X CHROMOSOME INACTIVATION

Spreading of Silencing

AGNESE LODA

ISBN 978-94-6299-453-9

Cover: A/C, info.ac.architettura@gmail.com

Layout: A/C, info.ac.architettura@gmail.com

Printing: Ridderprint BV, Ridderkerk, The Netherlands

The work performed in this thesis was performed at the Department of Developmental Biology at the Erasmus MC in Rotterdam, The Netherlands.

Printing of this thesis was financially supported by Erasmus University, Rotterdam and Department of Developmental Biology, Erasmus MC, Rotterdam.

X CHROMOSOME INACTIVATION

SPREADING OF SILENCING

X CHROMOSOOM INACTIVATIE

VERSPREIDING VAN HET STILLEGGEN

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 defense shall be held on
Friday 28 October 2016 at 9:30 hrs

by

Agnese Loda
born in Savona, Italy

Erasmus University Rotterdam

DOCTORAL COMMITTEE

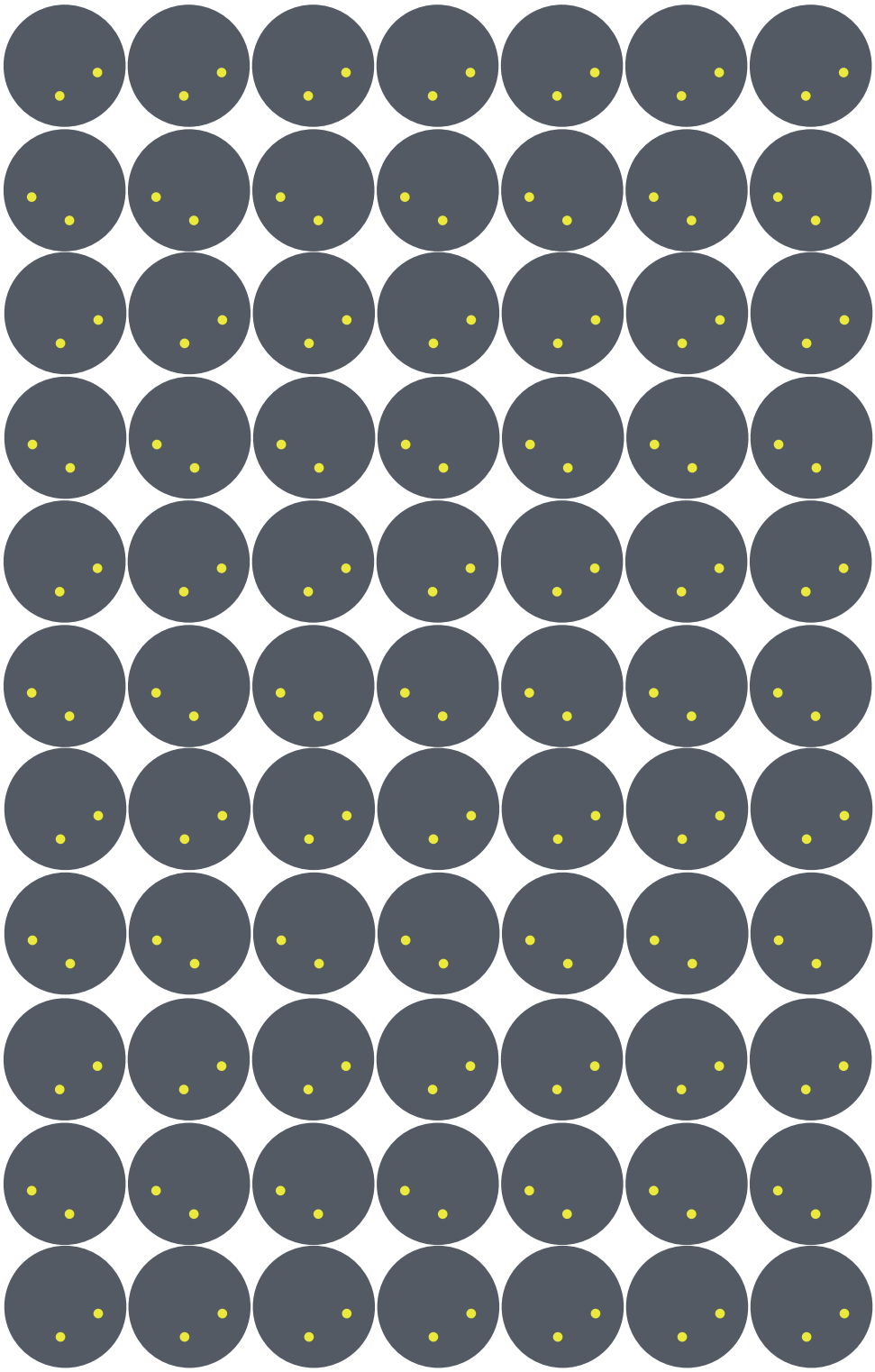
Promotor: Prof.dr. J.H. Gribnau

Other members: Prof.dr. E. Heard
Prof.dr. C.P. Verrijzer
Dr. R.A. Poot

TABLE OF CONTENTS

To my Father,

L'amor che move il sole e l'altre stelle (Alighieri, D. *Paradiso* XXXIII,145)



CHAPTER .1

GENERAL INTRODUCTION

Excerpts of this chapter adapted from:

Agnese Loda, Friedemann Loos, and Joost Gribnau (2015).
X Chromosome Inactivation in Stem Cells and Development.
In Stem Cell Biology and Regenerative Medicine, P. Charbord, and C. Durand, eds. (River Publishers).

SEX CHROMOSOMES AND GENE DOSAGE DIFFERENCES

The origin of sexual reproduction with all its advantageous effects on the fitness of offspring was accompanied by the evolution of complex sex determination systems. Sex determination can be regulated by environmental cues such as temperature, however, these systems are highly sensitive to changes in the natural environment, and do not guarantee an equal distribution of sexes. A more stable distribution of sexes is achieved by genetic sex determination, either by utilizing a specific sex-determining gene or by employing the sex chromosome to autosome ratio. However, the evolution of heteromorphic sex chromosomes resulted in potential gene dosage imbalances between sexes. In these heterogametic species across all taxa diverse gene dosage compensation mechanisms have evolved to counter the detrimental effects of haploinsufficiency of sex chromosome-linked genes (Figure 1) (Disteche, 2012; Livernois et al., 2013a). In mammals, for example, an XY system with a degenerated Y carrying a few functional genes including the single Y-linked sex-determining gene, *SRY*, led to the evolution of X chromosome inactivation (XCI). The fact that in humans only very few trisomies and no monosomies are viable clearly illustrates the potentially devastating effects of improper gene dosage (Disteche, 2012).

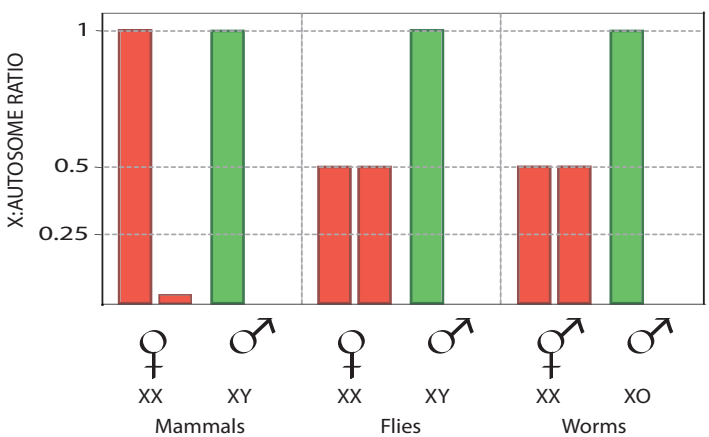


Figure 1. Dosage compensation strategies in different organisms.

In mammals, the loss of Y chromosome genes resulted in upregulation of X-encoded genes. In females, XCI equalizes X-linked gene expression between the sexes. In flies, the X linked genes are twofold overexpressed in males to match the total level of gene expression from the two female X chromosomes. In worms, X-linked gene expression is also upregulated but in hermaphrodite individuals gene expression of both X chromosomes is repressed to half the level observed in male.

1

DOSAGE COMPENSATION IN WORMS, FLIES AND BIRDS

Dosage compensation mechanisms have been investigated in only a couple of different species from relatively few taxa. Given the fact that mechanisms of dosage compensation vary extensively from the molecular to macroscopic level, it would come as no surprise to find new variants as sex chromosome research advances.

In the nematode *Caenorhabditis elegans* males are XO and hermaphrodites are XX. Sex is determined by the X to autosome ratio involving X-chromosomal numerators and autosomal denominators (Carmi et al., 1998; Powell et al., 2005) which lead to repression of the male-specifying gene *xol-1* in hermaphrodites (Miller et al., 1988). To compensate for the loss of one complete X chromosome, X-linked gene expression is up-regulated in male somatic cells (Deng et al., 2011). Although little is known about the mechanisms of X chromosome up-regulation, in XX hermaphrodites this up-regulation would be detrimental, and is counteracted by a mechanism that represses expression of both X chromosomes to approximately half the level observed in males. Chromosome wide down-regulation of gene expression in hermaphrodites involves a ten-protein dosage compensation complex (DCC), containing zinc-finger proteins and condensins which mediate repression of transcription (Meyer, 2010). In the absence of XOL-1 the DCC is activated and specifically targeted to the X chromosomes by a hierarchy of two different types of binding sites. The first type is capable of autonomously recruiting the DCC to the X by specific sequence motifs enriched on the X while the second type is dependent on the first and predominantly targets active promoters (Jans et al., 2009).

In the fruit fly *Drosophila melanogaster* males are heterogametic (XY). The X to autosome ratio determines sex by an intricate interplay of dosage-sensitive, autosomal and X chromosome-linked factors resulting in presence or absence of the sex-determining protein SXL, whose expression triggers female development (Cline, 1983; Salz et al., 1989). Dosage compensation has been studied extensively in this model organism. In males the single X chromosome is up-regulated approximately two-fold by a DCC to equalize X-linked gene expression between males and females and match X-linked gene expression with autosomal gene expression (Conrad and Akhtar, 2011; Gelbart and Kuroda, 2009; Straub et al., 2008). The Drosophila DCC, amongst other proteins, consists of the male-specific core protein MSL2 (Copps et al., 1998), the histone acetyltransferase MOF (Hilfiker et al., 1997), MSL3 and two X-linked non-coding RNAs, roX1 and roX2 (Meller et al., 1997). How exactly up-regulation is achieved remains an open debate. SXL is the key repressor of DCC, and only in male cells where SXL is absent DCC accumulates. The preferential binding of the DCC and accumulation of H4K16ac in gene bodies of active genes (Gilfillan et al., 2006) and global run-on sequencing showing higher RNA Polymerase II (Pol II) occupancy in the 3' region of X-linked genes (Larschan et al., 2011) suggest that enhanced transcriptional elongation causes X-specific up-regulation. Another study found slightly enhanced Pol II activity at X-linked promoters and concluded that transcription initiation is the most important determinant of up-regulation (Conrad et

al., 2012). In Drosophila, the DCC is initially recruited to the X by so-called high affinity sites (HAS), which are approx. 2-fold enriched on the X and show a weak GA-rich sequence motif (Alekseyenko et al., 2008; Straub et al., 2008). From HAS the DCC spreads into adjacent chromatin predominantly targeting active genes (Larschan et al., 2007), possibly by MSL3 binding to H3K36me3 (Sural et al., 2008).

In birds females are heterogametic (ZW) whereas males carry two Z chromosomes. Notably, although superficially similar, ZW chromosomes of birds and reptiles, and XY chromosomes in mammals are non-homologous chromosomes. In birds, *DMRT1*, a Z-linked gene, which is present in two copies in males, is involved in sex determination by dosage dependent initiation of testes development, while the W-linked genes *FET1* and *ASW* are necessary for female development (Nanda et al., 2002; 1999). Dosage compensation in birds appears to be incomplete. Examination of the Z to autosome ratio in ZW females revealed ratios between 0.6 and 0.8 arguing for partial but incomplete up-regulation of Z-linked genes (Wolf and Bryk, 2011). Z-linked gene expression ratios between male and female range between 1.2 and 1.6 (Ellegren et al., 2007; Itoh et al., 2007) and this ratio differs along the Z chromosome (Melamed and Arnold, 2007). Recent evidence suggests that dosage compensation in birds involves a mechanism that leads to inactivation of Z-linked, dosage-sensitive genes, in ZZ males on a gene-to-gene basis (Livernois et al., 2013b).

DOSAGE COMPENSATION BY X CHROMOSOME INACTIVATION

In 1949 Barr and Bertram observed a dense structure in the nucleus of neurons of a female calico cat that was absent in male neurons (Barr and Bertram, 1949), a structure nowadays called Barr body. This observation, together with the description of an X-linked locus conferring a mottled coat color in heterozygous females (Fraser, 1953), the fact that XO females survive and are fertile (Russell et al., 1959) and Ohno's proposal that the Barr body is actually a condensed X chromosome, led Mary Lyon to formulate her XCI theory (Lyon, 1961). She proposed that the heterochromatic Barr body could be randomly established on either the maternal or the paternal X chromosome and was clonally propagated through a near infinite number of cell divisions. Ohno postulated that dosage compensation evolved in two phases: A two-fold up-regulation from the X chromosome to compensate for the loss of one X in males was followed by inactivation of one X in females to account for gene dosage differences between sexes (Ohno, 1967). The first part of his hypothesis still is a matter of intense debate (Deng et al., 2011; Nguyen and Disteche, 2006; Xiong et al., 2010). Different studies report X chromosome to autosome expression ratios anywhere between 0.5 and 1 depending on which filters were used for the data analysis. Comparison of these different studies and methodologies indicates that Ohno's hypothesis holds true for highly expressed genes and a dosage-sensitive subset of genes, e.g. those coding for proteins present in complexes (Pessia et al., 2013; 2012).

XCI has mainly been studied in mice, which show two forms of XCI (Figure 6). The first wave of XCI is initiated at the 4-8-cell stage, and results in the exclusive inactivation of the paternally derived X chromosome (Xp) (Takagi and Sasaki, 1975). This so-called imprinted XCI (iXCI) is maintained in the extra-embryonic lineages. After reactivation of the paternal X chromosome in the inner cell mass (ICM) of the blastocyst (Mak et al., 2004; Okamoto et al., 2004), which gives rise to the embryo proper, a second wave of XCI takes place just after implantation at E5.5. XCI in the embryo is random XCI (rXCI) with respect to the parental X chromosome that is inactivated. After XCI is completed the inactive X chromosome (Xi) is stably transmitted to daughter cells. The consequence of this rXCI process is clearly visible in the calico cat, where the orange and black fur color are determined by an X-linked locus, but is also observed in several mouse strains with heterozygous X-linked marker genes (Hadjantonakis, 2001; Nesbitt and Gartler, 1971).

Even though most insights into XCI have come forth from studies in mice, a growing amount of data from other eutherian species suggest that the regulation and details of XCI might differ substantially between these species. Only rats and cattle have been found to initiate iXCI in the extra-embryonic lineages (Wake et al., 1976; Xue et al., 2002). In human, rabbit, monkey and horse iXCI has not been observed. Instead, the extra-embryonic tissues and embryo both display rXCI (Okamoto et al., 2011; Tachibana et al., 2012; Wang et al., 2012). In addition, the timing of XCI differs between species, with most species initiating XCI later during development than observed in mice.

