strictly concerted and rather stable in time. Interestingly, loss of an X chromosome severely affects the dynamics of both *Xist* and *Tsix* expression, and results in two different cell populations with semi-stable transcriptional states, absent in female ES cells. This indicates a regulatory role for the X:A ratio, regarding the nuclear concentration of X-encoded *trans*-acting factors. Similar semi-stable transcriptional states are observed in female

ES cells grown in medium supplemented with MEK and GSK3 inhibitors, displaying distinct XCI characteristics upon ES cell differentiation. Our findings suggest that XCI-activators are required to install a uniform transcriptional state of the *Xic* that allows proper up-regulation of *Xist* upon ES cell differentiation.

## Results

*Antagonistic roles for Xist and Tsix*

X chromosome inactivation (XCI) is orchestrated by *Xist* and *Tsix*, two non-coding RNA genes with antagonistic roles. *Xist* is essential for XCI to occur in *cis* (31, 32), while *Tsix* is a negative regulator of XCI (6, 33). Analysis of the regulation of *Xist* and *Tsix*, and their relationship during the onset of XCI is hampered by the architecture of the locus. *Tsix* entirely overlaps with *Xist*, is transcribed in antisense direction, and manipulation of one of the two genes always affects the antisense partner. To be able to follow and manipulate the activity of the *Xist* and *Tsix* promoters independently, we generated a series of reporter lines in murine ES cells (Fig. 1a). Exploiting BAC-mediated homologous recombination in polymorphic female 129/Sv-Cast/Ei ES cells (34), exons 1 of *Xist* and *Tsix*, located on the Cast/Eij X chromosome, were replaced with EGFP and mCherry coding sequences, respectively (Supplementary Fig. 1 and Supplementary Fig.2a-c). Expression of the reporters was thus controlled by the endogenous promoters of these two non-coding genes (Fig. 1b). Wild type female 129/Sv-Cast/Ei ES cells show preferential inactivation of the 129/Sv X chromosome in 70% of the cells, attributed to SNPs in regulatory elements that affect the regulation of *Xist* and *Tsix* throughout the ESC differentiation process. We found that the alleles behaved as full *Xist*/*Tsix* knockouts, resulting in complete skewing of XCI, because splice donor sites at the 3'-end of the targeted exons were removed and polyA signals downstream of the reporters terminated transcription (Supplementary Fig. 1 and Supplementary Fig. 2a-c). By successive rounds of targeting followed by cre-mediated removal of selection markers three ES cell lines were obtained: i) Xist promoter-EGFP knock-in (Xist-GFP) , ii) Tsix promoter-mCherry knock-in (Tsix-CHERRY) and iii) double knock-in on the same allele with Xist promoter-EGFP and Tsix promoter-mCherry (Xist-GFP/Tsix-CHERRY). Differentiation of these lines and expression of *Xist* and *Tsix* on the remaining wild-type 129/Sv allele was unperturbed (Supplementary Fig. 3a). Xist-GFP/Tsix-CHERRY cells displayed similar kinetics of Xist cloud formation as wild type cells, albeit with slightly reduced percentages as - probably due to stochastic initiation- expected from a full *Xist* knockout (Supplementary Fig. 3b). FACS analysis of EGFP and mCherry expression for all three ES cell lines showed faithful recapitulation of the behaviour of wild-type *Xist* and *Tsix* during the first days of differentiation (Fig 2a), which was not delayed by a half-life for EGFP and mCherry that ranged in the order of 11-14 hours (Supplementary Fig. 3c). As expected, comparison of Xist-GFP/Tsix-CHERRY ES cells, which allows independent tracking of *Xist*/*Tsix*, with Xist-GFP ES cells shows EGFP de-repression in undifferentiated cells when *Tsix* is deleted in *cis* (Fig.2b). Comparison of Xist-GFP/Tsix-CHERRY with Tsix-CHERRY revealed delayed down-regulation of the mCherry reporter in the double knockin cell line (Fig. 2c), indicating a role for *Xist* in silencing *Tsix*. The delay in mCherry down-regulation cannot be attributed to differences in mCherry expression/*Tsix* promoter activity between the Xi (in Tsix-CHERRY line) and the Xa (in Xist-GFP/Tsix-CHERRY line), suggesting that *Tsix* down-regulation on the future Xa is compromised upon ES cell differentiation in the absence of *Xist* (Supplementary Fig.3a). To verify that this effect is not due to the deletion of any DNA elements involved in the repression of *Tsix* in Tsix-CHERRY, we performed two-colour RNA FISH to distinguish between Xist and Tsix transcripts in

differentiating ES cells. Three independent *Xist* deletion lines, Xist-GFP, Xist[1lox] and ptet-Xist, harbouring an insertion of a doxycycline inducible promoter replacing the

endogenous Xist promoter ((35) and A. Loda, unpublished), show persisting *Tsix* transcription from Xa compared to wild-type cells (Fig. 2d-e). Taken together, these results show that *Xist* and *Tsix* display antagonistic roles, directly influencing the expression level of each other on the Xa during the early phases of ES cell differentiation. It also highlights the need to investigate the dynamics of their early genetic regulation on the uncoupled allele in Xist-GFP/Tsix-CHERRY.
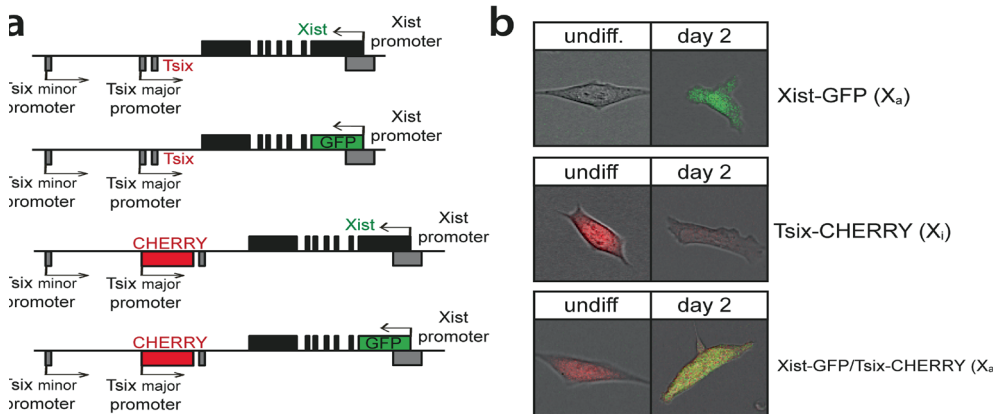


**Figure 1: Generation of the reporter alleles.**(a) Map of the *Xist*/*Tsix* locus showing design of the reporter cell lines. (b) Exemplary pictures of undifferentiated and differentiated cells.

*Dynamics of regulation of the Xic by live cell imaging*

To further analyze the dynamics of *Xist* and *Tsix* regulation, we performed live cell imaging of differentiating Xist-GFP/Tsix-CHERRY cells for extended periods of time by confocal microscopy (Fig. 3a). The integrated EGFP and mCherry fluorescence intensities (FI) of entire single cells were measured, resulting in semi-oscillating patterns due to accumulation of fluorescent reporters followed by dilution upon cell division (Fig. 3b). For each cell cycle, the slope of the linear regression of integrated FI over time gives an estimate of the activity of the *Xist* and *Tsix* promoters. Binning cell cycles with low, medium and high increase in EGFP FI into groups and comparing the corresponding values for mCherry confirms a concerted anti-correlated regulation independent of antisense transcription, with EGFP being up-regulated before down-regulation of mCherry (Fig. 3c). Next, we set a threshold for mean EGFP FI to estimate at which point EGFP FI rises above background noise. Low values for the slope of mCherry before, and high values after *Xist* activation argue that, in spite of concomitant anti-correlated regulation, *Xist* and *Tsix* are independently and stochastically regulated (Supplementary Fig. 4a).
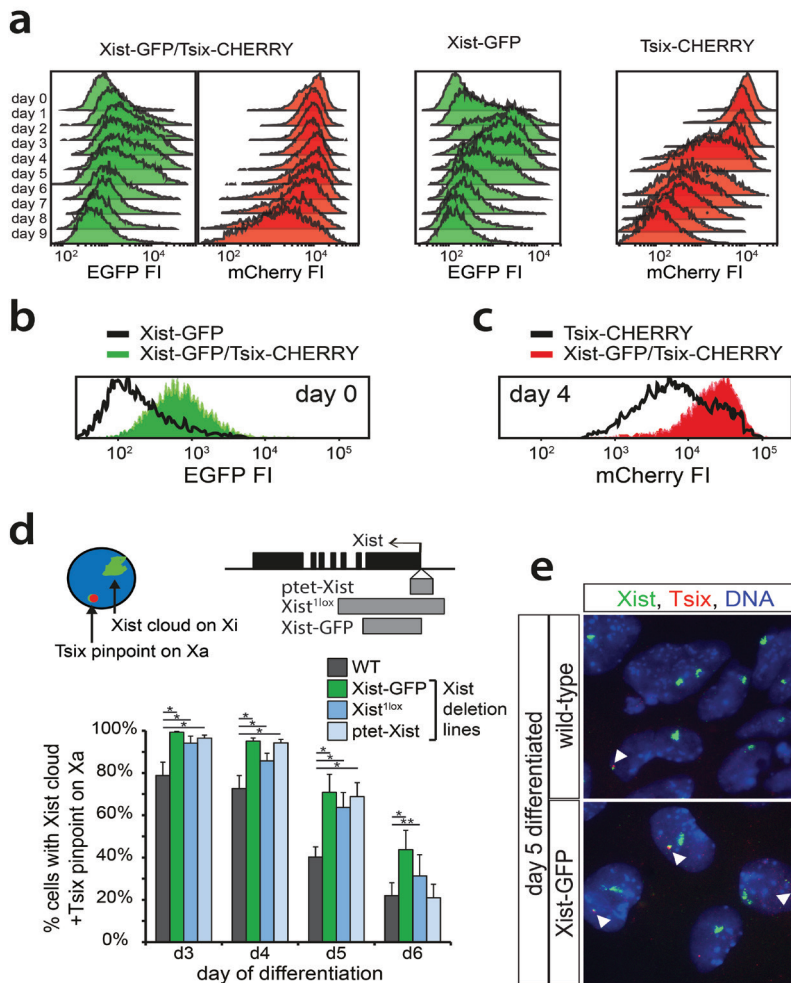
**Figure 2: *Xist* and *Tsix* Reporter Lines Reveal Antagonistic Roles for *Xist* and *Tsix*.** (a) Histograms of EGFP (green) and mCherry (red) FI distribution as determined by FACS analysis. Days 0 through 9 of differentiation are depicted for Xist-GFP, Tsix-CHERRY and Xist-GFP/Tsix-CHERRY. (b and c) Histograms of EGFP (b) and mCherry (c) FI distribution as determined by FACS analysis. Black outlines represent single knockin cell lines Xist-GFP undifferentiated (b) and Tsix-CHERRY at day four of differentiation (c). Solid colors represent FI distributions for Xist-GFP/Tsix-CHERRY. (d) Quantification of two-color RNA FISH detecting Xist and Tsix transcripts at different time points of differentiation. The proportion of cells with an Xist cloud, identifying the Xi, and a Tsix pinpoint from the Xa, is shown. Dark blue bars represent wild type female ES cells, green bar Xist-GFP line and light blue bars two independent *Xist* deletion lines. Top right panel shows exon-intron structure of *Xist*, grey bars indicate the deleted region of the respective deletion line. Error bars indicate 95% confidence interval, n > 150 for all time points and cell lines, asterisks indicate *P* < 0.05 (*) or P < 0.1(**) by two-proportion z-test. (e) Xist/Tsix

two-colour RNA-FISH of wild type and Xist-GFP cells. Green probe detects Xist and Tsix, red probe detects only Tsix.

Xi is identified by presence of Xist cloud, Tsix transcription from Xa by presence of separate two-color pinpoint in the same nucleus.

To unravel the relationship between activation of *Xist* and *Tsix* and establishment of the Xi we introduced a mTagBFP2-Ezh2 fusion gene into Xist-GFP/Tsix-CHERRY ES cells (Supplementary Fig 4b-c). Since we were not able to continually follow high numbers of cells until an Xi domain appeared, we instead scored cells at different time points of differentiation (Fig. 3d). The results show that high GFP levels almost never concur with an EZH2/Xi domain. RNA FISH analysis on day 2 differentiated FACS sorted EGFP low, intermediate, high, and very high cells, however, demonstrated that both *Xist* promoters become activated and that EGFP up-regulation and XCI initiation correlate (Fig. 3e). At day 3 the EGFP[high] and EGFP[veryhigh] FACS sorted fractions of cells contained less Xist clouds than the EGFP[intermediate] fraction. This suggests that EGFP[high] and EGFP[veryhigh] cells down-regulate EGFP before Xist clouds become detectable, but also indicates the presence of a sub-population of cells that strongly and consistently activate Xist-GFP without up-regulation of *Xist* on the wild-type X chromosome.

Live cell imaging also enabled us to follow single cells through mitosis and monitor the fate of daughter cells through successive rounds of cell division. Plotting the slope of EGFP/ mCherry FI for each generation confirms the previously described anti-correlation of *Xist* and *Tsix* activity for each given cell (Fig. 3f). Moreover, daughter cells display strikingly similar patterns of *Xist* and *Tsix* promoter activities, indicating that they generally follow the same fate. This implies that switches of *Xist* and *Tsix* activity occur rarely or slowly and that once a certain transcriptional state is established it is stably transmitted through cell division and relatively resistant to changes or reversal. Taken together, live cell imaging and fate mapping suggest that on an uncoupled allele, *Xist* and *Tsix* are antagonistically regulated in a developmentally concerted manner, even though up- and down-regulation of both genes per se are independent and probably stochastic.
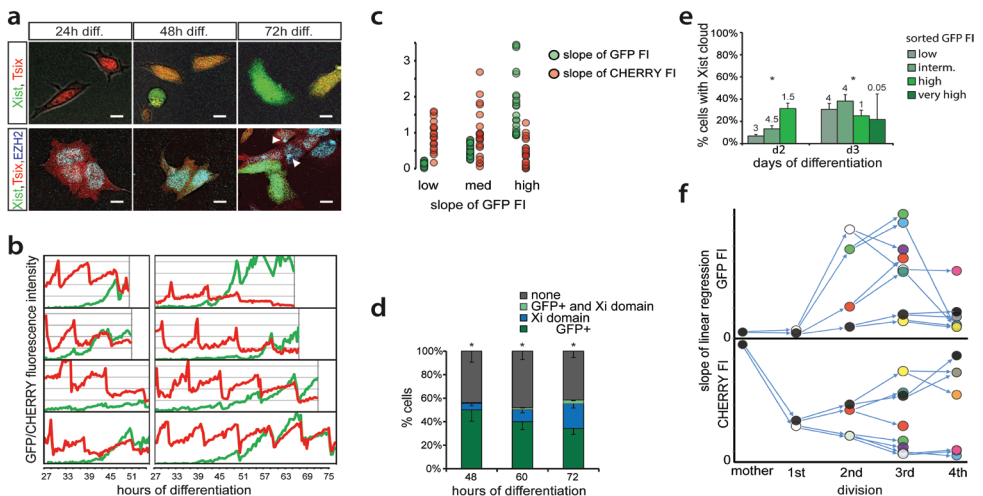


**Figure 3: Time-Lapse Imaging of Live Cells.** (a) Exemplary pictures of   Xist-GFP/Tsix-CHERRY

cells (top panels) and Xist-GFP/Tsix-CHERRY+Ezh2-Flag cells (bottom panels) taken at different time points of differentiation during time-lapse imaging. Scale bar is 5 μm. (b) Whole cell integrated FI values of EGFP (green) and mCherry (red) plotted over time for several exemplary cells during time-lapse imaging. (c) Linear regression of FI over time for each cell cycle was performed. Slope of linear regression as a proxy for promoter activity is plotted. Values for EGFP FI are binned into low (lowest tercile), medium (intermediate tercile) and high (highest tercile), and the corresponding values for mCherry are plotted right next to it. (d) Quantification of presence of mTagBFP2-Ezh2 focus/ Xi domain and/or high levels of EGFP at different time points of differentiation in Xist-GFP/Tsix-CHERRY+ mTagBFP2-Ezh2 cells. Error bars indicate 95% confidence interval, n =162 for 48 hours, n = 215 for 60 hours and n = 277 for 72 hours. (e) Day two and three differentiating Xist-GFP/Tsix-CHERRY cells were FACS-sorted into EGFP low, intermediate, high and very high fractions. Graphs show quantification of Xist RNA FISH in these fractions. The number on top of each fraction represents their relative abundance within the population before sorting. Error bars indicate 95% confidence interval, n > 250 for all time points and fractions. (f) Pedigree of an exemplary cell followed through four cell divisions. In top panel, slope of linear regression as described in (c) is shown for EGFP FI. In lower panel, slope of linear regression is shown for mCherry. Same colored dots represent the same cell, thus values for EGFP in top panel and mCherry in lower panel. Arrows connecting dots indicate mother cell to daughter cell relationship. Asterisks in (d) and (e) indicate $P < 0.05$ (*) as calculated by chi-square test.

*Effects of activators and inhibitors on XCI*
*RNF12 functions as a trans-activator of XCI (19) by targeting REX1, a repressor of XCI, for proteasomal degradation (20).* Previous work has indicated that REX1 might have a dual role in the activation of XCI by activating *Tsix* and repressing *Xist* (12, 16, 20). To dissect this XCI regulatory network and determine the role of these factors in the regulation of *Xist* and *Tsix* in ES cell lines harboring uncoupled *Xist/Tsix* alleles , we introduced Rnf12 and Rex1 transgenes into the three knock-in cell lines. Clones chosen for analysis consistently over-expressed *Rnf12* and *Rex1* two- to three-fold as compared to wild-type cells (Supplementary Fig. 5a). FACS analysis of differentiating Xist-GFP/Tsix-CHERRY ES cells showed that Rnf12 and Rex1 transgenes had a clear effect on the EGFP and mCherry reporters (Fig. 4a, b). REX1 strongly repressed the *Xist* and activated the *Tsix* promoter. Conversely, *Rnf12* overexpression resulted in increased activation of EGFP and reduced mCherry expression . This was also evident from quantitative analysis of RNA levels by qPCR. In the Xist-GFP/Tsix-CHERRY line, both *Xist* and EGFP were up-regulated by an Rnf12 transgene and down-regulated by a Rex1 transgene, while the opposite effect was observed for *Tsix* and mCherry (Supplementary Fig. 5b). Since we monitored the uncoupled allele in a comparatively well-preserved genomic context, we can exclude any indirect effects due to interference from the corresponding antisense partner. In the presence of the antisense partner, in the single knock-in Xist-GFP and Tsix-CHERRY lines, we observed that the effect of *Rnf12* and *Rex1* overexpression was strongly attenuated (Supplementary Fig. 5b). This finding indicates that antisense transcription or the antisense transcript represses transcription of the *Xist* and *Tsix* promoter located on the opposite strand, and that a balanced allele might be necessary for proper integration of regulatory signals.

The major difference between female cells that undergo XCI and male cells that do not is the X:A ratio. To better investigate the effects of changes in this X:A ratio on *Xist* and *Tsix* expression, we screened Xist-GFP/Tsix-CHERRY for subclones that had lost the wild-type 129 X chromosome by using an X-linked RFLP (Supplementary Fig. 5c). These XO lines showed a stable karyotype (Supplementary Fig. 5d), but comparison of these XO lines (XGTC-XO),

with the XX Xist-GFP/Tsix-CHERRY double knock in ES cell line indicated that the dynamics of both GFP and mCherry expression during ES cell differentiation was severely affected by loss of the wild type X chromosome (Fig.4c top panel, and 4d). In addition, the XO cells are present in two distinct mCherry-high and mCherry-low populations. This bimodal mCherry distribution was also observed for the XY Tsix-CHERRY only knock-in cells (Fig. 4c, bottom panel), indicating that the dynamics of these states is affected by the X:A ratio.

To test whether these effects are solely related to the *Rnf12* copy number we generated three independent XX Xist-GFP/Tsix-CHERRY Rnf12[+/-] heterozygous knockout cell lines where *Rnf12* was mutated on the 129/Sv allele (Supplementary Fig. 5e,g,h). Examination of these ES cell lines during differentiation, shows a severe reduction in upregulation of the Xist-GFP reporter allele (Supplementary Fig. 5i). However, the two sub-populations found in undifferentiated XGTC-XO, and Tsix-Cherry only ES cells are absent in Xist-GFP/Tsix-CHERRY Rnf12[+/-] cells (Fig. 4e), which show a similar FACS profile compared to Rex1 transgenic Xist-GFP/Tsix-CHERRY cells. A decrease in *Rnf12* levels, therefore, does not explain the reduced mCherry expression level throughout ES cell differentiation observed in XGTC-XO ES cells. In addition, comparison of *Tsix* RNA expression levels in male and female ES cell lines by qPCR analysis confirmed that lower levels of *Tsix* RNA are present in male ES cells (Supplementary Fig. 5f). These findings indicate that more X-encoded factors are involved in the regulation of XCI, and that the X:A ratio also directs the dose dependent activation of *Tsix*.
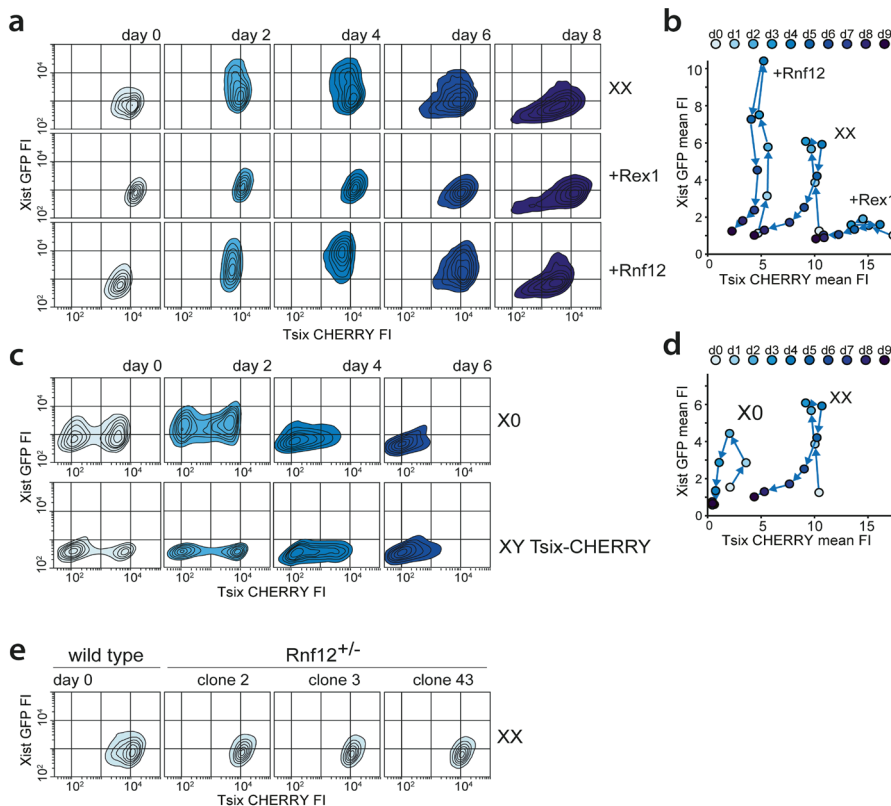
**Figure 4: Impact of the RNF12-REX1 regulatory network on *Xic* regulation.** (a) Contour plots of FACS analysis showing EGFP and mCherry FI at different time points of differentiation for Xist-GFP/Tsix-CHERRY (XX), Xist-GFP/Tsix-CHERRY+ Rex1 (+Rex1) and Xist-GFP/Tsix-CHERRY+Rnf12 (+Rnf12). For all experiments 100.000 cells were analyzed per time point. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events (logarithmic scale). (b) Same as in (a), but mean FI for EGFP and mCherry is plotted (linear scale). (c) Contour plots of FACS analysis showing EGFP and mCherry FI at different time points of differentiation for the XGTC-XO (top panels) and XY Tsix-CHERRY (bottom panels) lines. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events logarithmic scale). (d) Same as in (c), but mean FI for EGFP and mCherry is plotted for the XGTC-XO line (linear scale). (e) Contour plots of FACS analysis showing EGFP and mCherry FI in undifferentiated Xist-GFP/Tsix-CHERRY (XX) cells and for clones 2, 3, 43 of Xist-GFP/Tsix-CHERRY Rnf12$^{+/-}$ ES cell lines. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events logarithmic scale).

*Semi-stable transcriptional states of the Xic predict outcome of XCI*

The striking bimodal mCherry distribution of XGTC-XO ES cells indicates that in similar proportions of cells the *Tsix* promoter is either on or off. These two states switch, if at all, very slowly. This is evident from the presence of two distinct populations considering the half-life of mCherry, and the fact that recovery of the mixed population of mCherry positive and negative cells after FACS-sorting of one of the populations does not occur within two weeks (Fig. 5a). Moreover, seeding cells at a low density results in homogeneous colonies of either mCherry negative or positive cells (Fig. 5b). Also differentiation of sorted mCherry positive and negative XGTC-XO ES cells did not lead to an increase in switching between states (Fig. 5a).

Staining for the differentiation marker CD31 and alkaline phosphatase activity, specific for undifferentiated embryonic stem cells, did not reveal differences in cell differentiation between the different cell populations (Supplementary Fig. 6a). Also, bisulfite sequencing analysis of the *Tsix* promoter did not reveal differences between the mCherry high and low populations (Supplementary Fig. 6b). To find the basis of the difference between the two populations, RNA sequencing was performed on FACS-sorted mCherry positive and negative XGTC-XO cells. This analysis indicated that both populations have highly similar expression profiles (Pearson correlation coefficient (Pearson) r=0.9832; Supplementary Fig. 6c), and confirmed that expression of the pluripotency factors was indifferent between the two cell populations (Pearson r=0.999). Interestingly, close examination of expression levels of genes located in the Xic indicated several genes for which the expression level correlated or anti-correlated with *Tsix*-promoter driven mCherry expression (Pearson r=0.83, Fig. 5c). These differences were most prominent for genes located within the Tsix TAD (Pearson r=0.34), and suggest that the on-off switch of the *Tsix* promoter is based on distinct epigenetic states and/or the spatial conformation of the *Xic*.

Interestingly, in 2i+LIF conditions, that force ES cells to adopt a more naïve state, the two distinct XY Tsix-mCherry and XGTC-XO ES cell populations became uniform (Fig 5d, and data not shown), suggesting that tissue culture conditions have a severe impact on the transcriptional states of the *Xic*. Indeed, *Xist* qPCR analysis of wild type 129/Sv-Cast/Eij female ES cells indicates that *Xist* is more repressed in 2i versus serum+LIF conditions, but that during ES cell differentiation up-regulation of *Xist*, and skewing of XCI are indifferent between the two culture conditions (Supplementary Fig. 6d-e). Nevertheless, the 2i+LIF conditions did impact on the transcriptional states of the Xic in female Xist-GFP/Tsix-CHERRY cells now displaying two separable mCherry populations, absent in serum+LIF growth conditions (Fig.

6a). Again, after sorting mCherry low and high cells, recovery of the mixed population of cells did not occur in 2i+LIF or differentiation conditions in a time frame of two weeks (Fig. 6a). Intriguingly, the mCherry low population activates the Xist promoter-driven EGFP reporter much more strongly than the mCherry high population (Fig. 6a). This suggests that the potential to initiate XCI is determined by the state of the *Xic* already before differentiation. Xist RNA FISH performed on day 2 of differentiation on these cells moreover indicates that the mutant and wild-type allele co-exist with a high probability in the same state, because cells from the mCherry low population showed higher percentages of cloud formation (Fig. 6b). We also transferred the Xist-GFP/Tsix-CHERRY ES cells to serum+LIF to trigger a "primed" state (36). After 14 days culturing in this serum+LIF condition mCherry levels stay mostly stable, and preferential up-regulation of the Xist-GFP in the mCherry-medium cells is still observed (Fig. 6d). Similar to our findings with XGTC-XO ES cells, RNA sequencing of undifferentiated mCherry-low and -high Xist-GFP/Tsix-CHERRY ES cells, revealed marked differences between the two cell populations, of genes located within the Xist and Tsix TADs (Fig. 6e). Allele specific expression analysis of *Rnf12* showed increased *Rnf12* expression in mCherry low cells but no preference for expression from the 129/Sv or Cast/Eij alleles, indicating that transcriptional states are synchronized between the wild type and reporter alleles (Fig. 6c). Stabilization of these transcriptional states might be accomplished by feedforward and feedback loops involving *Rnf12* and *Rex1*. To test this we analysed wild type and *Rex1* and *Rnf12* transgenic Xist-GFP/ Tsix-CHERRY ES cells cultured in 2i+LIF. FACS analysis revealed that *Rex1* over-expression forces cells to adopt the mCherry-high state whereas *Rnf12* does the opposite, indicating that different transcriptional states are stabilized in trans by trans-acting factors (Fig. 6f,g). These findings argue that the on-off switch of the *Tsix* promoter is based on distinct epigenetic states and/or the spatial conformation of the *Xic* and also explains the observed *Xist* promoter activation on both alleles in the mCherry low population by increased levels of RNF12 (Fig. 6b,c). Our findings highlight the presence of differential epigenetic states, affected by extrinsic and intrinsic factors, capable of providing stable on-off switches for genes involved in XCI.
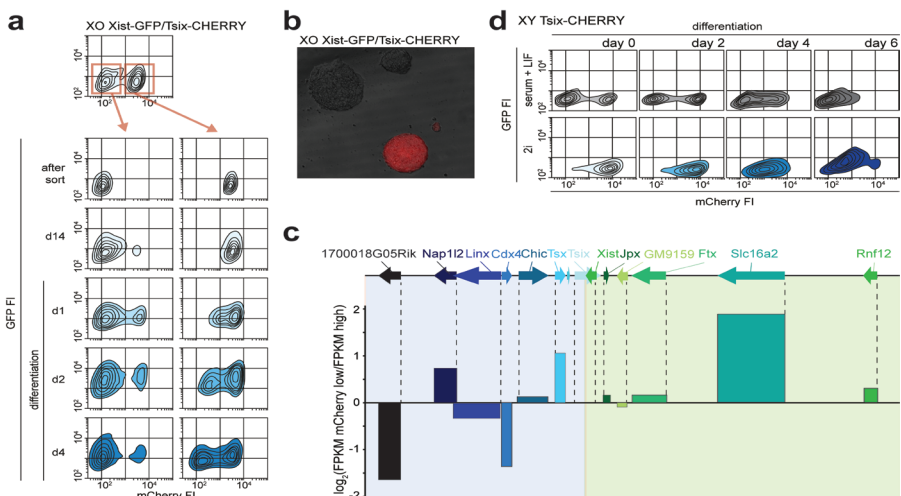


**Figure 5: Two Stable States of the *Xic* in XO cells.** (a) Contour plots of FACS analysis showing EGFP and mCherry FI for XGTC-XO line. Top panel depicts original population with bimodal mCherry

distribution, underneath the sorted mCherry low and high populations (as indicated by red bounding box and arrows) are shown directly after the sort, 14 days after the sort and upon differentiation. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events. (b) XGTC-XO ES cell clones after single cell plating. (c) Expression levels of genes located in the *Xic* as determined by RNA sequencing of XGTC-XO mCherry low and high populations. Top indicates location of genes along the X chromosome, bars show $\log_2$(FPKM mCherry low/FPKM mCherry high). (d) Contour plots of FACS analysis showing EGFP and mCherry FI for the XY Tsix-CHERRY ES line grown in serum+LIF (top panels, as shown in Fig. 4C) and 2i+LIF (bottom panels) conditions, prior to and at different timepoints after differentiation.
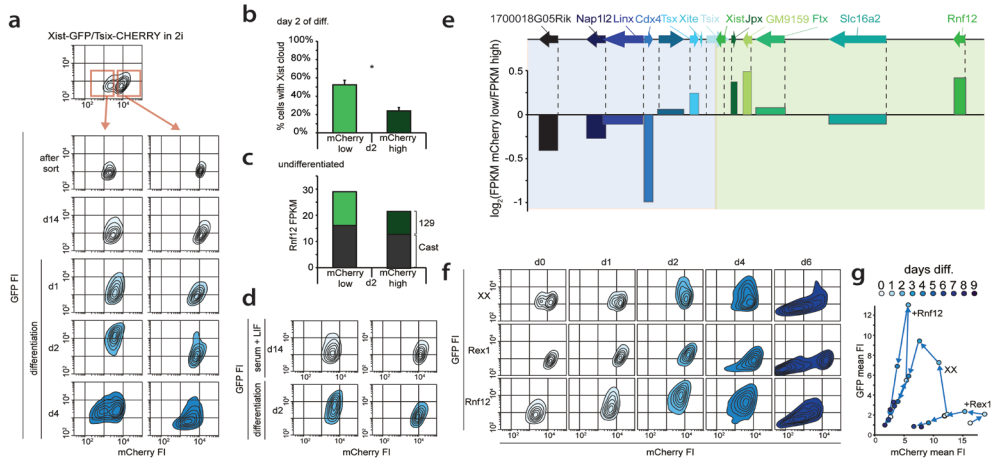
**Figure 6: Two Stable States of the *Xic* Predict XCI Potential in XX cells.** (a) Contour plots of FACS analysis showing EGFP and mCherry FI for the Xist-GFP/Tsix-CHERRY ES cell line grown in 2i+LIF. Top panel depicts original population with bimodal mCherry distribution, underneath the sorted mCherry low and high populations (as indicated by red bounding box and arrows) are shown directly after the sort, 14 days after the sort and upon differentiation. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events. (b) Quantification of Xist RNA FISH in female Xist-GFP/Tsix-CHERRY cellas at day two of differentiation after sorting mCherry low and high populations. Error bars indicate 95% confidence interval, n=313 for mCherry low and n=305 for mCherry high populations. Asterisk indicates $P < 0.05$ by two-proportion z-test. (c) Allele specific RNA expression analysis by RNA sequencing. Shown is the FPKM value and allele specific expression ratio (129/Sv:green, Cast:shaded). (d) Contour plots of 2i+LIF mCherry high and low Xist-GFP/Tsix-CHERRY ES cell populations 14 days after change from 2i+LIF to serum+LIF conditions (top panels), and two days later after start of differentiation. (e) Expression levels of genes located in the *Xic* as determined by RNA sequencing of Xist-GFP/Tsix-CHERRY mCherry low and high populations. Top indicates location of genes along the X chromosome, bars show $\log_2$(FPKM mCherry low/FPKM mCherry high). (f) Contour plots of FACS analysis showing EGFP and mCherry FI at different time points of differentiation for Xist-GFP/Tsix-CHERRY (XX), Xist-GFP/Tsix-CHERRY+ Rex1 (+Rex1) and Xist-GFP/Tsix-CHERRY+Rnf12 (+Rnf12) ES cells grown in 2i+LIF conditions. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events. (g) Same as in (F), but mean FI for EGFP and mCherry is plotted.

## Discussion

In mouse *Xist* and *Tsix* represent the key *cis*-regulatory players in proper execution of XCI. This sense-antisense transcribed gene couple fulfils antagonistic roles in the regulation of XCI, with the action of *Tsix* restricted locally as a negative regulator of *Xist*, whereas *Xist* acts over large distances silencing genes along the X chromosome. Our study confirms the repressive role of *Tsix* on *Xist* expression, although this effect appears most pronounced in undifferentiated ES cells. Xist upregulation is often interpreted to be the consequence of mono-allelic Tsix downregulation (33, 37). Interestingly, our study indicates that *Xist* acts locally facilitating the shutdown of *Tsix*, not only on the Xi but also on the future Xa, as we observed sustained *Tsix* expression comparing three different Xist knockout ES cell lines with wild type cells during ES cell differentiation. These findings indicate that *Xist* and *Tsix* are in a constant interplay, silencing of Tsix involves Xist dependent and independent mechanisms. Although this effect is likely mediated through Xist RNA instructed local recruitment of chromatin remodeling complexes, we cannot exclude a transcriptional interference mechanism to be involved.

Live cell imaging of XX cells harboring Xist/Tsix fluorescent reporters indicated that also in the absence of sense-antisense overlapping transcription expression of *Xist* and *Tsix* is anti-correlated. Nevertheless, this anti-correlation is not strict, and we find *Xist* up-regulation prior to *Tsix* down-regulation and vice versa. This suggests a mechanism of stochastic expression of both genes, where initiation of *Xist* expression is increased during differentiation until a level is reached which is sufficient to spread *in cis*, leading to *Tsix* silencing thereby providing a feed forward loop facilitating further *Xist* transcription initiation, accumulation and spreading.

The present live cell imaging studies indicate that regulation of *Xist* and *Tsix* is rather stable in time and that *Xist* and *Tsix* expression in daughter cells preferably adopt the same fate. This might be related to *Xic* locus intrinsic factors or to stable expression profiles of regulators of XCI. The studies involving XGTC-XO reporter cells grown in serum+LIF conditions and XX Xist-GFP/Tsix-CHERRY reporter cells cultured in 2i supplemented medium indicate that genes located within the *Xist* and *Tsix* TADs adopt different transcriptional fates, favoring expression of a subset of genes. These distinct transcriptional fates might represent semi-stable states of higher order chromatin structure that can be propagated through many cell divisions, and are different from reported X chromosome wide cohesion differences (38). A recently developed polymer model predicted such different states of higher order chromatin structure (39). These transcriptional states are maintained independent of *Tsix* promoter methylation (Supplementary Fig. 6b), and are likely independent of DNA methylation in general, which is nearly absent in 2i conditions (40). Switching between the different transcriptional states rarely occurs, but is more frequently observed upon ES cell differentiation, which might be related to the reported increased chromatin dynamics during the early stage of ES cell differentiation (37), possibly provoked by changes in regulators of the XCI process. In serum+LIF conditions no distinct sub-populations of XX ES cells are observed suggesting that switching between states happens at a much higher frequency, with a shifted equilibrium constant or that all cells adopt one and the same transcriptional state. Increased mobility of the Xic has also been reported during early ES cell differentiation and might reflect switching of transcriptional states described in this study (37). This does not necessarily mean that different transcriptional states as represented by the Tsix-mCherry-low and -high subpopulations are intrinsically stable. Rather, we favour a scenario in which chromatin conformation is fluctuating but exists preferentially in one or the other conformation (Fig. 7a). Our differentiation studies indicate that this

transcriptional state in XX ES cells under serum conditions responds more homogeneously to differentiation cues than ES cells grown in 2i conditions. Nevertheless, also in serum+LIF differentiated ES cells we observe cells that do not accumulate a PRC2 domain on the Xi, and continue to express the Xist-GFP reporter at high levels suggesting that these cells are locked in an epigenetic state that does not allow initiation of XCI on the wild type X chromosome. The results obtained with the 2i cells indicate that these transcriptional states can even predict the responsiveness of the *Xic* to XCI regulators prior to the initiation of this process, and that many cells do not initiate XCI at all. As Tsix-mCherry levels in serum+LIF are equal to the Tsix-mCherry high subpopulation in 2i conditions that is more refractory to XCI initiation, this indicates that different transcriptional states exist that cannot be fully separated by Tsix levels only.