Marsupials diverged from placental mammals about 148 million years ago, but share ancestral sex chromosomes which originated before the split. Similar to eutherians, marsupials have evolved XCI, although the marsupial form is less stable and more tissue-specific than in eutherians (Kaslow et al., 1987). In addition, XCI is imprinted – in embryonic and extra-embryonic tissues (Graves, 2006; Sharman, 1971). These findings have led to speculations of iXCI being the ancestral form of XCI. However, recent work implicating non-homologous molecular players in the process of XCI indicate that different dosage compensation mechanisms might have evolved independently in meta- and eutherians (Grant et al., 2012).

STEM CELLS AS A MODEL FOR XCI

Initiation of XCI is closely linked to loss of pluripotency and several pluripotent stem cell lines have been used to study XCI. Mouse embryonic carcinoma (EC) cells represented the first *in vitro* model for XCI (Martin GR, 1978). These cell lines, derived from female teratocarcinomas, can be clonally expanded and have two active X chromosomes, one of which is inactivated upon differentiation *in vitro*. However, a final game-changing discovery for the field of XCI was the derivation of embryonic stem (ES) cells from early mouse embryos (Evans and Kaufman, 1981; Martin, 1981), which resemble the pre-XCI cells in the developing embryo much better than the EC cells obtained from tumors. ES cells are derived from the ICM of a female blasto-

cyst, and capture the moment just after reactivation of the paternal X chromosome, containing two active X chromosomes, one of which is inactivated upon differentiation in a random fashion (Rastan and Robertson, 1985). ES cells thus constitute a perfect *ex vivo* model for the dissection of the molecular mechanisms underlying XCI (rXCI in particular) and much of the current knowledge of XCI has come forth from studies using ES cells.

THE X INACTIVATION CENTER

Soon after the initial X chromosome hypothesis had been established two key concepts of XCI were formulated. One concerned the number of X chromosomes being inactivated. Studies in humans with abnormal numbers of X chromosomes showed that all but one X chromosome were condensed and thus inactivated (Ferguson-Smith and Johnston, 1960; Fraccaro et al., 1960; Grumbach et al., 1963; Lyon, 1962) suggesting that cells were able to "count" the number of X chromosomes. These observations were further substantiated by experiments using female tetraploid mouse embryos which inactivate two X chromosomes (Copps et al., 1998; Takagi, 1993; Webb et al., 1992), indicating that each cell inactivates all X chromosomes except one per diploid genome. The finding of a Barr body in Klinefelter XXY patients also indicated that the regulation of sex-determination and dosage compensation involve separate mechanisms, contrasting the mechanisms driving sex-determination and dosage compensation in *C. elegans and Drosophila.*

A second concept concerned the choice of the X chromosome to inactivate, given the fact that during rXCI one of two identical X chromosomes is "chosen" to become inactivated. Initially, an X-linked locus involved in skewing of rXCI from an expected ratio of 50:50 was found by comparing the extent of position effect variegation in an X-to-autosome translocation model (Cattanach and Isaacson, 1967). Subsequently, crossings of inbred mouse strains showed that X chromosomes from certain strains are more resistant to XCI than X chromosomes from other strains, leading to the description of the X chromosome controlling element *(Xce)* as the element responsible for the observed skewing of XCI (Cattanach and Papworth, 1981; Cattanach et al., 1972). Since then, several studies have been trying to precisely define the extent of this locus on the X chromosome, with candidate regions ranging from 0.2Mb to 3.5Mb in length (Chadwick et al., 2006; Simmler et al., 1993; Thorvaldsen et al., 2012; Calaway et al., 2013). However, the mechanism(s) by which *Xce* mediates XCI choice remains elusive. Several other modifiers of choice have been identified, e.g. parent-of-origin effects (Chadwick and Willard, 2005) and autosomal loci (Percec et al., 2002), but similarly to the *Xce* element none of these factors' mode of action is understood and how they influence the outcome of XCI choice on a molecular level remains to be elucidated.

Based on studies of balanced X to autosome translocations showing that XCI could only spread into one of the two autosomal segments (Russell, 1963; Russell and Cacheiro, 1978), a single region on the X, overlapping with the *Xce*, has been proposed to control both counting

and choice. This region, termed X-inactivation center (*Xic* in mouse, *XIC* in human), would be responsible for the initiation of the silencing signal that would then spread over the entire X chromosome (Lyon et al., 1964; Russell, 1963). Using murine embryonic stem cells with truncated X chromosomes and somatic cells carrying a reciprocal X to autosome translocation, a 8cM (10-20Mb) region on the X chromosome, delineated by the T16H and the HD3 breakpoints, was shown to harbor the *Xic* (Figure 2) (Rastan and Robertson, 1985, Rastan, 1983, Eicher et al., 1972). As cells carrying only one *Xic* were never able to initiate XCI, these studies also demonstrated that two *Xic's* are necessary for XCI initiation and hinted at a mechanism of *trans*-communication involved in XCI. The human *XIC* was defined by a similar approach studying rearranged human X chromosomes, and revealed a region in Xq13 spanning ca. 1 Mb that is indispensable for XCI (Brown et al., 1991b).

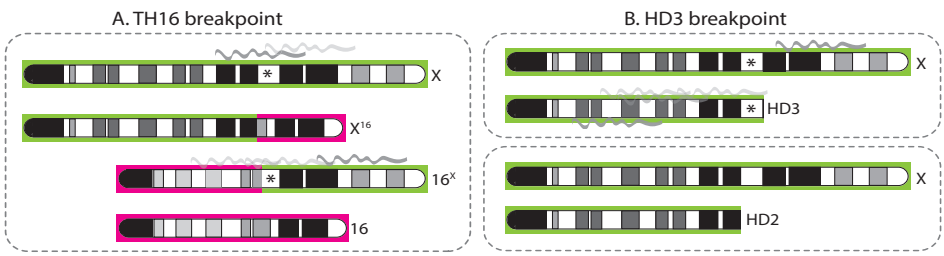


Figure 2. Mapping of the *Xic* on the mouse X chromosome.

Chromosomes that are competent to become inactivated *in cis* are indicated by asterixs.
(A) Schematic representation of the T(X;16)16H Searle's reciprocal transclocation. Although balanced Searle's translocation results in skewed XCI towards the wild type X chromosome, in unbalanced carrier embryos only the 16^{X} translocation product can be inactivated (Rastan 1983). Since the X^{16} product is also never inactivated in 40(X,X^{16},16,16) embryos, in which its transcriptional silencing would result in restored genetic balance (Tagaki 1980), the T16H breakpoints was suggested to define one of the *Xic* physical boundaries (Rastan 1983). (B) Schematic representation of two of the six different-sized deletions of the distal part of the X chromosomes that were studied to further map the *Xic* on the X chromosome (Rastan 1985). Upon differentiation of ES cells carrying the HD3 deletion either the wild type or the deleted X chromosome can be inactivated, whereas the HD2 deletion is never associated with XCI. Thus, the distal boundary of the *Xic* was mapped between HD3 and H2D breakpoints. Beside localizing the *Xic* on the X chromosome, these studies also showed that two *Xics* are necessary to trigger XCI. Indeed, unbalanced 40(X,X^{16},16,16) embryos and H2D differentiang ES cells never showed XCI.

Extensive genetic studies revealed a gene, *Xist/XIST* (X-inactive specific transcript in mouse/human), located within the *Xic/XIC*, which is exclusively expressed from the Xi (Borsani et al., 1991; Brockdorff et al., 1991; Brown et al., 1991a). Surprisingly, *Xist* encodes for a 15-17 kb long, spliced and poly-adenylated nuclear RNA lacking any conserved or significantly long ORFs (Brockdorff et al., 1992; Brown et al., 1992; Clemson et al., 1996). Together with the H19 RNA (Brannan et al., 1990), Xist constituted one of the first long non-coding RNAs to be

discovered. Even though overall low levels of *Xist* and *Xic* conservation suggest evolutionary constraints other than mere DNA sequence (Chureau et al., 2002; Nesterova et al., 2001a), the *Xist* gene in particular shows a similar structure and conserved repeats, while the *Xic* in general harbors several conserved genes (Brockdorff et al., 1992; Brown et al., 1992). The findings above and studies showing *Xist* up-regulation just before the onset of XCI during early mouse embryonic development (Kay et al., 1993) clearly implicated *Xist* in the process of XCI and made it a prime candidate for the *Xic*. Final proof for requirement of *Xist* for XCI to occur *in cis* came forth from targeted mutagenesis of *Xist* in mouse ES cells and mice resulting in primary non-random XCI of the wild type X chromosome (Marahrens et al., 1997; Penny et al., 1996). Paternal inheritance of the mutated *Xist* allele in mice is embryonic lethal due to loss of iXCI, and the incapability to activate the maternally inherited *Xist* allele. Thus, *Xist* is necessary for iXCI and rXCI to initiate XCI and establish the Xi, but it does not recapitulate all aspects of the Xic, because the rXCI counting process is not hampered by this mutation.

CIS- REGULATION OF XCI

Studies of an ESC line carrying a 65 kb deletion downstream of *Xist* indicated that this mutation invariably leads to inactivation of the mutated allele (Clerc and Avner, 1998), suggesting the presence of an element with a *cis*-repressive function on XCI. This element harbors the promoter of a second long non-coding gene, *Tsix*, which is transcribed antisense to *Xist*. *Tsix* spans 40kb and completely overlaps with the *Xist* transcriptional unit and its promoter. It is highly expressed in undifferentiated ES cells and becomes down-regulated upon differentiation (Lee et al., 1999b). Deletion of the *Tsix* promoter, or the *DXPas34 Tsix* regulatory element (Debrand et al., 1999), does not result in aberrant counting, but leads to up-regulation of *Xist in cis* and almost exclusive inactivation of the mutated allele *in vitro* and *in vivo*, supporting the idea that *Tsix* is repressing *Xist in cis*. However, the same *Tsix* promoter deletion did not result in aberrant XCI in undifferentiated cells, highlighting the involvement of additional developmentally regulated factors in the process of XCI (Lee et al., 1999a). Since maternal inheritance of *Tsix* mutations lead to early embryonic lethality supposedly derived from the erroneous inactivation of the two X chromosomes, *Tsix* has been also proposed to play a role in iXCI by preventing the maternal *Xist* allele to be up-regulated (Lee, 2000; Sado et al., 2001). However, *Tsix*-mediated repression of the maternal *Xist* allele has been recently reported to be dispensable during early embryonic development (Maclary et al., 2014). *TSIX*, the human homologue, appears to have a similar expression pattern as mouse *Tsix*, but its truncation raises questions about its role in XCI (Migeon et al., 2001). Both long non-coding RNAs, *Xist* and *Tsix*, are thus main players with opposing effects on the outcome of XCI initiation. How *Tsix* represses *Xist* remains an open question, however, several mechanisms have been proposed. *Tsix* may act through a transcriptional interference mechanism, supported by experiments showing that *Xist* repression requires *Tsix* transcription through the *Xist* promoter, but may also involve RNA

mediated recruitment of chromatin remodeling complexes (Shibata and Lee, 2004). Several studies indeed demonstrated that lack of *Tsix* antisense transcription compromises the establishment of repressive chromatin marks at the *Xist* promoter (Navarro et al., 2006; Ohhata et al., 2007; Sado et al., 2005; Sun et al., 2006). The RNA interference pathway has also been proposed to play a role in *Tsix*-mediated *Xist* regulation (Ogawa et al., 2008), but the reported effects in *Dicer* mutants appear indirect and mediated through DNA-hypomethylation (Nesterova et al., 2008). Several non-coding genes have been implicated in the *cis*-regulation of *Xist* and *Tsix* (Figure 3). *Tsx* and *Xite*, located proximal to *Tsix*, are positive regulators of *Tsix* (Anguera et al., 2011; Ogawa and Lee, 2003). Deletion of *Xite* or *Tsx* down-regulates *Tsix* expression and results in skewing toward XCI of the mutated allele, resembling the phenotype associated with *Tsix* mutants (Anguera et al., 2011; Ogawa and Lee, 2003; Stavropoulos et al., 2005). In contrast, deletion of *Jpx* and *Ftx*, which are located proximal to *Xist*, negatively affects *Xist* expression (Chureau et al., 2011; Sun et al., 2013; Tian et al., 2010).

Interestingly, chromatin conformation capture studies examining the higher order chromatin structure of the *Xic* indicates that *Tsix*, and its positive regulators *Xite* and *Tsx* reside in the same topological associated domain (TAD) which is flanking a distinct TAD that includes *Xist*, *Jpx* and *Ftx* (Nora et al., 2012; Spencer et al., 2011). The TAD-based partitioning of the genome has been observed both in mouse and human and is maintained across different cell types (Dixon et al., 2012). Importantly, several lines of evidence suggest that the high levels of chromatin interactions occurring within different TADs might play a role in mediating transcriptional regulation (Le Dily et al., 2014; Nora et al., 2012; Shen et al., 2012; Smallwood and Ren, 2013). These findings suggest a co-regulation mechanism for genes embedded within the same TAD through a yet unknown mechanism, and indicate that the *Xist* and *Tsix* TADs likely represent the maximum *cis*-regulatory region involved in XCI (Nora et al., 2012). In addition, structural fluctuations of chromatin conformation within the *Tsix* TAD have been proposed to ensure the monoallelic transcription of *Tsix* from one of the two alleles, which in turn leads to monoallelic *Xist* up-regulation and XCI initiation (Giorgetti et al., 2014).

INITIATION OF XCI

Several models have been proposed to explain female specific initiation of XCI. The blocking factor model relies on the action of a trans-acting blocking factor (BF) encoded from an autosome that binds a counting element on the X, thus preventing XCI to occur *in cis*. Every X chromosome within a nucleus could be potentially protected from XCI, however, since BF is expressed at a level that is just enough to protect one X chromosome from inactivation per diploid genome, all the extra Xs would undergo XCI by default (Rastan, 1983; Rastan and Robertson, 1985). To explain the finding that female ES cells carrying heterozygotous *Xist* mutations always inactivate the wild type allele instead of blocking XCI in half of the cell population as the model would predict, the two factor model was postulated (Lee, 2005; Lee

et al., 1999a; Marahrens et al., 1998). In addition to an autosomal BF, in this model the action of an X-linked competence factor (CF) is essential to initiate XCI. The X-linked CF initiates XCI by titrating away the BF only when the X chromosome to autosome ratio (X:A) is either 1 or higher (Lee, 2005; Sun et al., 2013). So far there is no evidence for the presence of a counting element, through which the CF and BF would exert there activity, and even removal of a 500kb region, which includes all known *cis*-acting elements in the *Xic*, does not affect XCI counting (Barakat et al., 2014).

The alternate states model relies on the hypothesis that the two X chromosomes in female ES cells are epigenetically different prior to XCI (Mlynarczyk-Evans et al., 2006). Although differences in sister chromatid cohesion between the two homologous X chromosomes seems to regulate the alternate states, the epigenetic marks that affect the choice of which X will be inactivated is not yet clear. Also transient *trans*-interaction between the two homologous X chromosomes has been proposed to regulate counting and choice. The *Xpr* region, together with *Tsix* and *Xite*, facilitate X-pairing of the two *Xics* at the onset of XCI (Augui et al., 2007; Xu et al., 2006). Interestingly, removal of all elements involved in X-pairing from one X chromosome in female cells does not affect counting, and analysis of XCI in heterokaryons obtained through fusion of a male and female cell does not reveal a preference for initiation of XCI in the female nucleus (Barakat et al., 2014). These findings indicate that X-pairing is likely the consequence of the XCI process, reflecting changes in gene activity leading to spatial movements.

All these models imply that the XCI process is a deterministic and mutually exclusive process, characterized by the exact number of X chromosomes always being inactivated in female cells. Interestingly, analysis of XCI in cells with a different X-autosome ratio revealed a direct relationship between this ratio and increased robustness of the XCI process, indicating the presence of X-linked XCI-activators driving the probability to initiate XCI (Monkhorst et al., 2008). According to this stochastic model, the probability of each X to be inactivated depends on the action of both the X-linked XCI-activators and the autosomally encoded XCI-inhibitors. The double dosage of X-linked activators in female cells is sufficient to generate a specific probability for *Xist* to be up-regulated, whereas the XCI inhibitors are involved in setting up a threshold that has to be overcome by *Xist* to accumulate. Because the activators are X linked, spreading of Xist will down-regulate the XCI activators *in cis*, and this feedback mechanism will prevent the inactivation of the second X chromosome. In male cells the levels of the XCI-activators will not be sufficient to overcome the threshold for XCI to initiate, and *Xist* will not be up-regulated. This model explains many of the experimental data obtained to date and recently several XCI-inhibitors and –activators have been described supporting this model.

1

TRANS- REGULATION OF XCI

Female specific initiation of XCI involves the tight orchestration of expression of both XCI-activators and -inhibitors during embryonic development. Importantly, the presence of two active X chromosomes within the same nucleus has been reported to stabilize the pluripotency network of female ES cells, whereas induction of ectopic XCI leads to down-regulation of pluripotency factors and triggers faster cell differentiation (Schulz et al., 2014). Thus, XCI has been proposed to work as a developmental check point upon embryonic development: if dosage compensation has not properly taken place, female development is delayed (Schulz et al., 2014). In line with this observation, the key pluripotency factors NANOG, OCT4, KLF4, REX1, SOX2, PRDM14 and the reprogramming factor C-MYC have been reported to act as negative regulators of XCI, thus highlighting the strong link between loss of pluripotency and initiation of XCI (Figure 3) (Donohoe et al., 2009; Ma et al., 2011; Navarro et al., 2008; 2010). NANOG, OCT4, SOX2 and PRDM14 bind to the first intron of *Xist* and have been shown to directly repress *Xist* in undifferentiated female and male ES cells (Ma et al., 2011; Navarro et al., 2008). However, female ES cell lines and mice carrying a heterozygous deletion of this region do not show an overt XCI phenotype, suggesting the presence of redundant mechanisms in the repression of XCI (Barakat et al., 2011; Minkovsky et al., 2013; Nesterova et al., 2011). Accordingly, several of the same pluripotency factors inhibit *Xist* expression indirectly, either by promoting *Tsix* up-regulation or by repressing the XCI activators. Indeed, OCT4 has been proposed to regulate *Tsix* expression through regulating *Xite* and in cooperation with CTCF by binding to the *Tsix* regulatory *DXPas34* element in ES cells (Donohoe et al., 2009), and a similar mechanism has been reported for REX1, KLF4 and C-MYC (Navarro et al., 2010). OCT4, SOX2, NANOG and PRDM14 have been reported to also act as negative regulators of *Rnf12*, the important *trans*-activator of *Xist* (Navarro et al., 2011; Payer et al., 2013). RNF12-mediated regulation of *Xist* involves the degradation of REX1, which acts as an inhibitor of *Xist* by binding and repressing *Xist* regulatory sequences but also through binding *Tsix* regulatory sequences involved in the activation of *Tsix* (Gontan et al., 2012). Finally, the MOF-containing MSL and NSL complexes, originally described in flies dosage compensation as the male-specific lethal and non-specific lethal complexes (see above), have also been proposed to play a role in XCI. The MSL complex has been shown to directly enhance *Tsix* expression whereas NSL proteins seem to mediate the maintenance of the pluripotency network (Chelmicki et al., 2014).

XCI activators act directly by up-regulating *Xist* or indirectly by repressing *Tsix* or suppressing the XCI inhibitors. The E3 ubiquitin ligase RNF12/RLIM is a key X-linked trans-acting activator of XCI (Jonkers et al., 2009). *Rnf12* is located 500 kb upstream of *Xist*, and its over-expression

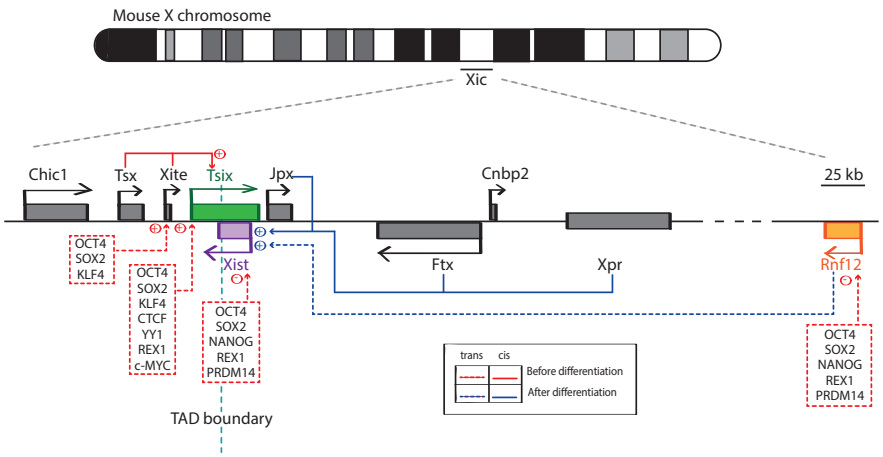


Figure 3. XCI regulatory network.

Schematic overview of the X inactivation centre *(Xic)*, its location on the X and XCI regulators. Dashed lines indicate *-trans* acting factors whereas *-cis* acting factors are indicated by continous lines. In undifferentiated ES cells XCI is repressed by the action of the XCI inhibitors repressing *Xist* and activating *Tsix* expression. Upon dierentiation, loss of XCI inhibitors concominant with the combined action of XCI activators triggers *Xist* upregulation.

triggers initiation of XCI on the single X chromosome in male cells and on both X chromosomes in female cells upon differentiation (Jonkers et al., 2009). Since RNF12 is X-encoded, the dose-dependent degradation of the autosomally encoded REX1 ensures female specific initiation of XCI (Gontan et al., 2012). Since *Rnf12*$^{-/-}$ ES cells completely fail to upregulate *Xist* upon differentiation, *Rnf12* has been suggested to be essential for the initiation of XCI (Barakat et al., 2011). However, a different *Rnf12* null allele has been associated with a milder XCI phenotype (Shin et al., 2010). One possible explanation for these contradictory observations is that the two knockout strategies result in different versions of the truncated protein, although differences in the genetic background of mouse ES cells are more likely to explain these results. *In vivo*, female embryos that maternally inherited an *Rnf12* null allele fail to initiate iXCI on the paternal X chromosome and die in utero, whereas paternal transmission of *Rnf12* mutations does not result in embryonic lethality (Shin et al., 2010). To bypass the embryonic lethality associated with *Rnf12* mutations thus being able to address the role of *Rnf12* in rXCI, a conditional system has been reported in which *Rnf12* depletion is exclusively achieved in the ICM of the developing embryo (Shin et al., 2014). In these *Rnf12*$^{-/-}$ female embryos, XCI has been reported to be unaffected, suggesting that RNF12 is dispensable for rXCI *in vivo* (Shin et al., 2014). Clearly, further studies are needed to better understand the observed discrepancies between *in vivo* and *in vitro* studies. Finally, heterozygous female ES cells carrying a single copy of *Rnf12* can still initiate XCI, although at a lower efficiency compared to wild type cells (Barakat et al., 2011). Since these *Rnf12*$^{+/-}$ female ES cells technically

have the same dosage of X-linked RNF12 as male ES cells, the ability of these cells to initiate XCI suggests the existence of one or more additional X-encoded XCI activators (Barakat et al., 2011). The zinc finger protein YY1, a transcription factor with both activator and repressor functions, has also been proposed to be an indispensable *trans*-acting activators of *Xist* (Makhlouf et al., 2014). YY1 competes with the XCI repressor REX1 for binding to a regulatory element of *Xist* located a few kb downstream of *Xist* TSS, and YY1 depletion upon ES cells differentiation results in impaired *Xist* upregulation (Makhlouf et al., 2014). However, since YY1 is encoded from an autosomal gene and the relative protein levels within female and male cells are supposedly similar, how YY1 would be able to mediate female-specific initiation of XCI is difficult to explain. *Ftx* and *Jpx*, both located within the ***Xic***, have been shown to act as positive regulators of XCI (Augui et al., 2007; Bacher et al., 2006; Chureau et al., 2011; Sun et al., 2010; Tian et al., 2010). *Ftx* and *Jpx* encode long non-coding RNA's, and deletion of both genes results in down-regulation of *Xist* (Chureau et al., 2011; Tian et al., 2010). *Jpx* has been proposed to act *in trans* by antagonizing CTCF-mediated *Xist* repression (Sun et al., 2013). Interestingly, deletion of a region including *Xist, Tsix, Jpx, Ftx* and *Xpr* does not result in loss of XCI on the wild type X chromosome, and further investigation of the function of the region encompassing *Jpx*, *Ftx* and *Xpr* indicates that this region acts ***in cis*** and is involved in the activation of *Xist* expression (Barakat et al., 2014). *Jpx* and *Ftx* therefore do not act as dose dependent *trans*-acting XCI-activators but are part of the ***cis***-regulatory region involved in the regulation of *Xist*. Further studies are therefore required to reveal additional XCI-activators.

CHROMOSOME-WIDE INACTIVATION OF THE X CHROMOSOME

The discoveries that *Xist* is essential for XCI (Marahrens et al., 1997; Penny et al., 1996) and that the Xist RNA "paints" the entire X chromosome (Clemson et al., 1996) spawned two intertwined key questions which still remain unresolved. How would a long non-coding RNA be able to "spread" along and coat an entire chromosome and how would this coating lead to actual silencing of that chromosome?

Xist's RNA *in cis* spreading

Xist RNA only spreads *in cis* on the chromosome from which it is transcribed and never diffuses *in trans* to other chromosomes within the same nucleus (Jonkers et al., 2008). Although human XIST RNA was observed to detach from the Xi during mitosis (Clemson et al., 1996; Brown et al., 1992), and live-cell imaging in mouse ES cells suggested that spreading might involve displacement and reappearance of Xist RNA before and after mitosis (Ng et al., 2011), detection of *Xist* by fluorescence in situ hybridization (FISH) confirmed that the retention of Xist RNA at the Xi territory is stable throughout the cell cycle (Jonkers et al., 2008; Duthie et al., 1999). Nonetheless, how Xist remains bound to the X chromosome is a mystery, and specific sequences that are both necessary and sufficient to allow Xist spreading have not yet

been identified.

One fascinating question that has been extensively addressed is whether Xist spreading relies on X-linked specific *cis*-acting elements. Limited spreading of Xist RNA into autosomal DNA was initially observed by variegation of color coat markers in X to autosomes translocation studies (Russell, 1963; Cattanach, 1961; Cattanach and Perez, 1970; Russell and Montgomery, 1970). For example, Cattanach's translocation Is(X;7)1ct cuts the mouse X chromosome into two segments separated by an insertion of a region of chromosome 7 and results in efficient silencing of X-linked color coat markers on both sites of the insertion, whereas the autosomal *albino (c)* locus does not seem to be affected by Xist mediated silencing (Cattanach, 1961; 1970; 1975). These observations led Gartler and Riggs to propose that spreading relies on X-linked specific "way stations" distributed along the X chromosome (Gartler and Riggs, 1983). Indeed, preferential spreading of Xist into X-linked DNA has been reported by several studies in which Xist RNA localization and gene inactivation were further assessed in both mouse and human cells carrying X to autosome rearrangements (Hall et al., 2002b; Duthie et al., 1999; Keohane et al., 1999; Sharp et al., 2002; Bala Tannan et al., 2014; Cotton et al., 2014). Since all these studies were performed in somatic cells, discriminating between inefficient Xist spreading upon early development from failure in maintenance of autosomal silencing was not feasible. However, attenuated spreading of Xist RNA into autosomal DNA at the onset of XCI was confirmed upon differentiation of mouse ES cells carrying a X;4 balanced translocation (Popova et al., 2006). In line with this sequence specific model for Xist spreading, Mary Lyon proposed long interspersed elements (LINE or L1) to work as the Gartler and Riggs "way stations" in conferring X-chromosome specificity (Lyon, 1998). Lyon's repeat hypothesis is supported by several lines of evidence including enrichment of LINE elements on both mouse and human X chromosomes compared to autosomes, higher density of LINE elements in proximity of the Xic and negative correlation between LINE enrichment and genes that escape XCI (Bailey et al., 2000; Boyle et al., 1990; Ross et al., 2005). Moreover, LINE elements facilitate the formation of a silent Xi compartment into which genes are displaced upon silencing (see below) (Chaumeil et al., 2006), and preferential spreading of Xist RNA in LINE-rich regions of autosomes have been observed in many X to autosome translocations both in mouse and human cells (Popova et al., 2006; Cotton et al., 2014; Bala Tannan et al., 2014). However, several studies in which ***Xist***-containing yeast artificial chromosomes (YAC) transgenes were integrated on different autosomes showed that Xist RNA ***per se*** is able to spread into and silence autosomal material (Heard et al., 1999; Lee and Jaenisch, 1997; Lee et al., 1996; 1999a), and similar approaches that exploited inducible ***Xist*** transgenes to study ***Xist's*** function confirmed Xist-mediated silencing of autosomal DNA (Tang et al., 2010; Wutz and Jaenisch, 2000; Chow et al., 2010; Minks et al., 2013; Ben-Nun et al., 2011; Hall et al., 2002a; Jiang et al., 2013; Chow et al., 2007). In addition, XCI occurs in some rodent species that lack any LINE activity (Cantrell et al., 2009; 2009) and inefficient autosomal silencing might be

rather attributed to selective disadvantages of cells inactivating autosomal genes than the intrinsic inability to silence those genes due to low density of LINE elements. Therefore, these data definitely rule out an absolute requirement of X chromosome-specific DNA elements for Xist spreading, but do not exclude that Xist RNA might preferentially silence X-linked rather than autosomal genes.