Interestingly, the present RNA-FISH studies on sorted 2i populations indicate cross talk between the *Xic's* with respect to this responsiveness, revealing significantly more cells initiating XCI on the wild type X in Tsix-mCherry low than high cells. This difference appears to be related to differences in the expression level of activators and inhibitors of XCI, coordinated with the transcriptional state of the *Xic*. A switch to a transcriptional state with a higher *Rnf12* transcription level on one allele will result in increased RNF12 mediated turnover of REX1 and *Xist* activation. In general, several pluripotency factors act as repressors of *Rnf12* (13, 42), and also reduced REX1 levels may therefore facilitate switching to a transcriptional state with higher *Rnf12* expression on the second X chromosome, providing a feed forward loop fixing in the transcriptional state (Fig 7b,c). Our results might explain previous results obtained with differentiating ES cells grown in 2i conditions, showing a high number of cells initiating XCI on both X chromosomes (43), and indicate that the 2i culture conditions are suboptimal for studying the XCI process.

The reporter lines generated for this study nicely recapitulate XCI. Nevertheless, RNA and protein stability, and differences in detection levels, clearly affect the our measurements. In our studies we removed exon 1 completely, as a previous attempt to generate a Xist-EGFP reporter allele failed because remaining *Xist* sequences prevented nuclear export of the RNA (Sado et al Mol Cell 2005). Removal of regulatory sequences and introduction of the reporters themselves might therefore have impacted on the regulation of *Xist* and *Tsix*. Previous work has implicated RNF12 in the regulation of random XCI by activation of *Xist* and repression of *Tsix*. ChIP analysis indicated two prominent REX1 binding peaks in both the *Xist* and *Tsix* intragenic regulatory elements. REX1 mediated repression of *Xist* involves competition of REX1 and YY1 binding for the same binding sites in the F-repeat region of Xist, YY1 being an activator of *Xist* expression (21). Despite the removal of this F-repeat region from our reporter allele, we still find clear effects of *Rnf12* and *Rex1* over-expression on *Xist* regulation, indicating a role for alternative binding sites, such as found in the *Xist* promoter, or indirect mechanisms to be instructive in *Xist* regulation. Our findings are supported by previous studies also showing an effect of changes in *Rnf12* expression on luciferase reporters linked to the minimal *Xist* promoter (16, 20). Although our results suggest a prominent role for the RNF12-REX1 axis in the regulation of XCI, the effects on *Xist* and *Tsix* transcription where much more prominent in the absence of overlapping transcription, indicating that activation of XCI requires a very balanced *cis-* and *trans-*acting environment for proper regulation. In addition, the severely reduced dynamics of Xist-GFP and Tsix-mCherry expression in XO reporter cell lines during ES cell differentiation, also indicates that more X-linked factors are involved in the regulation of XCI. Interestingly, these factors also boost *Tsix* expression, which might be a requirement for proper execution of a mutual exclusive

XCI process, providing a stable binary switch. XCI-activators therefore seem to act at two different levels, first by bringing the *Xic* to a transcriptional state that allows proper execution of XCI, and second by providing sufficient *Xist* promoter activity through direct and indirect mechanisms.
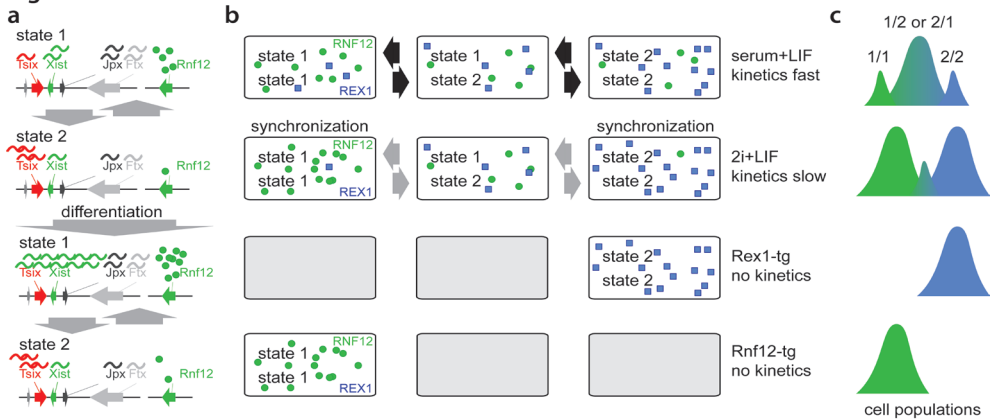


**Figure 7: Model for dynamics of Xic transcriptional states** (a) The Xic can adopt two distinct transcriptional states. State 1 is permissive whereas state 2 is refractive to Xist up-regulation upon differentiation. (b) In serum+LIF conditions female XX ES cells show rapid switching between different states, whereas in 2i+LIF conditions state switching dynamics is reduced leading two synchronization of states. Rex and Rnf12 overexpression forces cells to adopt one single transcriptional state. (c) The relative quantity of alleles adopting distinct combinations of transcriptional states.

## Methods

### Plasmids and Antibodies

Plasmids used for generation of transgenic cell lines were pCAG-Rex1-Flag, pCAG-Rnf12-Flag (20) and pCAG-mTagBFP2-Ezh2-Flag. The coding sequence of mTagBFP2 was inserted N-terminally to the EZH2 coding sequence amplified from mouse cDNA and cloned into pCAG-Flag to give pCAG-mTagBFP2-Ezh2-Flag. Antibodies used were against Flag-M2 (Sigma), REX1 (Abcam and Santa Cruz), RNF12 (Abnova), H3K27me3 (Diagenode) and CD31-FITC (BD Biosciences).

### Cell Lines

Standard ES cell culture conditions included serum+LIF, and both ES cell and differentiation conditions have been described (Barakat et al., 2011a). 2i+LIF conditions were: DMEM supplemented with 100 U/ml penicillin/streptomycin, 20% KnockOut Serum Replacement (Gibco), 0.1mM NEAA, 0.1mM 2-mercaptoethanol, 5000 U/ml LIF, 1μM MEK inhibitor PD0325901 (Stemgent) and 3M GSK3 inhibitor CH99021 (Stemgent). Transgenic ES cell lines were generated using wild-type female line F1 2-1 (129/Sv-Cast/Ei) and wild-type male line J1 (129/Sv). The BAC targeting strategy was used as has been described (34). In short, the Xist knockin was created as follows: an EGFP/neomycin-resistance-cassette flanked by lox sites was targeted by homologous recombination in bacteria to a BAC (34). 5'and 3'targeting arms were amplified from a BAC using primers 1 + 2 and 5 + 6, respectively. With the modified BAC wild-type ES cells were targeted, and the resistance cassette was removed by transient Cre

transfection, resulting in ES cell line Xist-GFP. A Xist ScrFI RFLP with primers 4 + 20 was used to screen drug-resistant clones for correct recombination events. The Tsix knockin was created as follows: a mCherry/neomycin-resistance-cassette flanked by lox sites was targeted by homologous recombination in bacteria to a BAC. 5'and 3'targeting arms were amplified from BAC using primers 25 + 27 and 29 + 30, respectively. With the modified BAC wild-type or Xist-GFP ES cells were targeted, and resistance cassettes were removed by transient Cre transfection, resulting in cell lines Tsix-CHERRY or Xist-GFP/Tsix-CHERRY, respectively. A Tsix PCR length polymorphism with primers 36 + 41 was used to screen drug-resistant clones for correct recombination events. Rex1, Rnf12 and mTagBFP2-Ezh2 transgenes were introduced by electroporation (Bio-Rad Gene Pulser Xcell) and subsequent puromycin selection. Over-expression of transgenes was verified by western blotting and qRT-PCR. The XGTC-XO ES cell line was generated by subcloning Xist-GFP/Tsix-CHERRY via single cell sorting on a FACSAria III platform. Single cell-derived subclones were screened for loss of the wild type X chromosome by an Pf1MI RFLP located in the X-linked gene *Atrx* using primers 68 + 69.

*FACS Analysis and Cell Sorting*
Single cell suspensions were prepared by TE treatment for 7 minutes at $37^0$C. Duplets were excluded by appropriate gating and dead/dying cells by Hoechst 33258 straining (1 μg/ml, Molecular Probes). Relative fluorescence intensities were determined for EGFP and mCherry. Cell analysis was performed on LSRFortessa and cell sorting on FACSAria III (BD Biosciences) with FacsDiva software. Statistical Analysis was performed in FlowJo.

*Expression Analysis*
RNA was isolated using Trizol reagent (Invitrogen) using manufacturer's instructions. DNAse I treatment was performed to remove genomic DNA, and cDNA was prepared using random hexamers and SuperScriptII (Invitrogen). Quantitative RT-PCR was performed on a CFX384 Real-Time PCR Detection System (Biorad) using Fast SYBR Green Master Mix (Applied Biosystems) and primers described in Table S1. Results were normalized to Actin, using the $\Delta C_T$ method and mostly represented as fold-change versus day 0 of differentiation.

*Live Cell Imaging and Image Analysis*
Cells were preplated to remove feeders and differentiation was initiated 12 hours prior to start of imaging.Cells were seeded at low density ($10^4$ cells/well) in a 6-well glass bottom dish (MatTek P06G-1.5-20-F) coated with human plasma fibronectin (Millipore). Imaging was performed on a Leica SP5 AOBS at $37^0$C and 5% $CO_2$ using adaptive focus control to keep cells in focus during the entire experiment. Pictures were taken every 20 minutes for a total of 68 hours. Tiled images were acquired and automatically stitched to record a large field of view at sufficient resolution to resolve subcellular structures and follow cells over time. Average projection of Z-stack was generated in Fiji (version 1.45b) and background corrected integrated fluorescence intensities for EGFP and mCherry were measured for single cells over the entire time frame that a given cell was clearly discriminable. Based on recorded values, linear regression by least squares method was performed to calculate the straight line that best fits the data. The slope of this function with fluorescence intensity being dependent on time was used as a proxy for *Xist* or *Tsix* promoter activity. Threshold for *Xist* activation was calculated by using 3.29 standard deviations (corresponding to 99.9% within confidence interval) of mean EGFP FI values measured in cells within the first six hours of time-lapse experiment.

*Fluorescent In Situ Hybridization and Immunofluorescence*

For Xist/Tsix RNA-FISH and immunofluorescence stainings cells were grown on or absorbed to poly-lysinated coverslips. For RNA-FISH, cells were fixed for 10 minutes with 4% paraformaldehyde (PFA)-PBS at room temperature, washed with 70% EtOH, permeabilized 4 minutes with 0.2% pepsin at 37°C and post-fixed with 4% PFA-PBS for 5 minutes at room temperature. Coverslips were washed twice with PBS and dehydrated in a gradient of 70%, 90%, and 100% EtOH. Nick-labeled DNA probes (digoxigenin for Xist/Tsix probe, biotin for Tsix probe) were dissolved in hybridization mixture (50% formamide, 2XSSC (1XSSC: 0.15 M NaCl, 0.015 M sodium citrate), 50 mM phosphate buffer (pH 7.0), 10% dextran sulfate) and 100 ng/µl mouse Cot DNA to a final concentration of 1 ng/µl. Probe was denatured for 5 min, pre-hybridized for 45 min at 37°C, and coverslips were incubated overnight in a humid chamber at 37°C. After hybridization, coverslips were washed once in 2XSSC, three times in 50% formamide-2X SSC, both at 37°C and twice in TST (0.1 M Tris, 0.15 M NaCl, 0.05% Tween 20) at room temperature. Blocking was done in BSA-TST for 30 minutes at room temperature. Detection was done by subsequent steps of incubation with anti-digoxigenin (Boehringer) and two FITC-labeled antibodies for Xist/Tsix RNA detection or anti-biotin (Roche) and two rhodamine-labeled antibodies for Tsix RNA detection in blocking buffer for 30 min at room temperature. Coverslips were washed twice with TST between detection steps and once finally with TS (0.1 M Tris, 0.15 M NaCl). Dehydrated coverslips were mounted with ProLong Gold Antifade with DAPI (Molecular Probes). For immunofluorescence, cells were fixed for 10 minutes at room temperature in 4% PFA-PBS followed by three washes in PBS and permeabilization in 0.25% Triton-X100-PBS. Blocking was done in blocking solution (0.5% BSA, 1% Tween20 in PBS) for 1 hour at room temperature. All antibody incubation steps were done for 1 hour at room temperature in blocking solution, followed by three washes in blocking solution. Primary antibodies were used at the following concentrations: anti-Flag-M2(1:1000), anti-H3K27me3 (1:500). Secondary antibodies used were conjugated to Alexa Fluor 488 or Alexa Fluor 546 (Molecular Probes; 1:500).

*RNA sequencing*

RNA samples were collected two days after FACS sorting different populations of undifferentiated ES cells, prepared with the Truseq RNA kit, sequenced according to the Illumina TruSeq v3 protocol on the HiSeq2000 with a single read 43 bp and 7bp index and mapped against the mouse mm10/GRCm38 reference genome using Tophat (version 2.0.10). Gene expression values were called using Cufflinks (version 2.1.1).

*Statistical Methods*

Confidence interval of 95% was calculated as:  to  , to with n for the number of cells counted and p for the percentage of Xist clouds scored. Standard deviation was calculated as: , with x for the sample mean and n for sample size. Linear regression was performed using the least squares method. Pearson product-moment correlation coefficient was calculated as: ,with x and y for values of paired data. Single-factor analysis of variance (ANOVA) using the F-distribution was used to test the null hypothesis that all of three or more groups of samples belong to populations with the same mean values. To test if the observed frequencies for three or more groups are equal to the expected frequencies, the chi-square test of independence was calculated by: , with $O_{ij}$ being the observed and $E_{ij}$ the expected frequencies. The two-proportion z-test

was calculated by: , with n for the number of cells analyzed, and and corresponding to the proportion and average proportion, respectively.

**Supplementary figures and table:**
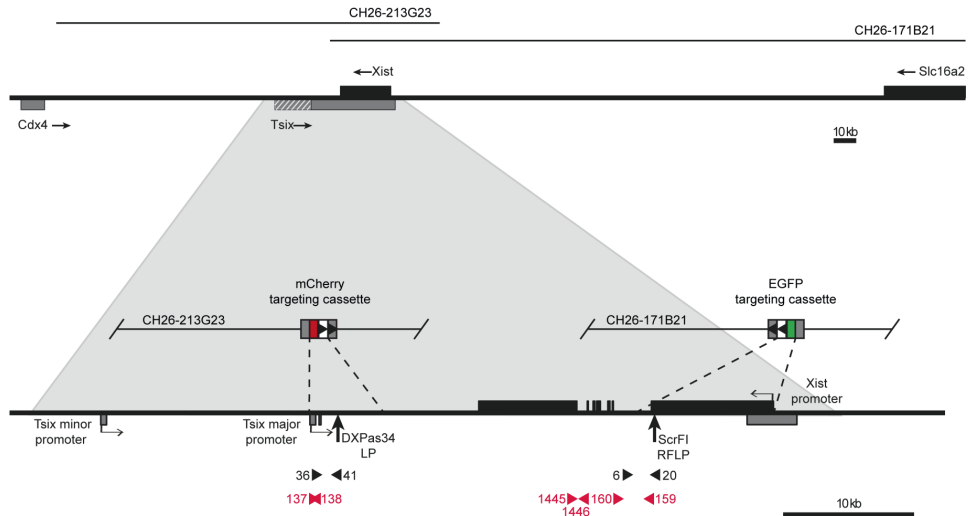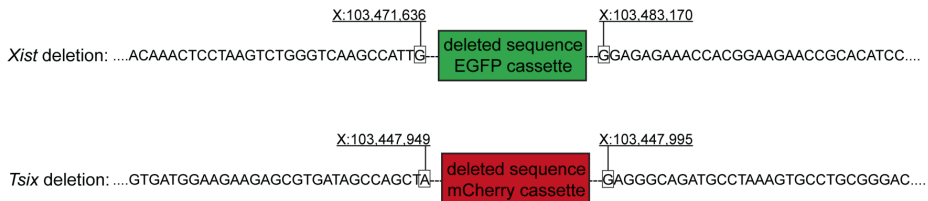


**Supplementary Figure 1. Targeting strategy.** (a) Targeting scheme for generation of Xist-GFP, Tsix-CHERRY and Xist-GFP/Tsix-CHERRY murine ES cell lines. BAC name, location and relative size are indicated on top of panel. Lower part of panel depicts mCherry and EGFP (pmCherry-C1 and pEGFP-C1, both Clontech) targeting cassettes, exon-intron structure of Tsix (grey) and Xist (black), and position of genotyping (black arrowheads) and phenotyping (red arrowheads) primers. Primer numbers as in description in Supplementary Table 1. Polymorphisms used for screening of correctly targeted clones are a length polymorphism (LP) in the DXPas34 region and an ScrFI restriction fragment length polymorphism (RFLP) in exon 1 of Xist. (b) Precise location of deleted *Xist*/*Tsix* sequences. Dots represent flanking sequences, dashes indicate location of the deletion. The exact coordinates of *Mus musculus* genome assembly GRCm38 are indicated for the last non-deleted bases.
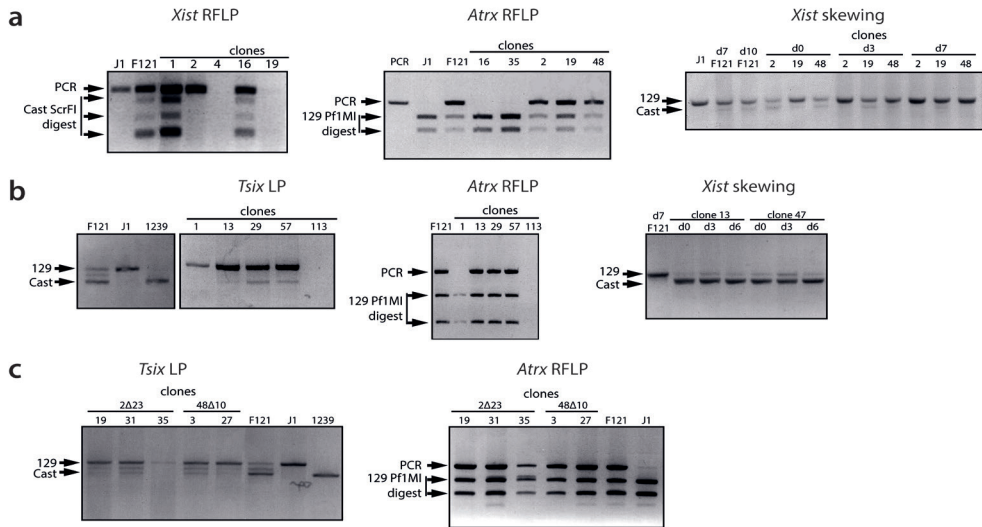
**Supplementary Figure 2: Targeting of Cell Lines.** (a) Targeting of EGFP to the *Xist* locus in female wild-type 129/Sv-Cast/Ei ES cell line. Left panel shows PCR amplification and ScrFI RFLP digest of PCR product to identify clones with a correctly targeted Cast/Ei *Xist* allele. Correct targeting of EGFP-cassette to Cast/Ei allele results in loss of Cast/Ei-specific restriction products, as shown for clone 2. Arrows on left indicate size of PCR product and size of ScrFI restriction fragments. J1 is a 129/Sv control, F121 the polymorphic 129/Sv-Cast/Ei mother cell line. Center panel shows PCR amplification and Pf1MI digest of an X-linked PCR product from the *Atrx* gene to verify presence of two X chromosomes. Arrows on left indicate size of PCR product and size of Pf1MI restriction fragments. J1 is a 129/Sv control, F121 the polymorphic 129/Sv-Cast/Ei mother cell line. For example, clones 16 and 35 had lost the Cast X chromosome. Right Panel shows PCR on cDNA over an *Xist* length polymorphism, demonstrating that in clone 2 only 129/Sv *Xist is* expressed upon differentiation (lower Cast band represents transcription read through only detectable in undifferentiated samples). Arrows on left indicate size of 129/Sv and Cast/Ei PCR products. (b) Targeting of mCherry to the *Tsix* locus in female wild-type 129/Sv-Cast/Ei ES cell line. Left panel shows PCR amplification of an Tsix length polymorphis on genomic DNA to identify clones with a correctly targeted Cast/Ei *Tsix* allele. Correct targeting of mCherry-cassette to Cast/Ei allele results in loss of Cast/Ei-specific band, as shown for clone 13. Arrows on left indicate size of PCR product for 129/Sv and Cast/Ei alleles. J1 is a 129/Sv control, 1239 is a Cast/Ei control and F121 is the polymorphic 129/Sv-Cast/Ei mother cell line. Center panel shows PCR amplification and Pf1MI digest on *Atrx* as in (A). Right panel shows PCR on cDNA over an *Xist* length polymorphism, demonstrating that in clone 13 *Xist* skewing is reversed and *Xist* is primarily expressed from Cas/Ei allele. (c) Targeting of mCherry to the *Tsix* locus in Xist-GFP ES cell line. Left and center panel as in (b), showing correct targeting in clone 2Δ23.
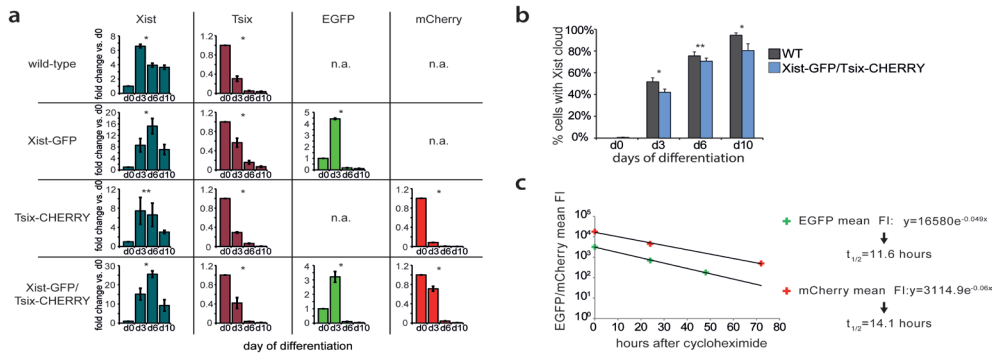
**Supplementary Figure 3: Behavior of Wild Type and Mutant Alleles of *Xist* and *Tsix*.** (a) Expression analysis of *Xist*, *Tsix*, EGFP, mCherry expression levels at different time points of differentiation by quantitative RT-PCR. Quantification is depicted as fold change as compared to undifferentiated cells. Of note, in wild type cells Xist and Tsix levels arise from both the future Xa and Xi; in Xist-GFP Xist arises from future Xi, Tsix from both future Xa and Xi and EGFP from future Xa; in Tsix-CHERRY Xist arises from both future Xa and Xi, Tsix from future Xa and mCherrry from future Xi; in Xist-GFP/Tsix-CHERRY Xist and Tsix arise from future Xi, while EGFP and mCherry arise from future Xa. Error bars represent SD of two or three independent experiments. (b) Quantification of Xist RNA FISH in differentiating wild type and Xist-GFP/Tsix-CHERRY cells. Error bars indicate 95% confidence interval, n = 100 for day 0, n = 350 for day 3 and 6, n = 150 for day 10. (c) Determination of half-life of EGFP and mCherry reporter proteins by cycloheximide chase and FACS analysis of mean FI values for EGFP and mCherry. Xist-GFP and Tsix-CHERRY cells were treated with 100µg/ml cycloheximide (Sigma) to stop protein synthesis and decay of fluorescent proteins was monitored over time. Values were fitted to a first order decay function to estimate the degradation rate constant k and half-life was calculated as: $t1/2 = \ln(2)/k$. Asterisks indicate $P < 0.05$ (*) or $P < 0.1$ (**) by single-factor analysis of variance (a) or two-proportion z-test (b).
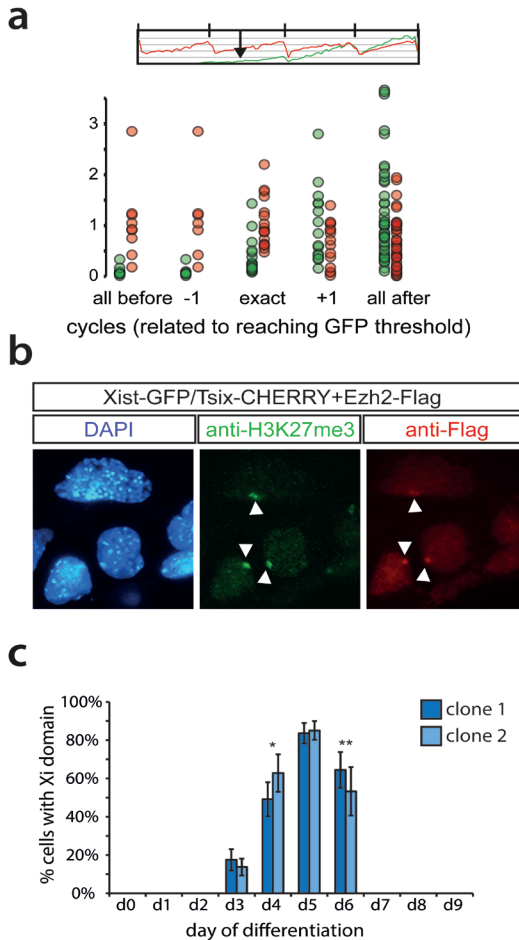
**Supplementary Figure 4: Life cell imaging of reporter lines.** (a) Linear regression of FI over time for each cell cycle was performed. Slope of linear regression as a proxy for promoter activity is plotted. Bins are chosen according to time point of *Xist* promoter activation. Threshold for *Xist* activation was set at 3.29 SDs (corresponding to 99.9% within confidence interval) of background mean EGFP FI measured within the first six hours of time-lapse experiment. Bins as depicted in cartoon on top of panel were chosen as follows: The exact cell cycle in which EGFP FI threshold is reached (exact), one cell cycle before or after threshold is reached (-1,+1) and all cell cycles before or after threshold is reached (all before, all after). (b) Immunofluorescence staining for H3K27me3 and Flag in Xist-GFP/Tsix-CHERRY line at day 3 of differentiation. White arrowheads indicate Xi domain as identified by H3K27me3 and Ezh2-Flag staining. (c) Quantification of Xist-GFP/Tsix-CHERRY cells showing transient enrichment of Ezh2-Flag on the Xi during differentiation determined by direct detection of fluorescence. Two different transgenic clones are shown. Error bars indicate 95% confidence interval, n > 150 for all time points showing Xi domains, n=100 for all time points without Xi domains. Asterisks indicate $P < 0.05$ (*) or $P < 0.1$(**) by two-proportion z-test.

**Supplementary Figure 5: Generation and analysis of Rnf12 and Rex1 transgenic and mutant and XO ES cell lines.** (a) Expression analysis of *Rnf12* and *Rex1* at different time points of differentiation by quantitative RT-PCR. Xist-GFP, Tsix-CHERRY and Xist-GFP/Tsix-CHERRY lines plus the corresponding Rnf12 and Rex1 transgenic lines are shown. Quantification is depicted as fold change as compared to undifferentiated cells without Rnf12 or Rex1 transgenes. Error bars represent SD of two independent experiments. (b) Expression analysis of *Xist*, *Tsix*, EGFP, mCherry expression levels at different time points of differentiation by quantitative RT-PCR. Xist-GFP, Tsix-CHERRY and Xist-GFP/Tsix-CHERRY lines plus the corresponding Rnf12 and Rex1 transgenic lines are shown. Quantification is depicted as fold change as compared to undifferentiated cells without Rnf12 or Rex1 transgenes. Error bars represent SD of two independent experiments, asterisks indicate P < 0.05 (*) or P < 0.1(**) by single-factor analysis of variance for RNF12/REX1 transgenic cell lines and their respective mother cell lines. (c) Screen to identify loss of wild type X chromosome in subclones of Xist-GFP/Tsix-CHERRY by utilizing an X-linked RFLP. PCR amplification and Pf1MI digest of an X-linked PCR product from the *Atrx* gene is shown. Arrows on left indicate size of PCR product and size of Pf1MI restriction fragments. F121 is the polymorphic 129/Sv-Cast/Ei mother cell line, Cast is pure Cast/Ei control. Four of 384 clones showed loss of an X chromosome including clone 76 which lost the wild type 129/Sv X chromosome. (d) Karyotype analysis of

XGTC-XO ES cells prior to FACS analysis. (e) Xist, Tsix and Rnf12 q-PCR expression analysis comparing day 3 differentiated control and three experimental Rnf12$^{+/-}$ ES cell lines. Asterisks indicate $P < 0.05$ (*) by Student's t-Test. (f) Targeting of Rnf12 in the Xist-GFP/Tsix-CHERRY ES cell line. Shown is PCR amplification of an RFLP on genomic DNA to identify clones with a correctly targeted Rnf12 allele. Correct targeting results in loss of the 129/Sv allele. Arrows on left indicate size of PCR product for 129/Sv and Cast/Ei alleles. Shown are 129/Sv-Cast/Ei (F1), 129/Sv (129) and Cast/Eij (cas) controls, and the starting Xist-GFP/Tsix-CHERRY (dki) and Rnf12$^{+/-}$ ES cell lines. (g) PCR amplification of DXMit65 length polymorphism on genomic DNA, to confirm presence of two X chromosomes. (h) Contour plots of FACS analysis showing EGFP and mCherry FI for the Xist-GFP/Tsix-CHERRY control and three Rnf12$^{+/-}$ ES cell lines at different time points of differentiation. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events. (i) Expression analysis of *Tsix* at different time points of differentiation by quantitative RT-PCR. Wild type female XX and male XY cell lines are shown. Quantification is depicted as fold change as compared to undifferentiated female cells.



**Supplementary Figure 6: RNA expression analysis of XGTC-XO, mCherry low and high subpopulations.** (a) FACS analysis of mCherry levels and pluripotency marker CD31 in XY Tsix-CHERRY (XY) and XGTC-XO (XO). Percentage of CD31$^+$ cells is shown for mCherry low and high populations, indicating that there is no difference in pluripotent state between the mCherry low and high populations. (b) Bisulfite sequencing analysis of the *Tsix* major promoter region in XO mCherry low and high populations (empty and filled circles depict unmethylated and methylated CpG sequences respectively). (c) RNA sequencing of XGTC-XO mCherry low and high populations. FPKM values for all genes are plotted, red dots are pluripotency factors, blue dots genes located in the *Xic*. From top to bottom zoom in is depicted as indicated on axes. Pearson correlation coefficient r=0.9832. (d) *Xist* qPCR expression analysis at different time points during ES cell differentiation of wild type 129/Sv:Cast ES cells cultured in serum+LIF and 2i conditions. (e) Allele specific expression analysis of *Xist* during ES cell differentiation indicates skewing of *Xist* expression throughout the XCI process.
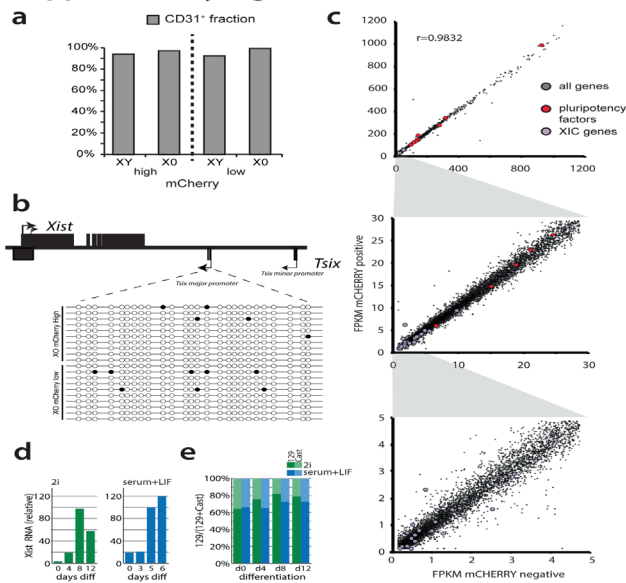
**Supplementary Table 1.** Primers used in this study as listed in the Materials and Methods section.

| # | SEQUENCE | DESCRIPTION |
|---|---|---|
| 1 | AGGTACCTCCCAAGGTATGGAGTCACC | Forward primer 5' targeting arm for Xist |
| 2 | TACCGGTAGGAGAGAAACCACGGAAGAA | Reverse primer 5' targeting arm for *Xist* |
| 5 | TAGTACTCAATGGCTTGACCCAGACTT | Forward primer 3' targeting arm for *Xist* |
| 6 | TAGTACTGTGCCAGAAGAGGGAGTCAG | Reverse primer 3' targeting arm for *Xist*; Reverse primer ScrFI RFLP in *Xist* |
| 20 | GCTGGTTCGTCTATCTTGTGG | Forward primer ScrFI RFLP in *Xist* |
| 25 | CTTTGGTCTCTGGGTTTCCA | Forward primer 5' targeting arm for *Tsix* |
| 27 | TACCGGTAGCTGGCTATCACGCTCTTC | Reverse primer 5' targeting arm for *Tsix* |
| 29 | GAGGGCAGATGCCTAAAGTG | Forward primer 3' targeting arm for *Tsix* |
| 30 | CGCAGGCATTTTACCTTCAT | Reverse primer 3' targeting arm for *Tsix* |
| 36 | AGTGCAGCGCTTGTGTCA | Forward primer *Tsix* length polymorphism, for DNA |
| 41 | TATTACCCACGCCAGGCTTA | Reverse primer *Tsix* length polymorphism, for DNA |
| 68 | TCCCCAATTA AAGGTGTTGA | Forward primer Pf1MI RFLP in *Atrx* |
| 69 | AATTCACGTTCTCCTCTTTCACT | Reverse primer Pf1MI RFLP in *Atrx* |
| 106 | AGGGCATCGACTTCAAGGAG | Forward primer EGFP expression |
| 107 | CACCTTGATGCCGTTCTTCTG | Reverse primer EGFP expression |
| 108 | CCCGTAATGCAGAAGAAGACC | Forward primer mCherry expression |
| 109 | CTTCAGCCTCTGCTTGATCTC | Reverse primer mCherry expression |
| 137 | GTGATGGAAGAAGAGCGTGA | Forward primer *Tsix* expression |
| 138 | GCTGCTTGGCAATCACTTTA | Reverse primer *Tsix* expression |
| 157 | AACCCTAAGGCCAACCGTGAAAAG | Forward primer *Actb* expression |
| 158 | CATGGCTGGGGTGTTGAAGGTCTC | Reverse primer *Actb* expression |
| 159 | GGATCCTGCTTGAACTACTGC | Forward primer *Xist* expression |
| 160 | CAGGCAATCCTT CTTCTTGAG | Reverse primer *Xist* expression |
| 1445 | ACTGGGTCTTCAGCGTGA | Forward primer *Xist* length polymorphism exon 6-7, for RNA |
| 1446 | GCAACAACGAATTAGACAACAC | Reverse primer *Xist* length polymorphism exon 6-7, for RNA |

**References**

1. Borsani G, Tonlorenzi R, Simmler MC, Dandolo L, Arnaud D, Capra V, Grompe M, Pizzuti A, Muzny D, Lawrence C, Willard HF, Avner P, Ballabio A. 1991. Characterization of a murine gene expressed from the inactive X chromosome. Nature 351:325-329.

2. Brockdorff N, Ashworth A, Kay GF, Cooper P, Smith S, McCabe VM, Norris DP, Penny GD, Patel D, Rastan S. 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. Nature 351:329-331.

3. Dixon-McDougall T, Brown C. 2016. The making of a Barr body: the mosaic of factors that eXIST on the mammalian inactive X chromosome. Biochem Cell Biol 94:56-70.

4. Moindrot B, Brockdorff N. 2016. RNA binding proteins implicated in Xist-mediated chromosome silencing. Semin Cell Dev Biol doi:S1084-9521(16)30029-5 [pii]10.1016/j.semcdb.2016.01.029.

5. Lee JT, Davidow LS, Warshawsky D. 1999. Tsix, a gene antisense to Xist at the X-inactivation centre. Nat Genet 21:400-404.

6. Lee JT, Lu N. 1999. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. Cell 99:47-57.

7. Navarro P, Page DR, Avner P, Rougeulle C. 2006. Tsix-mediated epigenetic switch of a CTCF-flanked region of the Xist promoter determines the Xist transcription program. Genes Dev 20:2787-2792.

8. Ohhata T, Hoki Y, Sasaki H, Sado T. 2008. Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. Development 135:227-235.

9. Sado T, Hoki Y, Sasaki H. 2005. Tsix silences Xist through modification of chromatin structure. Dev Cell 9:159-165.

10. Ma Z, Swigut T, Valouev A, Rada-Iglesias A, Wysocka J. 2011. Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. Nat Struct Mol Biol 18:120-127.

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

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

13. Payer B, Rosenberg M, Yamaji M, Yabuta Y, Koyanagi-Aoi M, Hayashi K, Yamanaka S, Saitou M, Lee JT. 2013. Tsix RNA and the germline factor, PRDM14, link X reactivation and stem cell reprogramming. Mol Cell 52:805-818.

14. Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT. 2007. Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. Mol Cell 25:43-56.

15. Sun S, Del Rosario BC, Szanto A, Ogawa Y, Jeon Y, Lee JT. 2013. Jpx RNA activates Xist by evicting CTCF. Cell 153:1537-1551.

16. Barakat TS, Gunhanlar N, Pardo CG, Achame EM, Ghazvini M, Boers R, Kenter A, Rentmeester E, Grootegoed JA, Gribnau J. 2011. RNF12 activates Xist and is essential for X chromosome inactivation. PLoS Genet 7:e1002001.

17. Minkovsky A, Barakat TS, Sellami N, Chin MH, Gunhanlar N, Gribnau J, Plath K. 2013. The pluripotency factor-bound intron 1 of Xist is dispensable for X chromosome inactivation and reactivation in vitro and in vivo. Cell Rep 3:905-918.

18. Nesterova TB, Senner CE, Schneider J, Alcayna-Stevens T, Tattermusch A, Hemberger M, Brockdorff N. 2011. Pluripotency factor binding and Tsix expression act synergistically to

repress Xist in undifferentiated embryonic stem cells. Epigenetics Chromatin 4:17.

19. Jonkers I, Barakat TS, Achame EM, Monkhorst K, Kenter A, Rentmeester E, Grosveld F, Grootegoed JA, Gribnau J. 2009. RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. Cell 139:999-1011.

20. Gontan C, Achame EM, Demmers J, Barakat TS, Rentmeester E, van IW, Grootegoed JA, Gribnau J. 2012. RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature 485:386-390.

21. Makhlouf M, Ouimette JF, Oldfield A, Navarro P, Neuillet D, Rougeulle C. 2014. A prominent and conserved role for YY1 in Xist transcriptional activation. Nat Commun 5:4878.

22. Shin J, Wallingford MC, Gallant J, Marcho C, Jiao B, Byron M, Bossenz M, Lawrence JB, Jones SN, Mager J, Bach I. 2014. RLIM is dispensable for X-chromosome inactivation in the mouse embryonic epiblast. Nature 511:86-89.

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

24. Tian D, Sun S, Lee JT. 2010. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. Cell 143:390-403.

25. Barakat TS, Loos F, van Staveren S, Myronova E, Ghazvini M, Grootegoed JA, Gribnau J. 2014. The Trans-Activator RNF12 and Cis-Acting Elements Effectuate X Chromosome Inactivation Independent of X-Pairing. Mol Cell 53:965-978.

26. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485:376-380.

27. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, Gribnau J, Barillot E, Bluthgen N, Dekker J, Heard E. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485:381-385.

28. Anguera MC, Ma W, Clift D, Namekawa S, Kelleher RJ, 3rd, Lee JT. 2011. Tsx produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. PLoS Genet 7:e1002248.

29. Ogawa Y, Lee JT. 2003. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. Mol Cell 11:731-743.

30. Schulz EG, Meisig J, Nakamura T, Okamoto I, Sieber A, Picard C, Borensztein M, Saitou M, Bluthgen N, Heard E. 2014. The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. Cell Stem Cell 14:203-216.

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

32. Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N. 1996. Requirement for Xist in X chromosome inactivation. Nature 379:131-137.

33. Stavropoulos N, Lu N, Lee JT. 2001. A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice. Proc Natl Acad Sci U S A 98:10232-10237.

34. Barakat TS, Rentmeester E, Sleutels F, Grootegoed JA, Gribnau J. 2011. Precise BAC targeting of genetically polymorphic mouse ES cells. Nucleic Acids Res 39:e121.

35. Csankovszki G, Panning B, Bates B, Pehrson JR, Jaenisch R. 1999. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. Nat Genet 22:323-324.

36. Marks H, Kalkan T, Menafra R, Denissov S, Jones K, Hofemeister H, Nichols J, Kranz A, Stewart AF, Smith A, Stunnenberg HG. 2012. The transcriptional and epigenomic foundations of ground state pluripotency. Cell 149:590-604.

37. Masui O, Bonnet I, Le Baccon P, Brito I, Pollex T, Murphy N, Hupe P, Barillot E, Belmont AS, Heard E. 2011. Live-cell chromosome dynamics and outcome of X chromosome pairing events during ES cell differentiation. Cell 145:447-458.

38. Mlynarczyk-Evans S, Royce-Tolland M, Alexander MK, Andersen AA, Kalantry S, Gribnau J, Panning B. 2006. X chromosomes alternate between two states prior to random X-inactivation. PLoS Biol 4:e159.

39. Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, Dekker J, Tiana G, Heard E. 2014. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. Cell 157:950-963.

40. Habibi E, Brinkman AB, Arand J, Kroeze LI, Kerstens HH, Matarese F, Lepikhov K, Gut M, Brun-Heath I, Hubner NC, Benedetti R, Altucci L, Jansen JH, Walter J, Gut IG, Marks H, Stunnenberg HG. 2013. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell 13:360-369.

41. Prissette M, El-Maarri O, Arnaud D, Walter J, Avner P. 2001. Methylation profiles of DXPas34 during the onset of X-inactivation. Hum Mol Genet 10:31-38.

42. Navarro P, Moffat M, Mullin NP, Chambers I. 2011. The X-inactivation trans-activator Rnf12 is negatively regulated by pluripotency factors in embryonic stem cells. Hum Genet 130:255-264.

43. Guyochin A, Maenner S, Chu ET, Hentati A, Attia M, Avner P, Clerc P. 2014. Live cell imaging of the nascent inactive X chromosome during the early differentiation process of naive ES cells towards epiblast stem cells. PLoS One 9:e116109.
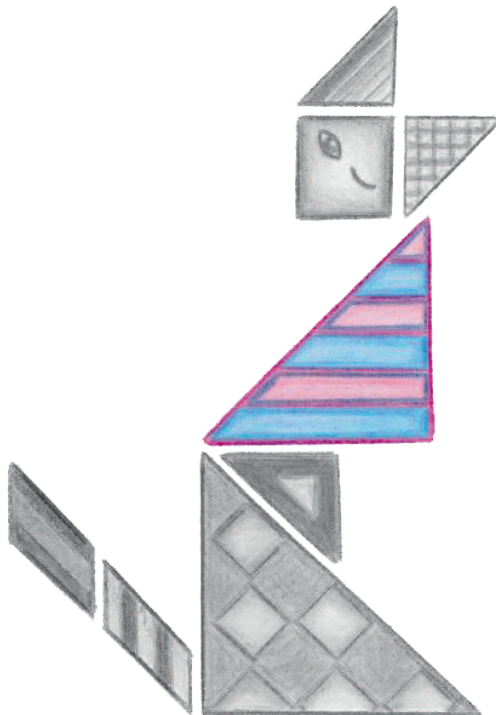
# Chapter 4

**Mega base deletion of the X chromosome**
**Finding a needle in a haystack**

# Chapter 4

*Maduro C.M, Rijsdijk G.*, Voskamp C.*, Rentmeester E. and Gribnau J.#*
*Department of Developmental Biology*
*Erasmus MC, University Medical Center, Rotterdam, The Netherlands*
*#Correspondence: j.gribnau@erasmusmc.nl*

*Work in progress*

## Abstract

Inactivation of one of the two X chromosomes in female eutherian cells is necessary to compensate for the X-encoded genes between XX females and XY males. X chromosome inactivation (XCI) is a tightly regulated process orchestrated by inhibitors and activators of XCI. XCI inhibitors set the threshold for XCI to occur whereas XCI activators, which are X-encoded, facilitate female specific initiation of XCI. XCI inhibitors have been well described, whereas *Rnf12* and *Uba1* have been identified as X-encoded activators. Although *Rnf12* and *Uba1* are important for XCI, remained initiation of XCI in *Rnf12$^{+/-}$:Uba1$^{+/-}$* female ES cells indicate that more factors are at play. Here we generated a 50Mb deletion of the X chromosome, telomeric to *Xist*, pinpointing a region containing novel activator(s) of XCI.

## Introduction

X chromosome inactivation (XCI) is a mechanism equalizing gene dosage of X-encoded genes between eutherian male and female cells by near complete silencing of one of the two X chromosomes in female somatic cells. Female specific initiation of XCI is achieved by an interplay between activators and inhibitors of XCI directing mono-allelic upregulation of the lncRNA X*ist*, when two X chromosomes are present in a diploid nucleus. XCI is a tightly regulated process ensuring one active X chromosome per diploid genome [1]. *Xist* spreads along the X chromosome in *cis,* recruiting different chromatin remodeling complexes to silence the X chromosome [2, 3]. XCI involves a plethora of epigenetic changes, such as DNA methylation, to ensure inactivation of the X chromosome and maintenance of the inactive X (Xi) throughout cell divisions. Even though maintenance of the Xi seems to be very stringent, some genes escape XCI. Genes escaping XCI include genes in the pseudo-autosomal regions (PAR) and individual genes escaping XCI, some of them with others without a Y homolog [4]. Genes in the PAR, the region homologous to the Y chromosome, and genes escaping XCI with a Y homolog are not differentially expressed between the sexes as males also have two copies of genes located in this region. In contrast, escapers without a functional Y homolog, most of them evolutionary recently added genes, are differentially expressed between sexes, thus hindering absolute equalization of gene dosage between the sexes by a dosage compensation mechanism which developed much earlier in evolution of the sex chromosomes [5].

All factors regulating XCI are located in the X inactivation center (Xic), which can be further defined into a cis-Xic and trans-Xic [6]. The cis-Xic comprises all *cis*-regulatory elements in XCI affecting *Xist* expression directly or indirectly. These elements include *Tsix* [7-9] and

*Xite,* which repress *Xist*, as well as *Jpx, Ftx* and *Xpr,* involved in activation of *Xist* in *cis* [10-13]. The trans-Xic comprises several elements and genes located on the X chromosome and on autosomes acting in *trans* to direct XCI by establishing the X chromosome number. Autosomally encoded XCI inhibitors repress *Xist* by increasing the threshold for XCI initiation. Female specific initiation is triggered by X-encoded activators of XCI, differentially expressed between male and female cells acting as dose dependent numerators. Several pluripotency factors, including PRDM14, SOX2, OCT4, NANOG and REX1 have been identified as XCI inhibitors, repressing XCI the early embryo and in embryonic stem cells (ESC), linking XCI to loss of pluripotency [14-17]. As female ESC differentiate, loss of this inhibitory activity and twofold expression of XCI activators leads to female exclusive XCI initiation. *Xist* RNA spreads and silences one of the X chromosomes, consequently silencing the X-encoded activators in *cis*, resulting in a feedback mechanism preventing the remaining X chromosome from being inactivated.

Two X-encoded XCI activators, *Rnf12* [18-20] and *Uba1* [Maduro 2016, Chapter 2], have been described. *Rnf12* and *Uba1* have been shown to be important for female specific XCI, and *Rnf12*$^{+/-}$*:Uba1*$^{+/-}$ compound heterozygous female ES cells display even more affected XCI as compared to *Rnf12*$^{+/-}$ female ES cells. However, some residual *Xist* upregulation is still observed in *Rnf12*$^{+/-}$*:Uba1*$^{+/-}$ ES cells, not observed in male cells [20]. This suggests that *Rnf12* and *Uba1* are not the only X-encoded activators involved in determination of the X chromosome number and initiation of XCI.

The advances in CRISPR/Cas9 genome engineering has not only allowed the possibility of generating small insertions/deletions, but also the ability to introduce mega base deletions/insertions [21-23]. Here we set out to identify a candidate region containing X-encoded activator(s) by generating a large deletion using CRISPR/Cas9 technology and identify a 50Mb region telomeric to *Xist* involved in initiation of XCI, indicating the presence of at least one activator located in this region.

**Experimental procedures**

*Plasmids*

The pSpCas9(BB)-2A-Puro vector (PX459) (Addgene plasmid # 48139) was used to insert the CRISPR guide RNAs [24].

*Cell lines*

Cell lines, culture media and culture conditions for ESC culture and differentiation were described previously [18,25]. CRISPR vectors were introduced in hybrid female F1 (129/Cast) mouse embryonic stem cells by electroporation. Cells were electroporated in a 0.2cm electroporation cuvette (BioRad) at 118kV, 1200μF and ∞Ω in a Gene Pulser Xcell Electroporation System (BioRad). After 24 hours of recovery, cells were grown on ES medium containing puromycin (1ug/ml) for 36 hours.

*Deletions with the CRISPR/Cas9 system*

CRISPR guides were designed by using the CRISPR design tool (http://crispr.mit.edu/). The designed CRISPR guide oligos with 5'- CACC and 3'- CAAA overhangs (Table1) were cloned into the pX459 CRISPR vector (Addgene) [24] by a simultaneous digestion-ligation reaction [26]. First, the pX459 vector was digested with *BbsI* (NEB) to allow the replacement of the restriction sites with direct insertion of the annealed oligo guides. Combinations of two guides (beginning and end) were used in each targeting and first screened as a pool to ensure the deletion to be present (Table 2), followed by single clone screens (Table 2 and 3).

*Xist RNA Fluorescent in situ hybridization*

Procedure, probe labelling and antibodies have been described previously [18, 25].

*RNA isolation from ES cells and cDNA synthesis*

TRIzol® (Invitrogen) was added to each sample and RNA was isolated according to manufacturer's instructions. RNA samples were diluted to contain 2µg RNA in 10µl. Before cDNA synthesis, DNase (Invitrogen) was added to the RNA samples, for 30 minutes at $37^0C$, to remove any contaminating genomic DNA. DNase was inactivated at $65^0C$ for 10 minutes. Random Hexamers (Applied BioSystems) and dNTP mix (Invitrogen) were added and incubated at $65^0C$ for 5 minutes and subsequently cooled on ice for 1 minute. 5x First Strand Buffer, 0.1M DTT and RNaseOUT (Invitrogen) were then added to the samples and incubated at $25^0C$ for 2 minutes before adding SuperscriptII Reverse transcriptase (Invitrogen). Samples were incubated at $25^0C$ for 10 minutes followed by 50 minutes at $42^0C$ and finally, 15 minutes at $70^0C$.

*Table 1: Oligo's for CRISPR guides for deletion of C77370-Klf8. 5'- CACC and 3'- CAAA overhangs are indicated in bold and the preferred G as a starting nucleotide is underlined.*

| Gene | Guide oligo |
| --- | --- |
| **C77370** | Fw: 5' **CACC**GTTATCTGGCCTAACAGTACC 3' |
| | Rv: 5' **AAAC**GGTACTGTTAGGCCAGATAAC 3' |
| | |
| **Klf8** | Fw: 5' **CACC**GTTGCAGGGGACTGTACCCCC 3' |
| | Rv: 5' **AAAC**GGGGGTACAGTCCCCTGCAAC 3' |

*Table 2: Primers for detection wildtype allele, deletion and allele check of the deletion of C77370-Klf8 and C77370-Klhl4*

| | Wildtype allele 5'-3' | PCR over deletion 5'- 3' |
| --- | --- | --- |
| **ΔC77370-Klf8** | Fw: TTGTTTCCTTGGTCTTGCTG | Fw: TTGTTTCCTTGGTCTTGCTG |
| | Rv: GATCCGTCTGCCTCTTTCTC | Rv: GGGGGAAAAGAGAGTGATGA |
| | | |

*Table 3: DXMit primers*

| | Location on X chromosome (cM) | Sequence 5'-3' | Size (bp) |
| --- | --- | --- | --- |
| **DXMit55** | 3.31 | Fw: CTGCTTCCAGAATATTATCACTACTCC | Cast: 115 |
| | | Rv: AAAACATCCATTTATGTTAACACACA | 129: 141 |
| | | | |
| **DXMit75** | 30.84 | Fw: ACCCTAGCCGGTTACTATCTCC | Cast: 144 |
| | | Rv: TCTCTCTCATTCCCCCCC | 129: 104 |
| | | | |
| **DXMit44** | 38.38 | Fw: TCTAAAAGCATGCCAAATTGG | Cast: 230 |

| | | Rv: TTCCTATCGCTCAGGTTTTTG | 129: 184 |
|---|---|---|---|
| | | | |
| **DXMit65** | 47.86 | Fw: ATATTAAGGGAGGTAACAAAGACCC | Cast: 176 |
| | | Rv: GGTTTCTGTGATTGCTATAGGACA | 129: 144 |
| | | | |
| **DXMit197** | 68.46 | Fw: TGTTCTATATTGCTTTGTTAGGTTTTG | Cast: 160 |
| | | Rv: GCAAAAAAGAAACCAACCCA | 129: 122 |
| | | | |
| **DXMit121** | 73.95 | Fw: CCTTTAAACATTGTGATTAGTGATGG | Cast: 104 |
| | | Rv: CCTAATTTTGAACAGTATCAGTTTTCA | 129: 150 |
| | | | |
| **DXMit184** | 78.70 | Fw: GTCTGGATTTGGGTTCTAAATATTT | Cast: 184 |
| | | Rv: AAGCACATACACACATACGTTCTC | 129: 150 |

## Results

*Candidate region for XCI activators*

XCI is initiated in female cells when the concentration of X-encoded activators is sufficient to overcome the threshold set by the XCI-inhibitors. The E3 ubiquitin ligase, *Rnf12* [10, 18, 20], and the E1 activating enzyme, *Uba1* [Maduro, Chapter 2], have been identified as important factors in activating XCI and female cells heterozygous for both genes display significantly affected XCI. However, slight upregulation of *Xist* expression was still observed in $Rnf12^{+/-}:Uba1^{+/-}$ cells, whereas male ES cells do not show this level of *Xist* upregulation or the presence of *Xist* clouds throughout differentiation. This observation suggests that there are other X-encoded activators involved in activation of XCI that differentiate these cells from male cells. Previous candidate gene approaches to find X-linked activators of XCI have proven laborious and precarious. We therefore decided to proceed with generating a large X chromosomal deletion of a candidate region using CRISPR/Cas9 technology.

XCI evolved early in evolution of mammalian sex chromosomes as a dosage compensation mechanism of X encoded genes. Therefore, XCI activators are expected to be located within an X conserved region (XCR), orthologous to chicken chromosome 4, representing the ancient X chromosome [27]. In addition, proper negative feedback of XCI activators is required to prevent both X chromosomes from being inactivated. Such a feedback mechanism would require XCI activators to be subject to XCI and silenced in early stages of XCI. Considering these points, a 50Mb candidate region telomeric to the cis-Xic locus was chosen as this region is part of the XCR, is most conserved in gene content between human and mouse [28], contains regions with genes which are methylated at early stages of XCI, and contains few escaping genes (Figure 1).
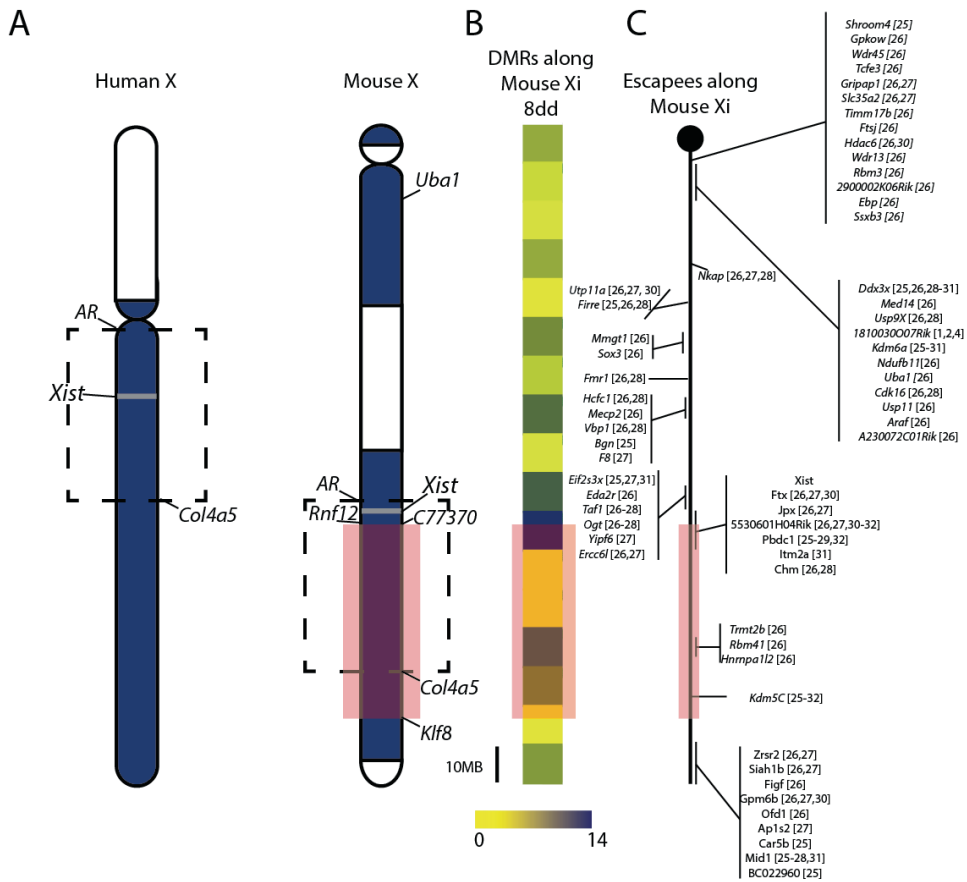
**Figure 1:** A) Comparison of human (left) and mouse (right) X chromosomes. Blue = X conserved region (XCR) corresponding to regions on the mammalian X chromosome that are found on chicken orthologous autosome 4 and the marsupial X chromosome; white = X added region (XAR) corresponding to regions on the mammalian X chromosome found on chicken orthologous autosome 1 added after divergence of eutherians and marsupials. The black dashed region (from *AR* to *Col4a5*) is conserved in gene content between human and mouse and only contains a 600kb inverted region around *Xist* (grey) [28]. The red region on mouse chromosome X depicts the 50Mb region studied herein (from *C77370 to Klf8*). *Uba1*, identified as an XCI activator (Chapter 2) is also depicted on the mouse X chromosome. B) Blocks of 10Mb with number of differentially methylated regions (2 kb flanking gene promoter), determined in cells at day 8 of differentiation (the DMR blocks are not corrected for gene number). C) XCI escape genes found in different studies in varying tissues ([29]=patski cells; [30]=NPCs; [31]=throphoblast cells; [32] = NPCs; [33]= neural stem cells; [34]= brain; [35]= brain; [36]= primary fibroblasts).

*A 50Mb deletion telomeric to Xist*

In order to determine whether the telomeric candidate region contained activators of XCI, we generated a CRISPR/Cas9 mediated deletion of this 50Mb region. The largest deletion reported so far was 30Mb, obtained with an efficiency of ~1% [37], suggesting a strategy to

delete such a large block of the XCR to be feasible. CRISPR guides were designed spanning *C77370* to *Klf8,* and used to target F1 hybrid Cast:129 *Rnf12*[+Cast/-129] cells [described in Chapter 2] (Figure 2A). If the deletion would encompass the most important factors involved in initiation of XCI, this would result in abolished initiation of XCI in these cells, mostly resembling male cells instead of wildtype female cells with respect to *Xist* expression and *Xist* cloud formation upon differentiation. To ensure the designed CRISPR guides were recognizing the correct loci and ensure a 50Mb deletion was feasible, we first screened all targeted cells as a pool using PCR primers spanning the deleted region. The PCR across *C77370* to *Klf8*, normally 50Mb apart, on the targeted pool showed correct targeting of the guides, indicating fusion of these loci by non-homologous end joining (NHEJ) (Figure 2B). We therefore proceeded to pick 96 clones which were screened with the same primers flanking the deleted region (Figure 2B). We identified 1 correct clone, containing the 50Mb deletion. The 50Mb deletion generated in the *Rnf12*[+/-] cells*,* now referred to as *Rnf12*[+/-]:⌐*50Mb*[+/-]*,* is heterozygous, and confirmed by sequencing (Figure 2B). Using DXMit marker primers along the X chromosome we could determine the targeted allele as well as the integrity of the rest of the X chromosome (Figure 2C) and the results clearly showed the presence of only one allele in the deleted region whereas two alleles were detected for the remainder of the X chromosome. The DXMit PCR products also showed that the deletion of *C77370* to *Klf8* is present on the Cast X chromosome, resulting in *trans* heterozygous *Rnf12*[+/-]:⌐*50Mb*[-/+] ES cells. Previous studies indicated that *Rnf12*[+/-] cells show preferential inactivation of the mutated allele [10], which would lead to complete loss of expression of genes located in the deleted region which might be unfavorable for the cell resulting in selection towards skewing of the Cast allele. However, as silencing is not expected to occur at the very early stages of XCI we decided to analyze the combined effects of these mutations on XCI initiation.
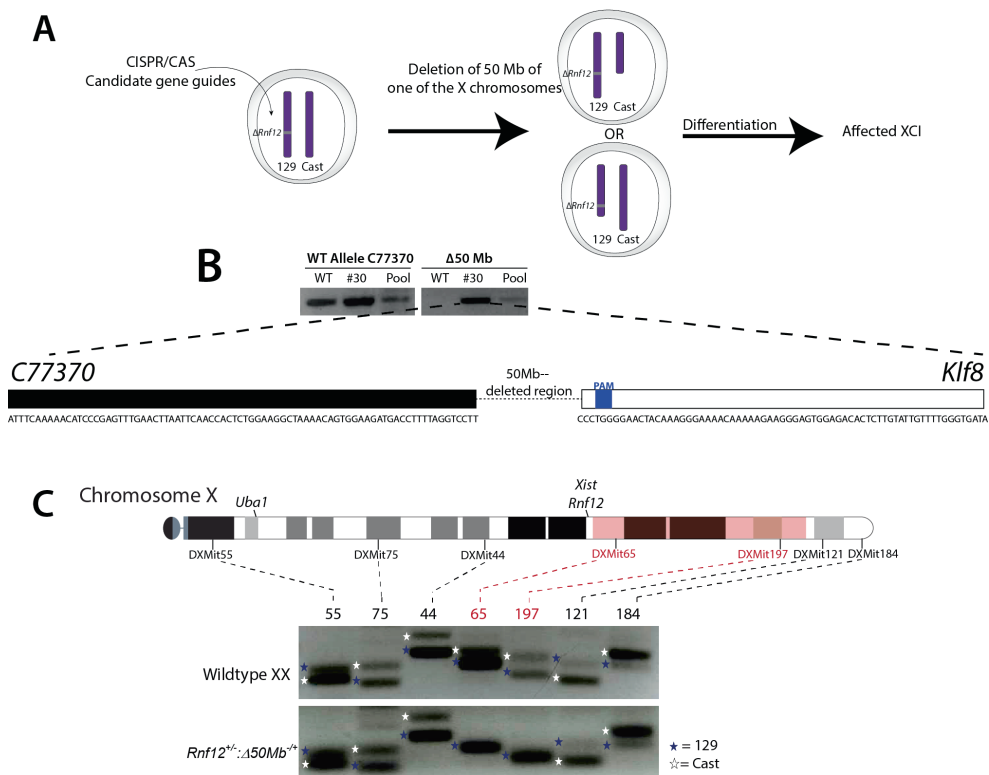
**Figure 2**: A) PCR of the wildtype *C77370* allele present in the wildtype, in the targeted clone 30 and in the targeted pool (top left figure). PCR analysis on the 50Mb deleted region present in the targeted clone 30 and in the targeted pool (top right figure). Bottom panel shows the sequenced PCR analysis on the deletion in clone 30. A) DXMit markers shown along the X chromosome (shaded red region represents the 50Mb deletion from *C77370 to Klf8*). DXMit65 and DXMit197 (red) are located inside the deletion. Wildtype 129/Sv:Cast/Eij ES cells reveal both 129/Sv (129) and Cast/Eij (cast) alleles for all DXMit markers tested, whereas for the *Rnf12$^{+/-}$:$\Delta$50Mb$^{/+}$* cells only the 129 allele is detected inside the deleted region, indicating a targeted Cast allele and the presence of two X chromosomes.

*Loss of XCI in Rnf12$^{+/-}$:$\Delta$50Mb$^{/+}$ cells.*

To test the effect of the 50Mb deletion on XCI, we performed *Xist* RNA FISH on day 4 and 8 differentiated ES cells (Figure 3A, E-F). The results indicated severely affected XCI which was more affected than *Rnf12$^{+/-}$* cells and did not recover after 8 days of differentiation, as observed in *Rnf12$^{+/-}$* cells. Loss of XCI in *Rnf12$^{+/-}$:50Mb$^{/-}$* cells was confirmed by *Xist* qPCR indicating severely reduced upregulation of *Xist* expression (Figure 3B). *Rnf12$^{+/-}$:$\Delta$50Mb$^{/+}$* cells showed very low *Xist* expression up to day 8 of differentiation only reaching the level of *Xist* expression obtained at day 2 of differentiation of wildtype cells. Previous work on *Rnf12$^{+/-}$* cells indicated preferential skewing towards inactivation of the mutated X chromosome, through failure to establish inactivation of the wildtype X chromosome due to loss of *Rnf12* expression. In our trans-compound knockout ES cells, inactivation of the X chromosome baring the deletion of *Rnf12*, as well as the X chromosome with the 50Mb deletion, will lead to dosage problems.

To better understand the cell selection process, we performed allele specific *Xist* RT-PCR analysis on our *Rnf12*$^{+/-}$*::Δ50Mb*$^{/+}$ cells over 8 days of differentiation as compared to wildtype cells and *Rnf12*$^{+/-}$ cells (Figure 3C). We find that skewing of XCI in *Rnf12*$^{+/-}$ cells becomes eminent at day 6 of differentiation and silencing also starts at this time point. In *Rnf12*$^{+/-}$*::Δ50Mb*$^{/+}$ cells, *Xist* expression is reduced from both alleles prior to cell selection and can therefore only be attributed to an XCI initiation defect.

We also analyzed *Tsix* expression in wildtype female and male, *Rnf12*$^{+/-}$ and *Rnf12*$^{+/-}$*::Δ50Mb*$^{/+}$ cells and found that both *Rnf12*$^{+/-}$*::Δ50Mb*$^{/+}$ and *Rnf12*$^{+/-}$ cells showed increased expression of *Tsix* as compared to wildtype female and male cells (Figure 3D). To ensure the *Rnf12*$^{+/-}$*::Δ50Mb*$^{/+}$ cells did differentiate properly, important for initiation of XCI to occur, *Oct4* (pluripotency marker) and *Gata6* (endoderm marker) expression was determined by qPCR. This analysis indicated no reduction in *Oct4* (Figure 3D) and but did show an increase in *Gata6* expression (Figure 3D) throughout ES cell differentiation indicating that these cells did differentiate into endoderm. *Tsix* upregulation was not related to a potential block in ES cell differentiation, but is most likely the consequence of loss of one copy of *Rnf12*, resulting in upregulation of REX1, a known activator of *Tsix*. Our findings therefore indicate that the deleted 50 Mb region covers genes or elements that play a crucial role in female specific initiation of XCI.
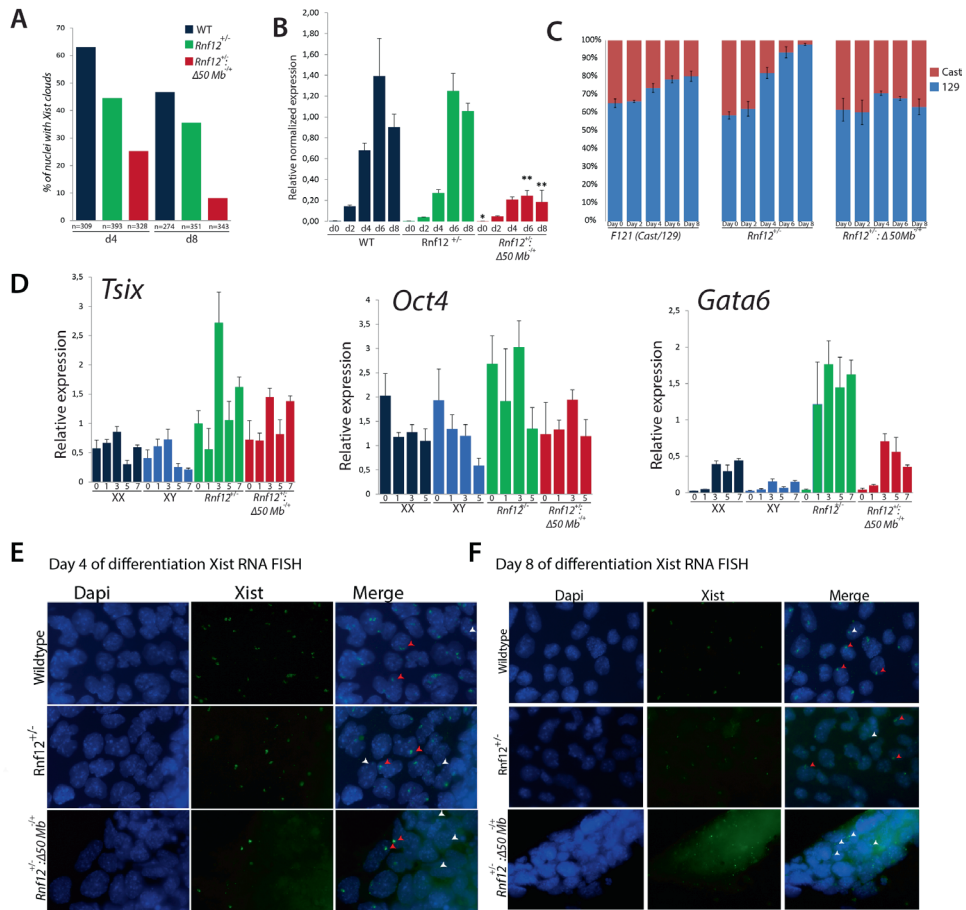
**Figure 3**: A) Xist RNA FISH results depicting percentage of cells with *Xist* clouds (dark blue) and *Xist*/*Tsix* pinpoints (light blue) at day 4 and day 8 of differentiation in wildtype, *Rnf12*[+/-] and *Rnf12*[+/-]::*50Mb*[/+] cells. *Rnf12*[+/-]::*50Mb*[/+] showed a reduced percentage of cells with *Xist* clouds B) Relative *Xist* expression normalized to β-actin day 0, 2, 4, 6 and 8 of differentiation in wildtype, *Rnf12*[+/-] and *Rnf12*[+/-]::*50Mb*[/+] cells (error bars indicate SEM and t-test is a comparison of Δ50Mb cells to *Rnf12*[+/-] cells * p< 0.05; ** p<0.0001). Δ50Mb cells have significantly reduced *Xist* expression levels at day 6 and day 8 of differentiation. C) Allele specific *Xist* RT-PCR (red = Cast; blue = 129) over 8 days of differentiation in wildtype (Cast/129), *Rnf12*[+/-] and *Rnf12*[+/-]::*50Mb*[/+] cells. D) QPCR analysis determining *Tsix* (left), *Oct4* (middle), and *Gata6* (right) expression at different time points in differentiation for female cells (XX), male cells (XY), *Rnf12*[+/-] and the *Rnf12*[+/-]::*50Mb*[/+] cells. E) Left: Xist RNA FISH on day 4 differnetiated wildtype, *Rnf12*[+/-] and *Rnf12*[+/-]::*50Mb*[/+] cells (*Xist*= FITC; DNA=Dapi). Right panels show (Xist clouds= red arrows); Xist/'Tsix pinpoints= white arrows). F) Xist RNA FISH on day 8 differentiated wildtype, *Rnf12*[+/-] and *Rnf12*[+/-]::*50Mb*[/+] cells (*Xist*= FITC; DNA=Dapi). Right panels show (Xist clouds= red arrows); Xist/'Tsix pinpoints= white arrows).

**Discussion and conclusion**

Activators of XCI play an important role in proper initiation of XCI and mono-allelic upregulation of *Xist* expression. Activators of XCI need to be located on the X chromosome in order to be expressed at higher concentrations in female cells as compared to male cells allowing a female cell, but not a male cell, to overcome the threshold for XCI to occur. In addition, activators of XCI have to be subject to XCI to prevent inactivation of the second X chromosome once initiation on one X chromosome has started. So far, only *Rnf12* and *Uba1* have been described as activators of XCI, but the presence of some remaining *Xist* expression and *Xist* cloud formation in differentiated heterozygous *Rnf12$^{+/-}$:Uba1$^{+/-}$* ES cells indicates that other activators are involved in initiation of XCI as well. Previous candidate gene approaches in an attempt to identify other activators of XCI have proven laborious and although candidate genes are based on careful selection, activators of XCI can still be falsely excluded. We have therefore deleted roughly one third of the telomeric X chromosome as a starting point to be able to identify the region(s) on the X chromosome involved in activation of XCI. The 50Mb deletion described here was generated in *Rnf12$^{+/-}$* ES cells to detect additive effects of the deletion on XCI. Analysis of *Rnf12$^{+/-}$:Δ50Mb$^{/+}$* cells upon differentiation revealed a very low percentage of cells with *Xist* clouds and significantly reduced *Xist* expression, which did not recover after 8 days of differentiation. The very low levels of *Xist* expression remaining and cloud formation in a low percentage of cells, indicates the presence of at least one more activator located outside the 50Mb region deleted in this study. As *Uba1* is not included in the deletion, this residual *Xist* expression and *Xist* cloud formation might be forced by differential expression of *Uba1*. In addition, the trans deletion of *Rnf12* and 50Mb deletion might have impacted on the robustness of XCI initiation and lack of recovery. Previous studies examining *Xist$^{+/-}$:Rnf12$^{+/-}$* cis-compound knockout ES cells indicate that forced inactivation of the remaining copy of *Rnf12* on the wildtype X chromosome, as might happen in our *Rnf12$^{+/-}$:Δ50Mb$^{/+}$* cells due to cell selection, leads to reduction of *Xist* cloud formation [10]. This effect will likely be less pronounced as in our *Rnf12$^{+/-}$:Δ50Mb$^{/+}$* cells as it will be forced by loss of expression of genes in the deleted region and not *Rnf12*. In addition, we show that the *Rnf12$^{+/-}$:Δ50Mb$^{/+}$* cells show reduced *Xist* expression from both alleles prior to cell selection. Taken together, our results suggest that XCI is regulated by at least three XCI activators of which one is located in the 50Mb deleted region of *C77370* to *Klf8*.