Further light has been shed on Xist spreading by a series of recent studies investigating chromatin states on the Xi by next generation sequencing techniques and novel pull-down assays. One study probed the Xi chromatin landscape in trophoblast stem cells and found an Xi-specific signature of DNase hypersensitive sites and H3K27me3 enrichment over transcription start sites of inactivated genes. No tight correlation between the position of an inactive or escaping gene within or outside of the Xist domain was observed (Calabrese et al., 2012b), suggesting that the mechanism of silencing might rather target specifically active genes than being the result of a general chromosome-wide exclusion from the transcription machinery. Another study inferring from chromatin states indirectly to binding and spreading of Xist observed that from initial ~150 Ezh2-binding sites PRC2 spreads into adjacent chromatin on a local scale (Pinter et al., 2012). These findings were supported by directly assessing Xist binding using oligo-based pull-down assays (Engreitz et al., 2013; Simon et al., 2013). The general emerging picture appears to be a two-step model in which Xist is first targeted to active genes or associated regulatory elements and subsequently spreads into adjacent intergenic chromatin, a model reminiscent of the spreading of the DCC in Drosophila (Maenner et al., 2012). Notably, no specific DNA sequence motifs were found in either study and LINE density is anti-correlated with Xist enrichment, contrary to earlier hypotheses mentioned above. Instead, the 3D conformation of the X chromosome via proximity of the *Xist* transcription locus (Engreitz et al., 2013) and the activity of a given gene (Calabrese et al., 2012a) are thought to promote initial Xist binding. Finally, the zinc finger protein YY1 has been implicated in tethering Xist RNA to the *Xist* locus by its ability to bind both Xist RNA and DNA via different domains. On the Xi, YY1 has been shown to bind to a trio of DNA sites in proximity of *Xist* Repeat F whereas the binding between YY1 and Xist RNA is mediated by the Xist repeat C (see below). Since YY1 knockdown resulted in dispersed Xist cloud formation, YY1 has been proposed to mediate the establishment of a "nucleation center" for Xist spreading (Jeon and Lee, 2011). However, inducible *Xist* transgenes lacking the YY1 binding sites manage to spread *in cis* with the same efficiency of wild type *Xist* (Wutz et al., 2002), and the *-trans* diffusion of Xist RNA from its transcription locus observed by Jeon and colleagues is most likely the consequence of massive Xist over-expression from multi-copy transgenes introduced into terminally differentiated cells.

Early epigenetic events on the Xi

Efforts to characterize the chromatin landscape of the Xi yielded a long list of chromatin fea-

tures which are specific for the Xi (Nora and Heard, 2010) and are good candidates for epigenetic transmission of the inactive state, since many of these features, remain on metaphase chromosomes during cell division (Chaumeil et al., 2002; Jeppesen and Turner, 1993; Jonkers et al., 2008; Mak et al., 2002). The earliest epigenetic event following Xist RNA spreading is the formation of a transcriptionally silent compartment consisting mainly of silent repeats from which most hallmarks of active transcription such as RNA Polymerase II and general transcription factors are excluded (Chaumeil et al., 2006). Several "active" histone modifications such as H3K4me2/3, H3K36 methylation, and H3/H4 acetylation (Chaumeil et al., 2002; 2006; Heard et al., 2001; Jeppesen and Turner, 1993; Keohane et al., 1996; O'Neill et al.) are excluded from the Xi, as well (Figure 5). Shortly after the loss of these active features, the Xi acquires many chromatin marks that are associated with silent chromatin (Chadwick and Willard, 2003; Heard et al., 2001; Peters et al., 2002), most notably H3K27me3 and H2AK119u1 whose deposition is catalyzed by the polycomb repressive complexes PRC2 and PRC1, respectively (de Napoles et al., 2007; Mak et al., 2002; Plath et al., 2003; Silva et al., 2003; Wang et al., 2001; Fang et al., 2004) (Figure 5). PRC2 and PRC1 are only transiently enriched on the Xi, at later stages of differentiation H3K27me3 and H2AK119u1 are maintained on the Xi without obvious enrichment of these complexes (Figure 5).

Polycomb complexes recruitment to the Xi

A first link between Xist localization and PRC2 recruitment was found when their co-localization was observed on Xi metaphase chromosomes (Mak et al., 2002). Subsequently, enrichment of PRC2 and H3K27me3 at the Xi has been shown to overlap with the initiation phase of XCI in both differentiating ES cells and upon *Xist* cDNA transgene induction (Plath et al., 2003; Silva et al., 2003; Kohlmaier et al., 2004). Similarly, lack of Xist RNA in fully differentiated cells results in loss of H3K27me3 enrichment at the Xi (Zhang et al., 2007; Kohlmaier et al., 2004), and high-resolution mapping of Xist binding sites along the X was shown to be consistent with the distribution of PRC2 core components obtained by allele-specific ChIP-seq analysis of differentiating mouse ES cells (Engreitz et al., 2013; Simon et al., 2013; Pinter et al., 2012). In addition, studies using truncated *Xist* transgenes showed that different domains of Xist confer different functions (Figure 4). A conserved 5' element, repeat A, is required for proper silencing (Hoki et al., 2009; Wutz et al., 2002), while several other domains, most prominently repeat C, seem to act co-operatively and/or redundantly in coating (Beletskii et al., 2001; Wutz et al., 2002; Hasegawa et al., 2010; Sarma et al., 2010). Additionally, deletion of repeat D in human *XIST* impairs its expression (Lv et al., 2016) and repeat F on *Xist* DNA has been proposed to function as a "nucleation center" for Xist RNA binding (Jeon and Lee, 2011) (see above). These observations not only provided direct evidence that Xist RNA's coating and silencing

Region	Function	Interacting factors	References
A	silencing	SPEN, WTAP EZH2, ATRX via the independent short RNA RepA	Wutz 2002, Zhang 2008, Sarma 2014, Chu 2015, Monfort 2015
B&F	PRC2 recruitment	JARID2	da Rocha 2014
C	localization	hnRNP U/SAF-A, YY1	Hasegawa 2010, Jeon 2011 Beletskii 2001, Sarma 2010
D&E	localization?		Senner 2011, Wutz 2002, Nesterova 2001
*	PRC2/PRC1 recruitment	hnRNP K	Chu 2015
**	localization	hnRNP U/SAF-A	Yamada 2014

Figure 4. Xist RNA functional domains.

Schematic representation of the *Xist* gene. Grey boxes indicate *Xist* tandem repeats. Asterisks indicates regions of *Xist* showed to interact with ***trans***-acting factors involved in mediating *Xist* function.

functions are uncoupled, but also suggested that Xist RNA might exploit its several domains to recruit multiple *trans* acting factors involved in gene silencing.

Biochemical analysis have suggested that a short RNA containing the A repeat region of the Xist transcript (RepA) directly recruits PRC2 to the Xi by binding EZH2, the catalytic subunit of PRC2 that mediates histone H3 lysine 27 methylation (Kaneko et al., 2010; Maenner et al., 2010; Zhao et al., 2008) (Figure 5). Moreover, the chromatin remodeler ATRX has been shown to enhance RepA/PRC2 interactions by directly binding to RepA RNA, whereas ATRX knock down results in defective Xist spreading and PRC2 recruitment (Sarma et al., 2014). It should be noted, however, that Xist RNA deleted for repeat A is still able to induce PRC2 and H3K27me3 enrichment on the Xi (Kohlmaier et al., 2004; Plath et al., 2003), thus suggesting that PRC2 recruitment cannot be exclusively mediated by RepA. Indeed, PRC2 cofactor *Jarid2* has been shown to transiently interact with Xist repeats B and F upon XCI initiation, and loss of *Jarid2* prevents efficient PRC2 and H3K27me3 enrichment to Xist-coated chromatin (Da Rocha et al., 2014). However, high-resolution microscopy data indicated significant spatial separation between Xist RNA and PRC2 core components, strongly arguing against a direct-interaction model for Xist-mediated PRC2 recruitment (Cerase et al., 2014). In line with these findings, none of the proteomic studies aimed to characterize Xist's interactome identified any of the PRC2 complex core components (see below). Thus, PCR2 recruitment to the Xi remains a

controversial issue, and further studies are definitely necessary to fit these apparently contradictory results into a common picture.

With the same kinetics observed for PRC2 and H3K27me3, PRC1 core components become enriched to the Xi upon XCI initiation (de Napoles et al., 2004; Plath et al., 2004; Schoeftner et al., 2006). Since recruitment of PRC1 to target sites has been proposed to be dependent on interactions between the chromo-domain CBX proteins and H3K27me3 (van Kruijsbergen et al., 2015; Fischle et al., 2003; Min et al., 2003; Wang et al., 2004), PRC1 localization to the Xi might rely on Xist-mediated recruitment of PRC2 itself. However, the PRC1 core component *Ring1B*, responsible for the catalysis of H2AK119Ub1, is efficiently recruited to the Xi upon differentiation of mouse ES cells lacking PRC2 (Schoeftner et al., 2006). Moreover, the existence of a PRC2-independent pathway for PRC1 recruitment has subsequently been confirmed by several studies (Terranova et al., 2008; Tavares et al., 2012; Bernstein et al., 2006). Therefore, PRC1 localization on the Xi most likely relies on two separate mechanisms: the CBX-PRC1 recruitment, which is linked to PRC2-mediated accumulation of H3K27me3, and the RYBP-PRC1 recruitment, which occurs in absence of PRC2 (Figure 5). Supporting this hypothesis, both RYBP and RING1B have been recently reported to directly bind to the Xist RNA (Chu et al., 2015).

Developmental regulation of Xist-mediated silencing

The developmental context constitutes a major component of *Xist's* capacity to induce silencing. The use of inducible *Xist* transgenes has shown that Xist is only able to trigger silencing in undifferentiated mouse ES cells and during early differentiation (Wutz and Jaenisch, 2000), and that Xist-dependent induction of H3K27me3 follows the same pattern (Kohlmaier et al., 2004). In addition, ectopic *Xist* expression leads to reversible silencing in undifferentiated mouse ES cells, while deletion of endogenous ***Xist*** (Csankovszki et al., 1999) or repression of an inducible *Xist* transgene after XCI has occurred (Wutz and Jaenisch, 2000) does not lead to reactivation of the Xi. Thus, maintenance of XCI is developmentally strictly regulated and appears to be independent of Xist RNA. Interestingly, an exception from ***Xist***-independent maintenance for iXCI is found in extraembryonic lineages (Ohhata et al., 2011), yet again emphasizing the plasticity and heterogeneity of XCI. To date, the only somatic cells in which ectopic ***Xist*** expression has been reported to be competent in inducing gene silencing is cancer cells derived from a lymphoma model (Agrelo et al., 2009). In this context, the chromatin organizer SATB1 has been proposed to be both necessary and sufficient for Xist-mediated silencing (Agrelo et al., 2009). Indeed, depletion of SATB1 in mouse ES cells results in impaired gene silencing and SATB1 overexpression in fully differentiated cells enhance Xist's silencing capacity (Agrelo et al., 2009). Moreover, SATB1 is expressed early during mouse ES cell differentiation, in a time window that perfectly matches the one in which XCI starts, and might help in relaying chromatin changes to actual silencing (Cai et al., 2003). However, female double

knockout mice for SATB1 and its closely related protein SATB2 are viable, and *Satb1*$^{-/-}$ and *Satb2*$^{-/-}$ null fibroblasts do not show impaired XCI, thus proving that SATB1 is certainly dispensable for XCI to occur (Nechanitzky et al., 2012).

Late epigenetic changes on the Xi

During the maintenance phase of XCI, CpG methylation at promoters of genes and other regulatory sequences, such as CpG islands, is acquired on the Xi (Lock et al., 1987; Brockdorff et al., 1991), suggesting that DNA methylation terminally locks silencing in place (Sado et al., 2000; Riggs and Xiong, 2004). The critical role of DNA methylation in maintaining gene silencing is supported by several lines of evidence. Deletion of the maintenance DNA methyltransferase *Dnmt1* is associated with partial Xi reactivation (Csankovszki et al., 2001), and instability of silencing has been associated with *de novo* mutations of the DNA methyltransferase DNMT3B in human (Hansen et al., 2000). In addition, depletion of histone H3K9 methyltransferase SETDB1 results in reactivation of X-linked reporter genes in somatic cells, and SETDB1 has been shown to belong to a pathway that couples DNA methylation and H3K9me3 deposition on the Xi (Yamada et al., 2015; Minkovsky et al., 2014; Keniry et al., 2016). Interestingly, the SmcHD1 protein appears to be necessary for the maintenance of DNA methylation on the Xi (Blewitt et al., 2008). This protein contains a SMC hinge domain normally found in core components of cohesion complexes, involved in *C. elegans* dosage compensation.

Enrichment of macroH2A histone variant at the Xi has also been associated with XCI maintenance in both human and mouse cells (Chadwick et al., 2001; Costanzi et al., 2000; Mietton et al., 2009) (Figure 5). MacroH2A has transcriptional repression activity (Doyen et al., 2006; Perche et al., 2000), and its recruitment to the Xi relies on Xist RNA (Wutz and Jaenisch, 2000), independently of its silencing function (Pullirsch et al., 2010). Since a high concentration of macroH2A has been observed during S phase, its enrichment on the Xi has been proposed to play a role in ensuring proper replication of the Xi (Chadwick and Willard, 2002). However, deletion of macroH2A does not result in re-activation of the Xi, thus questioning the role of this chromatin mark in XCI maintenance (Changolkar et al., 2007). Recently, depletion of the cohesin complex protein RAD21 or the CCCTC-binding factor CTCF have been associated with extensive Xi reactivation in somatic cells, thus proposing these factors mediating chromosome looping to play a role in the maintenance of XCI (Minajigi et al., 2015). However, gene reactivation from the inactivate Xi could be achieved exclusively by combining short hairpin RNAs (shRNAs) targeting RAD21 and CTCF with different epigenetic drugs previously reported to interfere with gene expression (Singh et al., 2013). Therefore, whether these factors are indeed necessary to maintain XCI remains under debate.

Xist RNA interactome and *trans*-acting factors involved in gene silencing

The recent discovery of several novel XCI-mediating factors led to significant advances in our understanding of the process (Mira-Bontenbal and Gribnau, 2016; Moindrot and Brockdorff, 2016; Cerase et al., 2015). Importantly, many of the proteins that have resulted to functionally contribute to *Xist's* function(s) have been independently identified by either proteomic or genetic strategies (Chu et al., 2015; McHugh et al., 2015; Minajigi et al., 2015; Monfort et al., 2015; Moindrot et al., 2015).

The proteomic approaches were based on paraformaldehyde (Chu et al., 2015) or UV light crosslinking (McHugh et al., 2015; Minajigi et al., 2015), followed by pull-down of Xist RNA together with its associated proteins. The Xist RNA interactome has been captured at different stages of XCI by performing screenings in several cell lines, including undifferentiated ES cells upon induction of transgenic Xist RNA, epiblast stem cells (EpiSC), trophoblast stem cells (TSCs) and fully differentiated somatic cells.

On the other hand, the two genetic approaches made use of mouse ES cells carrying *Xist* inducible transgenes inserted either on the single X chromosome of haploid ES cells (Monfort et al., 2015) or *in cis* close to an autosomal GFP reporter gene (Moindrot et al., 2015). The reporter ES cell line was used to screen with a lentiviral short harpin RNA (shRNA) library to identify factors whose depletion resulted in increased GFP signal upon XCI. Such an increase in GFP signal served as readout for impairment of *Xist's* silencing function. In the haploid system, insertional mutategenesis was exploited to screen for genes that, when mutated, enable ES clones to survive Xist RNA-triggered cell death that follows silencing of the single X chromosome.