We have narrowed down a region which must contain at least one activator of XCI, however 50Mb is still a large region and contains 202 annotated genes. We analyzed each of the 202 genes for different characteristics such as their molecular function, biological processes they are involved in, cellular localization, pathways involved in, or escape of XCI to obtain a list of candidate genes (Table 4). From this analysis, we obtained 7 candidate genes which are: subject to XCI, expressed in ES cells, are nuclear and silenced by DNA methylation in the early stages of XCI. The 7 genes which met all the requirements were: *C77370, Bex4, Morf4l2, Trap1a, Nup62cl, Tsc22d3* and *Nxt2* (Figure 4). *C77370* is located closest to *Rnf12* and the Xic (220kb) whereas the other 6 genes are located within 10Mb of each other and about 30Mb away from the Xic (Figure 4). Future targeted strategies removing individual genes in *Rnf12$^{+/-}$:Uba1$^{+/-}$* cells will need to be performed to establish a role in XCI next to other unbiased candidate strategies

Table 4: 202 genes located in the 50Mb deleted region ordered according to location. (Light green= subject to XCI; Dark green= expressed in ESCs; Blue= nuclear; Purple= early methylated)

| Gene | Molecular function | Biological process | Cellular component | Pathway | Escaping gene | Early methylated | Expressed in ESCs (FKPM >1) |
|---|---|---|---|---|---|---|---|
| C77370 | Nuclease; Exodeoxyribonuclease | Developmental Process | Nucleus | ND | No | Yes | Yes |
| Abcb7 | ATP-binding cassette (ABC) transporter | Cellular iron ion homeostasis; Transmembrane transport | Membrane | ND | No | No | Yes |
| Uprt | GTP binding; Uridine kinase activity | Pyrimidine nucleobase metabolic process | Cytoplasm; Nucleus | ND | No | No | No |
| Zdhhc15 | Acyltransferase; Transferase | Protein localization; Synaptic vesicle maturation | Membrane | Interferon-gamma signaling pathway | No | No | Yes |
| Magee2 | ND | ND | ND | ND | No | No | No |
| Pbdc1 | Phosphoprotein | ND | ND | ND | Yes | No | Yes |
| Magee1 | Tumor antigen | Ubl conjugation pathway | Cytoplasm; Nucleus; Cell membrane | Ubl conjugation pathway | No | No | No |
| Cypt2 | ND | ND | ND | ND | No | No | No |
| Fgf16 | Growth Factor | Cellular process; Developmental Process | Secreted | FGF signaling pathway | No | No | No |
| Atrx | DNA helicase | Biological regulation; Cellular component organization/biogenesis; Cellular process; Metabolic process | Nucleus | Wnt signaling pathway | No | No | Yes |
| Magt1 | Transporter | Establishment of localization | Endoplasmic reticulum; Plasma membrane | ND | No | No | Yes |
| Cox7b | Cytochrome-c oxidase | Central nervous system development; Hydrogen ion transmembrane transport | Mitochondrion; Membrane | ND | No | No | Yes |
| Atp7a | Cation transporter; Ion channel; Hydrolase | Copper transport; Ion transport; Transport | Cell membrane; Golgi apparatus; Membrane | ND | No | No | Yes |

4

| Gene | Molecular function | Biological process | Location | Pathway/Other | | | |
|---|---|---|---|---|---|---|---|
| Tlr13 | Receptor | Immunity; Inflammatory response; Innate immunity | Endosome; Membrane | ND | No | No | No |
| Pgk1 | Kinase; Transferase | Glycolysis | Cytoplasm | Glycolysis | No | No | Yes |
| Taf9b | Transcription factor; Nucleic acid binding | Biological regulation; Metabolic Process; Response to stimulus | Nucleus | General transcription regulation; Transcription regulation by bZIP transcription factor | No | No | No |
| Fnd3c2 | ND | ND | ND | ND | No | No | No |
| Fndc3c1 | ND | ND | Membrane | ND | No | No | No |
| Cysltr1 | G-protein coupled receptor; Receptor; Transducer | Calcium ion transport; Cell surface receptor signaling pathway | Membrane | ND | No | No | No |
| Gm5127 | ND | ND | ND | ND | No | No | No |
| Zcchc5 | ND | ND | ND | ND | No | No | No |
| Lpar4 | G-protein coupled receptor; Receptor; Transducer | ND | Membrane | ND | No | No | No |
| P2ry10 | G-protein coupled receptor; Receptor; Transducer | ND | Membrane | ND | No | No | No |
| Gpr174 | G-protein coupled receptor; Receptor; Transducer | ND | Membrane | ND | No | No | No |
| Itm2a | Beta-amyloid binding | Immunoglobulin production; Plasma cell differentiation; Negative regulation of amyloid precursor protein biosynthetic process ; Nervous system development | Extracellular; Golgi apparatus; Plasma membrane | ND | Yes | No | Yes |
| Tbx22 | Transcription factor; Nucleic acid binding | Biological regulation; Metabolic Process | Membrane | ND | No | No | No |

| Gene | Protein type/Function | Biological process | Cellular localization | Pathway | | | |
|---|---|---|---|---|---|---|---|
| Fam46d | ND | ND | ND | ND | No | No | No |
| Gm732 | Non-motor actin binding protein | Cellular component organization/biogenesis | Organelle; Cell part | ND | No | No | No |
| Brwd3 | ND | Cytoskeleton organization; Regulation of cell shape; Regulation of transcription from RNA polymerase II promoter | Nucleus | ND | No | No | Yes |
| Hmgn5 | Chromatin binding; DNA binding; poly(A) RNA binding | Transcription; Transcription regulation | Nucleus | ND | No | No | Yes |
| Sh3bgrl | ND | ND | Extracellular | ND | No | No | Yes |
| Gm6377 | ND | ND | ND | ND | No | No | No |
| Pou3f4 | Homeobox transcription factor | Biological regulation; Metabolic Process | Nucleus | ND | No | No | No |
| Cylc1 | Structural protein | ND | Cytoplasmic vesicle | ND | No | No | No |
| Rps6ka6 | Transferase | Protein metabolic process; Response to stimulus; Signaling | Nucleus; Organelle lumen; Mitochondrion | Interleukin signaling pathway ; PDGF signaling pathway, Ras pathway | No | | Yes |
| Hdx | Homeobox transcription factor; DNA binding protein | Biological regulation; Metabolic Process | Nucleus | ND | No | No | No |
| Tex16 | Nuclease; RNA binding | Cellular process; Metabolic Process | Nucleus | ND | No | No | No |
| 4933403O08Rik | Kinase modulator | Biological regulation; Metabolic Process | ND | ND | No | No | No |
| Apool | ND | Cristae formation | Mitochondrion inner membrane | ND | No | No | Yes |
| Sat1 | Acyltransferase; Transferase | ND | ND | ND | No | No | No |
| 2010106E10Rik | ND | ND | ND | ND | No | No | No |
| Znf711 | KRAB box transcription factor | Biological regulation; Metabolic Process | ND | ND | No | No | No |

| Pofib | ND | ND | Cell junction | ND | No | No | Yes |
|---|---|---|---|---|---|---|---|
| Chm | Acyltransferase; G-protein modulator | Biological regulation; Cellular process; Localization; Multicellular organismal process | Cytosol | ND | Yes | No | Yes |
| Dach2 | Transcription Factor | Developmental Process; Transcription; Transcription regulation | Nucleus | ND | No | No | No |
| Klhl4 | Transferase | ND | ND | ND | No | No | No |
| Ube2dnl1 | ND | ND | ND | Ubiquitin proteasome pathway | No | No | No |
| Cpxcr1 | ND | ND | ND | ND | No | No | No |
| Gm14920 | Histone | Cellular component organization/biogenesis; Cellular process; Metabolic Process | Nucleus; Organelle lumen | ND | No | No | No |
| Tgif2lx2 | DNA-directed RNA polymerase | ND | ND | ND | No | No | No |
| Pabpc5 | ND | ND | ND | ND | No | No | No |
| Pcdh11x | Cadherin | Cellular process; Developmental Process; Multicellular organismal process | Extracellular | Cadherin signaling pathway; Wnt signaling pathway | No | No | No |
| Nap1l3 | Phosphatase inhibitor | Apoptotic process; Biological regulation; Cellular component organization/biogenesis; Cellular process; Developmental process; Metabolic Process | Cytoplasm; Nucleus | ND | No | No | No |
| 3110007F17Rik | ND | ND | ND | ND | No | No | No |
| Vmn2r121 | ND | ND | ND | ND | No | No | No |
| Gm382 | ND | ND | ND | ND | No | No | No |
| 4921511C20Rik | ND | ND | ND | ND | No | No | No |
| 4930412D23Rik | ND | ND | ND | ND | No | No | No |

| Gene | Protein class | Biological process | Cellular component | Pathway | | | |
|---|---|---|---|---|---|---|---|
| Diap2 | Non-motor actin binding protein | Cellular process | Organelle; Cell part | Cytoskeletal regulation by Rho GTPase | No | No | Yes |
| Pcdh19 | Cadherin | Biological Adhesion; Cellular process; Developmental Process | Membrane | Cadherin signaling pathway, Wnt signaling pathway | No | No | No |
| Tnmd | Membrane-bound signaling molecule | Developmental Process | Membrane; Nucleus envelope | ND | No | No | No |
| Tspan6 | Membrane-bound signaling molecule; receptor; Cell adhesion molecule | Biological Adhesion; Cellular process; Immune system process; Multicellular organismal process; Reproduction; Response to stimulus | Extracellular; Plasma membrane | ND | No | No | Yes |
| Srpx2 | Hepatocyte growth factor binding; Identical protein binding; Receptor binding | Angiogenesis; Cell adhesion | Cell junction; Cytoplasm; Secreted; Synapse | ND | No | No | No |
| Sytl4 | Membrane trafficking regulatory protein | Biological regulation; Cellular component organization/biogenesis; Cellular process; Localization; Multicellular organismal process | Membrane; Organelle; Cell part | ND | No | No | No |
| Cstf2 | mRNA splicing factor | mRNA processing | Nucleus | ND | No | No | Yes |
| Nox1 | Oxidase; Transporter | Electron transport; Transport | Membrane | ND | No | No | No |
| Xkr1 | ND | ND | Membrane | ND | No | No | No |
| Arl13a | Small GTPase | Cellular process; Localization; Metabolic Process | Intracellular | ND | No | No | No |
| Trmt2b | RNA methyltransferase; RNA binding protein | tRNA processing | Mitochondrion | ND | Yes | No | Yes |
| Tmem35 | ND | ND | Peroxisome membrane; Cytoplasmic vesicle | ND | No | No | No |
| Cenpi | mRNA splicing factor | Metabolic process | Nucleus | ND | No | No | No |

| Drp2 | Non-motor actin binding protein | Cellular component organization/biogenesis; Cellular process | Organelle; Cell part | ND | No | No | Yes |
|---|---|---|---|---|---|---|---|
| Taf7l | Transcription Factor | Biological regulation; Metabolic Process | Nucleus | General transcription regulation; Transcription regulation by bZIP transcription factor | No | No | No |
| Timm8a1 | Metal ion binding | Protein transport; Translocation; Transport | Membrane; mitochondrion | ND | No | No | Yes |
| Btk | Kinase; Transferase; Tyrosine-protein kinase | Adaptive immunity, Apoptosis; Immunity, Innate immunity, Transcription; Transcription regulation | Cytoplasm; Membrane; Nucleus | B cell activation | No | No | No |
| Rpl36al | Structural constituent of ribosome | Translation | Cytoplasm | ND | No | No | Yes |
| Hnrnph2 | Ribosomal protein | ND | Nucleus | ND | No | No | Yes |
| Armcx4 | ND | ND | ND | ND | No | No | Yes |
| Armcx1 | ND | ND | Membrane | ND | No | No | Yes |
| Armcx6 | ND | ND | Membrane | ND | No | No | No |
| Armcx3 | ND | Cellular protein localization | Membrane | ND | No | Yes | No |
| Armcx2 | ND | ND | Membrane | ND | No | No | Yes |
| Nxf2 | RNA binding protein | Localization; Metabolic Process | Nucleus; organelle envelope | ND | No | Yes | No |
| Zmat1 | Zinc finger transcription factor; Nuclease | Apoptotic process; Biological regulation; Developmental Process; Metabolic Process | Nucleus | ND | No | No | No |
| AV320801 | ND | ND | ND | ND | No | No | No |
| Tceal6 | Transcription Factor | Nucleic acid-templated transcription | Nucleus | ND | No | No | No |
| Pramel3 | ND | ND | ND | ND | No | No | No |

| | RNA binding protein | Localization; Metabolic Process | | | | | |
|---|---|---|---|---|---|---|---|
| Nxf7 | ND | | Nucleus | ND | No | No | No |
| Prame | ND | ND | ND | ND | No | No | No |
| 1700008I05Rik | ND | ND | ND | ND | No | No | No |
| Tmsb15a | ND | ND | ND | ND | No | Yes | No |
| Armcx5 | ND | ND | ND | ND | No | No | Yes |
| Gprasp1 | ND | Endosome to lysosome transport; G-protein coupled receptor catabolic process | Cytoplasm | ND | No | No | Yes |
| Gprasp2 | Beta-amyloid binding; G-protein coupled receptor binding | ND | Cytoplasm; Nucleus | ND | No | No | Yes |
| Bhlhb9 | Protein homodimerization activity | Cell death; cell proliferation; Cellular component organization; System development | Extracellular; Nucleus | ND | No | No | Yes |
| Arxes2 | Peptidase activity | Fat cell differentiation; Signal peptide processing | Endoplasmic reticulum | Vasopressin synthesis | No | No | Yes |
| Arxes1 | Peptidase activity | Fat cell differentiation; Signal peptide processing | Endoplasmic reticulum | Vasopressin synthesis | No | No | Yes |
| Bex2 | RNA polymerase II activating transcription factor binding | Apoptosis; Cell cycle | Cytoplasm; Nucleus | ND | No | No | Yes |
| Bex4 | ND | ND | Cytoplasm; Nucleus | ND | No | Yes | Yes |
| Tceal8 | Transcription Factor | Transcription; Transcription regulation | Nucleus | ND | No | No | Yes |
| Tceal5 | Transcription Factor | Metabolic process; Transcription; Transcription regulation | Nucleus | ND | No | No | No |
| Bex1 | Developmental protein | Differentiation; Neurogenesis | Cytoplasm; Nucleus | ND | No | No | Yes |
| Tceal7 | Transcription Factor; Repressor | Metabolic process; Transcription; Transcription regulation | Nucleus | ND | No | Yes | No |
| Wbp5 | Transcription Factor | Transcription; Transcription regulation | Nucleus | ND | No | No | Yes |

111

| Gene | Molecular Function | Biological Process | Cellular Component | | | | |
|---|---|---|---|---|---|---|---|
| Ngfrap1 | Death receptor binding; Metal ion binding | Apoptosis | Cytoplasm; Nucleus | ND | No | No | Yes |
| Kir3dl2 | Membrane-bound signaling molecule Immunoglobulin receptor superfamily; Immunoglobulin receptor superfamily | Cellular process; Immune system process; Response to stimuli | ND | ND | No | No | No |
| Kir3dl1 | membrane-bound signaling molecule, Immunoglobulin receptor superfamily; Immunoglobulin receptor superfamily | Cellular process; Immune system process; Response to stimuli | Plasma membrane | ND | No | No | No |
| Tceal3 | Transcription Factor | Metabolic process | Nucleus | ND | No | No | No |
| Tceal1 | Transcription Factor | Metabolic process | Nucleus | ND | No | No | No |
| Morf4l2 | Transcription factor; Chromatin/chromatin-binding protein | Biological regulation; Metabolic Process | Nucleus | ND | No | Yes | Yes |
| BC065397 | ND | ND | ND | ND | No | No | Yes |
| Glra4 | Chloride channel; Ion channel; Ligand-gated ion channel; Receptor | Ion transport; Transport | Cell junction; Cell membrane | ND | No | No | No |
| Plp1 | Myelin protein | Cellular component organization; Cell differentiation; Lipid metabolic process; Response to stimuli; Signaling; system development | Plasma membrane | ND | No | No | Yes |
| Rab9b | GTP-binding; Nucleotide-binding | Protein transport; Transport | Cell membrane; Cytoplasmic vesicle; Membrane | ND | No | No | No |
| H2bfm | Histone | Cellular component organization/biogenesis; Cellular process; Metabolic Process | Nucleus; Organelle; Macromolecular complex; cell part | ND | No | No | No |

| Gene | Molecular Function | Biological Process | Cellular Component | Pathway | | | |
|---|---|---|---|---|---|---|---|
| Zcchc18 | ND | Biological regulation; Cellular process; Metabolic process; Response to stimulus | Organelle; cell part | ND | No | No | No |
| Slc25a53 | Amino acid transporter; Mitochondrial carrier protein; Transfer/carrier protein; Ribosomal protein; Calmodulin | Localization; Metabolic Process | Cell part | ND | No | No | No |
| Fam199x | ND | ND | ND | ND | No | No | No |
| Esx1 | Homeobox transcription factor; DNA binding protein | Biological regulation; Developmental Process; Metabolic Process; Multicellular organismal process | Nucleus | ND | No | No | No |
| il1rapl2 | Type I cytokine receptor | ND | Membrane | ND | No | No | No |
| Tex13a | ND | ND | ND | ND | No | No | No |
| Nrk | Kinase; Serine/threonine-protein kinase; Transferase | Activation of JNKK activity; Negative regulation of cell proliferation; | Cytoplasm | ND | No | Yes | No |
| Serpina7 | Enzyme regulator | Establishment of localization; Response to stimuli | Extracellular | ND | No | No | No |
| 4930513o06Rik | ND | ND | ND | ND | No | No | No |
| Mum1l1 | ND | ND | Extracellular | ND | No | No | No |
| Trap1a | Tumor antigen | ND | Nucleus | ND | No | Yes | Yes |
| Rnf128 | Ligase; Zinc ion binding | Ubl conjugation pathway | Cytoplasm | Ubl conjugation pathway | No | No | Yes |
| Tbc1d8b | Hydrolase; G-protein modulator | Biological regulation; Cellular component organization/biogenesis; Cellular process; Developmental process; Localization | Endomembrane system; Intracellular | ND | No | No | Yes |
| Ripply1 | Developmental protein; Repressor | Transcription; Transcription regulation | Nucleus | ND | No | Yes | No |
| Cldn2 | Identical protein binding; Structural molecule activity | Calcium-independent cell-cell adhesion via plasma membrane cell-adhesion molecules | Cell junction; Cell membrane | ND | No | No | No |

| Gene | Molecular Function | Biological Process | Cellular Component | Pathway | | | |
|---|---|---|---|---|---|---|---|
| Morc4 | Zinc ion binding | ND | Nucleus | ND | No | No | No |
| Rbm41 | Nucleotide binding, pre-mRNA intronic binding; U12 snRNA binding | Developmental process, mRNA splicing | U12-type spliceosomal complex | ND | Yes | No | Yes |
| Nup62cl | Lipid binding; transporter | Establishment of localization | Nucleus; organelle envelope | ND | No | Yes | Yes |
| E230019M04Rik | ND | ND | ND | ND | No | Yes | No |
| Frmpd3 | ND | ND | ND | ND | No | No | No |
| Prps1 | Kinase; Transferase | Nucleotide biosynthesis | ND | 5-phospho-alpha-D-ribose 1-diphosphate biosynthesis | No | No | Yes |
| Tsc22d3 | Transcription Factor | Biological regulation; Metabolic Process | Nucleus | ND | No | Yes | Yes |
| Mid2 | Cytoskeletal protein binding; ligase | Nucleic acid-templated transcription; Immune system process; Response to stimuli; Signaling | Cytoskeleton; extracellular; Non-membrane bound organelle | ND | No | No | No |
| Tex13 | ND | ND | ND | ND | No | No | No |
| Vsig1 | ND | Cell differentiation; Cellular component organization; Homeostatic process | Plasma membrane | ND | No | No | No |
| Psmd10 | Transcription factor binding | Apoptosis | Cytoplasm; Nucleus | ND | No | No | Yes |
| Col4a6 | ND | Cellular component organization; Response to stimuli; Signaling | Extracellular | Integrin signaling pathway | No | No | No |
| Col4a5 | ND | Cellular component organization; Response to stimuli; Signaling | Extracellular; Synapse | Integrin signaling pathway | No | No | No |
| Irs4 | insulin receptor binding; Phosphatidylinositol 3-kinase binding; Signal transducer activity | Biological regulation; Cellular process; Metabolic Process; Response to stimulus | Membrane; Cell part | Gonadotropin releasing hormone receptor pathway; Insulin/IGF pathway-mitogen activated protein kinase kinase/MAP kinase cascade | No | Yes | No |

| Gene | Protein class | Biological process | Localization | Pathway / Disease | | | |
|---|---|---|---|---|---|---|---|
| Gucy2f | Lyase | cGMP biosynthesis; Sensory transduction; Vision | Membrane | ND | No | No | No |
| Nxt2 | ND | mRNA transport; Protein transport; Transport | Cytoplasm; Nucleus | ND | No | Yes | Yes |
| Kcne1l | Voltage-gated potassium channel | Establishment of localization; Signaling | Plasma membrane | ND | No | No | No |
| Acsl4 | ATP-binding; Magnesium; Nucleotide-binding | Fatty acid metabolism; Lipid metabolism | Membrane | ND | No | No | Yes |
| Tmem164 | ND | | Membrane | ND | No | Yes | Yes |
| Ammecr1 | ND | | ND | ND | No | No | Yes |
| Rgag1 | ND | | ND | ND | No | No | No |
| Chrdl1 | Developmental protein | Differentiation; Neurogenesis; Osteogenesis | Secreted | ND | No | No | No |
| Pak3 | Developmental protein; Kinase; Serine/threonine-protein kinase; Transferase | Cell death; Cell proliferation; Cellular component organization; Protein metabolic process; Response to stimulus; System development | Endosome; Vacuole | Angiogenesis; Cytoskeletal regulation by Rho GTPase; Inflammation mediated by chemokine and cytokine signaling pathway, Ras pathway, T cell activation | No | No | No |
| Capn6 | Cysteine protease | Cellular component organization; Protein metabolic process | Cytoskeleton; Non-membrane bound organelle | Huntington disease | No | No | No |
| Dcx | Non-receptor serine/threonine protein kinase; non-receptor serine/threonine protein kinase; non-motor microtubule binding protein | Biological regulation; Cellular process; Metabolic process; Response to stimulus | Organelle; Cell part | ND | No | No | No |

| | Ubl conjugation pathway | Endoplasmic reticulum | Ubl conjugation pathway | | | |
|---|---|---|---|---|---|---|
| Alg13 | Glycosyltransferase; Hydrolase; Protease; Thiol protease; Transferase | Ubl conjugation pathway | Endoplasmic reticulum | Ubl conjugation pathway | No | No | Yes |
| Trpc5 | Calcium channel;Ion channel | Calcium transport; Ion transport; Transport | Membrane | Alzheimer disease-presenilin pathway | No | No | No |
| Gm15070 | ND | ND | ND | ND | No | No | No |
| Mageb16 | ND | ND | ND | ND | No | No | Yes |
| Zcchc16 | Nucleic acid binding; Zinc ion binding | ND | ND | ND | No | No | No |
| Amot | Angiostatin binding; Receptor activity | Cell differentiation; Cellular component organization; Response to stimulus; Signaling; system development | Cell junction; Tight junction | ND | No | No | No |
| Htr2c | G-protein coupled receptor; Receptor; Transducer | Behavior | Membrane | 5HT2 type receptor mediated signaling pathway | No | No | No |
| il13ra2 | Cytokine; Type I cytokine receptor; Defense/immunity protein | Cellular process; Developmental Process; Immune system process; Response to stimulus | Membrane; Secreted | Interleukin signaling pathway | No | No | No |
| Lrch2 | ND | ND | ND | ND | No | No | No |
| Gm15080 | ND | ND | ND | ND | No | No | No |
| Gm15107 | ND | ND | ND | ND | No | No | No |
| Gm15085 | ND | ND | ND | ND | No | No | No |
| Gm15127 | ND | ND | ND | ND | No | No | No |
| Luzp4 | ND | ND | ND | ND | No | No | No |
| Gm15093 | ND | ND | ND | ND | No | No | No |

| Gene | Molecular function | Biological process | Subcellular location | Pathway | | | |
|---|---|---|---|---|---|---|---|
| Gm10439 | ND | ND | ND | ND | No | No | No |
| Gm15097 | ND | ND | ND | ND | No | No | No |
| Gm15091 | ND | ND | ND | ND | No | No | No |
| Bmi1 | Chromatin regulator; Repressor | Transcription; Transcription regulation | Cytoplasm; nucleus | Oxidative Stress Induced Senescence; SUMOylation of DNA damage response and repair proteins and RNA binding proteins | No | No | Yes |
| Tmem29 | ND | ND | ND | ND | No | No | Yes |
| Apex2 | Endonuclease; Exonuclease; Hydrolase; Lyase; Nuclease | Cell cycle; DNA damage; DNA recombination; DNA repair | Cytoplasm; Mitochondrion; Nucleus | ND | No | No | Yes |
| Alas2 | Acyltransferase; Transferase | Heme biosynthesis | Mitochondrion | Heme biosynthesis | No | No | No |
| Pkfb1 | Carbohydrate phosphatase; Hydrolase; Kinase; Transferase | Carbohydrate derivative | Cytosol | Glycolysis; Gluconeogenesis | No | No | Yes |
| Tro | ND | Cellular component organization | Nucleus | ND | No | No | No |
| Maged2 | ND | ND | ND | ND | No | No | Yes |
| Gnl3l | GTP-binding; Nucleotide-binding | Ribosome biogenesis | Nucleus | ND | No | No | Yes |
| Fgd1 | Guanyl-nucleotide exchange factor | Biological regulation; Developmental Process | Cell projection; Cytoplasm; Cytoskeleton | NRAGE signals death through JNK; Rho GTPase cycle; G alpha (12/13) signaling events | No | No | Yes |
| Tsr2 | ND | ND | Nucleus | ND | No | No | Yes |

4

| | Molecular function | Biological process | Cellular component | Pathway | | | |
|---|---|---|---|---|---|---|---|
| Wnk3 | Kinase; Serine/threonine-protein kinase; Transferase | ND | Cytoplasm | Stimuli-sensing channels | No | No | No |
| Fam120c | poly(A) RNA binding | ND | Nucleus | ND | No | No | Yes |
| Phf8 | Activator; Chromatin regulator; Dioxygenase; Oxidoreductase | Cell cycle; Transcription; Transcription regulation | Nucleus | HDMs demethylate histones; Condensation of Prophase Chromosomes | No | No | Yes |
| Huwe1 | Ligase | Differentiation; DNA damage; DNA repair; Ubl conjugation pathway | Cytoplasm; Nucleus | Ubiquitin proteasome pathway; Antigen processing; Ubiquitination & Proteasome degradation | No | No | Yes |
| Hsd17b10 | Dehydrogenase; reductase; Lipid binding; RNA binding | Cellular component organization | Endoplasmic reticulum; mitochondrion; Non-membrane bound organelle; Organelle lumen | ND | No | No | Yes |
| Ribc1 | ND | ND | ND | ND | No | No | Yes |
| Smc1a | Chromatin/chromatin-binding protein; Hydrolase | Cellular component organization/biogenesis; Cellular process | Centromere; Chromosome; Nucleus | Meiotic synapsis; regulation of Sister Chromatid process; Cohesin Loading onto Chromatin; SUMOylation of DNA damage response and repair proteins | No | No | Yes |
| Iqsec2 | ND | Cellular component organization; Response to stimuli; Signaling | ND | ND | No | No | No |

| Gene | Protein class | Biological process | Cellular location | HDMs demethylate histones | | | |
|---|---|---|---|---|---|---|---|
| Kdm5c | Zinc finger transcription factor | Reproduction | Nucleus | HDMs demethylate histones | Yes | No | Yes |
| Tspyl2 | Phosphatase inhibitor; Chaperone; Chromatin regulator | Apoptotic process; Biological regulation; Cellular component organization/biogenesis; Cellular process; Developmental process; Metabolic Process | Cytoplasm; Nucleus | ND | No | No | Yes |
| Gpr173 | Gonadotropin-releasing hormone receptor activity | Negative regulation of neuron migration | Membrane | ND | No | No | No |
| Gm7157 | ND | ND | ND | ND | No | No | No |
| Shroom2 | Ligand-gated ion channel | ND | Cell junction; Cell membrane; Cytoplasm; Cytoskeleton; Membrane; Microtubule; Tight junction | ND | No | No | Yes |
| Gpr143 | G-protein coupled receptor; Receptor; Transducer | ND | Membrane | Amine ligand-binding receptors | No | No | No |
| Usp51 | Hydrolase | Nucleic acid templated transcription; Cellular component organization; Protein metabolic process | Nucleus, Organelle lumen | ND | No | No | No |
| Mageh1 | Tumor antigen | ND | Cytoplasm | ND | No | No | Yes |
| Foxr2 | Transcription factor DNA binding protein | Biological regulation; Cellular process; Metabolic process | Organelle; Cell part | ND | No | No | No |
| Rragb | GTP-binding; Nucleotide-binding; Hydrolase | Response to stimulus; Signaling | Endosome; Vacuole; Golgi apparatus; Nucleus | Macro autophagy; mTOR signaling; Energy dependent regulation of mTOR by LKB1-AMPK; TP53 Regulates Metabolic Genes | No | No | Yes |
| Klf8 | DNA binding | Transcription; Transcription regulation | Nucleus | ND | No | No | Yes |

4
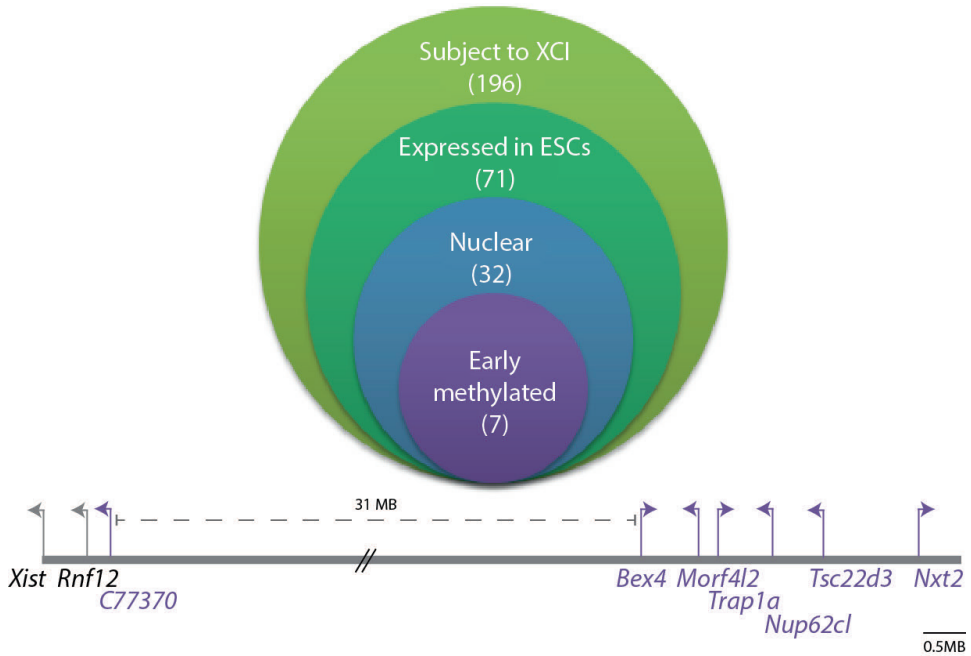
**Figure 4:** Selection of candidate genes in the 50Mb deleted region. First, genes were selected which were subject to XCI. Of these genes, genes were selected which were expressed in ESCs and were localized in the nucleus. And finally, we selected genes which were also silenced early in XCI as observed by DNA methylation. Seven genes met these criteria: *C77370, Bex4, Morf4l2, Trap1a, Tsc22d3* and *Nxt2*. The location of these genes relative to the Xic is depicted below.

## References

1. Monkhorst, K. et al. (2008) X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. Cell 132 (3), 410-21.

2. Brockdorff, N. (2013) Noncoding RNA and Polycomb recruitment. RNA 19 (4), 429-42.

3. Plath, K. et al. (2002) Xist RNA and the mechanism of X chromosome inactivation. Annu Rev Genet 36, 233-78.

4. Peeters, S.B. et al. (2014) Variable escape from X-chromosome inactivation: identifying factors that tip the scales towards expression. Bioessays 36 (8), 746-56.

5. Skaletsky, H. et al. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423 (6942), 825-37.

6. Gribnau, J. and Grootegoed, J.A. (2012) Origin and evolution of X chromosome inactivation. Curr Opin Cell Biol 24 (3), 397-404.

7. Barakat, T.S. et al. (2010) X-changing information on X inactivation. Experimental Cell Research 316 (5), 679-687.

8. Galupa, R. and Heard, E. (2015) X-chromosome inactivation: new insights into cis and trans regulation. Current Opinion in Genetics & Development 31, 57-66.

9. Lee, J.T. et al. (1999) Tsix, a gene antisense to Xist at the X-inactivation centre. Nat Genet 21 (4), 400-4.

10. Barakat, T.S. et al. (2014) The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. Mol Cell 53 (6), 965-78.

11. Tian, D. et al. (2010) The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. Cell 143 (3), 390-403.

12. Chureau, C. et al. (2011) Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. Hum Mol Genet 20 (4), 705-18.

13. Sun, S. et al. (2010) Characterization of Xpr (Xpct) reveals instability but no effects on X-chromosome pairing or Xist expression. Transcription 1 (1), 46-56.

14. Navarro, P. et al. (2008) Molecular coupling of Xist regulation and pluripotency. Science 321 (5896), 1693-5.

15. Navarro, P. and Avner, P. (2009) When X-inactivation meets pluripotency: an intimate rendezvous. FEBS Lett 583 (11), 1721-7.

16. Navarro, P. and Avner, P. (2010) An embryonic story: analysis of the gene regulative network controlling Xist expression in mouse embryonic stem cells. Bioessays 32 (7), 581-8.

17. Navarro, P. et al. (2011) The X-inactivation trans-activator Rnf12 is negatively regulated by pluripotency factors in embryonic stem cells. Hum Genet 130 (2), 255-64.

18. Barakat, T.S. et al. (2011) RNF12 activates Xist and is essential for X chromosome inactivation. PLoS Genet 7 (1), e1002001.

19. Gontan, C. et al. (2012) RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature 485 (7398), 386-90.

20. Jonkers, I. et al. (2009) RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. Cell 139 (5), 999-1011.

21. Kraft, K. et al. (2015) Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice. Cell Rep.

22. Li, J. et al. (2015) Efficient inversions and duplications of mammalian regulatory DNA elements and gene clusters by CRISPR/Cas9. J Mol Cell Biol 7 (4), 284-98.

23. Zhang, L. et al. (2015) Large genomic fragment deletions and insertions in mouse using

CRISPR/Cas9. PLoS One 10 (3), e0120396.

24. Ran, F.A. et al. (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8 (11), 2281-308.

25. Jonkers, I. et al. (2008) Xist RNA is confined to the nuclear territory of the silenced X chromosome throughout the cell cycle. Molecular and Cellular Biology 28 (18), 5583-94.

26. Cong, L. et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339 (6121), 819-23.

27. Shevchenko, A.I. et al. (2007) Genes flanking Xist in mouse and human are separated on the X chromosome in American marsupials. Chromosome Res 15 (2), 127-36.

28. Boyd, Y. et al. (2000) A phenotype map of the mouse X chromosome: models for human X-linked disease. Genome Res 10 (3), 277-92.

29. Yang, F. et al. (2010) Global survey of escape from X inactivation by RNA-sequencing in mouse. Genome Res 20 (5), 614-22.

30. Marks, H. et al. (2015) Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. Genome Biol 16, 149.

31. Calabrese, J.M. et al. (2012) Site-specific silencing of regulatory elements as a mechanism of X inactivation. Cell 151 (5), 951-63.

32. Splinter, E. et al. (2011) The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes Dev 25 (13), 1371-83.

33. Li, S.M. et al. (2012) Transcriptome-wide survey of mouse CNS-derived cells reveals monoallelic expression within novel gene families. PLoS One 7 (2), e31751.

34. Berletch, J.B. et al. (2015) Escape from X inactivation varies in mouse tissues. PLoS Genet 11 (3), e1005079.

35. Wu, H. et al. (2014) Cellular resolution maps of X chromosome inactivation: implications for neural development, function, and disease. Neuron 81 (1), 103-19.

36. Lopes, A.M. et al. (2011) Clustered transcripts that escape X inactivation at mouse XqD. Mamm Genome 22 (9-10), 572-82.

37. Essletzbichler, P. et al. (2014) Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. Genome Res 24 (12), 2059-65.

**4**
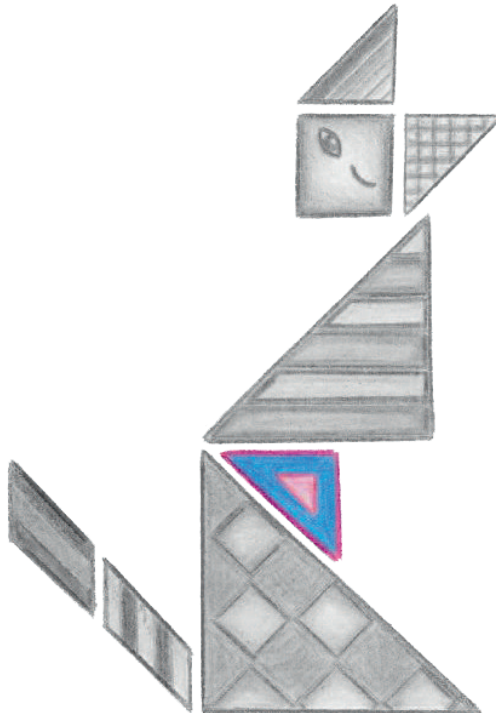
# Chapter 5

**Characterization of histone modifications associated with inactive X-chromosome in Trophoblast Stem Cells, eXtra-embryonic Endoderm cells and in in vitro derived undifferentiated and differentiated Epiblast Like Stem cells**

# Chapter 5

*Catherine Dupont, Cheryl M Maduro, Hannah Den Braanker, Ruben G Boers, Dorota Kurek, Joost Gribnau.*

*Erasmus MC*
*Department of Developmental Biology*
*Wytemag 80*
*3015 CN Rotterdam*
*The Netherlands*

**Abstract**

In mouse, X-chromosome inactivation (XCI) can either be imprinted or random. Imprinted XCI (iXCI) is considered unstable and depending on Polycomb Repressive Complexes, whereas random XCI (rXCI) is classified as stable and depending on DNA methylation. Here we have systematically examined epigenetic modifications associated with the inactive X-chromosome (Xi) in Trophoblast Stem cells, eXtra-Embryonic Endoderm Cells, undifferentiated and differentiated Epiblast Stem Cells in order to understand intrinsic differences in epigenetic mechanisms involved in silencing of the inactive X-chromosome in lineages presenting iXCI and rXCI. Whereas euchromatic histone modifications are predominantly lost from the Xi territory in all cell types, the accumulation of heterochromatic modifications diverges in between the analysed cell lineages. Particularly, only the Xi of multipotent Trophoblast (iXCI) and Epiblast stem cells (rXCI).display a visible accumulation of Polycomb Repressive Complexes (PRCs), in contrast to the Xi in differentiated Epiblast Stem Cells and eXtra-embryonic Endoderm cells. Despite this, the histone modifications catalysed by PRCs; ubH2AK119 and H3K27me3, remain the best heterochromatic markers for the Xi in all assessed lineages. Heterochromatic chromatin modifications associated with the Xi are a reflection of the epigenetic landscape of the entire genome of the assessed cell regardless whether XCI is imprinted or random.

**Introduction**

Sex determination in mammals is determined by the presence of the male sex determining gene SRY located on the Y-chromosome. Female mammals have two X-chromosomes, whereas males possess one X-chromosome and one Y-chromosome. This imbalance is compensated through inactivation of one of the two X-chromosomes in all female cells. The actual silencing of the X chromosome in placental mammals is a highly dynamic and complex process. A crucial player in initiation of silencing of the X is the long non-coding RNA *Xist* [1, 2]. *Xist* RNA

recruits specific protein complexes, which trigger, a cascade of epigenetic events resulting in the inactivation of the *Xist*-expressing X-chromosome [3]. A widely used animal model to study X-inactivation (XCI) is the mouse. In the female mouse embryo, *Xist* starts to be expressed during early embryogenesis from the 2-cell stage onwards, leading to silencing in cis. This form of XCI is referred as imprinted XCI (iXCI), as it exclusively leads to XCI of the paternally derived X-chromosome. Whereas all developing extra-embryonic lineages maintain iXCI, lineages that will form the embryo proper characteristically erase iXCI and re-establish XCI in a random manner (rXCI) [4]. *In vitro* differentiation of embryonic stem (ES) cells derived from the inner cell mass (ICM) has provided quite detailed information on the sequence of epigenetic events assisting in the inactivation of one of the X-chromosomes in embryonic tissues [5-11]. In differentiating ES cells the first epigenetic event following the accumulation of *Xist* is the loss of euchromatic marks such as methylation of histone H3K4 and acetylation of H3K9. Subsequently, characteristic repressive histone modifications like methylation of H3K27, H3K9 and H4K20 and ubiquitination of H2A can be detected on the Xi. XCI in extra-embryonic tissues is, in contrast to fully differentiated embryonic tissues, considered unstable [12-16]. In order to understand how and why XCI is stable or unstable and if epigenetic events differ between rXCI and iXCI, a full characterization of chromatin modifications in lineages of differing origin is necessary.

In this study, we have systematically characterized histone modifications associated with the inactivated X-chromosome (Xi) in trophoblast stem (TS) cells, eXtra-embryonic Endoderm (XEN) cells, *in vitro* derived Epiblast Stem Cells (EpiLCs) and to mesoderm differentiated EpiLCsEpiLCs. The obtained data were completed with reported data of chromatin modifications on the Xi in pre-implantation embryos (Table 1) and cell lineages directly derived from the pre- and early post-implantation embryo (Table 2). This study has generated a comprehensive overview of the epigenetic landscape of the Xi in different cell lineages presenting either iXCI or rXCI.

**Table 1.** *Chromatin Marks associated with the Xi in pre-implantation embryos*

| Chromatin Marks Xi | 4-cell | 8-cell | Morula | Blastocyst |
|---|---|---|---|---|
| H3K27me3 | | Absent [17, 18] | Present [17, 18] Absent [6] | Present [17, 18] Present [6] |
| H4K20me1 | | | | |
| H3K9me2 | | Absent [17, 18] | Absent [17, 18] | Present [17, 18] |
| ubH2A | | | | |
| PRC1 | | | | |
| PRC2 | | Absent [17, 18] | Present [17, 18] Absent [6] | Present [17, 18] Present [6] |
| RNA Pol II exclusion | Present [17, 18] | Present [17, 18] | Present [17, 18] | Present [17, 18] |
| H3K4me2 exclusion | Absent [17, 18] | Present [18]; | Present [17, 18] | Present [17, 18] |
| H3K9ac exclusion | Absent [17, 18] | Present [18] | Present [17, 18] | Present [17, 18] |
| H4ac exclusion | | | Present [17, 18] | Present [17, 18] |
| H4K16Ac exclusion | | | | |

**Table 2.** *Chromatin Marks associated with the Xi cell lineages*

| Chromatin Marks Xi | ES Differentiation | MEF | TS Cell | Trophoblast and differentiated TSC | XEN |
|---|---|---|---|---|---|
| H3K27me3 | Transient [5-7] | Absent [6] | Present [5, 6, 19, 20] | Transient [5, 6]<br><br>Stable [14] | Present [21]<br><br>Absent [19, 20] |
| H4K20me1 | Transient [7] | | Present [19] | Transient [14] | |
| H3K9me2 | Stable [7-10]<br><br>Absent<br><br>[7] | Stable [8]<br><br>Absent [7] | Minorly present [20] | | Minorly present [20] |
| ubH2A | Transient [5] | | Present [5, 19] | Transient [5] | |
| Polycomb Repressive Complex 1 | Transient [5] | | Present [5, 19] | Transient [5] | |
| Polycomb Repressive Complex 2 | Transient [5-7] | Absent [6] | Present [5, 6, 19, 20, 22] | Transient (Ezh2 and Eed1)<br><br>[5, 6]<br><br>Stable (Eed1)<br><br>[14] | Absent [19, 20] |
| RNA Polymerase II exclusion | Stable [11] | | | Absent to partially present [14] | |
| H3K4me2 exclusion | Stable [7-9] | Present [8] | Present [19, 20] | Minorly present [14] | Present [20] |
| H3K9ac exclusion | Stable [8, 9] | Present [8] | Present [20] | | Present [20] |
| H4ac exclusion | Stable [8, 9, 23] | Present [8] | Present [19] | Absent to partially present [14] | |
| H4K16Ac exclusion | Stable [9] | | | Absent [14] | |
| H3Ac exclusion | | | Present [19] | | |

## Results

Despite the wealth of experiments, a complete and comprehensive overview of all histone modifications associated with the Xi in cell types of different embryonic lineages is lacking. We therefore generated TS, XEN, and ES cells from pre-implantation embryos with the same genomic back ground, and differentiated the ES cell lines into EpiLCsEpiLCs, that were further differentiated towards the mesodermal lineage using WNT3 and BMP4 ligands. For our studies we examined Xi and *Xist* associated histone modifications in TS and XEN cell lines, and in undifferentiated and differentiated EpiLCs with an embryonic origin. The obtained results were compared to available data in the literature (reviewed in Table 1 and 2).

**Loss of euchromatic marks on the Xi**

Previous studies indicate that the first epigenetic changes observed on the *Xist* coated X are related to loss of histone modifications, H3K4me2, H3K9ac, H4ac, H4K16ac and RNA polymerase II, all associated with active chromatin. To test whether these markers were depleted throughout our panel of cell lines we performed RNA FISH for *Xist* RNA in combination with immunohistochemistry for these histone modifications on TS (Fig S1A), XEN cells (Fig S2A), EpiLCs (Fig S3A) and differentiated EpiLCs (Fig S4A). To quantify the results, 53 to 354 cells were counted and the percentage of cells displaying *Xist* clouds with and without co-localization of lost euchromatic marks was determined (Fig 1 and 2). Although the detection varied per cell type, loss of euchromatic marks is a feature that is present in a high percentage of cells in all lineages, indicating that the loss of euchromatic marks is detected in lineages that are both independent (differentiated EpiLCs) and fully dependent on *Xist* expression (TS and XEN) for maintenance of XCI (Fig 3).
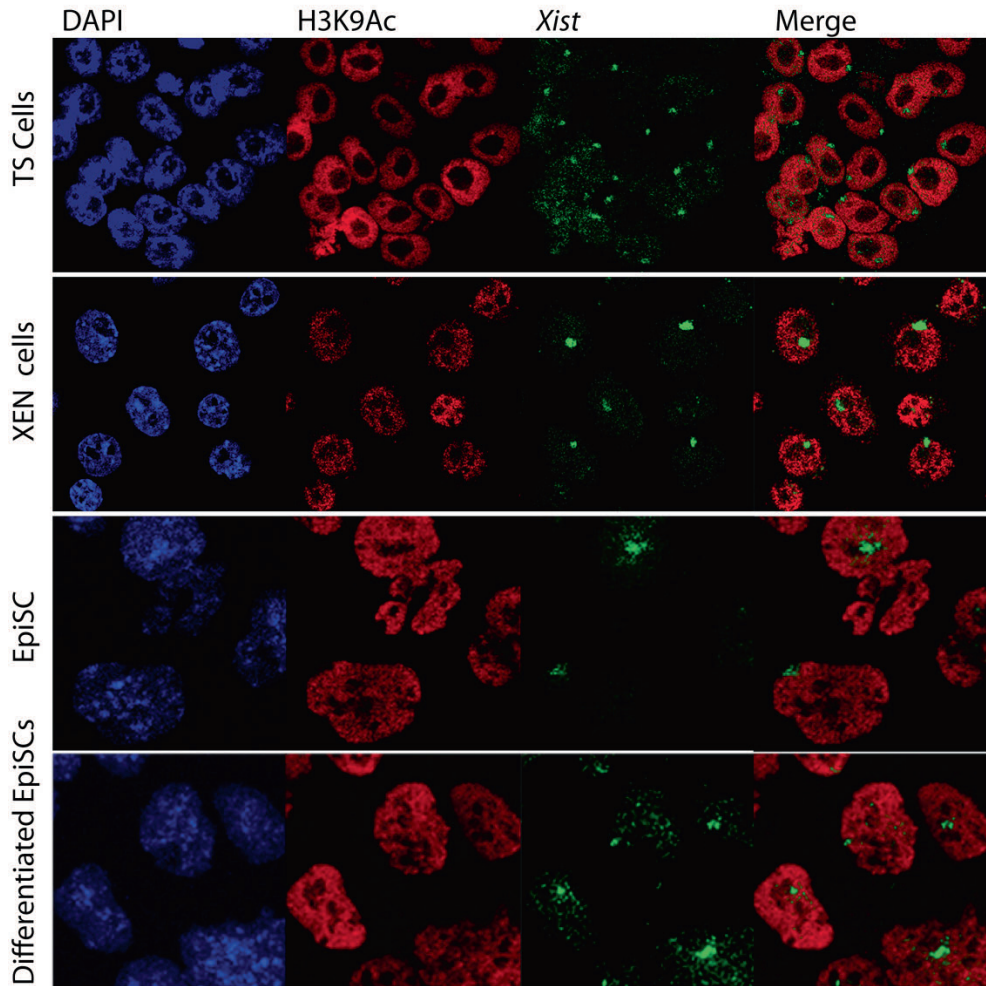
**Figure 1.** Immuno-RNA FISH detecting H3K9Ac (Rhodamine red) and *Xist* (FITC) on TS cells, XEN cells, EpiLCs and differentiated EpiLCs (DAPI is DNA).
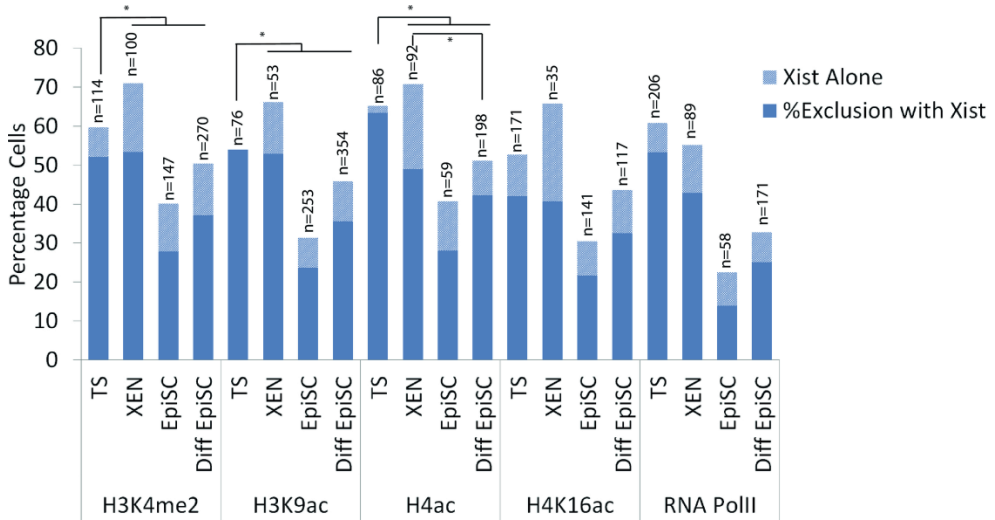
**Figure 2.** Percentage of cells accumulating either *Xist* alone or showing *Xist* together with exclusion of the euchomatic marks H3K4me2, H3K9ac, H4ac, H4K16ac, and RNAPolII. Statistical significance ($p<0,05$) shown in the Figure and tested via z-test for proportion independent groups.

*Polycomb repressive complexes*

Silencing of the X chromosome is thought to proceed via the recruitment of polycomb repressive complexes (PRC) 1 and 2. While each complex consists of several proteins, for our studies only RNF2/RING1B has been assessed from the PRC1 complex, and JARID2, EZH2 and EED were evaluated as representative of the PRC2 complex, using the same panel of TS, XEN, EpiLC and differentiated EpiLC lines. RNA-immuno FISH was performed detecting *Xist* in combination with one of the PRC complex members. The number of cells counted per staining is displayed in Fig 4. As previously reported [5, 6, 19, 20], TS cell lines distinctly displayed accumulation of PRC2 on the Xi (Fig S4A and S4B). Whereas PRC2 associated proteins to some extent also clearly accumulated on the Xi in undifferentiated EpiLCs, Xi associated accumulation of JARID2, EZH2 and SUZI12 was not detected in XEN cells or in differentiated EpiLCs (Fig 5A, 5B, S1B, S2B, S3B and S4B). JARID2 did show accumulation in XEN cells, but the accumulation was not associated with *Xist*. PRC2 catalyses trimethylation of lysine 27 on histone H3 (H3K27me3). Immuno-RNA-FISH analysis detecting H3K27me3 and *Xist* confirms that accumulation of this modification is present at the Xi in all cell types and indicates that even when the catalysing complex is below the detection limit, as found in XEN and differentiated EpiLC cells, its downstream modification is maintained (Fig 5A and 5B). Nevertheless, the percentage of XEN cells and differentiated EpiLC cells with *Xist* clouds together with H3K27me3 accumulation is reduced compared to the other cell lines (Fig 4). The percentage of cells displaying an accumulation of H3K27me3 in differentiated EpiLCs was much higher when only immunostaining was performed. This was observed in all cell types studied and must be related to the more stringent conditions used for *Xist* RNA-FISH.

RING1B-*Xist* immuno-RNA FISH indicated clear Xi associated accumulation of RING1B in TS but not XEN cells, whereas this analysis failed for EpiLCs and differentiated EpiLCs. For XEN cells, EpiLCs and differentiated EpiLCs we therefore performed double

immunohistochemistry of RING1B and ubH2AK119, as ubH2AK119 as the catalytic product of this enzyme is the best marker for the Xi in XEN, EpiLC and differentiated EpiLC (see data below). This approach indicated that there was no clear accumulation of RING1B on the Xi in XEN cells, EpiLCs and differentiated EpiLCs; although sporadically an accumulation could be observed (Fig 6). UbH2AK119, however, could be observed in all analysed cell lineages (Fig 4, 6), again indicating that the detection of Xi associated enzyme complexes is not a requirement for the maintained presence of a corresponding chromatin modification. Similarly to H3K27me3, an immunostaining of ubH2AK119 in differentiated EpiLCs detected a higher percentage of cells presenting the modification.

| | XEN | TS | EpiSC | Diff EpiSC |
|---|---|---|---|---|
| **RNA PolII** | 77,8 | 87,7 | 61,9 | 76,5 |
| **H4Ac** | 69,2 | 97,3 | 68,9 | 82,9 |
| **H3K9Ac** | 80,0 | 100,0 | 75,9 | 77,8 |
| **H4K16Ac** | 61,9 | 79,8 | 71,1 | 74,6 |
| **H3K4me2** | 75,1 | 87,2 | 69,3 | 73,7 |

**Figure 3**.Percentage of *Xist* positive cells showing exclusion of euchromatic modifications at the Xi.
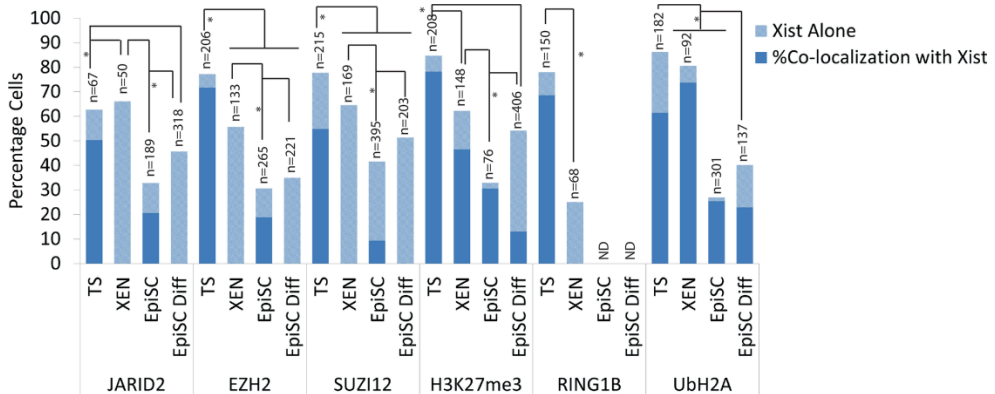


**Figure 4.** Percentage of cells accumulating either *Xist* alone or *Xist* together with accumulated members of the PRC1 and 2 complexes and their catalysed modifications ubH2AK119 and H3K27me3. Statistical significance (p<0,05) shown in the Figure and tested via z-test for proportion independent groups.

### Accumulation of other heterochromatic markers

H3K9me2 and H4K20me1 are both histone modifications associated with the Xi but not as well characterized as H3K27me3 and ubH2AK119. Several lysine 9 histone H3 methyltransferases, euchromatic and heterochromatic, have been characterized but it is unclear which enzyme catalyzes this modification on the Xi. We found that H3K9me2 accumulation, associating with

*Xist* clouds, could only be observed in XEN and TS cells, but was not as abundant as H3K27me3 and ubH2AK119 (Fig 7A, 8, 9). Immuno-RNA FISH indicates that H4K20me1, likely catalysed by SET8/PR-Set7 [24, 25], is a better marker for the Xi than H3K9me2, and also accumulates to a minor extent in EpiLCs (Fig 7B, 8 and 9). In conclusion, we found that accumulation of PRC1 and PRC2 is very variable on the Xi, but that the modifications, ubH2A119 and H3K27me3, catalysed by these complexes are better Xi associated chromatin markers compared to H3K9me2 and H4K20me1.
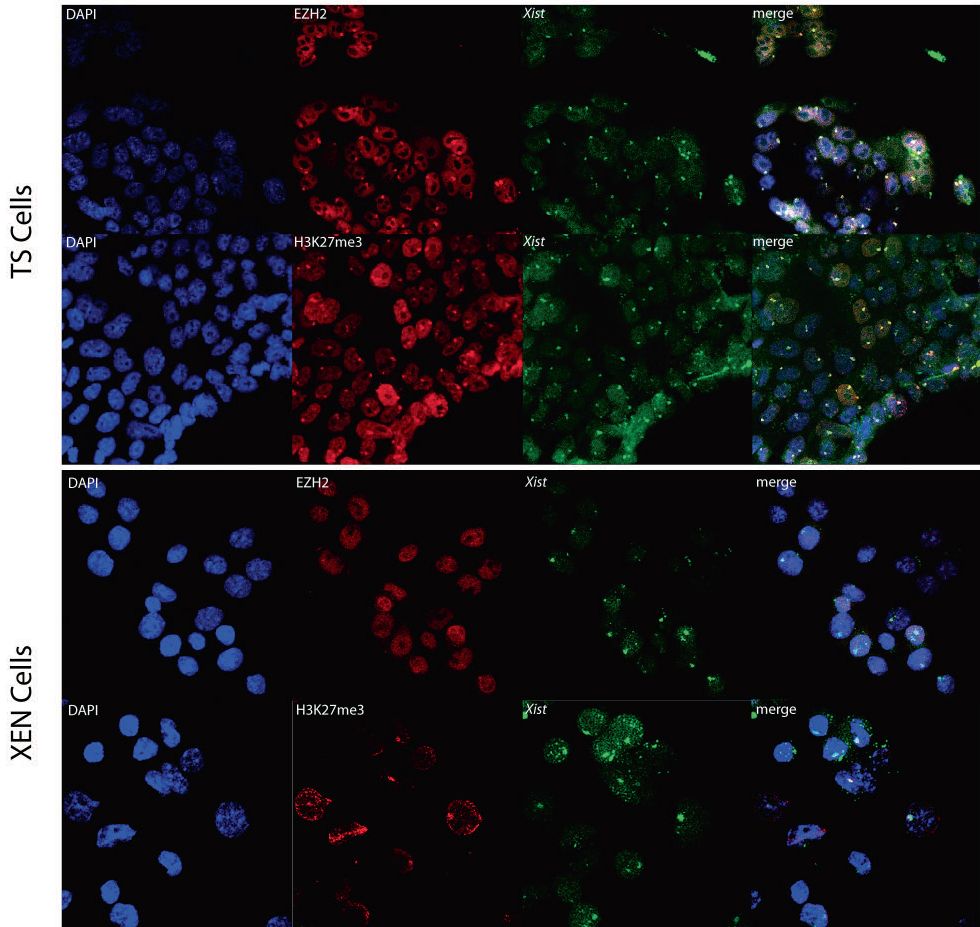


**Figure 5A**. Immuno-RNA FISH detecting EZH2 or H3K27me3 (Rhodamine red) together with *Xist* (FITC) on TS cells and XEN cells (DAPI is DNA).
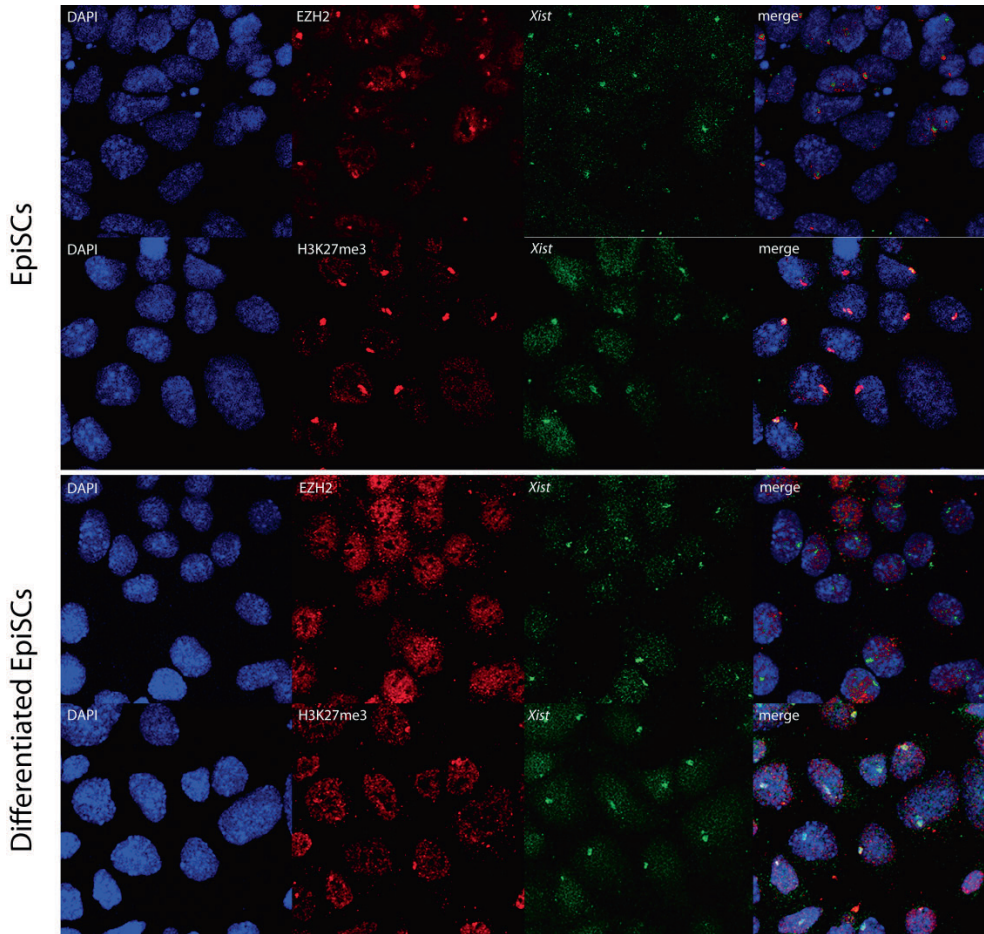
**Figure 5B**. Immuno-RNA FISH detecting EZH2 or H3K27me3 (Rhodamine red) together with *Xist* (FITC) on TS EpiLCs and differentiated EpiLCs (DAPI is DNA).
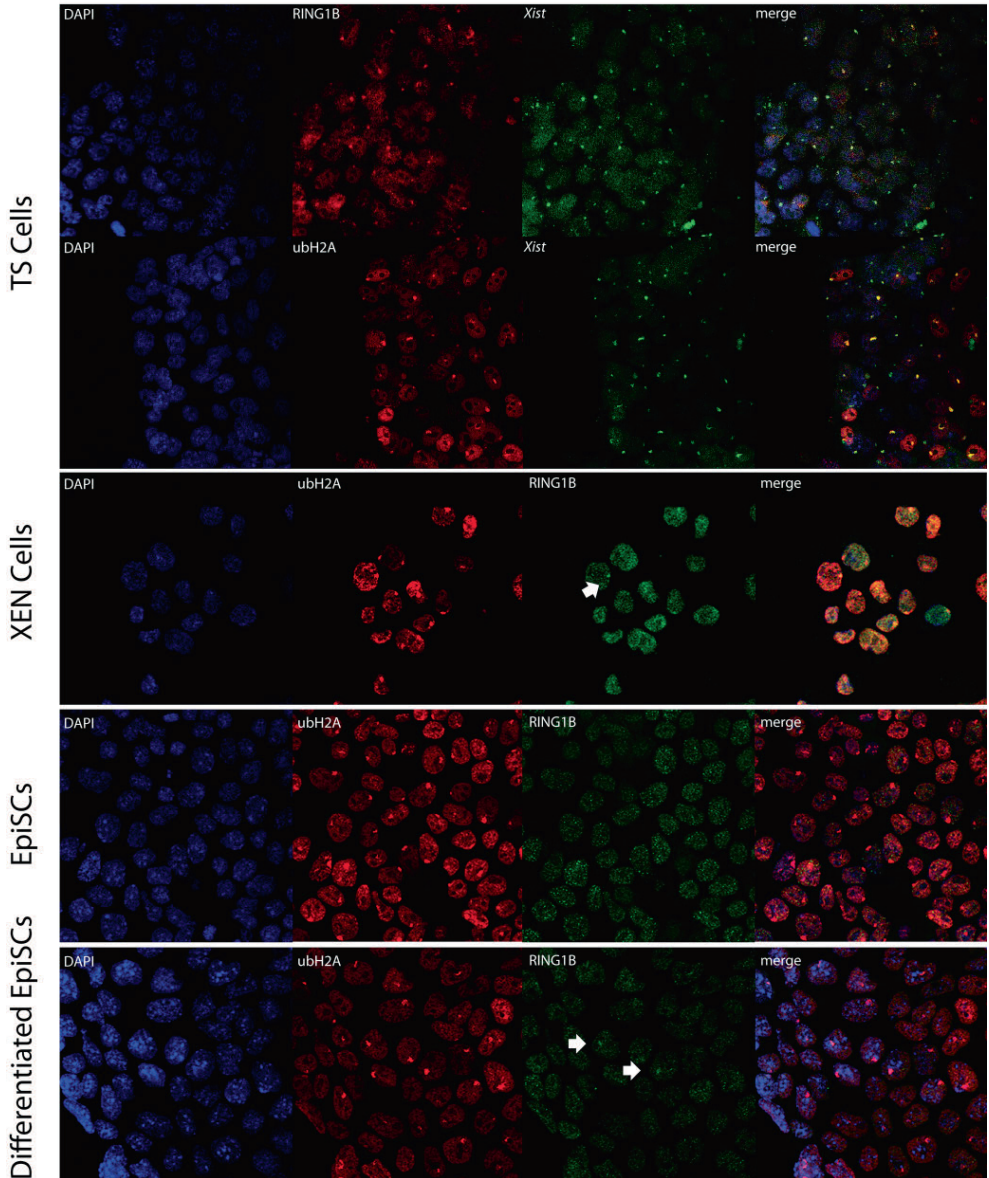
**Figure 6.** Immuno-RNA FISH detecting ubH2AK119 (Rhodamine red) together with RING1B or *Xist* (FITC) on TS cells, XEN cells, EpiLCs and differentiated EpiLCs (DAPI is DNA).
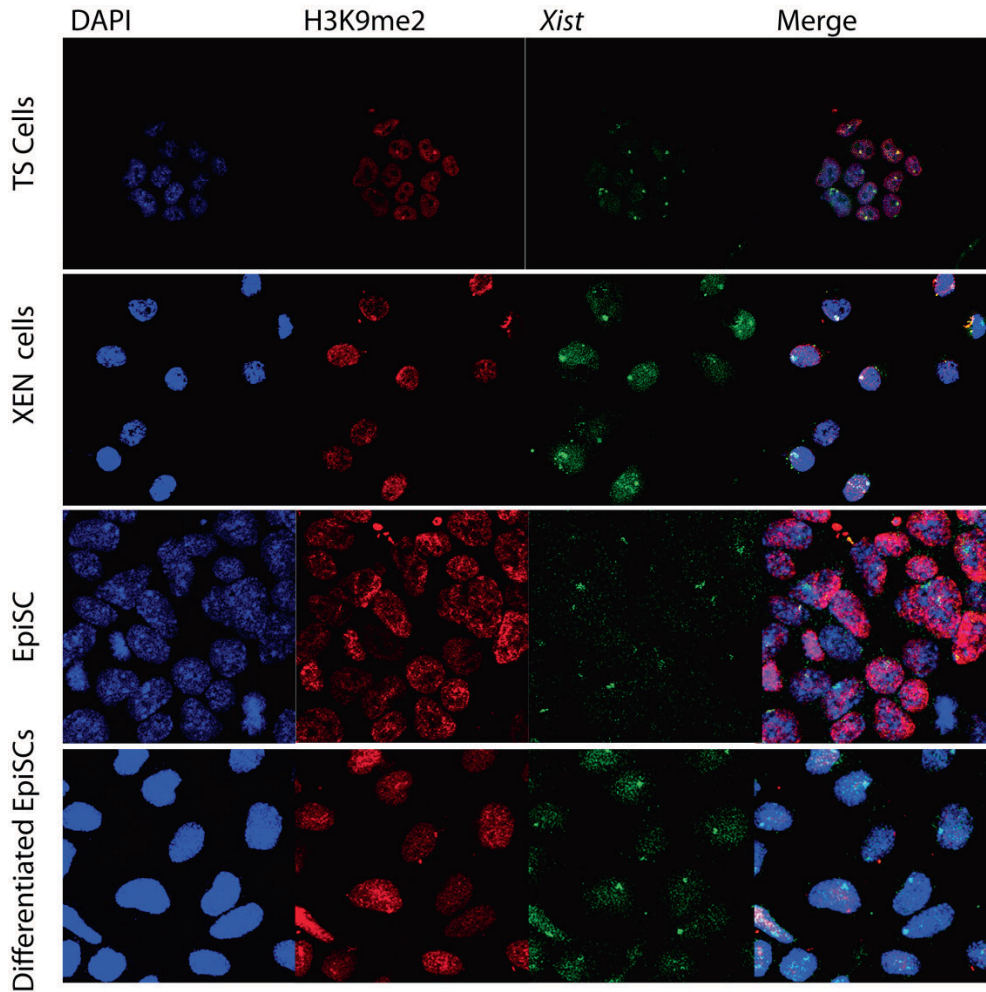
**Figure 7A**. Immuno-RNA FISH detecting H3K9me2 (Rhodamine red) together with *Xist* (FITC) on TS cells, XEN cells, EpiLCs and differentiated EpiLCs (DAPI is DNA).
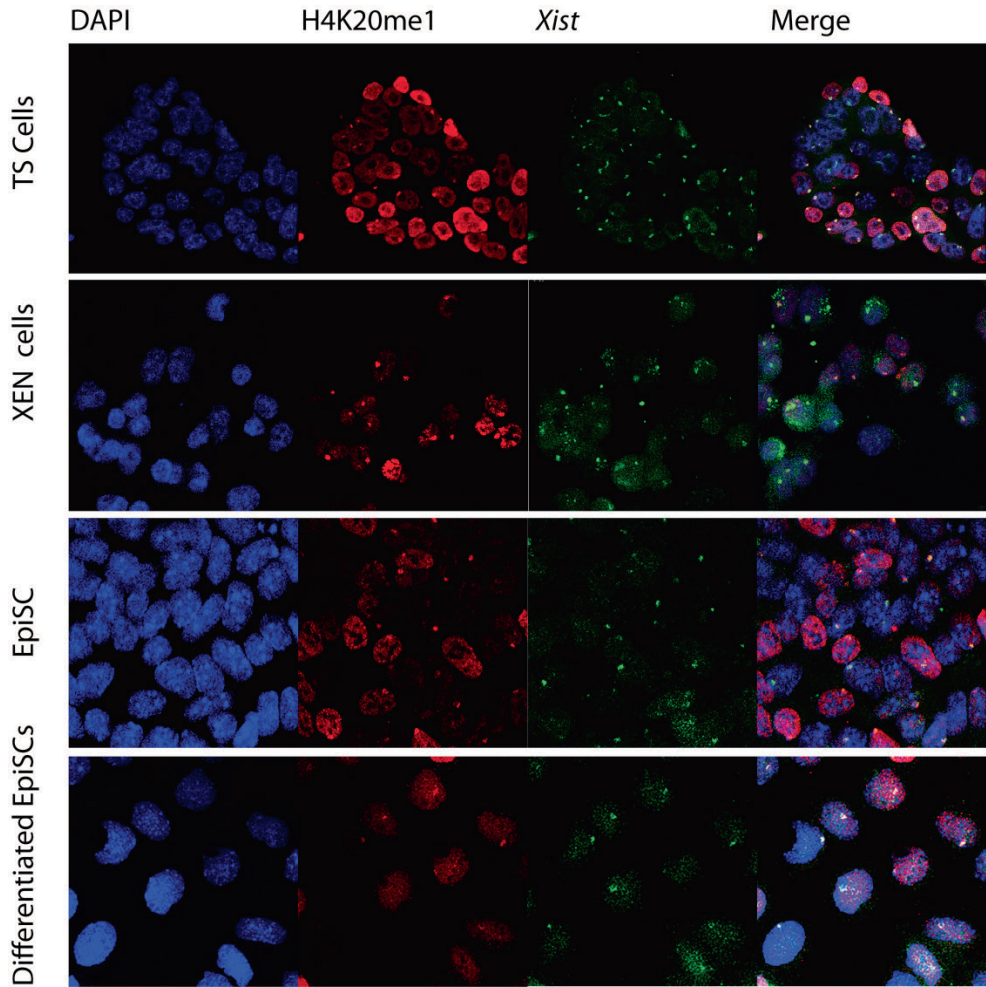
**Figure 7B**. Immuno-RNA FISH detecting H4K20me1 (Rhodamine red) together with *Xist* (FITC) on TS cells, XEN cells, EpiLCs and differentiated EpiLCs (DAPI is DNA).
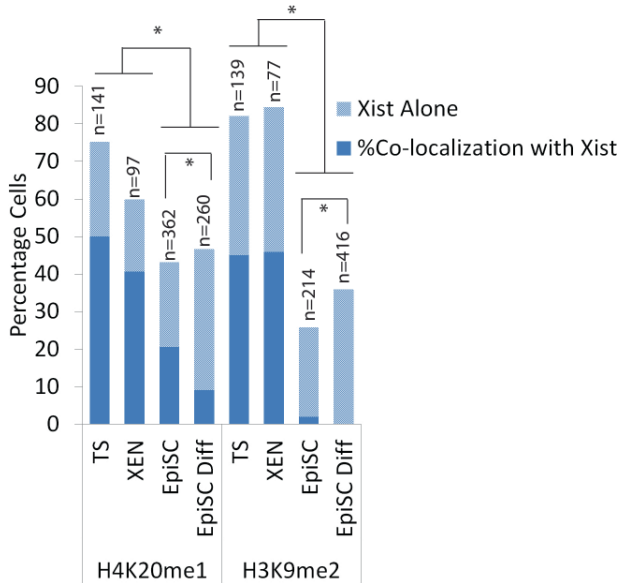
**Figure 8.** Percentage of cells accumulating either *Xist* alone or *Xist* together with accumulated histone modifications H4K20me1 and H3K9me2. Statistical significance (p<0,05) shown in the Figure and tested via z-test for proportion independent groups.

| | XEN | TS | EpiSC | Diff EpiSC |
|---|---|---|---|---|
| **JARID2** | 0,0 | 80,2 | 62,7 | 0,0 |
| **EZH2** | 0,0 | 92,9 | 61,7 | 0,0 |
| **SUZI12** | 0,0 | 70,5 | 22,5 | 0,0 |
| **H3K27me3** | 74,9 | 92,4 | 92,9 | 24,1 |
| **RING1B** | 0,0 | 87,9 | N.D | N.D |
| **UbH2A119** | 91,7 | 71,2 | 94,5 | 56,8 |
| **H4K20me1** | 67,8 | 66,5 | 47,4 | 19,3 |
| **H3K9me2** | 54,3 | 54,9 | 7,6 | 0,0 |

**Figure 9**. Percentage cells accumulating Xi associated heterochromatic modifications in *Xist* expressing cells.

## Discussion

Random and imprinted XCI are respectively depicted as stable and unstable. Since silencing of the X-chromosome relies on histone modifications, we tried to understand whether the stability of XCI could be simply conveyed to the presence or absence of histone modifications associated with the Xi. To address this question, reported chromatin modifications (Table 1 and 2) associated with the Xi were assessed in all early cell lineages presenting either iXCI or rXCI. XEN and TS cells representing cell lineages with iXCI, whereas undifferentiated and differentiated EpiLC portrayed lineages with rXCI. This is the first study that provides a comprehensive overview of such a wide range of chromatin modifications associated with the Xi in different embryonic and extra-embryonic cell types with the same genetic background.

Upon ES cell differentiation, among the first epigenetic changes detectable after *Xist* accumulation are the loss of euchromatic features including RNA polymerase II, H3K9Ac, H4Ac, H4K16Ac. Our study indicates that this loss of euchromatic marks is a general characteristic of the Xi in a large percentage of cells in all lineages examined, and is central to the iXCI and rXCI processes. Recently, SPEN family members and co-factors have been identified as interactors of *Xist*. SPEN mediated recruitment of HDAC3 has been implicated in triggering histone deacetylation in differentiated ES cells [26-30]. Whether iXCI also involves SPEN mediated recruitment of HDAC3, and if H3K4 demethylases are actively recruited to the Xi to trigger the loss of RNA polymerase II needs further investigation [31].

The loss of active histone marks observed during ES cell differentiation is soon followed by the accumulation of histone modifications associated with facultative heterochromatin. In contrast to the absence of euchromatic marks we found that these heterochromatic features associated with the Xi were very heterogenic (Fig.10). Although H3K27me3 could be detected in all cell types, the PRC2 subunits EZH2, SUZ12 and JARID2, catalysing this modification, were only present on the Xi in both TS cells and EpiLCs. Similarly, ubH2AK119 accumulation associated with the Xi, could be detected in all cell types, despite the absence of Xi associated RING1B in XEN cells and undifferentiated and differentiated EpiLCs. The failure to detect PRC members in some tissues in spite of the presence of their catalysed histone modifications does not exclude accumulation of the PRC complexes, which may be present under the technical detection level, or the tested proteins may be substituted with other paralogs such as EZH1 [32] and RING1A [33]. The results, however, imply that a visible accumulation of PRC complexes on the Xi are merely an extrapolation of the plasticity of the genome of EpiLCs (rXCI) and TS cells (iXCI) rather than whether XCI is random or imprinted.

XCI in extra-embryonic tissues is classified as unstable in contrast to XCI in embryonic lineages. This is highlighted by the fact that forced expression of *Tsix*, a negative regulator of *Xist* and leading to *Xist* downregulation, results in reactivation of the Xi in extra-embryonic tissues [34]. Interestingly, also EpiLCs still have the capacity to initiate rXCI upon differentiation [35], a feature that is lost upon differentiation of this cell type, highlighting the plasticity of this cell type with respect to XCI. The presence of such a meta-stable state of XCI in these three cell types might be attributed to reduced levels of CpG methylation. In trophoblast and XEN cells CpG methylation is reported to be very low in comparison to embryonic tissues [36, 37]. Interestingly, despite high expression levels of *Dnmts* [36], the DNA methylation level of CpG islands on the X-chromosome in female EpiLCs is also very low [36], and CpG islands only become methylated upon EpiLC differentiation (R.G. Boers, and J Gribnau personal communication). Therefore, the strong accumulation of H3K27me3, and possibly H4K20me1, and ubH2AK119

therefore might antagonize or precede the accumulation of DNA CpG methylation, reminiscent of findings showing a switch in the transcriptional control of repressed promoters via from H3K27me3 to DNA methylation during ES cell differentiation and carcinogenesis [38-40].