The only *trans*-acting factor that was consistently found in all five the above-mentioned studies is the repressive transcriptional factor SPEN (also known as SHARP). SPEN contains four RNA-binding domains and is able to recruit SMRT, a component of the transcriptional co-repressor complex that activates the histone deacetylase HDAC3, thus leading to repression of transcription (Ariyoshi and Schwabe, 2003; Guenther et al., 2001; Shi et al., 2001; Mikami et al., 2013). Knock-down and knock-out experiments extensively validated SPEN's role in XCI by showing abrogation of Xist-mediated silencing upon SPEN depletion (Moindrot et al., 2015; Monfort et al., 2015; Chu et al., 2015; McHugh et al., 2015). Moreover, SPEN is unable to bind Xist RNA lacking repeat A (Chu et al., 2015), the indispensable element for *Xist's* silencing function, and SPEN's RNA binding domains have been shown to interact with the A repeat in vitro (Monfort et al., 2015). Since SPEN depletion also results in reduction of PRC2 and PRC1 on the Xi, SPEN was suggested to play a role in polycomb recruitment as well (McHugh et al., 2015; Monfort et al., 2015). However, H3K27me3 levels were unaffected in one SPEN knock down study (Moindrot et al., 2015), and the enrichment of both H3K27me3 and H2AK119ub to the Xi has been proven to be at least partially independent of *Xist* repeat A (see above) (Kohlmaier et al., 2004; Plath et al., 2003; Schoeftner et al., 2006), confirming that polycomb

recruitment relies on multiple redundant mechanisms (see above). Thus, SPEN mediates transcriptional silencing most likely by recruiting HDAC3 activity via SMRT binding. Indeed, knock down of either SMRT or HDAC3 mimics silencing defects observed in the absence of SPEN (McHugh et al., 2015). Nevertheless, global hypoacetylation of histone H4 on the Xi has been shown to be unaffected upon repeat A deletion (Pullirsch et al., 2010), again highlighting redundancies in the system.

Although identified only in one study, the hnRNP *K* component of the heterogeneous nuclear ribonucleoprotein (hnRNP) complex also seems to play a role in *Xist's* silencing function (Chu et al., 2015). HnRNP K depletion resulted in impaired gene silencing and significantly reduced H3K27me3 and H2AK119ub accumulation on Xi. Moreover, hnRNP K has been shown to bind Xist RNA downstream of the repeat F region, supporting the idea that repeat F might indirectly recruit polycomb complexes to the Xi (Da Rocha et al., 2014; Chu et al., 2015). Both the proteomic and the genetic approaches identified several additional factors, including the SPEN family member RBM15, the core subunit of the m6A RNA methyltransferase complex, WTAP, SWI/SNF chromatin remodeling factors, topoisomerases and cohesins. However, functional studies are needed to validate their functional role in the XCI process.

Nuclear localization of the Xi

The observation that the Xi is frequently found in the nuclear periphery (Rego et al., 2008) and in close proximity to the nucleolus (Zhang et al., 2007), both regions that consist mainly of heterochromatin, led to the proposition that the sub-nuclear localization and organization of the Xi might be involved in initiation and/or maintenance of silencing. Indeed, the Xi seems to form a specialized compartment, evidenced from the fact that removal of DNA does not affect the Xi associated nuclear matrix nor the Xist RNA domain (Smeets et al., 2014; Clemson et al., 1996; Sarma et al., 2010; Hasegawa et al., 2010). Thus, Xist RNA has been proposed to have a structural role in stabilizing the Xi domain within the nucleus, and the matrix attachment protein hnRNP U/SAF-A has been shown to contribute to this function. HnRNP U is enriched on the Xi and is required for Xist RNA localization to the Xi territory (Helbig and Fackelmayer, 2003; Hasegawa et al., 2010; Pullirsch et al., 2010). A direct interaction between Xist RNA and hnRNP U is supported by super-resolution microscopy analysis and seems to be mediated by the *Xist* repeat C element, previously showed to be indispensible for Xist RNA localization (Smeets et al., 2014; Hasegawa et al., 2010; Sarma et al., 2010) (Figure 5). In addition, deletions of *Xist* exon 7 have result in aberrant Xist RNA localization, similar to hnRNP U depletion, and *Xist* exon 7 has been reported to be involved in the Xist RNA/ hnRNP U interaction (Yamada et al., 2015). Interestingly, all the proteomic screenings mentioned above have identified hnRNP U as an Xist-interacting factor (Chu et al., 2015; McHugh et al., 2015; Minajigi et al., 2015), and two independent knock-down studies functionally confirmed its role in Xist RNA localization

(Chu et al., 2015; McHugh et al., 2015). Although the role of hnRNP U in mediating Xist RNA localization has been proven, hnRNP U appeared to be recruited to the Xi relatively late upon

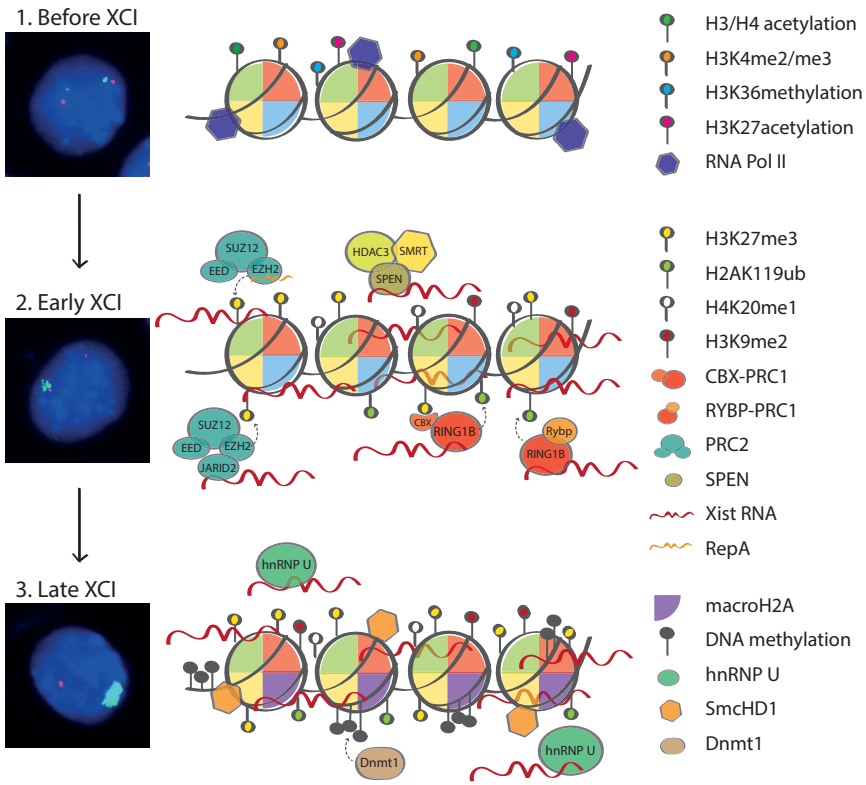


Figure 5. Xist-mediated chromosome-wide gene silencing upon ES differentiation.

(Left) two-color RNA FISH analysis for Xist RNA and the X-linked gene *Mecp2* performed at different time points upon ES cells differentiation. *Xist*, FITC; *Mecp2*, Rhodamin, DNA, DAPI. In undifferentiated ES cells both X chromosomes are active whereas following monoallelic Xist RNA up-regulation upon ES cells differentiation *Mecp2* becomes transcriptionally inactivated. (Middle and Right) an overview of the epigenetic events that occur upon XCI is shown. In undifferentiated ES cells, X-linked chromatin is enriched for euchromatic marks (H3K4me2/3, H3K36 methylation and H3 and H4 acetylation). Later or upon XCI gene silencing is established and chromatin modifying complexes are recruited to the X chromosome, leading to the enrichement of facultative heterochromatin marks (H3K27me3, H2AK199ub, H3K9me2 and H4K20me1). At the latest stage of XCI, macroH2A is incorporated, together with the establishment of DNA methylation and recruitment of hnRNP U to the Xi.

ES cells differentiation, at the time when macroH2A is enriched (see above) (Pullirsch et al., 2010). This observation could be either related to technical limitations in detecting low levels of hnRNP U at the onset of XCI, or might indicate that the general function of hnRNP U in

maintaining the nuclear organization is initially needed for the Xi domain to form, whereas the actual Xist RNA/ hnRNP U binding might occur only at a later stage of XCI.

However, it has to be emphasized that despite the extensive knowledge of chromatin signatures present on the Xi, and the growing list of *trans*-acting factors mediating *Xist's* function, the causal relationship between histone marks, Xist RNA binding proteins, and transcriptional silencing remains largely elusive. A major open question is to undoubtedly understand which factor(s) are necessary for the establishment of transcriptional silencing during XCI. To date, only Xist-deleted female embryos show early embryonic lethality associated with loss of Xist-mediated silencing upon imprinted XCI (Marahrens et al., 1997), whereas several of the above mentioned factors have been shown to be dispensable for silencing establishment: (I) PRC2 knock out embryos (*Eed*$^{-/-}$) are unable to maintain XCI in extra-embryonic tissues but show unaffected initiation of XCI (Wang et al., 2001; Kalantry and Magnuson, 2006), (II) random XCI is not impaired upon differentiation of mouse ES cells carrying homozygous deletions of genes encoding PRC1 components (Leeb and Wutz, 2007), (III) SPEN$^{-/-}$ null mutations result in embryo lethality around E12.5, a time point at which XCI has already taken place (Kuroda et al., 2003), (IV) Xist RNA lacking repeat A is still able to trigger several epigenetic changes including enrichment of H3K27me3 (Kohlmaier et al., 2004), H2AK119ub (Schoeftner et al., 2006) and macroH2A, recruitment of hnRNP U and hypomethylation of histone 4 (Pullrisch et al., 2010). Finally, Xist RNA itself is capable of creating a transcriptionally silent domain independently of gene silencing (Chaumeil et al., 2006), suggesting that the epigenetic features that have been so far described as early XCI events might indeed belong to the maintenance phase of XCI, whereas factors directing the establishment of gene silencing or the mechanism(s) by which Xist RNA itself initiates transcriptional inactivation, still need to be discovered.

ESCAPE FROM XCI

Although XCI leads to chromosome-wide gene silencing of one entire X chromosome in female cells, 12-20% and 3-7% of human and mouse X-linked genes remain transcriptionally active from both the active (Xa) and inactive (Xi) X chromosomes within the same nucleus (Balaton and Brown, 2016). Both in human and in mouse, escape from XCI can be either stable or variable between different tissues, individuals and developmental stages (Schoeftner et al., 2009; Lingenfelter et al., 1998; Yang et al., 2010). Generally, genes escaping XCI are highly enriched at the pseudoautosomal regions (PAR) of the X chromosome. Since PAR regions represent the only region showing homology between chromosomes X and Y, genes belonging to these regions will always be biallelically expressed from the two sex chromosomes in both females and males (Berletch et al., 2011). However, several escape genes are located outside of the PAR regions and are either organized in discrete clusters that reach several Mb in size on the human X chromosome, or are scattered as single genes along the entire length of the silent Xi in mouse (Berletch et al., 2011). Importantly, the higher female-specific expression of

escaping genes that do not retain a functional Y-linked copy is responsible for the phenotypic differences between females and males (Xu and Disteche, 2006). For example, different expression levels of *Jarid1c (Kdm5c)* and *Utx (Kdm6a)* in brains of male and females mice were proposed to explain differences in brain function between the sexes (Xu et al., 2008). Similarly, escaping genes are involved in sex-specific susceptibility to X-linked diseases. Mutations in *Ddx3x* and *Usp9x*, which resist XCI in both human and mouse (Berletch et al., 2015; Carrel and Willard, 2005; Li et al., 2016; Marks et al., 2015), have been associated with intellectual disability in human female carriers (Snijders Blok et al., 2015; Reijnders et al., 2016). *Utx* mutations have been reported to cause Kabuki syndrome, characterized by both developmental delay and intellectual disability (Banka et al., 2015; Miyake et al., 2013), and mutations of *Jarid1c* result in disability and autism (Adegbola et al., 2008; Jensen et al., 2005). Furthermore, hypermutation of the Xi has been shown to be a frequent feature of female cancer genomes (Jäger et al., 2013). If somatic mutations of inactivated genes are unlikely to have an impact on tumorigenesis, mutations of escaping genes may function as cancer driver events as it has been reported for *Ddx3x* mutations in female tumors (Cheng et al., 2015).

Although the functional implications of XCI escape have been extensively studied, the molecular mechanism(s) by which escaping genes are able to resist Xist-mediated silencing is largely unknown. Within the transcriptionally silent inactive X chromosome (Xi), escaping genes retain several euchromatic features such as the active histone marks H3K4 di- and tri-methylation (Goto and Kimura, 2009; Sadreyev et al., 2013) and H3K27- (Kelsey et al., 2015; Cotton et al., 2013), H3K9- acetylation (Goto et al., 2009). Moreover, escapees lack Xist RNA coating (Simon et al., 2013; Engreitz et al., 2013; Murakami et al., 2009) and are depleted of H3K27me3 and macroH2A repressive chromatin marks (Simon et al., 2013, Chalgolkar et al., 2010). One attractive hypothesis is that *cis* acting elements may protect escaping genes from spreading heterochromatin, thus allowing active and inactive domains to co-exist within the Xi. Indeed, BAC transgenes carrying the escaping gene *Jarid1c* and its flanking inactivated genes retain proper XCI regulation when integrated at different loci on the X chromosome that are subjected to XCI (Li and Carrel, 2008). However, truncated versions of the same BAC transgenes lead to inappropriate escaping of neighboring genes up to 350 kb downstream the transgene integration site (Horvath et al., 2013).

These observations suggest that escaping genes are intrinsically able to resist XCI but also that XCI escape is likely driven by dominant elements. Several lines of evidence suggested that the chromatin insulator protein CTCF plays a role in mediating XCI escape. CTCF is enriched at the transition regions between the escaping genes *Jarid1c*, *Eif2s3x* and their neighboring inactivated genes, and allele-specific CTCF binding on the Xi cluster at escaping loci (Filippova et al., 2005; Berletch et al., 2015). However, integration of multiple CTCF binding sites in the vicinity of a X-linked reporter gene is not sufficient to prevent its transcriptional silencing, questioning the role of CTCF alone as boundary element between active and inactive

chromatin regions on the Xi (Ciavatta et al., 2006). On the other hand, escaping genes have been shown to be located outside of the Xist RNA domain (Chaumeil et al., 2006, Chow et al., 2006), suggesting that CTCF might work as an anchor that allows looping out of active escaping domains from the heterochromatic Barr body (Heard and Bickmore, 2007). In this context, the advent of the chromosome conformation capture techniques (the "C" methods), which allow to measure the frequency of physical interactions between any locus of the genome, led to major advances in understanding the nuclear ultrastructure of active and silent genes on the Xi (Dekker, 2014; Dekker and Heard, 2015). Allele-specific 4C revealed that genes on the active X chromosome (Xa) tend to interact with other active regions both *in cis* and *in trans*, whereas on the Xi only genes that escape XCI are engaged in long-range contacts with each other (Splinter et al., 2011). The greater number of specific contacts between escaping genes compared to inactivated genes have been confirmed by allele specific Hi-C (Deng et al., 2015). Similar approaches in both human and mouse revealed that the Xi is devoid of topologically associated domains (TAD) that have been defined along the entire genome (Rao et al., 2014; Dixon et al., 2012; Gibcus and Dekker, 2013, Giorgetti et al., 2016). Rather, the Xi is organized into a specific bipartite structure in which two large superdomains are separated by a boundary region that is located proximal to the *DXZ4/Dxz4* locus both in human and mouse (Minajigi et al., 2015; Deng et al., 2015; Rao et al., 2014; Giorgetti et al., 2016). Since clusters of genes escaping XCI have been reported to overlap with specific TADs defined in ES cells (Marks et al., 2015), one attractive hypothesis is that XCI escape relies on the ability of escaping genes to retain a TAD-like organization within the topologically unorganized Xi (Giorgetti et al., 2016). However, understanding whether the euchromatic environment and the 3D spatial organization of escaping genes enable them to resist XCI or vice versa whether these features are the consequence of their active transcriptional state remains an open fascinating challenge.