This study describes many undocumented chromatin modifications on the Xi in XEN cells, *in vitro* derived EpiLCs and differentiated EpiLCs. Most of our findings fit very well with documented observations, although we also found some noticeable exceptions. For instance, the stable association of H3K27me3 on the Xi in XEN cells contrasts previous findings [19, 20]. Also, the failure to demonstrate a strong accumulation of H3K9me2 on the Xi in EpiLCs and differentiated EpiLCs is in conflict with earlier described studies in embryonic lineages [7-10]. This discrepancy might be explained through differences in detection efficiency, but could also be related to differences in cell types that have been used.
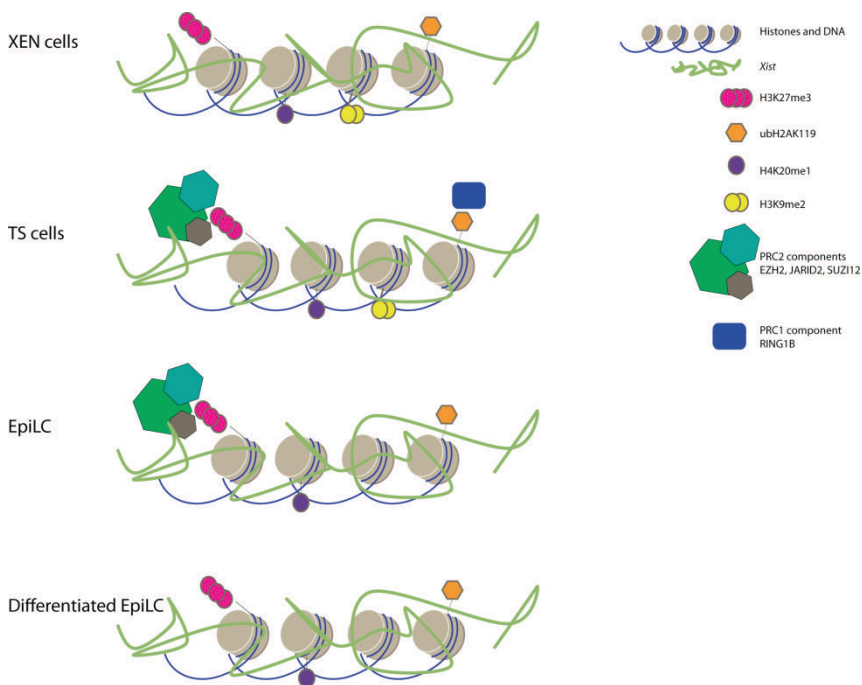
**Figure 10**. Comprehensive overview of accumulation of RING1B, EZH2, SUZI12, JARID2, H4K20me1, H3K9me2, H3K27me3 and ubH2AK119 (positive if found in more than 10% of the cells).

**Conclusion**

In conclusion, what governs stability of XCI cannot be simply determined by the assessment of the presence or absence of histone modifications or classified depending on whether XCI is random or imprinted. XCI is initiated and maintained by a multitude of epigenetic layers consisting of histone modifications and DNA CpG methylation. The deposition and maintenance of histone modifications on the Xi is *Xist* dependent whereas the accumulation of DNA CpG methylation makes XCI maintenance irreversible and *Xist* independent. In a specific epigenetic

context, XCI is relatively stable regardless whether XCI is random or imprinted. Problems in XCI stability may arise when XCI maintenance is still *Xist* dependent and the changing epigenetic context during differentiation and development cannot be implemented fast enough to create other epigenetic layers on the Xi. Which heterochromatic chromatin modifications can be found on the Xi is dependent on the epigenetic context and plasticity of the entire genome of the assessed cell. Despite the observed differences in histone modifications associated with the Xi in different lineages, our results indicate that both ubH2AK119 and H3K27me3 are the best heterochromatic markers for the Xi in all assessed lineages.

## Materials and Methods

### Embryo Collection and Cell Culture

*Ethics statement and embryo collection*

All animal experiments were in accordance with the legislation of the Erasmus MC Animal Experimental Commission. Except for the TS cells, all analysed lineages were hybrid, derived from embryos from crossings of 129Sv or C57Bl6 female mice with Cast/Eij or males. The analysed TS cells were both hybrid and inbred. The ovarian cycle of females was synchronized by intra-peritoneal injections of Folligonan and Chorullon (both 150 μl, 50U/ml, both one injection with 48 hours in between). Following the last injection each female mouse was mated with a male mouse. About 20 hours after the last injection, mating was evaluated by looking for the presence of a white plug in the vagina. Females that had mated were euthanized at E3.5 for ES or XEN cell derivation or killed at E6.5 for the creation of TS cells.

*XEN cell line derivation and culture conditions*

E3.5 embryos were flushed with M2 medium and placed in M16 medium at 37°C and 5% $CO_2$ before being placed in XEN derivation culture conditions. For XEN derivation, E3.5 embryos were placed into gelatinized wells covered with MEFs and cultured in XEN-medium (RPMI 1640; with 20% FBS, 1 mM sodium pyruvate, 2 mM L-glutamine, PS and 100μ M β-mercaptoethanol). Every 1-2 days, fresh medium was added until embryos were attached to the gelatinized surface of the well. After the embryos were attached and showed significant outgrowth, the XEN cells were split with 0.25% trypsin/EDTA and cultured in standard XEN culture conditions. For combined immunohistochemistry RNA-FISH, cells were grown onto plastic slide chamber flasks.

*TS cell line derivation and culture conditions*

TS cells were derived from E6.5 embryos. The extra embryonic part of E6.5 egg cylinders were plated in six-well plates on irradiated MEFs and maintained in XEN-medium supplemented

with 25 ng/ml human recombinant fibroblast growth factor 4 (hrFGF4) and heparin sulfate. After the embryos were attached and there was significant outgrowth, the TS cells were split with 0.25% trypsin/EDTA and further cultured in standard TS culture conditions. For combined immunohistochemistry RNA-FISH, cells were grown onto plastic slide chamber flasks.

*ES derivation and culture conditions*

For ES derivation, E3.5 blastocysts were placed into culture dishes coated with gelatin (0.2%) and irradiated MEFs in ES cell medium containing DMEM, 15% fetal calf serum (FCS), PenStrep (PS), 1 mM non-essential amino acids (NEAA), 50mM β-mercaptoethanol, leukaemia inhibitory factor (LIF), MEK inhibitor (PD98059, 4μM) and GSK3inhibitor (CHIR99021, 3.3 μM). Approximately one week after embryo recovery the outgrowth of the ICM was enzymatically split and plated in the same culture conditions as previously described. After a few passages, inhibitors for MEK and GSK3 were removed from the culture medium.

*ES differentiation towards EpiLC*

ES cells were trypsinized with 0.25% trypsin/EDTA and cultured in EpiLC conditions [41] for 3 passages (with collagenase) before being analysed. For combined immunohistochemistry RNA-FISH, cells were grown onto plastic slide chamber flasks.

*EpiLC differentiation towards mesoderm*

EpiLC cells were for combined immunohistochemistry RNA-FISH cells passaged onto plastic slide chamber flasks and grown for an extra 48 to 72 hours in EpiLC medium devoid of IWP2 put in the presence of WNT3 and BMP4.

Immunohistochemistry and RNA-FISH

*DNA Xist probe*

The *Xist* probe used was a 5kb cDNA BglII fragment covering exons 3-7 [42]. The probe was directly labeled by random priming. A total of 500ng DNA was dissolved in a total volume of 23 μl. 20μl of random primers 2.5x was added and denaturation was then performed at 95°C for 5 min. Immediately following the denaturation, the probe was cooled on ice and dNTPs, labeled dUTP and Klenow fragment were added. This mix was incubated at 37°C for 2 hours. For precipitation, Cot1-DNA, salmon sperm (SS) DNA, tRNA, sodiumacetate 3M and EtOH 100% were added to the labelled DNA. The mixture was frozen for 20 min in -20°C. To obtain the DNA pellet, the tube was centrifuged at 13200 rpm for 20-30 min and the supernatant was carefully removed. The pellet was thoroughly resuspended in 70% EtOH by vortexing and centrifuging at 13200 rpm for 5-10 min. Supernatant was carefully removed and the pellet was air-dried for 10 min. The labeled probe was dissolved with 50+ Hybridization mix and stored at -20°C.

*Antibodies*

| Antibody | Dilution | Host | Supplier |
|----------|----------|------|----------|
| H3K27me3 | 1:500 | Rabbit | Diagenode CS-069-100 |
| H4K20me1 | 1:500 | Rabbit | Abcam 16974 |
| H3K9me2 | 1:200 | Mouse | Cosmo Bio (MCA-MABI0007-20-EX) |
| ubH2A | 1:200 | Rabbit | Cell signalling (8240) |
| H3K4me2 | 1:1500 | Rabbit | Upstate 07-030 |
| H3K9ac | 1:1000 | Rabbit | Sigma H9286 |
| H4K16ac | 1:100 | Rabbit | Abcam 1240-100 |
| H4ac | 1:100 | Rabbit | Upstate 06-598 |
| RNApolII | 1:600 | Mouse | Abcam 817-100 |
| RING1B | 1:50 | Mouse | Generous gift from Dr. H. Koseki |
| SUZ12 | 1:100 | Rabbit | Diagenode pAb-029-050 |
| EZH2 | 1:200 | Rabbit | Leica Microsystems (NCL-L-EZH2) |
| JARID2 | 1:500 | Rabbit | Abcam (48137) |

*RNA FISH*

Cells cultured on plastic slide chamber flasks were fixed in 4% paraformaldehyde (PFA) for 10 min at room temperature and subsequently 3 times rinsed in PBS. The coverslips were incubated with 0.2% pepsin dissolved in water and incubated at 37°C in a water bath. After 4 min exposure, the pepsin was removed and the coverslips were rinsed in water. The coverslips were post-fixed in 4% PFA for 5 min and again 3 times washed with PBS. To dehydrate the cells, an ethanol gradient was used with 70%, 90% and 100% EtOH. Cells were hybridized overnight at 37°C with the denatured *Xist* probe (10min at 99°C followed by 45min at 37°C). The next day the coverslips were 3 times washed with 0.05x Saline-Sodium Citrate (SSC) in a pre-heated water bath at 40°C. The cells were mounted with Vectashield/DAPI.

*RNA FISH and immunohistochemistry*

Cells cultured on plastic slide chamber flasks were fixed in 3% PFA for 10 min at room temperature and 3 times rinsed in PBS. Permeabilization was performed with PBS containing 20% Triton-X 100 and 20 µM Vanadyl Ribonucleoside Complex (VRC)(New England Biolabs S1402S). After rinsing 3 times in PBS, preparations were blocked in a blocking solution containing bovine serum albumin (BSA) (Biolabs, B90015) and 20 µM VRC in PBS for 30 min. The primary antibody, diluted in blocking solution, was applied to the wells of the slide chamber flasks and incubated in a humid box for 1 hour at room temperature. After 3 times 10 min washes with PBS, the secondary antibody, diluted in the blocking solution was applied to the wells of the slide chamber flask and incubated for 1 hour at room temperature.

The secondary antibody, was removed by 3 washes in PBS for 5min each. Following removal

of the plastic chamber, the remaining plastic slides were accordingly post-fixed in 4% PFA for 10 min at room temperature, rinsed in PBS and washed two times with 2x SSC before being air dried. The denatured *Xist* probe was applied on the slides, coverslips were placed and glued with rubber cement onto the slide and incubated for 15-20 hours in a dark box at 37°C. After 3 washes with 0.05x SSC at 38-40°C, DNA was counterstained for 2 minutes in 10μl DAPI. A Leica TCS SP 5 confocal microscope and Adobe Photoshop CS 6 and Illustrator were used for image acquisition.

**Author´s Contributions**

C.D. concept, performed stem cell derivation, cell culture, all stainings, visualization, drafted manuscript; C.M. performed all countings, H.D.B. tests staining procedures and selection antibodies; R.B. DNA methylation assays; D.K. assisted with epiblast cultures; J.G. funding, concept and correction manuscript

**Supplementary Figures**
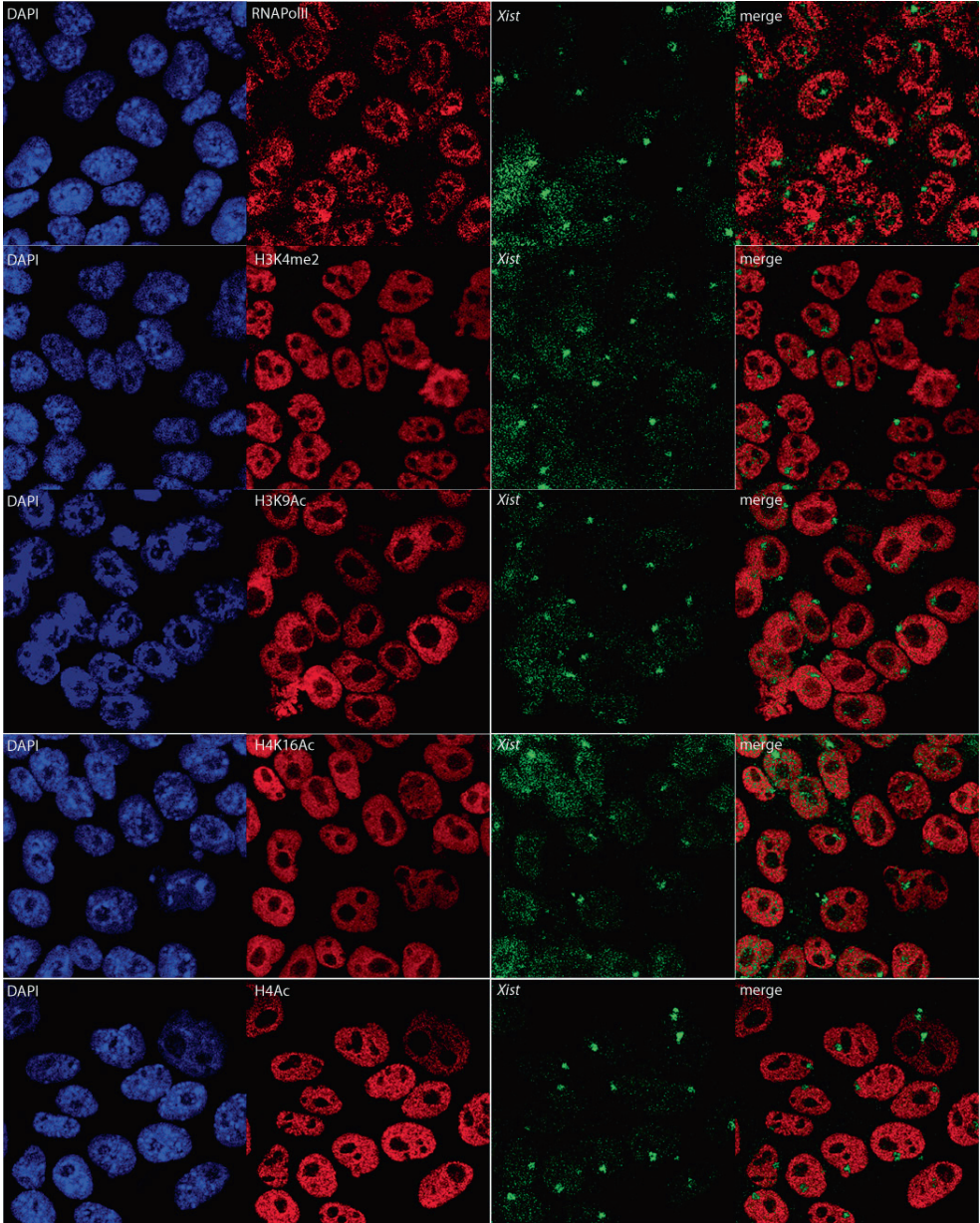
**Figure S1A:** Immuno-RNA FISH detecting RNA pol II, H3K9Ac, H4K16Ac and H4Ac (Rhodamine red) together with *Xist* (FITC) on TS cells (DAPI is DNA).
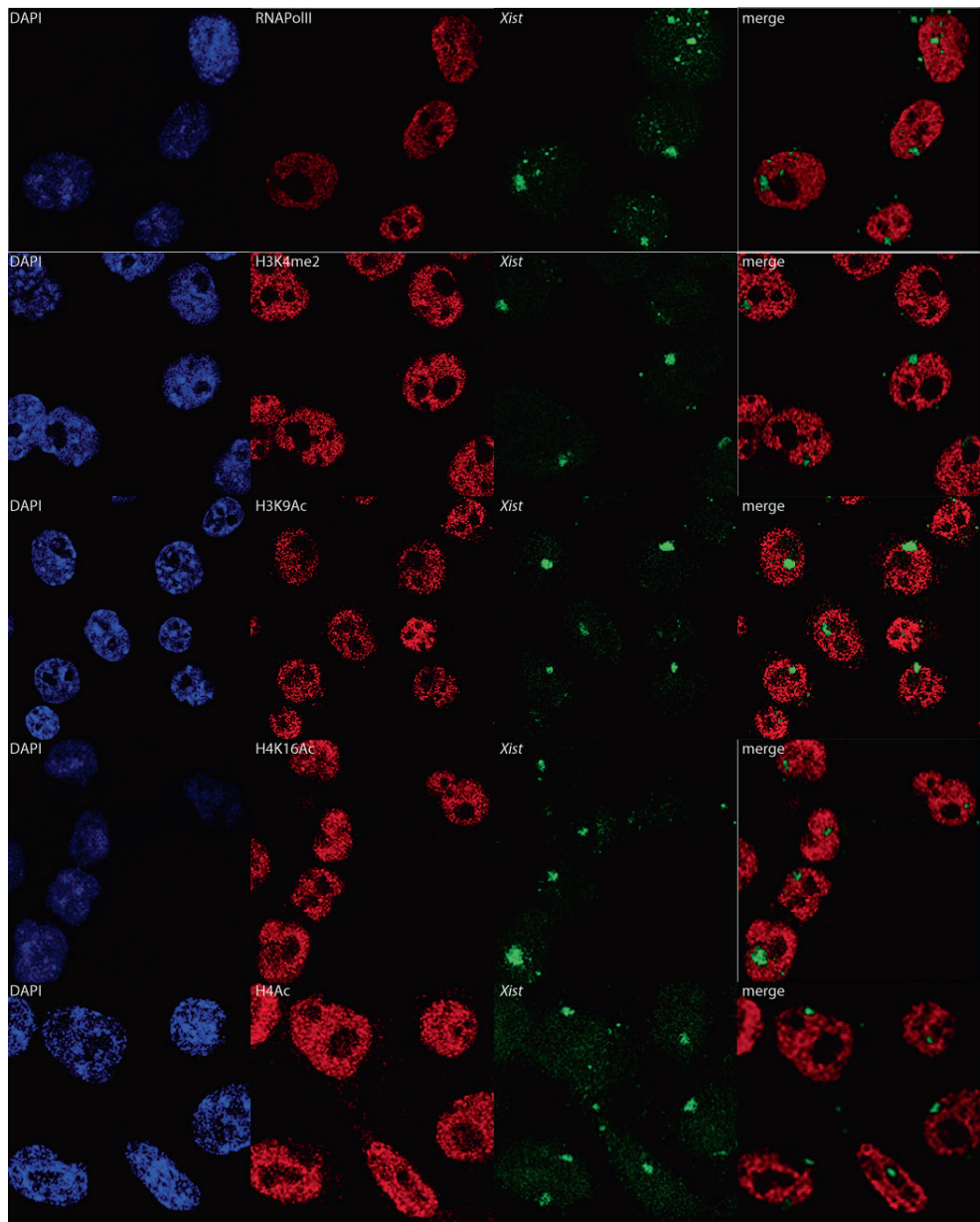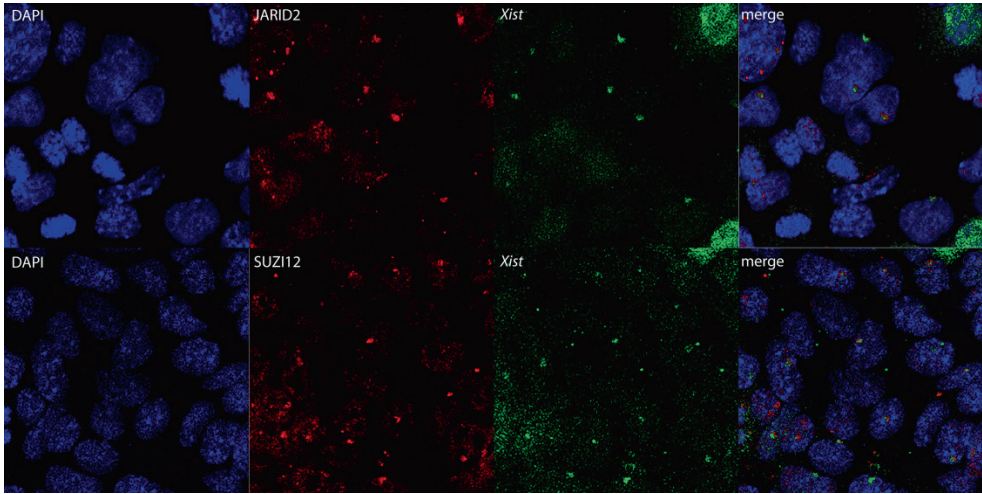
**Figure S1B:** Immuno-RNA FISH detecting JARID2 and SUZ12 (Rhodamine red) together with *Xist* (FITC) on TS cells (DAPI is DNA).
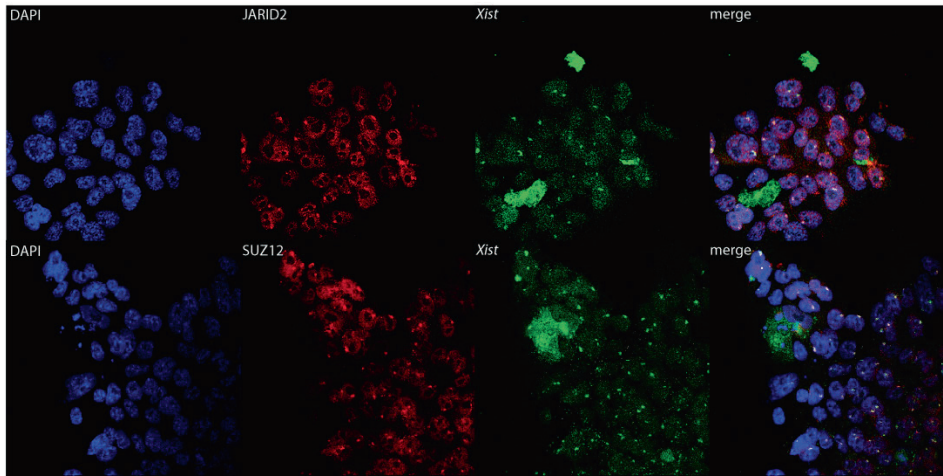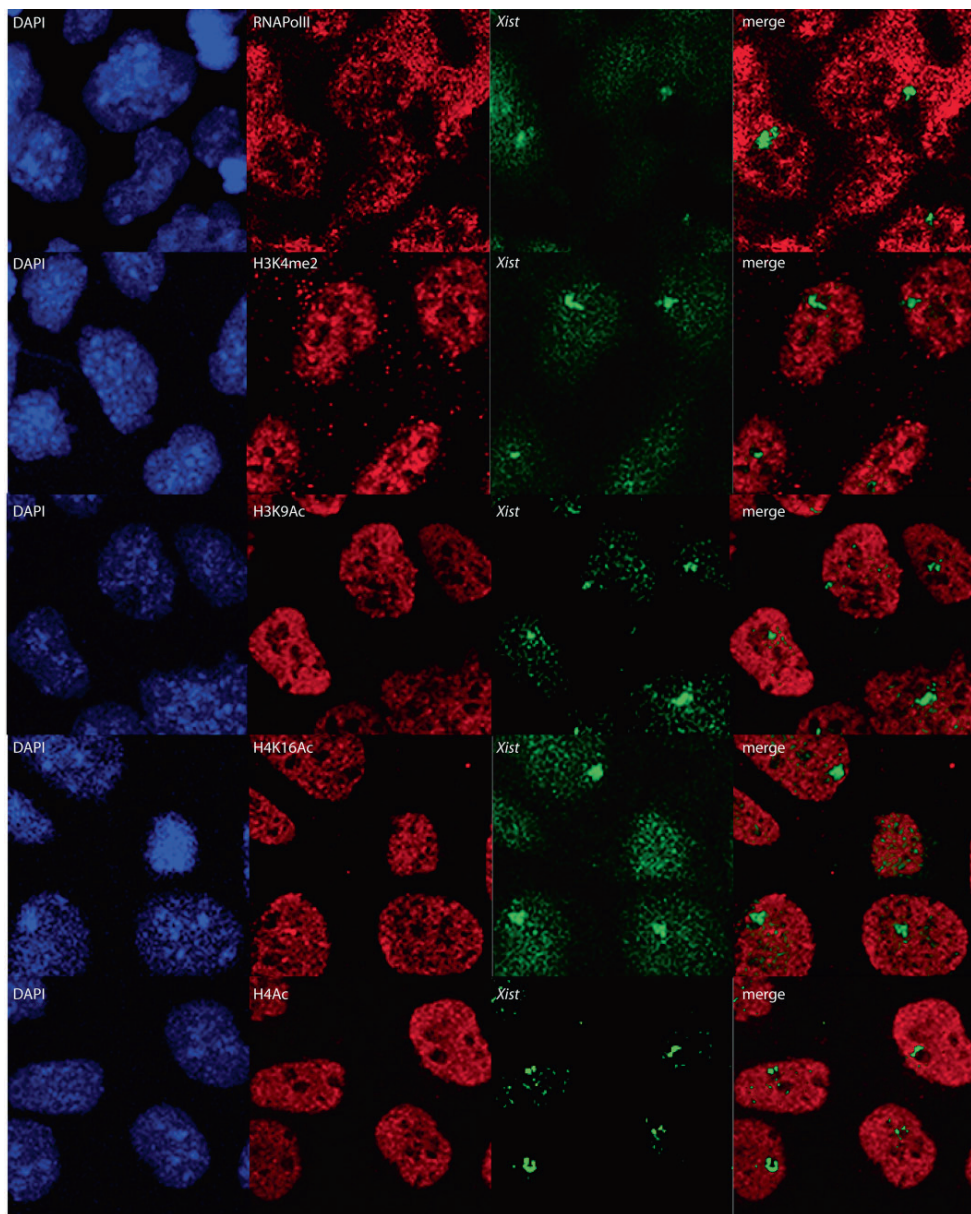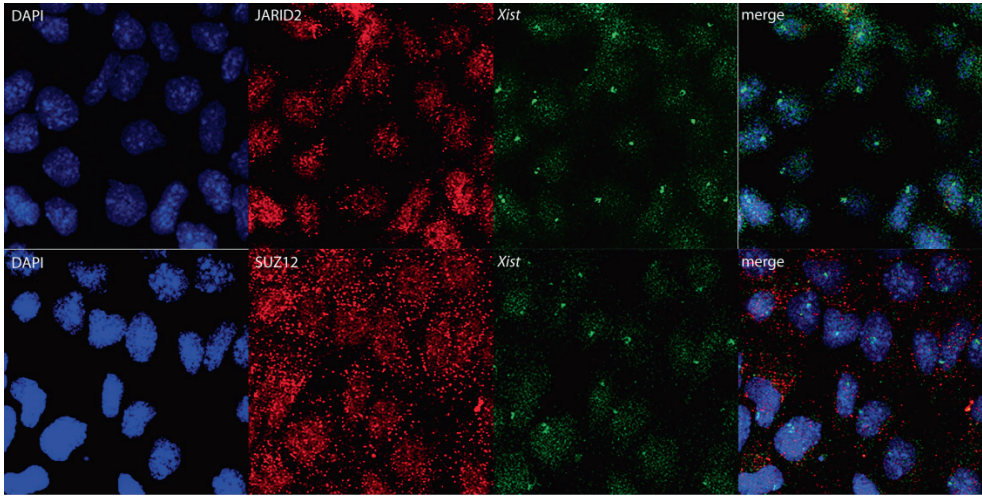
**Figure S2A:** Immuno-RNA FISH detecting RNA pol II, H3K9Ac, H4K16Ac and H4Ac (Rhodamine red) together with *Xist* (FITC) on XEN cells (DAPI is DNA).
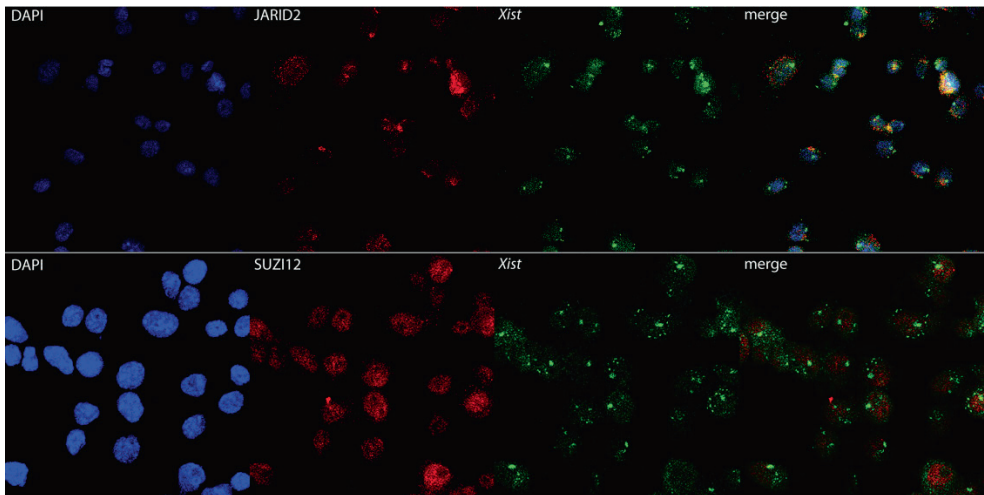


**Figure S2B:** Immuno-RNA FISH detecting JARID2 and SUZ12 (Rhodamine red) together with *Xist* (FITC) on XEN cells (DAPI is DNA).
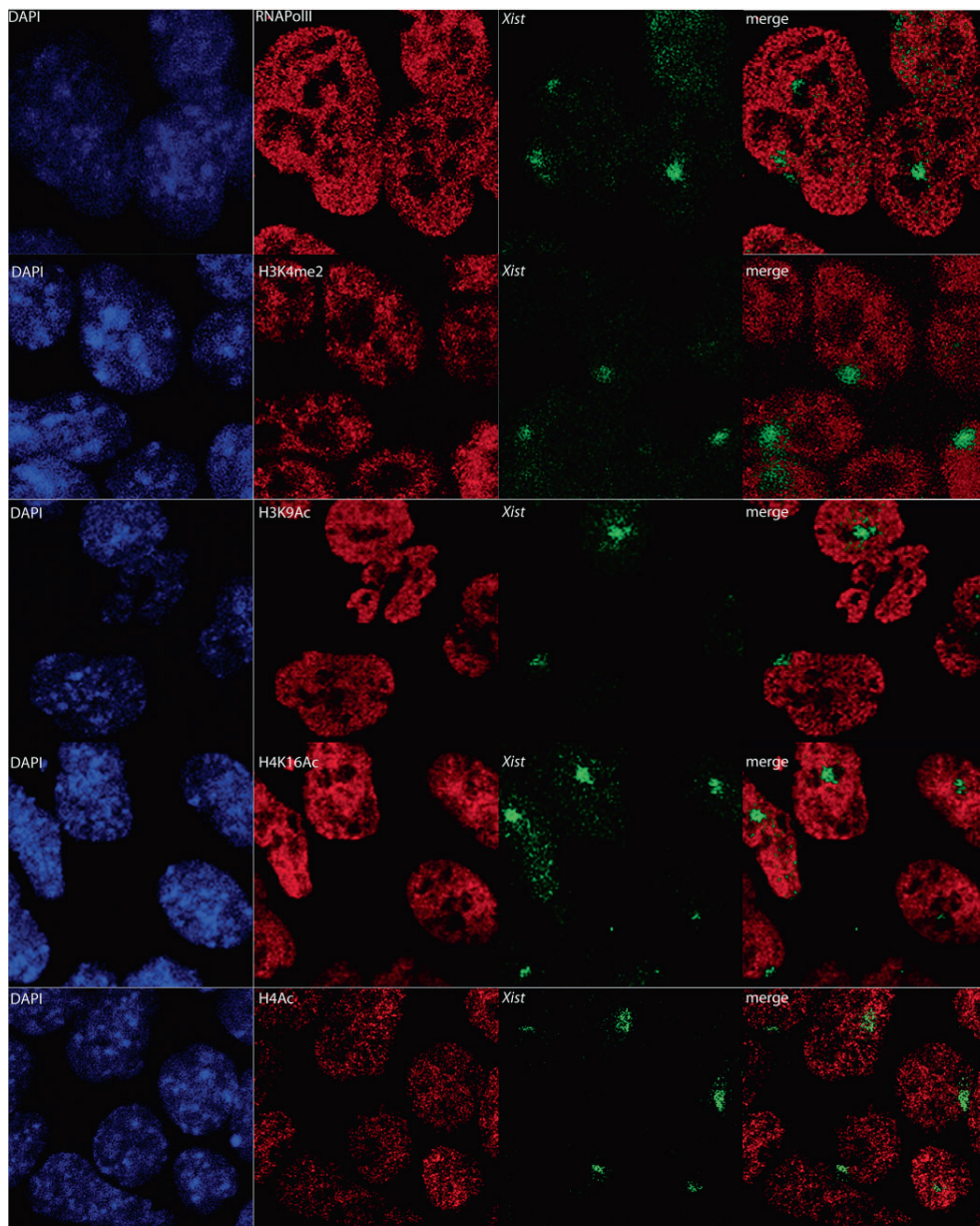
**Figure S3A:** Immuno-RNA FISH detecting RNA pol II, H3K9Ac, H4K16Ac and H4Ac (Rhodamine red) together with *Xist* (FITC) on EpiSCs (DAPI is DNA).

**Figure S3B:** Immuno-RNA FISH detecting JARID2 and SUZ12 (Rhodamine red) together with *Xist* (FITC) on EpiSCs (DAPI is DNA).

**Figure S4A:** Immuno-RNA FISH detecting RNA pol II, H3K9Ac, H4K16Ac and H4Ac (Rhodamine red) together with *Xist* (FITC) on diferentaited EpiSCs (DAPI is DNA). **Figure S4B:** Immuno-RNA FISH detecting JARID2 and SUZ12 (Rhodamine red) together with *Xist* (FITC) on differentiated EpiSCs (DAPI is DNA).

**Figure S4B:** Immuno-RNA FISH detecting JARID2 and SUZ12 (Rhodamine red) together with *Xist* (FITC) on differentiated EpiSCs (DAPI is DNA).

## References

1.      Borsani, G., et al., *Characterization of a murine gene expressed from the inactive X chromosome.* Nature, 1991. **351**(6324): p. 325-9.

2.      Brockdorff, N., et al., *Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome.* Nature, 1991. **351**(6324): p. 329-31.

3.      Okamoto, I. and E. Heard, *Lessons from comparative analysis of X-chromosome inactivation in mammals.* Chromosome Res, 2009. **17**(5): p. 659-69.

4.      Rastan, S., *Timing of X-chromosome inactivation in postimplantation mouse embryos.* J Embryol Exp Morphol, 1982. **71**: p. 11-24.

5.      Fang, J., et al., *Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation.* J Biol Chem, 2004. **279**(51): p. 52812-5.

6.      Plath, K., et al., *Role of histone H3 lysine 27 methylation in X inactivation.* Science, 2003. **300**(5616): p. 131-5.

7.      Kohlmaier, A., et al., *A chromosomal memory triggered by Xist regulates histone methylation in X inactivation.* PLoS Biol, 2004. **2**(7): p. E171.

8.      Heard, E., et al., *Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation.* Cell, 2001. **107**(6): p. 727-38.

9.      Chaumeil, J., et al., *Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells.* Cytogenet Genome Res, 2002. **99**(1-4): p. 75-84.

10.     Mermoud, J.E., et al., *Histone H3 lysine 9 methylation occurs rapidly at the onset of random X chromosome inactivation.* Curr Biol, 2002. **12**(3): p. 247-51.

11.     Chaumeil, J., et al., *A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced.* Genes Dev, 2006. **20**(16): p. 2223-37.

12.     Patrat, C., et al., *Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice.* Proc Natl Acad Sci U S A, 2009. **106**(13): p. 5198-203.

13.     Hadjantonakis, A.K., et al., *An X-linked GFP transgene reveals unexpected paternal X-chromosome activity in trophoblastic giant cells of the mouse placenta.* Genesis, 2001. **29**(3): p. 133-40.

14. Corbel, C., et al., *Unusual chromatin status and organization of the inactive X chromosome in murine trophoblast giant cells.* Development, 2013. **140**(4): p. 861-72.

15. Kratzer, P.G., et al., *Differences in the DNA of the inactive X chromosomes of fetal and extraembryonic tissues of mice.* Cell, 1983. **33**(1): p. 37-42.

16. Krumlauf, R., et al., *Differential expression of alpha-fetoprotein genes on the inactive X chromosome in extraembryonic and somatic tissues of a transgenic mouse line.* Nature, 1986. **319**(6050): p. 224-6.

17. Mak, W., et al., *Reactivation of the paternal X chromosome in early mouse embryos.* Science, 2004. **303**(5658): p. 666-9.

18. Okamoto, I., et al., *Epigenetic dynamics of imprinted X inactivation during early mouse development.* Science, 2004. **303**(5658): p. 644-9.

19. Kalantry, S., et al., *The Polycomb group protein Eed protects the inactive X-chromosome from differentiation-induced reactivation.* Nat Cell Biol, 2006. **8**(2): p. 195-202.

20. Kunath, T., et al., *Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts.* Development, 2005. **132**(7): p. 1649-61.

21. Merzouk, S., et al., *Lineage-specific regulation of imprinted X inactivation in extraembryonic endoderm stem cells.* Epigenetics Chromatin, 2014. **7**: p. 11.

22. Mak, W., et al., *Mitotically stable association of polycomb group proteins eed and enx1 with the inactive x chromosome in trophoblast stem cells.* Curr Biol, 2002. **12**(12): p. 1016-20.

23. Keohane, A.M., et al., *X-Inactivation and histone H4 acetylation in embryonic stem cells.* Dev Biol, 1996. **180**(2): p. 618-30.

24. Fang, J., et al., *Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase.* Curr Biol, 2002. **12**(13): p. 1086-99.

25. Nishioka, K., et al., *PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin.* Mol Cell, 2002. **9**(6): p. 1201-13.

26. McHugh, C.A., et al., *The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3.* Nature, 2015. **521**(7551): p. 232-6.

27. Moindrot, B., et al., *A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors*

*Required for Xist RNA-Mediated Silencing.* Cell Rep, 2015. **12**(4): p. 562-72.

28.    Monfort, A., et al., *Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells.* Cell Rep, 2015. **12**(4): p. 554-61.

29.    Minajigi, A., et al., *Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation.* Science, 2015. **349**(6245).

30.    Chu, C., et al., *Systematic discovery of Xist RNA binding proteins.* Cell, 2015. **161**(2): p. 404-16.

31.    Hughes, C.M., et al., *Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus.* Mol Cell, 2004. **13**(4): p. 587-97.

32.    Shen, X., et al., *EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency.* Mol Cell, 2008. **32**(4): p. 491-502.

33.    de Napoles, M., et al., *Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation.* Dev Cell, 2004. **7**(5): p. 663-76.

34.    Ohhata, T., et al., *Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification.* Development, 2008. **135**(2): p. 227-35.

35.    Gayen, S., et al., *A Primary Role for the Tsix lncRNA in Maintaining Random X-Chromosome Inactivation.* Cell Rep, 2015. **11**(8): p. 1251-65.

36.    Senner, C.E., et al., *DNA methylation profiles define stem cell identity and reveal a tight embryonic-extraembryonic lineage boundary.* Stem Cells, 2012. **30**(12): p. 2732-45.

37.    Sado, T., et al., *X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation.* Dev Biol, 2000. **225**(2): p. 294-303.

38.    Mohn, F., et al., *Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors.* Mol Cell, 2008. **30**(6): p. 755-66.

39.    Schlesinger, Y., et al., *Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer.* Nat Genet, 2007. **39**(2): p. 232-6.

40.    Ohm, J.E., et al., *A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing.* Nat Genet, 2007. **39**(2): p. 237-42.

41.    ten Berge, D., et al., *Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast*

**5**

*stem cells.* Nat Cell Biol, 2011. **13**(9): p. 1070-5.

42.     Monkhorst, K., et al., *X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator.* Cell, 2008. **132**(3): p. 410-21.

5
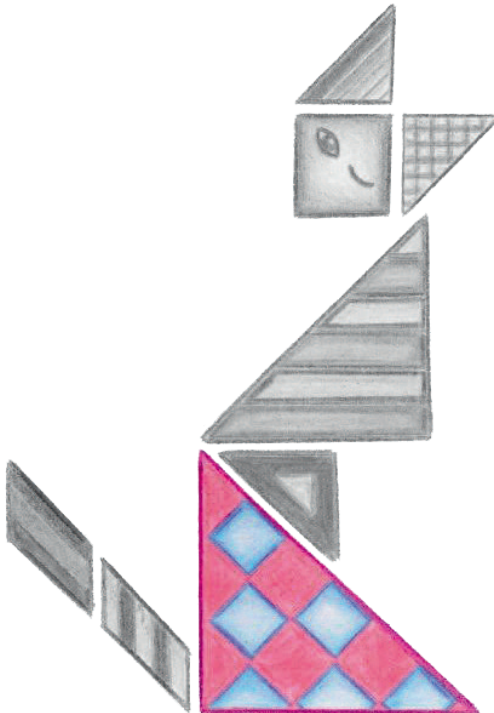
# Chapter 6
*General Discussion*

# Chapter 6

*Maduro C.M and Gribnau J.*

In mammals, males are the heterogametic sex having an XY sex chromosome pair, whereas females are homogametic having an XX sex chromosome pair. The Y chromosome is small and harbors only a few genes mainly involved in male sex determination and spermatogenesis [1-3], whereas the X chromosome contains over a thousand genes important for both male and female development and homeostasis [1,2]. XX females have double the amount of X encoded genes as compared to XY males and this imbalance requires a dosage compensation mechanism to equalize expression of these X encoded genes in the two sexes. Dosage compensation in mammals is achieved through a process called X chromosome inactivation (XCI), in which one of the two X chromosomes in female cells is inactivated. In the following sections, regulation and timing of the different phases of XCI are discussed and how, together and through redundant processes, activation of XCI is ensured and maintained throughout cell divisions.

**Initiation of XCI**

Random XCI is a stochastic and tightly regulated process in which either the maternal or the paternal X chromosome can be inactivated. XCI ensures one active X chromosome per diploid genome by regulation through one or more pathways to ensure proper inactivation and circumvent inactivation of both or continued activity on both X chromosomes [4]. Regulation of XCI is achieved by a counterbalance of inhibitors and activators of XCI acting on the key regulators of XCI, *Xist* and *Tsix*, which will determine whether XCI occurs and which X chromosome will be inactivated.

*Key regulators of XCI: Xist and Tsix*

Key regulators of XCI are the two lncRNAs, *Xist* [5] and *Tsix* [6], involved in activation and repression of XCI respectively. *Xist* is expressed exclusively from the inactive X chromosome (Xi), spreads and coats the entire X chromosome whilst recruiting chromatin remodeling factors to render the X inactive [7,8]. *Tsix* is a negative regulator of *Xist* and promotes repression of XCI, keeping the X chromosome active. Regulation of *Xist* and *Tsix* is achieved through activators and inhibitors. However, analysis of independent *Xist* and/or *Tsix* regulation is hampered as a result of overlapping antisense transcription of the two lncRNAs [9]. Effects on XCI as a result of a deletion of the activators *Jpx/Ftx/Xpr/Rnf12* in *cis* for example, can be rescued by a deletion of *Tsix* in *cis* [10] and in the same way XCI can be dysregulated in normal differentiated female cells (XiXa) by a deletion of *Tsix* on the active X (Xa) chromosome resulting in ectopic *Xist* expression from the Xa [11]. Similarly, a mutation of *Tsix* in undifferentiated female ES cells results in ectopic upregulation of *Xist* [12,13]. Thus, *Tsix* expression interferes with and masks regulatory effects exerted on the *Xist* regulatory network and vice versa, highlighting the complication in studying regulation of *Xist* and *Tsix* without interference of the antisense partner. To study regulation of *Xist* and *Tsix* on an uncoupled allele, we replaced the first exons of *Xist* with an EGFP and *Tsix* with an mCherry cassette, generating 3 different cell lines with either Xist-EGFP, Tsix-mCherry or both reporters on the same allele (Chapter 3). Expression of the reporters remained

under control of the endogenous promoters with a poly-A signal terminating transcription downstream of the reporter, preventing overlapping antisense transcription. On an uncoupled allele we confirmed that *Tsix* is involved in repression of *Xist*, but also show that *Xist* is involved in down regulation of *Tsix* expression on the Xi as well as on the Xa. In addition, we found that expression of *Xist* and *Tsix* is mostly anti-correlated but rather than a strict anti-correlation, they appear regulated by *trans*-acting factors, inhibitors and activators, in a stochastic fashion. An XCI effect of RNF12 (activator) or REX1 (inhibitor) was more pronounced in cells with both Xist-EGFP:Tsix-mCherry reporters as compared to cells with either Xist-GFP or Tsix-Cherry alone, highlighting that antisense transcription masks the regulatory impact of XCI activators or inhibitors and emphasizing the importance of these cells with both Xist-EGFP:Tsix-mCherry in identification of novel activators and inhibitors of XCI. In addition, the double Xist-EGFP:Tsix-mCherry cells offer an easy readout and enable distinction between direct and indirect effects of regulators on the promoters of *Xist* and *Tsix*.

With these Xist-EGFP:Tsix-mCherry cells we were also able to observe distinct populations of cells showing different levels of *Tsix* expression in cells grown in 2i medium and cells which have lost their wildtype X chromosome. These two distinct populations also showed differences in expression of other genes in the Xic, indicating that these semi-stable transcriptional states reflect fluctuating chromatin conformations that can be stabilized in *trans* by *trans*-acting factors. Our data suggests the presence of differential epigenetic states established by intrinsic and extrinsic factors, capable of providing stable, mutually exclusive, on/off switches of genes in the Xic and predicts the responsiveness of the Xic to its regulators prior to initiation of XCI. A polymer model, which can deconvolve sub-TAD contact frequencies measured by 5C into single-cell chromatin configurations, also predicted two main classes of conformations of the *Tsix* TAD [14]. In addition, this study shows that the *Tsix* TAD of the two X chromosomes in a single nucleus differ in transcriptional activity which is related to the chromatin conformation of the TAD from which they are expressed. Moreover, intrinsic fluctuations in conformation of the *Tsix* TAD, coupled to variation in transcription, may play a role in transcriptional asymmetry between the two Xics which ensures responsiveness leading to upregulation of *Xist* on only one X chromosome already prior to initiation of XCI. Furthermore, considering the alternate states model [15], which showed that the two Xics of the X chromosomes differ prior to XCI in ES cells and is no longer observed in differentiated cells, we cannot exclude that XCI encompasses multiple ways to ensure XCI to occur properly. Nevertheless, these states, as determined in our study, are only observed in male cells or upon addition of 2i medium. In serum conditions switching occurs more frequently, resulting in one homogeneous population, facilitating the dynamics required for the mutual exclusive XCI process.

In Chapter 3, we also observe semi-stable promoter activity of *Xist* and *Tsix* from mother cell to daughter cells. The stability of transcriptional states observed represent distinct states of transcriptional activity and likely plays a role in epigenetic memory, which is clonally propagated to daughter cells for many cell divisions.

*Activation of XCI: Xist activators and XCI activators*
Initiation of XCI is regulated by a balance of autosomally encoded XCI inhibitors setting a threshold for XCI to occur, and the X encoded XCI activators, allowing XX female cells to overcome the threshold to initiate XCI [16]. Activation of XCI is achieved by activation of *Xist* and/or repression of *Tsix* by XCI activators. Factors exclusively activating *Xist*, which are autosomally encoded, are termed *Xist* activators. YY1 has been previously described as an *Xist* activator

which competes with the XCI inhibitor REX1 for binding sites at *Xist* downstream regulatory elements [17]. *Xist* activators activate *Xist* in *trans* but are not involved in the counting process as these factors are equally expressed in male and female cells. In Chapter 2, we describe another *Xist* activator, *Rnf6*. *Rnf6* is highly homologous to *Rnf12* and a deletion of *Rnf6* in *Rnf12*[+/-] ES cells has an increased effect on XCI. No additive effect of deleting the second copy of *Rnf6* on XCI was observed, suggesting that deleting an *Xist* activator affects the threshold but is not able to lower the threshold further when completely absent as XCI activators are the key players in overcoming the threshold for XCI to occur. In *vitro* ubiquitination studies indicate that *Rnf6* most likely acts on XCI through proteasomal degradation of REX1. *Xist* activators are different from XCI activators as the latter are X-encoded, allow a cell to count the number of active X chromosomes in a cell through mechanisms involving indirect or direct activation of *Xist* and/ or repression of *Tsix*, to initiate XCI. Moreover, the XCI activator will be silenced in the process of XCI, allowing feedback to prevent the remaining X chromosome from being inactivated. The only X encoded XCI activator reported so far was the E3 ubiquitin ligase *Rnf12* [4,10,18]. In Chapter 2 we describe an additional XCI activator, the E1 activating enzyme *Uba1*. In XCI, *Rnf12*, *Rnf6* possibly along with *Uba1,* function by targeting the XCI inhibitor and pluripotency factor REX1 for proteasomal degradation through the ubiquitin pathway.

*Ubiquitination and XCI*

In Chapter 2 we explored ubiquitination and aspects of the RNF12 mediated degradation of REX1[19] in the context of XCI in more detail. The outcome of ubiquitination of a target protein depends on the function of the E1, E2 and E3 enzyme and the manner in which ubiquitin is loaded onto the target protein as well as the lysines which are preferentially ubiquitinated within the target protein. We showed that the main and X-encoded E1 activating enzyme, UBA1, is an activator of XCI. As UBA1 is an E1 activating enzyme, UBA1 could be the start-point and part of an activation cascade such as that of the targeted degradation of REX1 by RNF12 and RNF6. In a cell, E1 levels are mostly saturated, making the reaction velocity independent on the substrate or enzyme present (Michaelis-Menten equation) [20,21]. However, at lower E1 concentrations, when the E1-ub formation is still linear, this could indeed have an effect on the outcome of the reaction. When we consider this pathway in activation of XCI, we propose a model involving an activation cascade ultimately leading to proteasomal degradation of the XCI inhibitor REX1, relieving the brake on *Xist* expression and allowing XCI to be initiated. In this activation cascade, the E1 activating enzyme UBA1 activates ubiquitin and transfers this to an E2 conjugating enzyme, which subsequently binds either RNF12 or RNF6, which in turn allows transfer of active ubiquitin molecules to REX1, targeting it for proteasomal degradation. However, further protein data, analyzing an effect of *Uba1* on proteasomal degradation of REX1, is required to establish a role for *Uba1* in this activation cascade. It appears that in XCI, REX1 is central, as initiation of XCI involves the proteasomal degradation of REX1 by the ubiquitin cascade as well as regulation by the *Xist* activator, *Yy1*, which competes with REX1 for binding sites at the *Xist* promoter. However, knockout mice for REX1 [22] or RNF12 [23] are viable and a deletion of the REX1/YY1 binding sites at the *Xist* promoter does not affect XCI [24]. This raises the question on how important this pathway really is *in vivo*. It seems likely that XCI is regulated by redundant and linked pathways to ensure that XCI does indeed progress. Similarly, to ensure dosage compensation in *D. melanogaster* and *C. elegans*, for example, several X-encoded numerator genes are required and together regulate the key regulatory genes *Sxl* and *Xol-1*, respectively [25]. Since deletion of *Uba1* in *Rnf12*[+/-] ES cells resulted in significantly affected XCI,

but still showed a fraction of cells initiating XCI, this suggests presence of other X-encoded XCI activators. Recently, a study proposed that the factors involved in the counting process of XCI, are escaping genes which further activate *Xist* and help trigger X-linked gene silencing [11]. However, in initiation of XCI, initial activation cues resulting in upregulation of *Xist* need to occur prior to *Xist* spreading and silencing the X chromosome and feedback through silencing of the activator in *cis*, is required to prevent the other X chromosome from being inactivated, which is difficult to achieve if these activating factors where escaping genes. It cannot be excluded that genes escaping XCI do have a role in further promoting the silent state, but a role for escaping genes in initiation of XCI and the counting process, seems unlikely.

**Identification of factors regulating of XCI**

In Chapter 2, Chapter 4 and Addendum 1, we set out to identify novel factors involved in activation of XCI. In Chapter 2 and Addendum 1, we carefully selected candidate genes based on known functions/interactions and used two different approaches to analyze for an effect on XCI. We overexpressed candidate genes in male cells, which in case of an activator would result in ectopic *Xist* cloud formation and inactivation of the single X chromosome in male cells as opposed to a heterozygous deletion of a candidate gene in female cells, which in case of an activator would significantly affect or abolish XCI. Overexpression of XCI activators in male cells results in cell death, which makes analysis more difficult, as was found for *Rnf12*. Two of the five candidate genes displayed ectopic *Xist* cloud formation, nonetheless only in a small percentage of cells. This required further analysis of *Cited1* (Addendum A) and *Rnf6* (Chapter 2) in female cells and is also how we screened two additional candidate genes, *Uba1* (Chapter 2) and *Phf6* (Addendum A). This second strategy, consisted of generating heterozygous knockouts for the candidate genes in female *Rnf12*[+/-] ES cells using the CRISPR/Cas9 technology. As *Rnf12*[+/-] cells show reduced XCI initiation (Chapter 2 and [4]), a heterozygous deletion of an activator would result in further affected XCI, which is what we observed for *Uba1* and *Rnf6*. With both candidate gene approaches, two out of seven candidate genes appeared to have a function in initiation of XCI. Nevertheless, these two approaches and careful selection of candidate genes are very laborious and precarious as activators can easily be falsely excluded. We therefore initiated a large-scale deletion of approximately one third of the X chromosome in a semi-unbiased approach in which a 50Mb candidate region, telomeric to the Xic, containing about 202 genes, was deleted in *Rnf12*[+/-] cells. We observed near complete abolished XCI, of which the slight residual *Xist* expression may be the attributed by of the presence of two copies of *Uba1* or another XCI activator, located outside of the deleted region. Taken together, our results indicated the presence of at least three XCI activators involved in activation of XCI.

Comparing the three different strategies used, we can conclude that firstly, generating heterozygous knockouts in female cells generates a better readout than male cells, which die upon inactivation of the X chromosome. In addition, in female cells all the known activators can be deleted in one cell to allow a better readout of factors with a lower impact on activation on XCI which would be missed otherwise. Second, when we compare the candidate gene approach versus the candidate region approach, the latter seems favorable as we now have a 50Mb candidate region, which would be practical to narrow down to a smaller region(s) followed by a candidate selection and CRISPR/Cas9 technology to pinpoint candidate genes involved in initiation of XCI.

## Establishment and maintenance of the Xi

When XCI is initiated and *Xist* RNA spreads along the X chromosome, euchromatic marks such as H3K4 methylation and H3K9 actetylation are lost and repressive marks such as H3K9me2 accumulate [7,8]. Polycomb complex 1 and 2 (PRC1 and PRC2) are recruited [26-29] catalyzing H2A119ub [28] and H3K27me3 [8], macroH2A is incorporated [30,31] and finally CpG island methylation accumulates [32,33] further ensuring the silent state. In the mouse, two forms of XCI exist: imprinted XCI (iXCI) and random XCI (rXCI). Besides difference in timing, it has been reported that iXCI is less stable with regard to the inactivated X as compared to rXCI [34-38]. In order to understand the difference in stability of the Xi in these two forms of XCI, we generated TS, XEN (imprinted XCI), ES and EpiSCs (random XCI) which were also further differentiated towards the mesodermal lineage, all with the same genomic background to obtain a comprehensive overview of all the histone modifications associated with the Xi in these cell types representing different embryonic lineages.

Loss of euchromatic marks such as H3K4me2, H3K9ac, H4ac, H4K16Ac and RNA PolII hallmarks the initial shift from the active state to silencing of the X chromosome. This is followed by accumulation of H3K27me3 and H2AK11Ub indicating early chromatin changes to the silent state, and was also observed in all lineages. Even though all cell types displayed accumulation of H3K27me3 and H2AK119Ub on the Xi, marking the X transcriptionally inactive, subunits of the PRC2 (EZH2, SUZ12 and JARID2) and PRC1 (RING1B) complex, catalyzing these modifications, were not always detected. This indicates that either these complexes were below the detection limit of our immunostaining, or as previously reported, PRC2 association even though constant may be very transient [39] making detection of its accumulation more difficult by immunostaining. As for PRC1, only subunit RING1B was analyzed and the combined immune-FISH did not work for all linages, we can therefore not completely exclude accumulation of PRC1 on the Xi in specific cell types from these results.

Presence of H3K9 methylation on the Xi has been controversial, however more recently, H3K9 methylation has been found enriched at intergenic and gene poor regions of the Xi and interspersed between H3K27me3 domains most likely catalyzed by *Setdb1* [40]. In addition a role for H3K9 methylation in late establishment and early maintenance of silencing by silencing of repeats and so facilitating alterations to the conformation of the Xi, thereby initiating the maintenance phase has been proposed. In Chapter 5, we show accumulation of H3K9me2 only on the Xi in the extraembyonic TS and XEN cells, whether or not this mark is present in the other cell lineages can be debated. Keniry et al report H3K9me2 in a study using ChIP-seq and not observing enrichment of H3K9 methylation on the Xi in our study could be related to H3K9 detection being more difficult by immunofluorescence[40]. To be able to conclude the presence of differential H3K9 methylation in iXCI versus rXCI we need to look at this chromatin mark on the Xi in more detail than capable with immunofluorescence. Another chromatin mark associated with the Xi, H4K20me1, was observed in TS, XEN and EpiSCs and is most likely catalyzed by *Pr-Set7* or *Set8,* however involvement of these methyltransferases in XCI is unclear [41-43] and might be a general mark associated with inactive chromatin rather than involved in the stability of the inactive state as no differential H4K20 methylation was observed in the cell types displaying iXCi or rXCI.

Taken together, our results indicate only few distinct differences in chromatin modifications present on the Xi in iXCI as opposed to rXCI. The stability of the inactive state depends rather on the plasticity of a cell. As the X chromosome transits from an active to an inactive state, H3K4me2, H3K9ac, H4ac, H4K16Ac and RNA PolII are lost and directly followed

by H3K27me3 and H2AK119Ub accumulation on the Xi as *Xist* RNA spreads along the X chromosome and further silences repeat regions by methylation of H3K9. DNA methylation will lock down the inactive state for the next phase in XCI, maintenance of the Xi, which is most likely less pronounced or even absent in iXCI which could attribute to a less stable silencing found in iXCI [38,44].

Regulation of XCI in initiation, establishment and maintenance of the Xi rely on several different and sometimes redundant pathways to ensure proper execution of XCI. The initiation phase and activation of XCI depends on the upregulation of *Xist*, which as the different chapters in this thesis explain, can be achieved in different ways. In addition, inhibitors and activators do not appear to be strictly autosomal or X-encoded as was previously believed, rather multiple layers and levels of regulation of *Xist*, together, all ensure proper activation of XCI. Multiple layers of epigenetic modifications are also instrumental in the establishment phase and maintenance phase, these layers are different between iXCI and rXCI to facilitate embryonic requirements of plasticity related to cell fate decisions. This makes XCI an intriguing yet difficult epigenetic phenomenon to unravel.

## References

1       Graves, J. A., Gécz, J. & Hameister, H. Evolution of the human X--a smart and sexy chromosome that controls speciation and development. *Cytogenet Genome Res* **99**, 141-145, doi:71585 (2002).

2       Graves, J. A. Evolution of the testis-determining gene--the rise and fall of SRY. *Novartis Found Symp* **244**, 86-97; discussion 97-101, 203-106, 253-107 (2002).

3       Lahn, B. T. & Page, D. C. Functional coherence of the human Y chromosome. *Science* **278**, 675-680 (1997).

4       Jonkers, I. *et al.* RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. *Cell* **139**, 999-1011, doi:S0092-8674(09)01360-9 [pii] 10.1016/j.cell.2009.10.034 (2009).

5       Brown, C. J. *et al.* A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**, 38-44, doi:10.1038/349038a0 (1991).

6       Lee, J. T., Davidow, L. S. & Warshawsky, D. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* **21**, 400-404, doi:10.1038/7734 (1999).

7       Heard, E. *et al.* Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell* **107**, 727-738, doi:S0092-8674(01)00598-0 [pii] (2001).

8       Rougeulle, C. *et al.* Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* **24**, 5475-5484, doi:10.1128/mcb.24.12.5475-5484.2004 (2004).

9       Horvath, J. E. *et al.* Comparative analysis of the primate X-inactivation center region and reconstruction of the ancestral primate XIST locus. *Genome Res* **21**, 850-862, doi:10.1101/gr.111849.110 (2011).

10      Barakat, T. S. *et al.* The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. *Mol Cell* **53**, 965-978, doi:S1097-2765(14)00120-8 [pii] 10.1016/j.molcel.2014.02.006 (2014).

11      Gayen, S., Maclary, E., Hinten, M. & Kalantry, S. Sex-specific silencing of X-linked genes by Xist RNA. *Proc Natl Acad Sci U S A* **113**, E309-318, doi:10.1073/pnas.1515971113

(2016).

12    Luikenhuis, S., Wutz, A. & Jaenisch, R. Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. *Mol Cell Biol* **21**, 8512-8520, doi:10.1128/MCB.21.24.8512-8520.2001 (2001).

13    Lee, J. T. & Lu, N. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* **99**, 47-57, doi:S0092-8674(00)80061-6 [pii] (1999).

14    Giorgetti, L. *et al.* Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* **157**, 950-963, doi:10.1016/j.cell.2014.03.025 (2014).

15    Mlynarczyk-Evans, S. *et al.* X chromosomes alternate between two states prior to random X-inactivation. *PLoS Biol* **4**, e159, doi:10.1371/journal.pbio.0040159 (2006).

16    Barakat, T. S., Jonkers, I., Monkhorst, K. & Gribnau, J. X-changing information on X inactivation. *Experimental Cell Research* **316**, 679-687, doi:10.1016/j.yexcr.2010.01.015 (2010).

17    Makhlouf, M. *et al.* A prominent and conserved role for YY1 in Xist transcriptional activation. *Nat Commun* **5**, 4878, doi:10.1038/ncomms5878 (2014).

18    Barakat, T. S. *et al.* RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS Genet* **7**, e1002001, doi:10.1371/journal.pgen.1002001 (2011).

19    Gontan, C. *et al.* RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature* **485**, 386-390, doi:nature11070 [pii] 10.1038/nature11070 (2012).

20    Wee, K. E. *et al.* Steady-state kinetic analysis of human ubiquitin-activating enzyme (E1) using a fluorescently labeled ubiquitin substrate. *J Protein Chem* **19**, 489-498 (2000).

21    Siepmann, T. J., Bohnsack, R. N., Tokgöz, Z., Baboshina, O. V. & Haas, A. L. Protein interactions within the N-end rule ubiquitin ligation pathway. *J Biol Chem* **278**, 9448-9457, doi:10.1074/jbc.M211240200 (2003).

22    Masui, S. *et al.* Rex1/Zfp42 is dispensable for pluripotency in mouse ES cells. *BMC Dev Biol* **8**, 45, doi:10.1186/1471-213X-8-45 (2008).

23    Shin, J. *et al.* Maternal Rnf12/RLIM is required for imprinted X-chromosome inactivation in mice. *Nature* **467**, 977-981, doi:10.1038/nature09457 (2010).

24    Wutz, A., Rasmussen, T. P. & Jaenisch, R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* **30**, 167-174, doi:10.1038/ng820 ng820 [pii] (2002).

25    Cline, T. W. & Meyer, B. J. Vive la différence: males vs females in flies vs worms. *Annu Rev Genet* **30**, 637-702, doi:10.1146/annurev.genet.30.1.637 (1996).

26    Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. & Lee, J. T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322**, 750-756, doi:322/5902/750 [pii] 10.1126/science.1163045 (2008).

27     Plath, K. *et al.* Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. *J Cell Biol* **167**, 1025-1035, doi:10.1083/jcb.200409026 (2004).

28     de Napoles, M. *et al.* Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell* **7**, 663-676, doi:10.1016/j.devcel.2004.10.005 (2004).

29     Schoeftner, S. *et al.* Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J* **25**, 3110-3122, doi:10.1038/sj.emboj.7601187 (2006).

30     Costanzi, C., Stein, P., Worrad, D. M., Schultz, R. M. & Pehrson, J. R. Histone macroH2A1 is concentrated in the inactive X chromosome of female preimplantation mouse embryos. *Development* **127**, 2283-2289 (2000).

31     Mietton, F. *et al.* Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome. *Mol Cell Biol* **29**, 150-156, doi:10.1128/MCB.00997-08 (2009).

32     Grant, M., Zuccotti, M. & Monk, M. Methylation of CpG sites of two X-linked genes coincides with X-inactivation in the female mouse embryo but not in the germ line. *Nat Genet* **2**, 161-166, doi:10.1038/ng1092-161 (1992).

33     Kaslow, D. C. & Migeon, B. R. DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation. *Proc Natl Acad Sci U S A* **84**, 6210-6214 (1987).

34     Hadjantonakis, A. K., Cox, L. L., Tam, P. P. & Nagy, A. An X-linked GFP transgene reveals unexpected paternal X-chromosome activity in trophoblastic giant cells of the mouse placenta. *Genesis* **29**, 133-140 (2001).

35     Corbel, C., Diabangouaya, P., Gendrel, A. V., Chow, J. C. & Heard, E. Unusual chromatin status and organization of the inactive X chromosome in murine trophoblast giant cells. *Development* **140**, 861-872, doi:10.1242/dev.087429 (2013).

36     Dubois, A. *et al.* Spontaneous reactivation of clusters of X-linked genes is associated with the plasticity of X-inactivation in mouse trophoblast stem cells. *Stem Cells* **32**, 377-390, doi:10.1002/stem.1557 (2014).

37     Prudhomme, J. *et al.* A rapid passage through a two-active-X-chromosome state accompanies the switch of imprinted X-inactivation patterns in mouse trophoblast stem cells. *Epigenetics Chromatin* **8**, 52, doi:10.1186/s13072-015-0044-2 (2015).

38     Sado, T. *et al.* X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. *Dev Biol* **225**, 294-303, doi:10.1006/dbio.2000.9823 (2000).

39     Sunwoo, H., Wu, J. Y. & Lee, J. T. The Xist RNA-PRC2 complex at 20-nm resolution

**6**

reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proc Natl Acad Sci U S A* **112**, E4216-4225, doi:1503690112 [pii] 10.1073/pnas.1503690112 (2015).

40      Keniry, A. *et al.* Setdb1-mediated H3K9 methylation is enriched on the inactive X and plays a role in its epigenetic silencing. *Epigenetics Chromatin* **9**, 16, doi:10.1186/s13072-016-0064-6 (2016).

41      Kohlmaier, A. *et al.* A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. *PLoS Biol* **2**, E171, doi:10.1371/journal.pbio.0020171 (2004).

42      Nishioka, K. *et al.* PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Mol Cell* **9**, 1201-1213 (2002).

43      Fang, J. *et al.* Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. *Curr Biol* **12**, 1086-1099 (2002).

44      Payer, B. Developmental regulation of X-chromosome inactivation. *Semin Cell Dev Biol* **56**, 88-99, doi:10.1016/j.semcdb.2016.04.014 (2016).

**6**
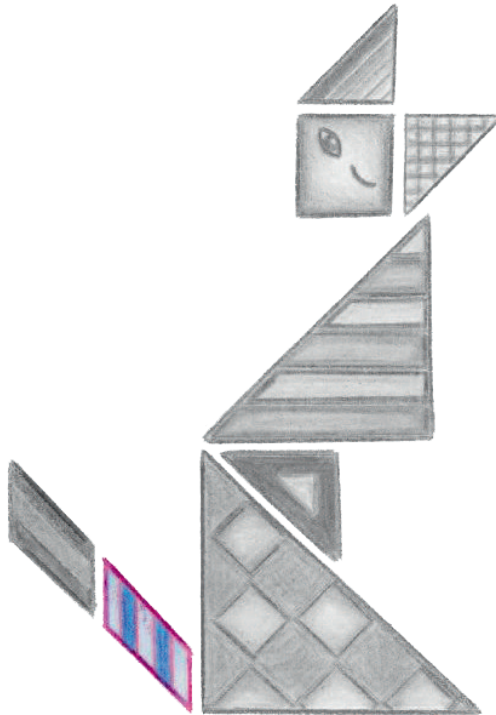
# Addendum 1

## Genes not involved in initiation of XCI

# Addendum 1

**Abstract**

Proper initiation of X chromosome inactivation (XCI), to equalize gene dosage between XX females and XY males, requires upregulation of X-encoded activators to allow a female cell to overcome the threshold for XCI to occur. So far, *Rnf12* and *Uba1* have been reported as activators of X chromosome inactivation, nevertheless a heterozygous deletion of both genes in female cells does not lead to full abrogation of XCI upon differentiation and suggests the presence of other activator(s). Here we report several candidate genes which we have tested to determine whether they function in initiation of XCI. We found that *Nshp1, Dax1, Pja1, Cited1* can be excluded as regulators of XCI initiation and indicate a role for *Phf6* at later stages of XCI.

**Introduction**

X chromosome inactivation (XCI) evolved as a dosage compensation mechanism ensuring equal gene dosage of sex-chromosomes between eutherian XX females and XY males. XCI is mostly studied in mouse embryonic stem (ES) cells, as these cells recapitulate random XCI upon differentiation. Random XCI *in vivo* occurs early in development of the female embryo and implies that, in a female cell it is either the maternal or the paternal X chromosome which can be inactivated [1]. It is the balance between X-encoded activators and autosomally encoded inhibitors of XCI as well as the stochasticity of the process which determines the probability to initiate XCI. In addition, feedback through inactivation of XCI activators in *cis* prevents inactivation of both X chromosomes [1, 2]. Studies on X/A translocations indicated that initiation of cis-inactivation depends on the presence of a unique region on the X chromosome, the X inactivation center (Xic) [3]. In addition, studies on female cells with truncated X chromosomes suggested that at least two copies of the Xic are required for XCI to occur [4]. The Xic has therefore been described as the minimum region both necessary and sufficient to initiate XCI. More recently it has been proposed that two distinguishable Xics exist: the cis-Xic and the trans-Xic [5]. The cis-Xic contains the three key players of XCI, which are *Xist* and its inhibitors *Tsix* and *Xite*, but also *cis* activators of *Xist* such as *Jpx, Ftx* and *Xpr* [6]. The trans-Xic contains all the factors that act in *trans* to activate *Xist* either directly or indirectly. The trans-Xic comprises the autosomally encoded inhibitors, which set up a threshold for XCI to occur, and the X encoded activators which allow a cell to determine the number of active X chromosome in a cell and initiate XCI when this number is more than 1 per diploid genome. Two X-linked activators have been described, *Rnf12* [2, 6-8] and *Uba1* (Chapter 2). Studies with $Rnf12^{+/-}:Uba1^{+/-}$ ES cells have shown that *Rnf12* and *Uba1* are not the only activators of XCI, as these cells still initiated XCI in a percentage of cells, whereas wildtype male cells do not.

Here we report studies determining the function of five candidate genes in initiation of XCI using two different strategies. Candidate genes were selected based on a relevant known function, have to be subject to XCI to allow feedback and also based on their location in close proximity to *Xist*. Candidate genes included: *Nshp1, Dax1, Pja1, Cited1* and *Phf6*. *Nspb1* affects chromatin structure and function and can exert nucleosomal binding and transcriptional activation [13, 14]; *Dax1* plays a role in sex reversal and was shown to downregulate *Oct4* expression [15, 16] and bind RNF12 [Gontan, unpublished]; *Cited1* is a CBP/p300 trans-activator which

also interacts with *Smad5*, estrogen receptors α and β and TFAP2 [17-20]; and *Pja1* is an E3 ubiquitin ligase, also named *Rnf70*, and part of the same family as *Rnf12* [21]. As for *Phf6*, unpublished knockdown studies (Van Vlierberghe) resulted in downregulated *Xist* expression.

## Experimental procedures

### BACs and plasmids

The BACs (CHORI) used in this study were RP23-346H6 (*Dax1*), RP23-312H18 (*Pja1*), RP23-204G22 (*Cited1*) and RP23-332L18 (*Nsbp1*). Recombination between the BAC and an ori-less recombination vector to introduce a kanamycin/neomycin resistance cassette, allowing selection in ES cells, into the BAC backbone was described previously [9]. The pSpCas9(BB)-2A-Puro (PX459) (Addgene plasmid # 48139) was used to insert the CRISPR guide RNAs [10].

*Culturing ES cells*

Cell lines, culture media and culture conditions for ESC culture and differentiation were described previously [7, 11]. The BACs (linearized) were introduced in F1 wildtype male ES cells and CRISPR vectors were introduced in hybrid F1 129sv/Cast female ES cells with a deletion of *Rnf12* on the 129 allele (Chapter 2), by electroporation in a 0.2cm electroporation cuvette (BioRad) at 118kV, 1200μF and ∞Ω in a Gene Pulser Xcell Electroporation System (BioRad). After 24 hours of recovery, cells were grown on ES medium containing neomycin (270μg/ml) or puromycin (1ug/ml) for seven days or 48 hours for BACs and the CRISPR vectors respectively.

*Deletions with the CRISPR/CAS system*

CRISPR guides were designed by using the CRISPR design tool (http://crispr.mit.edu/). The designed CRISPR guide oligos with 5'- CACC and 3'- CAAA overhangs (Table 1) were cloned into the pX459 CRISPR vector (Addgene) [10] by a simultaneous digestion-ligation reaction [12]. First, the pX459 vector was digested with BbsI (NEB) to allow the replacement of the restriction sites with direct insertion of the annealed oligo guides and used to transform heat competent bacteria. A combination of two guides was used in each targeting of ES cells for which the cells were screened as a pool to ensure the deletion was present, followed by single clone screens with different primers (Table 2).

*Xist RNA Fluorescent in situ hybridization*

Procedure, probe labelling and antibodies have been described previously [7, 11].

*RNA isolation from ES cells and cDNA synthesis*

TRIzol® (Invitrogen) was added to each sample and RNA was isolated according to manufacturer's instructions. RNA samples were diluted to contain 2μg RNA in 10μl. Before cDNA synthesis, DNase (Invitrogen) was added to the RNA samples, for 30 minutes at 37ºC, to remove any contaminating genomic DNA after which the DNase was heat-inactivated at 65ºC for 10 minutes. Random Hexamers (Applied BioSystems) and dNTP mix (Invitrogen) was added and incubated at 65ºC for 5 minutes and subsequently cooled on ice for 1 minute. 5x First Strand Buffer, 0.1M DTT and RNaseOUT (Invitrogen) were then added to the samples and incubated at 25ºC for 2 minutes before adding SuperscriptII Reverse transcriptase (Invitrogen). The samples were then incubated at 25ºC for 10 minutes followed by 50 minutes at 42ºC and finally, 15 minutes at 70ºC. The obtained cDNA was diluted with 200μl $H_2O$, of which 3-5μl was used in a Q-PCR reaction.