X CHROMOSOME REACTIVATION

In mouse embryonic development, the inactive state of Xi is reversed twice, in the ICM of the blastocyst and during specification of female primordial germ cells (PGCs). To become reactivated, the Xi needs to somehow reverse the extremely stable multi-layer of epigenetic features that have ensured its complete silencing. Although X chromosome reactivation (XCR) clearly represents a powerful tool for understanding developmentally regulated epigenetic changes and chromatin dynamics, the small amount of cells that undergo XCR in vivo constitutes an obvious technical limitation. In this context, reprogramming of female mouse somatic cells into ES cell-like induced pluripotent stem cells (iPSCs), which results in reactivation of the Xi, provides a powerful in vitro alternative to study the dynamics of XCR.

X chromosome reactivation in embryonic development

In mouse, the first wave of XCI is initiated between the 4 to 8 cells stage of embryonic development and is subject to imprinting (iXCI), with the paternal X chromosome (Xp) being exclusively inactivated. Understanding whether the imprint is of maternal (Okamoto et al., 2005) or paternal (Namekawa et al., 2010; Sun et al., 2015) origin is still a matter of debate, however, several studies support the existence of a maternal imprint which is located within the *Xist* locus and protects maternally inherited X chromosomes (Xm) from inactivation (Kay et al., 1994; Goto and Takagi, 1998; 2000; Okamoto et al., 2005; Nesterova et al., 2001b). Although the nature of this repressive maternal imprint is not yet known and DNA methylation has been excluded to play a central role (Chiba et al., 2008), enrichment of H3K9me3 at the *Xist* promoter has been recently proposed to ensure proper repression of the maternal *Xist* allele (Fukuda et al., 2014). However, the maternal imprint on Xm appears to be established at a late stage of oocyte development (Tada et al., 2000), whereas the enrichment of H3K9me3 at the *Xist* promoter on Xm has resulted to be unchanged between non-growing and fully-grown oocytes, thus questioning H3K9me3 to be the imprint mark on Xm (Fukuda et al., 2015). iXCI was initially thought to occur in a lineage-specific manner, with Xp being inactivated exclusively in extra embryonic tissues and random XCI occurring in cells originating from the inner cell mass (ICM). However, the observation of Xp reactivation in the ICM of pre-implantation embryos highlights the plasticity of XCI (Mak et al., 2004; Okamoto et al., 2004) (Figure 6). In developing female embryos, Xist RNA initially starts to accumulate on Xp at the 4 cells stage. Subsequently, around day E3.5, every cell of the early stage blastocyst shows a silent Xp domain enriched for Xist RNA and H3K27me3 (Mak et al., 2004). One day later, at the late blastocyst stage, Xist clouds and PRC2 foci start to disappear from cells of the ICM, thus triggering Xp reactivation. Contrarily, iXCI is maintained in cells that will contribute to the extra-embryonic tissues. Even though a few X-linked genes on the Xp have been reported to be reactivated prior to loss of Xist coating and depletion of H3K27me3 (Williams et al., 2011), for most genes Xp reactivation in the ICM is strictly dependent on *Xist* repression. Thus, iXCI strongly differs from random XCI, where *Xist* silencing after establishment of the Xi does not lead to XCR (Csankovszki et al., 1999; Wutz and Jaenisch, 2000). Notably, XCR in the ICM appears to be restricted to NANOG and PRDM14 positive cells suggesting that the establishment of a pluripotent ground state is necessary for XCR to take place. Indeed, NANOG is required for establishment of the naïve pluripotent state of ES cells, whereas PRDM14 is involved in maintenance of this state, by protecting the ICM to differentiate towards extra-embryonic endoderm (Payer et al., 2013). In addition, PRDM14 is indispensable for primordial germ cell (PGC) development, representing the only cell lineage where the inactive state of Xi is reversed for the second time during mouse embryonic development (Yamaji et al., 2008) (Yamaji et al., 2008). Both *Nanog* and *Prdm14* knockout embryos show impaired XCR of the paternal X chromosome, showing persistent H3K27me3 foci at day E4.5 of blastocyst development (Payer et al., 2013; Silva et al.,

2009). However, knockout and transgene studies indicate that binding of these factors to the *Xist* intron 1 region plays a minor role in the regulation of *Xist* and XCR (Minkovsky et al., 2013; Nesterova et al., 2011). Interestingly, ChIP-seq studies reveal binding of PRDM14 and NANOG in *Rnf12* regulatory regions, and *Rnf12* has been shown to be up-regulated in *Prdm14* knockout ESCs, suggesting that PRDM14 and possibly NANOG act on XCR through repression of *Rnf12* (Payer et al., 2013). In addition, Xp specific induction of *Tsix* in the developing blastocyst has also been reported to repress *Xist* expression thus promoting XCR (Ohhata et al., 2011). However, in *Tsix* knockout mice XCR seems to be delayed but is not completely abrogated, indicating that *Tsix*-mediated repression of *Xist* is unlikely to be the sole mechanism involved in Xp reactivation (Payer et al., 2013). Thus, a RNF12 mediated mechanism likely acts in parallel with *Tsix* mediated repression to faithfully initiate XCR in the pre-implantation embryo.

Primordial germ cell (PGC) specification starts at day E6.5 of embryonic development, with a cluster of cells from the post implantation epiblast undergoing major epigenetic changes in order to repress the epiblast somatic program, and to promote re-acquisition of pluripotency and initiate genome-wide DNA de-methylation (Saitou and Yamaji, 2012). By day E7.5, around 40 PGCs expressing *Blimp1*, *Stella* and *Prdm14* start migrating and colonize the genital ridges at E12.5 (Ohinata et al., 2005). In female mouse embryos, epigenetic reprogramming in PGCs is accompanied by reactivation of Xi (Figure 6). *Prdm14* knockout mice fail to develop functional primordial germ cells (PGCs) (Yamaji et al., 2008). In absence of PRDM14, PGCs are initially specified but fail to undergo epigenetic reprogramming thus losing their identity around day E8.5, showing de-regulation of genes indispensable for germ cell specification together with down-regulation of pluripotency factors and up-regulation of the DNA methyltransferases *Dnmt3b/3a* (Grabole et al., 2013). Whereas reactivation of the paternal X chromosome in the ICM of the embryos occurs in 24 hours, XCR in PGCs is a slower process and requires several days (Sugimoto and Abe, 2007). Xist RNA FISH analyses show a heterogeneous pattern of *Xist* expression during PGCs development, with few cells that have lost *Xist* appearing at day E7.0. This number increases during PGC development to reach a complete loss of Xist clouds at day E10.5 (Sugimoto and Abe, 2007). *Tsix* is not expressed at any stage of PGC development (de Napoles et al., 2007; Sugimoto and Abe, 2007) suggesting that *Tsix* is not required for repression of *Xist* and is dispensable for XCR. The Xi specific markers EED and H3K27me3 disappear between day E9.5 and 11.5 of PGCs specification (Chuva de Sousa Lopes et al., 2008; de Napoles et al., 2007), and bi-allelic expression of X-linked genes is first detected around day E10.5 (Sugimoto and Abe, 2007). Interestingly, between days E10.5 and E14.5 several X-linked genes showed mono-allelic expression in absence of *Xist*. Therefore, although a limited number of X-linked genes have been tested, *Xist* repression does not seem to be the only mechanism responsible for X chromosome reactivation in PGCs. This observation contrasts with XCR in the ICM, underscoring the difference in epigenetic states of the

Xi. Discrepancies in silencing reversibility might be associated with the cell lineage. Xist-mediated silencing in the ICM is not yet stabilized by DNA methylation, whereas silencing of Xi associated genes in cells of the post-implantation epiblast that give rise to PGCs may involve

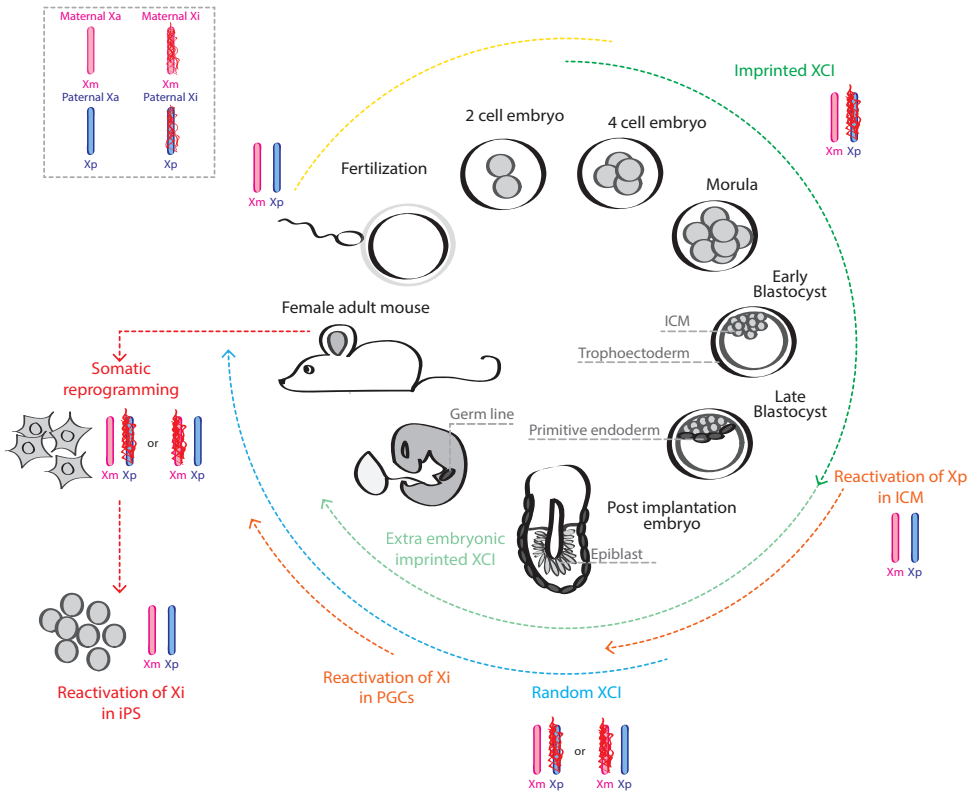


Figure 6. The cycle of life and XCI.

In female mouse embryos, both X chromosomes are active after fertilization. Around the 4 cell stage imprinted XCI is initiated leading to selective inactivation of the paternal X chromosome (Xp). The inactive state of the Xp is maintained in the extra-embryonic tissues. In cells of the ICM the Xp is reactivated and random XCI is subsequently initiated in cells of the post implantation epiblast, leading to the random inactivation of one X chromosome. Random XCI is reversed in PGCs, to allow the establishment of epigenetic instructions required for proper initiation of imprinted XCI. X chromosome reactivation also occurs during somatic reprogramming.

DNA methylation and other histone modifications that fix in the inactive state (Sado et al., 2000). Therefore, the drastic epigenetic reprogramming events that characterize PGC specification, including up-regulation of pluripotency factors, genome wide DNA de-methylation and erasure of histone modifications, are most likely the key factors able to trigger reactiva-

tion of the DNA-methylation dependent inactive state of Xi in PGCs.

***In vitro* X chromosome reactivation: somatic cells reprogramming**

In vitro XCR can be achieved by several methods. For example, mouse embryos generated by somatic cell nuclear transfer (SCNT) show reactivation of an X-linked GFP reporter that was silenced in the donor somatic cell nucleus (Eggan et al., 2000). Also, in hybrid cells obtained by fusion of somatic cells with pluripotent cells such as embryonic carcinoma (EC) and ES cells, the Xi of the somatic cells become reactivated (Tada et al., 2001; Takagi et al., 1983; Ying et al., 2002). These methods both exploited XCR to assess the erasure of the somatic cell's epigenetic memory, but neither of the two has provided an efficient model for studying the dynamics of XCR in vitro. Contrarily, the iPSCs technology enormously contributed to the generation of a robust and controllable system to study XCR *in vitro*. Four genes encoding the transcription factors, KLF4, SOX2, c-MYC and OCT4, appeared to be the minimal necessary requirements to reprogram somatic cells to pluripotent stem cells (Takahashi et al., 2007). Mouse iPSCs are functionally indistinguishable from ES cells and have been reported to efficiently contribute to every tissue of chimeric mice including the germ line (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007), and to give rise to "all iPSCs" mice in a tetraploid complementation assay, the most stringent test known for pluripotency (Boland et al., 2009; Stadtfeld et al., 2010; Zhao et al., 2009). Reprogramming of mouse somatic cells into iPSCs results in XCR accompanied by loss of Xist clouds and H3K27me3 foci, together with re-acquisition of bi-allelic expression of Tsix and other X-linked genes. Moreover, iPSCs undergo random XCI upon differentiation with the same dynamics described for females ES cells (Maherali et al., 2007). Importantly, XCR occurs relatively late during reprogramming, when iPSCs become independent of the exogenous expression of the reprogramming factors (Stadtfeld et al., 2008; Pasque et al., 2014). In line with this observation, *Xist* repression upon reprogramming has been reported to occur exclusively in those cells in which reactivation of the endogenous *Nanog* has taken place (Pasque et al., 2014), thus confirming the tight link between pluripotency and *Xist* regulation (see above). Indeed, *Prdm14* greatly accelerates reprogramming of EpiSCs into iPSCs when expressed in combination with *Klf2*, and is required for self-renewal of iPSCs (Sasaki et al., 2015; Gillich et al., 2012; Irie et al., 2015; Sugawa et al., 2015), whereas *Prdm14* null iPSCs partially fail to down-regulate *Xist* during reprogramming (Payer et al., 2013). REX1 plays a role in XCI initiation as a target of RNF12 (Gontan et al., 2012), and in cells undergoing somatic reprogramming the coinciding up-regulation of *Rex1* and down-regulation of *Rnf12* may indeed trigger XCR.

Interestingly, several XCR associated epigenetic changes on the Xi have been shown to occur in the inverse order of what has been observed upon XCI (Pasque et al., 2014). For example, although the PRC2 complex is supposedly not required for XCR to occur, EZH2 becomes transiently enriched on the Xi, suggesting that PRC2 recruitment to the Xi can only take place

in a specific intermediate stage of somatic reprogramming, which most probably resembles the developmental transcriptional/chromatin state at which PRC2 is enriched on the Xi upon XCI (Pasque et al., 2014). Contrarily, macroH2A and DNA methylation, which are normally established in the mantainance phase of XCI, have been reported to be lost from the Xi at a very late stage of reprogramming, thus suggesting that only some of the Xi's epigenetic features strictly follow the differentiation state of the cells undergoing somatic reprogramming. However, although *in vitro* XCR is a slow process that needs several cell divisions and requires DNA de-methylation, further studies are necessary to understand whether the XCR achieved upon somatic reprogramming can accurately mimic the *in vivo* process. Importanly, new *in vitro* systems that allow germ line derivation from both mouse (Hayashi et al., 2012; Nakaki et al., 2013) and human (Sugawa et al., 2015; Sasaki et al., 2015; Irie et al., 2015) stem cells have been recently described, thus providing additional tools for a better understanding of the XCR process upon PGCs specification.

Since the first derivation of human iPSCs in 2007 (Takahashi et al., 2007; Yu et al., 2007), the possibility of using these cells to study human XCI appeared very promising. Nonetheless, generation of human pluripotent stem cells that faithfully recapitulate XCR and subsequently initiate XCI upon differentiation in vitro turned out to be challenging. In contrast to female mouse ESCs and iPSCs where XCR leading to two Xa's is related to the naïve pluripotent state, this scenario is extremely controversial in human stem cells. *In vitro*, human embryonic stem cells (hESCs) show very heterogeneous patterns of *XIST* expression and X-linked gene silencing (Hall et al., 2008; Shen et al., 2008). Based on their XCI phenotypes, hESCs have been grouped in three different classes (Silva et al., 2008). Class I hESCs resemble mouse ES cells, have two active Xa's, lack *XIST* and H3K27me3 foci and undergo random XCI upon differentiation. Class II hESCs carry one XIST cloud and show X-linked gene silencing in undifferentiated state, whereas in class III hESCs, the silent state of Xi is maintained but XIST is no longer present, along with loss of accumulation of H3K27me3. Human iPSCs (hiPSCs) and hESCs are morphologically similar and re-express the endogenous pluripotency factors NANOG, OCT4 and SOX2 (Takahashi et al., 2007; Yu et al., 2007). Different studies examinating a wide range of hiPSCs indicate that similar to hESCs, hiPSCs show heterogeneous XCI patterns and hIPSCs even tend to lose *XIST* expression after prolonged passaging, a process called erosion of XCI (Mekhoubad et al., 2012). Even though *XIST* loss does not seem to trigger consistent Xi reactivation, further studies at chromosome-wide level are necessary to confirm these results both in hESCs and hiPSCs (Hall et al., 2008; Lengner et al., 2010; Shen et al., 2008). These discrepancies in X chromosome epigenetic features between mouse ESCs/iPSCs and human ESCs/iPSCs may arise from intrinsic differences between mouse and human early development, with human ESCs/iPSCs resembling mouse EpiSCs, which besides carrying one silent X chromosome, share many more molecular features (Hanna et al., 2010; Lagarkova et al., 2006; Nichols and Smith, 2009; Rossant, 2008; Thomson and Marshall, 1998; Vallier et al., 2005). Cells

of the human ICM have been reported to carry two active X chromosomes both coated by XIST RNA *in vivo* (Okamoto et al., 2011). This finding suggests that regulation of XCI is different between human and mouse and that lack of Xi reactivation in hiPSCs might reflect this difference. Nevertheless, recently hESCs and hiPSCs lines with two active X chromosomes have been described, suggesting that lack of XCR might be overcome by changing culture conditions which push the hESCs/iPSCs to a more naïve state (Hanna et al., 2010; Lengner et al., 2010; Silva et al., 2008; Ware et al., 2009). In two studies, naïve hESCs and hiPSCs have been efficiently generated and maintained with combinations of cytokines and small molecule inhibitors independently of the constitutive expression of exogenous factors (Gafni et al., 2013; Hanna et al., 2010). These naïve hESCs/hiPSCs appear to maintain two active X chromosomes and upon differentiation *XIST* expression is up-regulated suggesting that rXCI is initiated. However, since X-linked gene expression has not been assessed, whether the loss of epigenetic features corresponds to robust XCR remains unclear. Importantly, although these studies might provide us with an *in vitro* model to explore XCI in humans, a better understanding of the *in vivo* XCI process in human embryos is necessary to be able to faithfully mimic human XCI initiation and maintenance upon differentiation of pluripotent cells.

AIM OF THIS THESIS

Although considerable advances have been made in XCI research, many questions remain unsolved. The growing list of XCI activators and inhibitors is not complete yet. The mechanisms that direct Xist-mediated gene silencing still need to be resolved, and the cues that ensure the developmental regulation of XCI and its irreversibility upon differentiation await identification. Furthermore, whether the XCI key regulatory factors identified in mouse are conserved across different mammalian species has not yet been addressed, and how escaping genes resist chromosome-wide inactivation also remains a mistery.

In this thesis, we set out to explore different levels of *Xist* regulation and Xist RNA function in a specific developmental context. Our experimental stategies are based on the generation of genetically modified mouse ES cells. First, in **chapter 2**, to study the interplay between *Xist* and *Tsix*, we uncoupled their regulation by replacing both non-coding genes with fluorescent reporter genes. This approach allowed us to follow the dynamics of *Xist* and *Tsix* expression in undifferentiated ES cells and upon ES differentiation, demonstrating a strong antagonistic role between the two non-coding genes. Also, we identified two semi-stable transcriptional states of the *Tsix* allele that might correspond to different higher order chromatin conformations of the Xic and might play a role in directing the initiation of XCI. Second, in **chapter 3**, to explore the mechanisms directing Xist's silencing function, we developed a novel inducible *Xist* expression system in ES cells. By inducing ectopic XCI from X-linked and autosomal loci, we were able to address whether the genomic environment from which Xist RNA starts to spread is instructive in determining the efficiency of gene silencing. Furthermore, we com-

pared the ectopic inactivation of X-linked genes from different loci on chromosome X with endogenous XCI. Also, we explored the role of LINE elements and specific chromatin signatures in facilitating transcriptional silencing of both X-linked and autosomal genes. Finally, we were able to assess whether CTCF binding is implicated in XCI escape. Third, in **chapter 4**, to provide new insights into the specific developmental phase into which XCI takes place, we generated a reporter plasmid *cis*-linked to an inducible antisense promoter. By following reporter expression after induction of antisense transcription at different time points upon ES cells differentiation, we were able to discriminate between the epigenetic features that direct reversible and irreversible antisense-mediated transcriptional silencing. Next, in **chapter 5**, to explore the dynamics of XCI in a model organism different from the mouse, we established a novel *in vitro* differentiation protocol for rat ES cells. In this study, we were able to identify the key culture conditions that allow XCI to occur *in vitro*, and we could follow several XCI-related epigenetic events upon differentiation of rat ES cells. Lastly, in **chapter 6**, I discuss our findings in the light of our current XCI knowledge, and I speculate on how we could potentially address several of the questions that still need to be aswered in the future.

1

REFERENCES

Adegbola, A., Gao, H., Sommer, S., and Browning, M. (2008). A novel mutation in JARID1C/SMCX in a patient with autism spectrum disorder (ASD). American journal of medical genetics. Part A 146A, 505-511.

Agrelo, R., Souabni, A., Novatchkova, M., Haslinger, C., Leeb, M., Komnenovic, V., Kishimoto, H., Gresh, L., Kohwi-Shigematsu, T., Kenner, L., et al. (2009). SATB1 defines the developmental context for gene silencing by Xist in lymphoma and embryonic cells. Dev Cell 16, 507-516.

Alekseyenko, A. A., Peng, S., Larschan, E., Gorchakov, A. A., Lee, O. K., Kharchenko, P., McGrath, S. D., Wang, C. I., Mardis, E. R., Park, P. J., et al. (2008). A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. CELL 134, 599-609.

Anguera, M. C., Ma, W., Clift, D., Namekawa, S., Kelleher, R. J. 3., and Lee, J. T. (2011). Tsx produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. PLoS Genet 7, e1002248.

Ariyoshi, M., and Schwabe, J. W. R. (2003). A conserved structural motif reveals the essential transcriptional repression function of Spen proteins and their role in developmental signaling. Genes & Development 17, 1909-1920.

Augui, S., Filion, G. J., Huart, S., Nora, E., Guggiari, M., Maresca, M., Stewart, A. F., and Heard, E. (2007). Sensing X chromosome pairs before X inactivation via a novel X-pairing region of the Xic. Science 318, 1632-1636.

Bacher, C. P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., and Heard, E. (2006). Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat Cell Biol 8, 293-299.

Bailey, J. A., Carrel, L., Chakravarti, A., and Eichler, E. E. (2000). Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc Natl Acad Sci U S A 97, 6634-6639.

Bala Tannan, N., Brahmachary, M., Garg, P., Borel, C., Alnefaie, R., Watson, C. T., Thomas, N. S., and Sharp, A. J. (2014). DNA methylation profiling in X;autosome translocations supports a role for L1 repeats in the spread of X chromosome inactivation. Human Molecular Genetics 23, 1224-1236.

Balaton, B. P., and Brown, C. J. (2016). Escape Artists of the X Chromosome. 1-12.

Banka, S., Lederer, D., Benoit, V., Jenkins, E., Howard, E., Bunstone, S., Kerr, B., McKee, S., Lloyd, I. C., Shears, D., et al. (2015). Novel KDM6A (UTX) mutations and a clinical and molecular review of the X-linked Kabuki syndrome (KS2). Clinical genetics 87, 252-258.

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

Barakat, T. S., Loos, F., van Staveren, S., Myronova, E., Ghazvini, M., Grootegoed, J. A., and Gribnau, J. (2014). The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. Molecular Cell 53, 965-978.

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

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

Ben-Nun, I. F., Montague, S. C., Houck, M. L., Tran, H. T., Garitaonandia, I., Leonardo, T. R., Wang, Y., Charter, S. J., Laurent, L. C., Ryder, O. A., et al. (2011). Induced pluripotent stem cells from highly endangered species. Nat Methods 8, 829-831.

Berletch, J. B., Ma, W., Yang, F., Shendure, J., Noble, W. S., Disteche, C. M., and Deng, X. (2015). Escape from X Inactivation Varies in Mouse Tissues. PLoS Genet 11, e1005079.

Berletch, J. B., Yang, F., Xu, J., Carrel, L., and Disteche, C. M. (2011). Genes that escape

from X inactivation. Hum Genet 130, 237-245.

Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol Cell Biol 26, 2560-2569.

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

Boland, M. J., Hazen, J. L., Nazor, K. L., Rodriguez, A. R., Gifford, W., Martin, G., Kupriyanov, S., and Baldwin, K. K. (2009). Adult mice generated from induced pluripotent stem cells. Nature 461, 91-94.

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

Boyle, A. L., Ballard, S. G., and Ward, D. C. (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, 7757-7761.

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

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

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

Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R., and Willard, H. F. (1991a). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature 349, 38-44.

Brown, C. J., Lafreniere, R. G., Powers, V. E., Sebastio, G., Ballabio, A., Pettigrew, A. L., Ledbetter, D. H., Levy, E., Craig, I. W., and Willard, H. F. (1991b). Localization of the X inactivation centre on the human X chromosome in Xq13. Nature 349, 82-84.

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

Cai, S., Han, H., and Kohwi-Shigematsu, T. (2003). Tissue-specific nuclear architecture and gene expression regulated by SATB1. Nature Genetics 34, 42-51.

Calabrese, J. M., Sun, W., Song, L., Mugford, J. W., Williams, L., Yee, della, Starmer, J., Mieczkowski, P., Crawford, G. E., and Magnuson, T. (2012a). Site-Specific Silencing of Regulatory Elements as a Mechanism of X Inactivation. CELL 151, 951-963.

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

Cantrell, M. A., Carstens, B. C., and Wichman, H. A. (2009). X chromosome inactivation and Xist evolution in a rodent lacking LINE-1 activity. PLoS One 4, e6252.

Carmi, I., Kopczynski, J. B., and Meyer, B. J. (1998). The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex. Nature 396, 168-173.

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

Cattanach, B. M. (1961). A chemically-induced variegated-type position effect in the mouse. Zeitschrift für Vererbungslehre 92, 165-182.

Cattanach, B. M. (1975). Control of chromosome inactivation. Annu Rev Genet 9, 1-18.

Cattanach, B. M. (1970). Controlling elements in the mouse X-chromosome. 3. Influence upon both parts of an X divided by rearrangement. Genet Res 16, 293-301.

Cattanach, B. M., Wolfe, H. G., and Lyon, M. F. (1972). A comparative study of the coats of chimaeric mice and those of heterozygotes for X-linked genes. Genet Res 19, 213-228.

Cattanach, B. M., and Isaacson, J. H. (1967). Controlling elements in the mouse X chromosome. Genetics 57, 331-346.

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

Cattanach, B. M., and Perez, J. N. (1970). Parental influence on X-autosome translocation-induced variegation in the mouse. Genet Res 15, 43-53.

Cerase, A., Smeets, D., Tang, Y. A., Gdula, M., Kraus, F., Spivakov, M., Moindrot, B., Leleu, M., Tattermusch, A., Demmerle, J., et al. (2014). Spatial separation of Xist RNA and polycomb proteins revealed by superresolution microscopy. Proceedings of the National Academy of Sciences 111, 2235-2240.

Cerase, A., Pintacuda, G., Tattermusch, A., and Avner, P. (2015). Xist localization and function: new insights from multiple levels. Genome Biology 16, 166.

Chadwick, B. P., Valley, C. M., and Willard, H. F. (2001). Histone variant macroH2A contains two distinct macrochromatin domains capable of directing macroH2A to the inactive X chromosome. Nucleic Acids Research 29, 2699-2705.

Chadwick, B. P., and Willard, H. F. (2003). Barring gene expression after XIST: maintaining facultative heterochromatin on the inactive X. Semin Cell Dev Biol 14, 359-367.

Chadwick, B. P., and Willard, H. F. (2002). Cell cycle-dependent localization of macroH2A in chromatin of the inactive X chromosome. The Journal of Cell Biology 157, 1113-1123.

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

Chadwick, L. H., and Willard, H. F. (2005). Genetic and parent-of-origin influences on X chromosome choice in Xce heterozygous mice. Mamm Genome 16, 691-699.

Changolkar, L.N., Singh G, Cui K, Berletch J.B., Zhao K, Disteche C.M., Pehrson J.R. (2010). Genome-wide distribution of macroH2A1 histone variants in mouse liver chromatin. Mol Cell Biol. Dec;30(23):5473-83.

Changolkar, L. N., Costanzi, C., Leu, N. A., Chen, D., McLaughlin, K. J., and Pehrson, J. R. (2007). Developmental changes in histone macroH2A1-mediated gene regulation. Mol Cell Biol 27, 2758-2764.

Chaumeil, J., Le Baccon, P., Wutz, A., and Heard, E. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev 20, 2223-2237.

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

Chelmicki, T., Dündar, F., Turley, M. J., Khanam, T., Aktas, T., Ramírez, F., Gendrel, A., Wright, P. R., Videm, P., Backofen, R., et al. (2014). MOF-associated complexes ensure stem cell identity and Xist repression. eLife 3, e02024.

Cheng, F., Liu, C., Lin, C., Zhao, J., Jia, P., Li, W., and Zhao, Z. (2015). A Gene Gravity Model for the Evolution of Cancer Genomes: A Study of 3,000 Cancer Genomes across 9 Cancer Types. PLoS computational biology 11, e1004497.

Chiba, H., Hirasawa, R., Kaneda, M., Amakawa, Y., Li, E., Sado, T., and Sasaki, H. (2008). De novo DNA methylation independent establishment of maternal imprint on X chromosome in mouse oocytes. Genesis (New York, N.Y. : 2000) 46, 768-774.

Chow, J. C., Hall, L. L., Baldry, S. E. L., Thorogood, N. P., Lawrence, J. B., and Brown, C. J. (2007). Inducible XIST-dependent X-chromosome inactivation in human somatic cells is reversible. Proc Natl Acad Sci U S A 104, 10104-10109.