*Table 1: Oligo's for CRISPR guides for deletion of Cited1 and Phf6. 5'- CACC and 3'- CAAA overhangs are indicated in bold and the preferred G as a starting nucleotide is underlined.*

| Gene | | | Guide oligo |
|------|---|---|-------------|
| **Cited1** | Combination 1: | Fw 1: | 5'CACC**G** TCTCATTGATTCTGACCCGG 3' |
| | | Rv 1: | 5'AAAC CCGGGTCAGAATCAATGAGA **C**3' |
| | | Fw 3: | 5'CACC**G** AGCCCCGCGCCGATAGGTTG 3' |
| | | Rv3: | 5'AAAC CAACCTATCGGCGCGGGGCT **C**3' |
| | Combination 2: | Fw2: | 5'CACC**G** CCAATGAGCTTCCCGAGCTG 3' |
| | | Rv2: | 5'AAAC CAGCTCGGGAAGCTCATTGG **C**3' |
| | | Fw 3: | 5'CACC**G** AGCCCCGCGCCGATAGGTTG 3' |
| | | Rv3: | 5'AAAC CAACCTATCGGCGCGGGGCT **C**3' |
| | Combination 3: | Fw2: | 5'CACC**G** CCAATGAGCTTCCCGAGCTG 3' |
| | | Rv2: | 5'AAAC CAGCTCGGGAAGCTCATTGG **C**3' |
| | | Fw4: | 5'CACC**G** TACTTACGGAGGGCCCGAGT 3' |
| | | Rv4: | 5'AAAC ACTCGGGCCCTCCGTAAGTA **C**3' |
| **Phf6** | Combination 1: | Fw2: | 5'CACC**G** GCCGCCCCTGCCCGTTGTTA3' |
| | | Rv2: | 5'AAAC TAACAACGGGCAGGGGCGGC **C**3' |
| | | Fw3: | 5'CACC**G** TGGGAGGTTAACTAATGATG 3' |
| | | Rv3: | 5'AAAC CATCATTAGTTAACCTCCCA C3' |

*Table 2: Primers for detection wildtype allele, deletion and allele check of the deletion of Cited1 and Phf6*

| | Wildtype allele 5'-3' | PCR over deletion 5'- 3' | Allele check 5'-3' |
|---|---|---|---|
| **ΔCited1 Fw** | AACAGAATCGGTGGCTTTTT | AACAGAATCGGTGGCTTTTT | AAACACCCTCCTCAAACTGG |
| **Rv** | CCAACCTTGGAGTGAAGGAT | GGGCGAGAAGTGACCATAAT | GAGGTAAAGGACGCTTGGAG |
| | | | AscI SNP on 129 |
| **ΔPhf6 Fw** | CTCTCCTCGAATGAAAGGAA | TCTCTCCTCGAATGAAAGGAA | TGGCTTTCTAATGCAGATGG |
| **Rv** | TTTTCTTTTTCTCCCCCTCA | CGCTGTTCTACCAGCTACAGA | GTGGCTCACAACATGGAATC |
| | | | ScaI on CAS |

## Results

*Additional copies of candidate gene Cited1 induce ectopic cloud formation in male cells*

Activation of XCI depends on the ability of the X-encoded XCI activators to overcome the threshold set by the XCI inhibitors. X-encoded XCI activators allow a female cell with two X chromosomes to initiate XCI, whereas males with one X chromosome cannot. In order to test candidate genes as potential XCI activators, these candidate genes were introduced in extra copies in male cells. If the candidate gene is indeed an XCI activator, the male cells will initiate XCI on their single X chromosome as was shown for *Rnf12* (Figure 1A). Introduction of a BAC containing *Rnf12* into male cells, resulted in ectopic *Xist* cloud formation in ~9% and ~21% of the cells with one and two extra copies of *Rnf12* respectively [2]. Using a BAC transgene approach, we introduced extra copies of *Nshp1, Dax1, Pja1* and *Cited1* in male ES cells and analyzed for

ectopic *Xist* cloud formation by *Xist* RNA FISH (Figure 1A). Only when extra copies of *Cited1* were introduced in male cells ectopic *Xist* cloud formation was observed in different clones, but not in male control cells (Figure B and C). However, unlike *Rnf12*, ectopic clouds were only observed in up to 3% of male cells with extra copies of *Cited1* (Figure 1D). Moreover, no clear dosage correlation between copy number and cloud formation was observed in males with extra copies of *Cited1*, even though the percentage of cells with *Xist* clouds appeared to increase with copy number (Figure 1D).
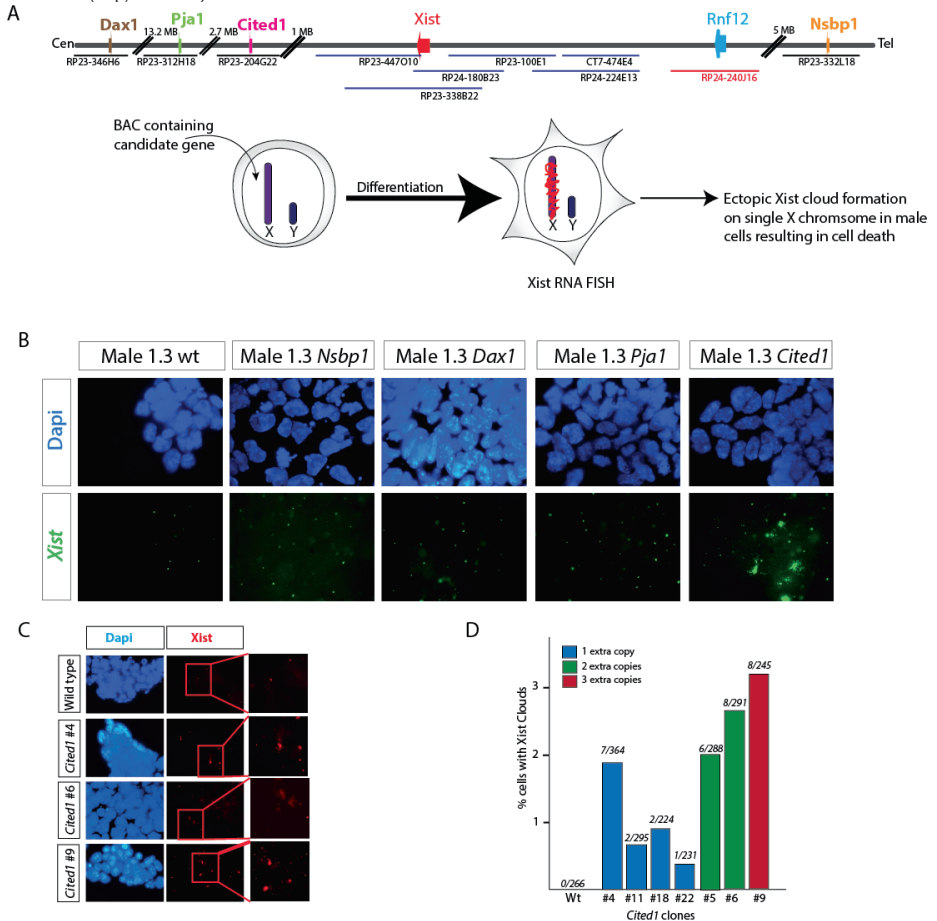
**Figure 1:** A) Top: Region on the X chromosome showing location of the candidate genes *Dax1, Pja1, Cited1* and *Nsbp1* and the corresponding BACs used are shown below (black). BACs shown in blue were used in the screen to identify *Rnf12* of which the BAC shown in red resulted in ectopic *Xist* cloud formation in male cells [2]. Bottom: These BACs were introduced in male cells and if the BAC would contain an activator of XCI, ectopic Xist cloud formation would be observed after differentiation of these cells, followed by cell death when the single X chromosome is inactivated. B) Xist RNA FISH on male 129sv/Cast cells at day 3 of differentiation with extra copies of the candidate genes *Nsbp1, Dax1, Pja1* and *Cited1*. Only when extra copies of *Cited1* was introduced in male cells, *Xist* clouds (green) were observed. C) Xist RNA FISH on three *Cited1* clones showing *Xist* clouds (red). D) *Cited1* clones with 1 (blue), 2 (green) and 3 (red) extra copies of *Cited1* with the corresponding percentage of cells observed with *Xist* clouds depicted for each clone.

*Loss of Cited1 does not result in reduced Xist expression*

In order to further test a role for *Cited1* in XCI, we generated clones with a ~4.6 kb deletion of *Cited1* on the 129sv allele, the Cast allele or on both alleles in female $Rnf12^{+Cast/129-}$ cells (described in Chapter 2), using the CRISPRS/Cas9 technology (Figure 2A). If *Cited1* would be involved in initiation of XCI, a heterozygous deletion would further affect XCI in these $Rnf12^{+/-}$ cells visualized as a reduction in the percentage of cells with *Xist* clouds by Xist RNA FISH and reduced *Xist* expression by q-PCR (Figure 2A bottom figure). We generated 8 clones with a deletion of *Cited1* (Figure 2B), of which two clones were mixed clones (Figure 2C clone 2 and clone 3). The analysis was continued with 3 clones with the deletion of *Cited1* on the 129 allele (clone 1, 5 and 7), one clone with the deletion on the Cast allele (clone 8) and one clone which had the deletion on both alleles (clone 6). These 5 clones, now referred to as ΔCited 129-1, ΔCited 129-2, ΔCited 129-3, ΔCited Cas and Cited KO, still retained two X chromosomes (Figure 2D) and were analyzed by q-PCR for *Cited1* and *Xist* expression after 4 and 6 days of differentiation. In the Cited KO cells, no *Cited1* was observed after 4 or 6 days of differentiation as expected, whereas the ΔCited Cas and ΔCited 129-1 clones showed similar *Cited1* levels after both 4 and 6 days of differentiation and the ΔCited 129-2, ΔCited 129-3 clones showed an increase in *Cited1* expression from day 4 to day 6. Nevertheless, on both days *Cited1* expression was reduced in all clones as compared to wildtype levels (Figure 2E). When analyzing *Xist* expression after 4 days of differentiation, there was no difference in *Xist* levels between any of the clones with the $Rnf12^{+/-}$ cells, however, after day 6 of differentiation, *Xist* levels in ΔCited 129-1, ΔCited Cas and Cited KO were significantly increased as compared to the $Rnf12^{+/-}$ cells (Figure 2F). In these 3 clones, no *Cited1* upregulation was observed from day 4 to day 6, whereas in ΔCited 129-2 and ΔCited 129-3 clones showing no significant difference compared to $Rnf12^{+/-}$ cells, this upregulation was absent. Taken together, the results show that *Cited1* is not an XCI activator and might even be involved in repression of *Xist*.
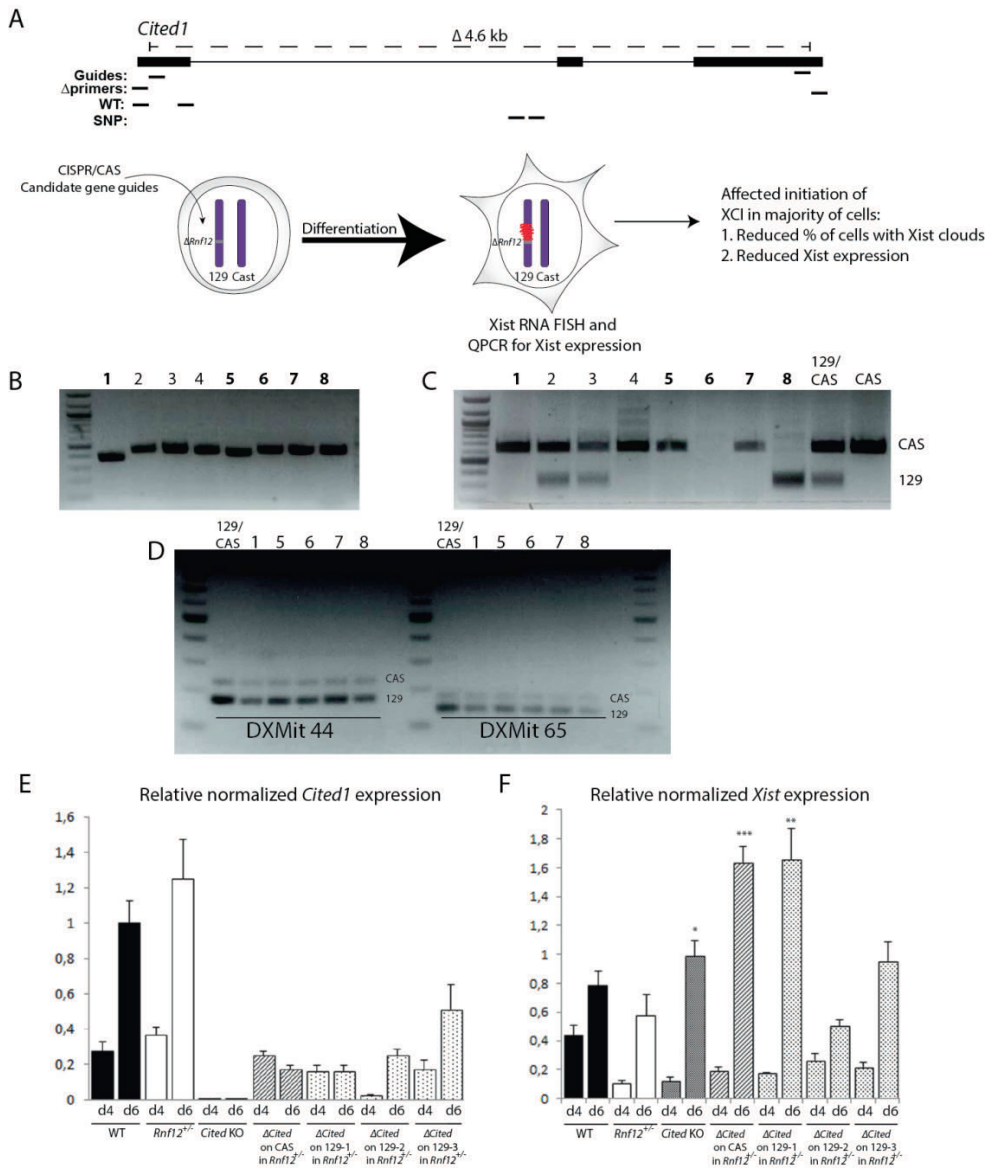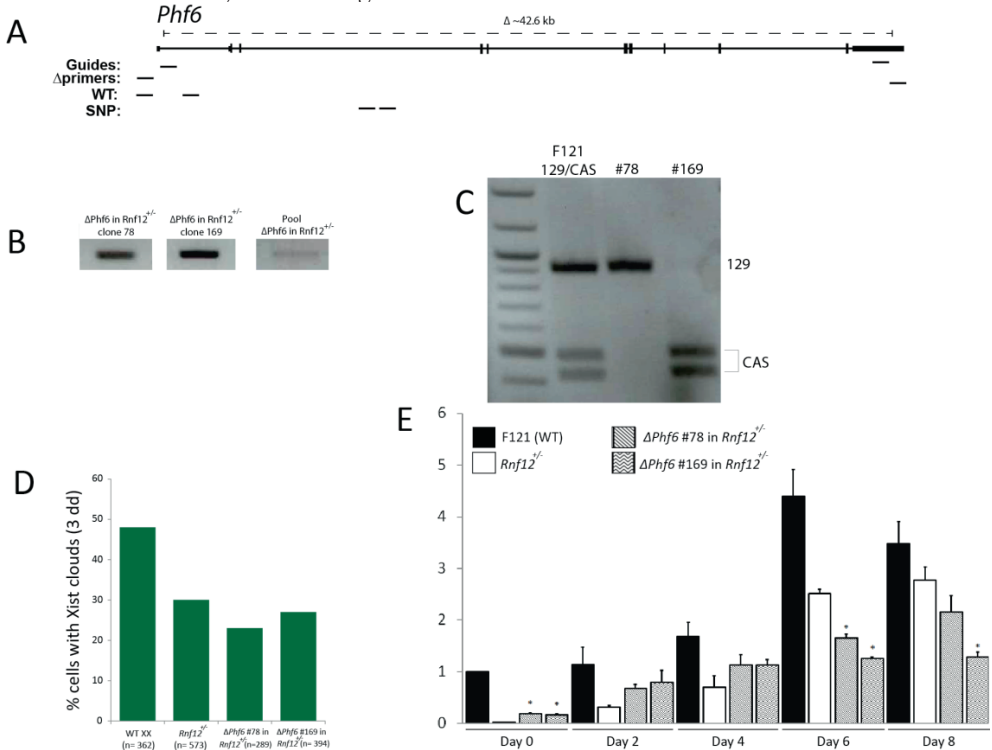
**Figure 2**:A) *Cited1* gene with deletion guides, generating a deletion of approximately 4.6 kb, and location of primers used to screen the obtained clones. The guides were introduced in $Rnf12^{+/-}$ cells and analyzed by Xist RNA FISH and Xist QPCR for an effect on XCI after differentiation. B) PCR over the deleted region in 8 clones. C) Allele specific PCR of *Cited1* allele identifying the targeted allele in the 8 clones. D) PCR over DXMit44 and DXMit65 markers to determine whether the clones retained two X chromosomes. E) Relative *Cited* expression normalized to β-actin in wildtype, $Rnf12^{+/-}$, Cited KO, ΔCited Cas, ΔCited 129-1, ΔCited 129-2 and ΔCited 129-3 cells (all clones are in a $Rnf12^{+/-}$ background). (Error bars indicate SEM). Cited KO cells have no *Cited1* expression at day 4 or 6 as expected and the heterozygous clones have reduced *Cited1* expression as compared to wildtype and $Rnf12^{+/-}$ cells. F)

Relative *Xist* expression normalized to β-actin in wildtype, *Rnf12*[+/-] , Cited KO, ΔCited Cas, ΔCited 129-1, ΔCited 129-2 and ΔCited 129-3 cells (all clones are in a *Rnf12*[+/-] background). (Error bars indicate SEM; t-test comparing *Rnf12*[+/-] to *Cited1* clones, * p<0.05; ** p<0.001 and *** p< 0.0001). *Xist* levels in ΔCited 129-1, Cited Cas and Cited KO are significantly increased as compared to the *Rnf12*[+/-] cells.

*Loss of Phf6 affects later stages of XCI*

Knockdown of *Phf6,* a chromatin adapter protein, using shRNAs in human female CD34+ T-cell precursors followed by gene expression profiling showed *Xist* as one of the strongest downregulated genes as compared to control cells (Van Vlierberghe, unpublished). We therefore generated a 42.6 kb deletion of *Phf6* (Figure 3A), in female *Rnf12*[+/-] cells with a deletion of *Rnf12* on the 129 allele (described in Chapter 2). We generated two clones (Figure 3B) of which one clone, clone 78, obtained the deletion on the Cast allele and the other clone, clone 169, obtained the deletion on the 129 allele (Figure 3C). The cells were differentiated for 3 days and analyzed by *Xist* RNA FISH. At day 3 of differentiation, no apparent difference was observed in the percentage of cells with *Xist* clouds between the two *Phf6* clones and the *Rnf12*[+/-] cells (Figure 3D). Analysis of *Xist* expression and reviewing more time points with longer differentiation, resulted in a significant increased *Xist* expression at day 0 in both clones as compared to *Rnf12*[+/-] cells followed by significantly decreased *Xist* expression at day 6 for both clones. As the effects of the *Phf6* mutation are different throughout the differentiation time course, a role for *Phf6* in XCI remains inconclusive as in the early stages of XCI the two clones show no affected XCI, however a role for *Phf6* in later stages needs to be further elucidated.

**Figure 3:** A) *Phf6* gene with deletion guides, generating a deletion of approximately 42.6 kb, and location of primers used to screen the obtained clones.B) PCR over the deleted region of *Phf6* in clone 78, 169

and the targeted pool. C) Allele check *Phf6*. Clone 78 obtained the deletion on the CAS allele and clone 169 has obtained the deletion on the 129 allele. D) Xist RNA FISH results depicting the percentage of cells with *Xist* clouds in wildtype, *Rnf12*[+/-] cells and the two *Phf6* clones at day 3. E) Relative *Xist* expression (normalized to β-actin) of wildtype, *Rnf12*[+/-] cells and the two *Phf6* clones at day 0, 2,4,6 and 8 of differentiation (error bars represent standard deviation; t-test comparison between *Rnf12*[+/-] cells with clone 78 and clone 169; * p< 0,05.

**Discussion and conclusion**

In order to identify XCI activators we used the candidate gene approach in which we carefully selected *Nsbp1, Dax1, Pja1, Cited1* and *Phf6*. However, all these genes did not appear to behave as activators of XCI, which shows that a candidate approach can be laborious and precarious. *Cited1* did induce ectopic *Xist* cloud formation in male cells, but had the opposite effect expected from an XCI activator when deleted in female cells. One explanation could be that the BAC used to introduce extra copies in male cells contained other genes besides *Cited1* responsible for the formation of these ectopic clouds. Besides *Cited1* the BAC contained two predicted genes and one miRNA and this can be looked further into. *Cited1* has been reported to be an immediate target of *Wnt* signaling and an important regulator of the Wnt pathway [22], which when affected can result in increased in proliferation. Moreover, in the Wnt pathway *Cited1* has been described to function in a 'just right' model, which implies that *Cited1* over-expression and loss of expression can lead to the same result (e.g. tumorigenesis), which is also what we observe, as both overexpression and loss resulted in more *Xist* (clouds or expression) as compared to control cells. Even though dosage of *Cited1* expression requires proper regulation, we exclude *Cited1* as an XCI activator initiating XCI.

A heterozygous deletion of *Phf6* in *Rnf12*[+/-] cells resulted in downregulated *Xist* expression but only after the time window of reversible XCI, whereas within the time window of XCI the clones seemed to behave better than the *Rnf12*[+/-] cells which suggests that *Phf6* is not an XCI activator involved in initiation of XCI. However, at day 6 and day 8 XCI seemed affected which could indicate a role for *Phf6* in the establishment phase involving changes to the chromatin and fits with a known role of *Phf6* as a chromatin adapter protein. Loss of *Phf6* could then affect the establishment of the Xi leading to e.g. loss of *Xist* spreading along the Xi and loss of XCI in a proportion of cells. Nonetheless, we do not have sufficient data to support such a role for *Phf6* and this needs to be further elucidated.

In this study we used two different approaches to screen candidate genes, introducing extra copies of a candidate gene in male cells in which addition of an activator would result in ectopic *Xist* cloud formation, as opposed to a deletion of a candidate gene in female *Rnf12*[+/-] cells, in which a heterozygous loss of an activator would lead to affected XCI. The latter provides a better readout since in male cells identification of XCI activators with a relatively small activity would be difficult. Moreover, inactivation of the single X chromosome in male cells would result in cell death and ectopic clouds formed are different from the Xi in female cells [23]. When using female cells on a hybrid background the targeted allele can be easily identified in addition to which X chromosome is inactivated and are therefore a more convenient system to identify XCI activators. To screen more efficiently, the region containing XCI activators can be narrowed down first in *Rnf12*[+/-]:*Uba1*[+/-] cells (described in Chapter 2), in which a candidate region would result in abolished XCI, followed up by a candidate gene approach of genes located in this region.

# References

1. Monkhorst, K. et al. (2008) X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. Cell 132 (3), 410-21.

2. Jonkers, I. et al. (2009) RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. Cell 139 (5), 999-1011.

3. Brown, C.J. et al. (1991) Localization of the X inactivation centre on the human X chromosome in Xq13. Nature 349 (6304), 82-4.

4. Augui, S. et al. (2011) Regulation of X-chromosome inactivation by the X-inactivation centre. Nat Rev Genet 12 (6), 429-42.

5. Gribnau, J. and Grootegoed, J.A. (2012) Origin and evolution of X chromosome inactivation. Curr Opin Cell Biol 24 (3), 397-404.

6. Barakat, T.S. et al. (2014) The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. Mol Cell 53 (6), 965-78.

7. Barakat, T.S. et al. (2011) RNF12 activates Xist and is essential for X chromosome inactivation. PLoS Genet 7 (1), e1002001.

8. Gontan, C. et al. (2012) RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature 485 (7398), 386-90.

9. Barakat, T.S. and Gribnau, J. (2015) Generation of knockout alleles by RFLP based BAC targeting of polymorphic embryonic stem cells. Methods Mol Biol 1227, 143-80.

10. Ran, F.A. et al. (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8 (11), 2281-308.

11. Jonkers, I. et al. (2008) Xist RNA is confined to the nuclear territory of the silenced X chromosome throughout the cell cycle. Molecular and Cellular Biology 28 (18), 5583-94.

12. Cong, L. et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339 (6121), 819-23.

13. Rochman, M. et al. (2009) The interaction of NSBP1/HMGN5 with nucleosomes in euchromatin counteracts linker histone-mediated chromatin compaction and modulates transcription. Mol Cell 35 (5), 642-56.

14. Rochman, M. et al. (2010) HMGN5/NSBP1: a new member of the HMGN protein family that affects chromatin structure and function. Biochim Biophys Acta 1799 (1-2), 86-92.

15. Jadhav, U. et al. (2011) Hypogonadotropic hypogonadism in subjects with DAX1 mutations. Mol Cell Endocrinol 346 (1-2), 65-73.

16. Sun, C. et al. (2009) Dax1 binds to Oct3/4 and inhibits its transcriptional activity in embryonic stem cells. Mol Cell Biol 29 (16), 4574-83.

17. Howlin, J. et al. (2006) CITED1 homozygous null mice display aberrant pubertal mammary ductal morphogenesis. Oncogene 25 (10), 1532-42.

18. Rodriguez, T.A. et al. (2004) Cited1 is required in trophoblasts for placental development and for embryo growth and survival. Mol Cell Biol 24 (1), 228-44.

19. Yahata, T. et al. (2000) The MSG1 non-DNA-binding transactivator binds to the p300/CBP coactivators, enhancing their functional link to the Smad transcription factors. J Biol Chem 275 (12), 8825-34.

20. Yahata, T. et al. (2001) Selective coactivation of estrogen-dependent transcription by CITED1 CBP/p300-binding protein. Genes Dev 15 (19), 2598-612.

21. Saha, T. et al. (2006) RING finger-dependent ubiquitination by PRAJA is dependent on TGF-beta and potentially defines the functional status of the tumor suppressor ELF. Oncogene 25 (5), 693-705.

22. Méniel, V. et al. (2013) Cited1 deficiency suppresses intestinal tumorigenesis. PLoS Genet 9 (8), e1003638.

23. Smeets, D. et al. (2014) Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. Epigenetics Chromatin 7, 8.

**&-1**
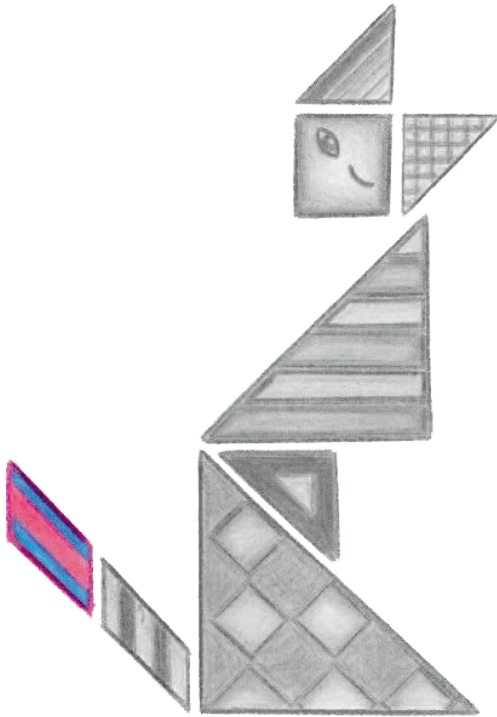
# Addendum 2

# Addendum 2

**Summary**

Dosage compensation of X encoded genes in placental mammals is achieved by near complete inactivation of one of the two X chromosomes in female cells. Even though X chromosome inactivation (XCI) has been discovered over 50 years ago, the complete mechanism ensuring proper inactivation of only one X chromosome and maintenance of the inactive X (Xi) is not yet fully understood. This thesis focusses on the activation of XCI, the initiation phase, in which XCI is established on one X chromosome and also touches upon maintenance of the inactive state. The two key players *Xist* and *Tsix*, both lncRNAs, ensure one X chromosome is in the proper state prior to XCI by differential expression and chromatin conformation of the two X chromosomes. When a cell initiates XCI, autosomally encoded XCI inhibitors, acting as de-nominators, will be down regulated and X encoded XCI activators, acting as numerators, will be upregulated. XCI inhibitors, inhibiting *Xist* or promoting *Tsix* expression, consist mainly of pluripotency factors and link XCI to loss of pluripotency. XCI activators, promoting *Xist* and inhibiting *Tsix* expression, have been less well characterized. XCI activators, as they are X encod-ed, are differentially expressed between males and females and allow female specific initiation of XCI. Initiation of XCI on one of the two X chromosomes is a tightly regulated but stochastic process in which each X chromosome has an equal probability of being inactivated. Regulation of XCI is achieved by the counterbalance of the XCI inhibitors and XCI activators. *Rnf12,* an E3 ubiquitin ligase, has been identified as a potent XCI activator and downregulates XCI inhibitor REX1 by proteasomal degradation. In this thesis we report *Uba1*, an E1 activating enzyme, as a novel XCI activator which could along with *Rnf12* be involved in a cascade in the proteasomal degradation of REX1. REX1 functions most likely by competing with the *Xist* activator *YY1* for binding sites at the *Xist* promoter and thereby inhibiting *Xist* expression. *Xist* activators differ from XCI activators as they exclusively activate *Xist* and are autosomally encoded. In this thesis we identified autosomally encoded *Rnf6*, a close homolog of *Rnf12*, as a novel *Xist* activator and propose a function for *Rnf6* in XCI by aiding in the downregulation of REX1. The interaction between UBA1, RNF12, RNF6 and REX1 is a good example of the possible interplay between activation and inhibition of XCI. Even though we have identified novel players in initiation of XCI, our data also shows there must be other factors at play which remain to be identified. Once XCI is initiated and *Xist* RNA spreads and coats the X chromosomes it recruits different chro-matin remodeling factors to further induce the silent state. At the same time, downregulation of XCI activators by inactivation of the X chromosome allows proper feedback and prevents inactivation of the second X chromosome. Ultimately DNA methylation appears to be the final lock required to maintain the inactive state throughout cell divisions. All things considered, XCI remains an intriguing epigenetic phenomenon which appears to involve multiple intertwined and redundant pathways and regulators.

**Samenvatting**

In zoogdieren worden mogelijke concentratieverschillen in X gecodeerde genproducten wordt gelijk gemaakt tussen XX vrouwelijke en XY mannelijke cellen bereikt door één van de twee X chromosomen in vrouwelijke cellen uit te schakelen. Dit X chromosoom inactivatie (XCI) proces is meer dan 50 jaar geleden ontdekt. Echter het mechanisme dat ervoor zorgt dat slechts één X chromosoom wordt geïnactiveerd, en de processen betrokken bij het onderhouden van de geïnactiveerde X (Xi) zijn nog steeds niet volledig begrepen. Deze thesis richt zich op het activeren van XCI, de beginfase waarin XCI op één X chromosoom wordt geïnitieerd. Daarnaast zijn mechanismen betrokken bij het onderhouden van de geïnactiveerde staat onderzocht. De twee hoofdrolspelers *Xist* en *Tsix*, die allebei lange niet coderende RNAs afschrijven, zorgen ervoor dat alleen één X chromosoom wordt uitgezet. Initiatie van XCI wordt bepaald door autosomaal gecodeerde remmers van XCI en X gecodeerde activators van XCI. Veel factoren die een belangrijke rol spelen bij het in stand houden van de ongedifferentieerde 'pluripotente' staat van embryonale stamcellen vervullen de functie van XCI remmers. Dit doen zij door *Xist* expressie te onderdrukken en *Tsix* expressie te bevorderen. Hierdoor wordt XCI gekoppeld aan ontwikkeling en cel differentiatie. Omdat XCI activators X gecodeerd zijn komen deze genen verschillend tot expressie in mannen en vrouwen, en dit leidt tot exclusieve vrouwelijke XCI initiatie. Initiatie van XCI is een sterk gereguleerd maar stochastisch proces waarin elk X chromosoom een even grote kans heeft om geïnactiveerd te worden. Regulatie van XCI wordt dus bereikt door het tegengewicht van de XCI remmers en XCI activators. *Rnf12* is een E3 ubiquitin ligase, geïdentificeerd als een krachtige XCI activator en functioneert door het geprogrammeerd afbreken van de XCI remmer REX1 door het proteasoom. In deze thesis tonen wij aan dat *Uba1,* een E1 activerend enzym dat ook een rol speelt bij de afbraak van eiwitten, als een nieuwe XCI activator, samen met *Rnf12,* betrokken is in activatie van XCI. In XCI functioneert REX1 hoogstwaarschijnlijk door te concurreren met de *Xist* activator YY1 voor bindingsplaatsen in de Xist regulatoire gebieden en onderdrukt daarmee *Xist* expressie. *Xist* activators verschillen van XCI activators, omdat *Xist* activators uitsluitend *Xist* activeren en autosomaal gecodeerd zijn. In dit proefschrift hebben wij het autosomaal gecodeerde *Rnf6* geïdentificeerd, een homoloog aan *Rnf12,* als nieuwe *Xist* activator, en stellen een functie voor *Rnf6* in XCI voor bij het helpen van de afbraak van REX1. De wisselwerking tussen UBA1, RNF12, RN6 en REX1 is een prachtig voorbeeld van de complexe interactie tussen remmers en activators van XCI. Ook al hebben we nieuwe spelers in initiatie van XCI geïdentificeerd, laat dit onderzoek ook zien dat er andere factoren meespelen die nog geïdentificeerd moeten worden. Zodra XCI wordt geïnitieerd en *Xist* RNA over de X chromosomen spreidt, werft *Xist* verschillende chromatine hermodelleringsfactoren om de inactieve staat te induceren. Tegelijkertijd geeft het uitschakelen van XCI activators door inactivatie van het X chromosoom goede feedback en voorkomt uitschakeling van het tweede X chromosoom. DNA methylering is hoogstwaarschijnlijk het benodigde laatste slot te zijn om de inactieve staat te behouden. Alles bij elkaar genomen, blijft XCI een intrigerend epigenetisch fenomeen dat als krachtig model dient om gen regulatie te bestuderen in een brede context.

**Curriculum Vitae**

Name:        Cheryl Mandy Maduro
Address:     Den Haag
Date of birth:   June 7 1987
Email:
Phone number:
Nationality:   Dutch

---
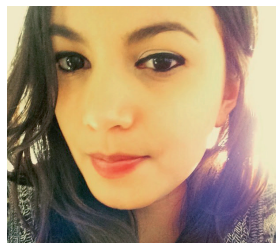
o        **Education**

| | |
|---|---|
| 2011-Now | PhD in Biomedical science |
| | Erasmus MC, Rotterdam |
| | |
| 2009-2011 | Msc. Molecular Medicine |
| | Erasmus MC, Rotterdam |
| | |
| 2005-2009 | Bsc. Life Science, passed with distinction |
| | Major: Biochemistry |
| | Minor: Medical microbiology and Bioprocess technology |
| | HAN university of applied science, Nijmegen |

---

o        **Professional training**

September  2011 – Now
*PhD project: Department of reproduction and development, Erasmus MC Rotterdam*
• Understanding the molecular mechanism of X chromosome inactivation
• Supervision of 3 Master students

September 2010 - August 2011
*Internship: Department of reproduction and development, Erasmus MC Rotterdam*
• Identification of novel activators involved in X chromosome inactivation

October 2009 - July 2010
*Internship: Department of cell biology, Erasmus MC Rotterdam*
• Identification of the role of LGI4 in mouse intestinal stem cells and development of the gut

January 2009 - June 2009
*Internship: Food & Biobased Research department, Wageningen UR*
• Investigate hydrogen production of *C. saccharolyticus* cultured in different food waste products
• Involved in an international project: Hyvolution

September 2008 – January 2009
*Internship: HAN Biocentre, Nijmegen*
• Design of a gene expression system for filamentous fungi (*Aspergillus*)
• Supervision of 1 Master student

## Courses/ Skills/ Languages

**Courses**

| | |
|---|---|
| 2014 | Employability outside academia |
| 2014 | Research Based Business: Business development & Technology Transfer |
| 2014 | Biostatistics |
| 2013 | Research Based Business: Opportunities |
| 2012 | Genetics |
| 2011 | Cell and Developmental Biology |
| 2011 | Biochemistry and Biophysics |

**Software skills**

Microsoft Office programs (Word, Excel, PowerPoint)
Adobe (Photoshop, Illustrator)
Ability to work with operating systems Windows and Mac OSX

**Languages**

| | |
|---|---|
| Dutch | Fluent in reading, writing and speaking |
| English | Fluent in reading, writing and speaking |
| Papiamento | Fluent in reading, writing and speaking |
| Spanish | Fair/basic in reading, writing and speaking |

## Interests

**Extracurricular activities**

| | |
|---|---|
| 2013 | Pantomime Committee member for Cluster 15 |
| 2011-2012 | Activity Committee member for Cluster 15 |

**Other activities**

Interior design, fitness and creative journaling

# List of Publications

**Maduro C.**, de Hoon, B., and Gribnau, J. "Fitting the Puzzle Pieces: the Bigger Picture of XCI". *Trends in Biochemical Sciences 41, 138-147*

**Maduro C.M.**, Rentmeester E., Gontan C., Mulugeta E., Sleddens-Linkels E., Gribnau J., "Activation of XCI: A role for ubiquitination", *Manuscript in preparation*

Loos F., **Maduro C.***, Loda A.*, Lehmann J., Kremers G., ten Berge D., Grootegoed J.A., Gribnau J., "Xist and Tsix 1 transcription dynamics is regulated by the X-to -autosome ratio and semi-stable transcriptional states'. *Mol Cell Biol. 2016 Nov 1; 36(21): 2656–2667*

Dupont C., **Maduro C.M.**, den Braanker H., Boers R.G., Kurek D., Gribnau J., "Chromatin modifications associated with inactive X-chromosome in murine early cell lineages reflect developmental potential". *Under (minor) revision PLOS ONE*

**Maduro C.M**, Rijsdijk G.*, Vossenkamp C.*, Rentmeester E., Gribnau J., "Megabase deletions of the X chromosome: Finding a needle in an haystack". *Work in progress*

## PhD Portfolio Summary
### Summary of PhD training and teaching activities

| | |
|---|---|
| Name PhD student: Cheryl Mandy Maduro<br>Erasmus MC Department: Developmental Biology<br>Research School: Biomedical Sciences | PhD period: 2011-2016<br>Promotor: Prof.dr. J. Gribnau<br>Supervisor: Prof.dr. J. Gribnau |

| 1. PhD training | Year |
|---|---|
| **Specific courses** | |
| Biochemistry and Biophysics | 2011 |
| Cell and Developmental Biology | 2011 |
| Genetics | 2012 |
| Literature course | 2012 |
| Epigenetic regulation in health and disease | 2012 |
| Biostatistical Methods I: Basic | 2013 |
| Safely working in the lab | 2014 |
| **Presentations** | |
| Developmental Biology/Biochemistry/Clinical Genetics work discussions | 2011-2016 |
| Winter school "Chromatin Changes in Differentiation and Malignancies", Kleinwalsertal, Austria | 2011 |
| Joint XCI retreats (with Heard laboratory) | 2012, 2014 |
| **(Inter)national conferences** | |
| Winterschool "Chromatin Changes in Differentiation and Malignancies", Kleinwalsertal, Austria | 2011 |
| 3rd X-inactivation meeting: 50 years of X-inactivation , Oxford, UK (poster) | 2011 |
| 5th International Society of Differentiation Conference, Stem Cells, Development and Regulation (Amsterdam) | 2012 |
| "Chromatin changes in differentiation and malignancies", Egmond aan zee , Netherlands | 2013 |
| **Seminars and workshops** | |
| The 21th MGC Symposium | 2011 |
| 19th MGC PhD Workshop in Dusseldorf | 2012 |
| Joint XCI retreat (with Heard laboratory), Greece Spetses | 2013 |
| The 20th MGC-PhD workshop in Luxembourg (poster) | 2013 |
| Joint Dutch Chromatin Meeting & NVBMB Fall Symposium | 2014<br>2014 |
| The 24th MGC Symposium | 2014 |
| Joint XCI retreat (with Heard laboratory), Rotterdam | 2014 |
| PhD-day 2014: Your PhD Profile for Success! | 2014 |
| Where your PhD can take you: Career opportunities | |

| Other | 2013 |
|---|---|
| Research Based Business: Opportunities | 2014 |
| PCDI course: Employability outside academia | 2014 |
| PCDI course: Plantum Matchmaking Day | 2014 |
| Research Based Business: Technology Transfer | 2014 |
| Research Based Business: Business Development | |
| **2. Teaching activities** | **Year** |
| **Supervising practicals** | |
| Junior Science Program, Erasmus MC, 2 high school students | 2014 |
| Teaching assistant in "VO Vroege embryos klonen en transgenese" | 2015 |

&-2

Lieve Ma en Pa. Bedankt voor al jullie unconditional support en liefde. Zonder jullie was ik niet waar ik nu ben. Ik ben jullie forever greatful. Melissa, my sister and best friend. Jij bent mijn support pilaar geweest door al de hard times en ik kon (en kan) altijd op je rekenen, albeit 1000 miles apart. Michael en Jeff, onze epic adventures en random moments zal ik nooit vergeten en jullie konden me altijd helpen see the brighter side of things, zelfs als jullie het niet eens door hadden. Bedankt voor je hulp met InDesign Jeff ☺. Ik hou van jullie allemaal en kon niet vragen voor een beter gezin♥. Patrick, danki tambe pa bo support tur e añanan sin mester di un explanation for any situation ;) Y hopi danki pa semper pone nos tur sinti na cas serca boso cu bo home cooked meals. Kleine lieve Terrick, zonder dat je het weet kan jij me als geen ander helemaal opvrolijken. Ik ben een trotse tante ☺. Michelle, ik ben heel erg blij met de mozaïek-kat die je voor mij getekend hebt en het maakt mijn thesis compleet. Heel erg bedankt hiervoor en ook dat ik altijd bij jullie (tante Jane, oom Willem, Bianca, Damien, Melvin en lieve Jessie) terecht kan.

Dear Tom, it has been 3 years now since we first met and seems like a lifetime ago. We have been through so much together and experienced our individual dreams come true and now have to make our dream together come true ☺ Thank you for sticking with me through my worst and celebrating with me at my best. I am very proud of what you have achieved since we first met and look forward to a new chapter together in London.

There are too many people I want to thank and much more that I want to say to everyone than I can write down, but thank you all so much from the bottom of my ♥.

A PhD will give you more than in depth knowledge on a subject. When you put in hard work, long hours and endure failing experiments, rejected manuscripts and p values greater than 0.05 with your head held high, a smile on your face and no regrets you will be ready to take on the next challenge in life.

Cheryl