Chow, J. C., Ciaudo, C., Fazzari, M. J., Mise, N., Servant, N., Glass, J. L., Attreed, M., Avner, P., Wutz, A., Barillot, E., et al. (2010). LINE-1 activity in facultative het-

erochromatin formation during X chromosome inactivation. CELL 141, 956-969.

Chu, C., Zhang, Q. C., Da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., Magnuson, T., Heard, E., and Chang, H. Y. (2015). Systematic discovery of Xist RNA binding proteins. CELL 161, 404-416.

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

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

Chuva de Sousa Lopes, S. M., Hayashi, K., Shovlin, T. C., Mifsud, W., Surani, M. A., and McLaren, A. (2008). X chromosome activity in mouse XX primordial germ cells. PLoS Genet 4, e30.

Ciavatta, D., Kalantry, S., Magnuson, T., and Smithies, O. (2006). A DNA insulator prevents repression of a targeted X-linked transgene but not its random or imprinted X inactivation. Proc Natl Acad Sci U S A 103, 9958-9963.

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

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

Cline, T. W. (1983). The interaction between daughterless and sex-lethal in triploids: a lethal sex-transforming maternal effect linking sex determination and dosage compensation in Drosophila melanogaster. Dev Biol 95, 260-274.

Conrad, T., Cavalli, F. M., Vaquerizas, J. M., Luscombe, N. M., and Akhtar, A. (2012). Drosophila dosage compensation involves enhanced Pol II recruitment to male X-linked promoters. Science 337, 742-746.

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

Copps, K., Richman, R., Lyman, L. M., Chang, K. A., Rampersad-Ammons, J., and Kuroda, M. I. (1998). Complex formation by the Drosophila MSL proteins: role of the MSL2 RING finger in protein complex assembly. EMBO J 17, 5409-5417.

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

Cotton, A. M., Ge, B., Light, N., Adoue, V., Pastinen, T., and Brown, C. J. (2013). Analysis of expressed SNPs identifies variable extents of expression from the human inactive X chromosome. Genome Biology 14, R122.

Cotton, A. M., Chen, C. Y., Lam, L. L., Wasserman, W. W., Kobor, M. S., and Brown, C. J. (2014). Spread of X-chromosome inactivation into autosomal sequences: role for DNA elements, chromatin features and chromosomal domains. Human Molecular Genetics 23, 1211-1223.

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

Csankovszki, G., Nagy, A., and Jaenisch, R. (2001). Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. The Journal of Cell Biology 153, 773-784.

Da Rocha, S. T., Boeva, V., Escamilla-Del-Arenal, M., Ancelin, K., Granier, C., Matias, N. R., Sanulli, S., Chow, J., Schulz, E., Picard, C., et al. (2014). Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. Molecular Cell 53, 301-316.

de Napoles, M., Nesterova, T., and Brockdorff, N. (2007). Early loss of Xist RNA expression and inactive X chromosome associated chromatin modification in developing primordial germ cells. PLoS One 2, e860.

de Napoles, M., Mermoud, J. E., Wakao, R., Tang,

Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M., et al. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev Cell 7, 663-676.

Debrand, E., Chureau, C., Arnaud, D., Avner, P., and Heard, E. (1999). Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression. Mol Cell Biol 19, 8513-8525.

Dekker, J. (2014). Two ways to fold the genome during the cell cycle: insights obtained with chromosome conformation capture. 7, 1-12.

Dekker, J., and Heard, E. (2015). Structural and functional diversity of Topologically Associating Domains. FEBS letters 589, 2877-2884.

Deng, X., Ma, W., Ramani, V., Hill, A., Yang, F., Ay, F., Berletch, J. B., Blau, C. A., Shendure, J., Duan, Z., et al. (2015). Bipartite structure of the inactive mouse X chromosome. Genome Biology 16, 67.

Deng, X., Hiatt, J. B., Nguyen, D. K., Ercan, S., Sturgill, D., Hillier, L. W., Schlesinger, F., Davis, C. A., Reinke, V. J., Gingeras, T. R., et al. (2011). Evidence for compensatory upregulation of expressed X-linked genes in mammals, Caenorhabditis elegans and Drosophila melanogaster. Nature Genetics 43, 1179-1185.

Disteche, C. M. (2012). Dosage compensation of the sex chromosomes. Annu Rev Genet 46, 537-560.

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

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

Doyen, C., An, W., Angelov, D., Bondarenko, V., Mietton, F., Studitsky, V. M., Hamiche, A., Roeder, R. G., Bouvet, P., and Dimitrov, S. (2006). Mechanism of polymerase II transcription repression by the histone variant macroH2A. Mol Cell Biol 26, 1156-1164.

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

Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W. 3., Yanagimachi, R., and Jaenisch, R. (2000). X-Chromosome inactivation in cloned mouse embryos. Science 290, 1578-1581.

Eicher, E. M., Nesbitt, M. N. & Francke, U. (1972). Cytological identification of the chromosomes involved in Searle's translocation and the location of the centromere in the X chromosome of the mouse. Genetics 71, 643–648.

Ellegren, H., Hultin-Rosenberg, L., Brunstrom, B., Dencker, L., Kultima, K., and Scholz, B. (2007). Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. BMC Biol 5, 40.

Engreitz, J. M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E. S., et al. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science 341, 1237973.

Evans, M. J., and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature 292, 154-156.

Fang, J., Chen, T., Chadwick, B., Li, E., and Zhang, Y. (2004). Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. The Journal of biological chemistry 279, 52812-52815.

Ferguson-Smith, M. A., and Johnston, A. W. (1960). Chromosome abnormalities in certain diseases of man. Ann Intern Med 53, 359-371.

Filippova, G. N., Cheng, M. K., Moore, J. M., Truong, J., Hu, Y. J., Di Kim Nguyen, Tsuchiya, K. D., and Disteche, C. M. (2005). Boundaries between Chromosomal Domains of X Inactivation and Escape Bind CTCF and Lack CpG Methylation during Early Development. Developmental Cell 8, 31-42.

Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes & Development 17, 1870-1881.

Fraccaro, M., Kaijser, K., and Lindsten, J. (1960). A child with 49 chromosomes. Lancet 2, 899-902.

Fraser, A. S. S. S. A. S. C. C. (1953). Mottled: a sex-modified lethal in the house mouse. J. Genet. 51, 217-221.

Fukuda, A., Mitani, A., Miyashita, T., Umezawa, A., and Akutsu, H. (2015). Chromatin condensation of Xist genomic loci during oogenesis in mice. Development 142, 4049-4055.

Fukuda, A., Tomikawa, J., Miura, T., Hata, K., Nakabayashi, K., Eggan, K., Akutsu, H., and Umezawa, A. (2014). The role of maternal-specific H3K9me3 modification in establishing imprinted X-chromosome inactivation and embryogenesis in mice. Nature communications 5, 5464.

Gafni, O., Weinberger, L., Mansour, A. A., Manor, Y. S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature 504, 282-286.

Gartler, S. M., and Riggs, A. D. (1983). Mammalian X-chromosome inactivation. Annu Rev Genet 17, 155-190.

Gelbart, M. E., and Kuroda, M. I. (2009). Drosophila dosage compensation: a complex voyage to the X chromosome. Development 136, 1399-1410.

Gibcus, J. H., and Dekker, J. (2013). The Hierarchy of the 3D Genome. Molecular Cell 49, 773-782.

Gilfillan, G. D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P. B. (2006). Chromosome-wide gene-specific targeting of the Drosophila dosage compensation complex. Genes Dev 20, 858-870.

Gillich, A., Bao, S., Grabole, N., Hayashi, K., Trotter, M. W., Pasque, V., Magnusdottir, E., and Surani, M. A. (2012). Epiblast stem cell-based system reveals reprogramming synergy of germline factors. Cell Stem Cell 10, 425-439.

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

Giorgetti, L., Lajoie, B. R., Carter, A. C., Attia, M., Zhan, Y., Xu, J., Chen, C., Kaplan, N., Chang, H. Y., Heard, E., et al. (2016). Structural organization of the inactive X chromosome in the mouse. Nature 535, 575-579.

Gontan, C., Achame, E. M., Demmers, J., Barakat, T. S., Rentmeester, E., van IJcken, W., Grootegoed, J. A., and Gribnau, J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature 485, 386-390.

Goto, Y., and Kimura, H. (2009). Inactive X chromosome-specific histone H3 modifications and CpG hypomethylation flank a chromatin boundary between an X-inactivated and an escape gene. Nucleic Acids Research 37, 7416-7428.

Goto, Y., and Takagi, N. (2000). Maternally inherited X chromosome is not inactivated in mouse blastocysts due to parental imprinting. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology 8, 101-109.

Goto, Y., and Takagi, N. (1998). Tetraploid embryos rescue embryonic lethality caused by an additional maternally inherited X chromosome in the mouse. Development 125, 3353-3363.

Grabole, N., Tischler, J., Hackett, J. A., Kim, S., Tang, F., Leitch, H. G., Magnusdottir, E., and Surani, M. A. (2013). Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. EMBO Rep 14, 629-637.

Grant, J., Mahadevaiah, S. K., Khil, P., Sangrithi, M. N., Royo, H., Duckworth, J., McCarrey, J. R., VandeBerg, J. L., Renfree, M. B., Taylor, W., et al. (2012). Rsx is a metatherian RNA with Xist-like properties in X-chromosome inactivation. Nature 487, 254-258.

Graves, J. A. (2006). Sex chromosome specialization and degeneration in mammals. CELL 124, 901-914.

Grumbach, M. M., Morishima, A., and Taylor, J. H. (1963). Human Sex Chromosome Abnormalities in Relation to DNA Replication and Heterochromatinization. Proc Natl Acad Sci U S A 49, 581-589.

Guenther, M. G., Barak, O., and Lazar, M. A. (2001). The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. Mol Cell Biol 21, 6091-6101.

Hadjantonakis, K. (2001). Green fluorescent tortoiseshell mice. Curr Biol 11, R544.

Hall, L. L., Byron, M., Sakai, K., Carrel, L., Willard, H. F., and Lawrence, J. B. (2002a). An ectopic human XIST gene can induce chromosome inactivation in postdifferentiation human HT-1080 cells. Proc Natl Acad Sci U S A 99, 8677-8682.

Hall, L. L., Clemson, C. M., Byron, M., Wydner, K., and Lawrence, J. B. (2002b). Unbalanced X;autosome translocations provide evidence for sequence specificity in the association of XIST RNA with chromatin. Human Molecular Genetics 11, 3157-3165.

Hall, L. L., Byron, M., Butler, J., Becker, K. A., Nelson, A., Amit, M., Itskovitz-Eldor, J., Stein, J., Stein, G., Ware, C., et al. (2008). X-inactivation reveals epigenetic anomalies in most hESC but identifies sublines that initiate as expected. Journal of Cellular Physiology 216, 445-452.

Hanna, J., Cheng, A. W., Saha, K., Kim, J., Lengner, C. J., Soldner, F., Cassady, J. P., Muffat, J., Carey, B. W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc Natl Acad Sci U S A 107, 9222-9227.

Hansen, R. S., Stöger, R., Wijmenga, C., Stanek, A. M., Canfield, T. K., Luo, P., Matarazzo, M. R., D'Esposito, M., Feil, R., Gimelli, G., et al. (2000). Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant. Human Molecular Genetics 9, 2575-2587.

Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., Tsutui, K., and Nakagawa, S. (2010). The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. Dev Cell 19, 469-476.

Hayashi, K., Ogushi, S., Kurimoto, K., Shimamoto, S., Ohta, H., and Saitou, M. (2012). Offspring from oocytes derived from in vitro primordial germ cell-like cells in mice. Science 338, 971-975.

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

Heard, E., Mongelard, F., Arnaud, D., and Avner, P. (1999). Xist yeast artificial chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies. Mol Cell Biol 19, 3156-3166.

Heard, E., and Bickmore, W. (2007). The ins and outs of gene regulation and chromosome territory organisation. Current Opinion in Cell Biology 19, 311-316.

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, 173-182.

Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A., and Lucchesi, J. C. (1997). mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in Drosophila. EMBO J 16, 2054-2060.

Hoki, Y., Kimura, N., Kanbayashi, M., Amakawa, Y., Ohhata, T., Sasaki, H., and Sado, T. (2009). A proximal conserved repeat in the Xist gene is essential as a genomic element for X-inactivation in mouse. Development 136, 139-146.

Horvath, L. M., Li, N., and Carrel, L. (2013). Deletion of an X-inactivation boundary disrupts adjacent gene silencing. PLoS Genet 9, e1003952.

Irie, N., Weinberger, L., Tang, W. W. C., Kobayashi, T., Viukov, S., Manor, Y. S., Dietmann, S., Hanna, J. H., and Surani, M. A. (2015). SOX17 is a critical specifier of human primordial germ cell fate. CELL 160, 253-268.

Itoh, Y., Melamed, E., Yang, X., Kampf, K., Wang, S., Yehya, N., van Nas, A., Replogle, K., Band, M. R., Clayton, D. F., et al. (2007). Dosage compensation is less effective in birds than in mammals. J Biol 6, 2.

Jans, J., Gladden, J. M., Ralston, E. J., Pickle, C. S., Michel, A. H., Pferdehirt, R. R., Eisen, M. B., and Meyer, B. J. (2009). A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome. Genes Dev 23, 602-618.

Jäger, N., Schlesner, M., Jones, D. T. W., Raffel, S., Mallm, J., Junge, K. M., Weichenhan, D., Bauer, T., Ishaque, N., Kool, M., et al. (2013). Hypermutation of the inactive X chromosome is a frequent event in cancer. CELL 155, 567-581.

Jensen, L. R., Amende, M., Gurok, U., Moser, B., Gimmel, V., Tzschach, A., Janecke, A. R., Tariverdian, G., Chelly, J., Fryns, J., et al. (2005). Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. Am J Hum Genet 76, 227-236.

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

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

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

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

Jonkers, I., Monkhorst, K., Rentmeester, E., Grootegoed, J. A., Grosveld, F., and Gribnau, J. (2008). Xist RNA is confined to the nuclear territory of the silenced X chromosome throughout the cell cycle. Mol Cell Biol 28, 5583-5594.

Kalantry, S., and Magnuson, T. (2006). The Polycomb group protein EED is dispensable for the initiation of random X-chromosome inactivation. PLoS Genet 2, e66.

Kaneko, S., Li, G., Son, J., Xu, C. F., Margueron, R., Neubert, T. A., and Reinberg, D. (2010). Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. Genes Dev 24, 2615-2620.

Kaslow, D. C., Migeon, B. R., Persico, M. G., Zollo, M., VandeBerg, J. L., and Samollow, P. B. (1987). Molecular studies of marsupial X chromosomes reveal limited sequence homology of mammalian X-linked genes. Genomics 1, 19-28.

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

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

Kelsey, A. D., Yang, C., Leung, D., Minks, J., Dixon-McDougall, T., Baldry, S. E. L., Bogutz, A. B., Lefebvre, L., and Brown, C. J. (2015). Impact of flanking chromosomal sequences on localization and silencing by the human non-coding RNA XIST. Genome Biology 16, 208.

Keniry, A., Gearing, L. J., Jansz, N., Liu, J., Holik, A. Z., Hickey, P. F., Kinkel, S. A., Moore, D. L., Breslin, K., Chen, K., et al. (2016). Setdb1-mediated H3K9 methylation is enriched on the inactive X and plays a role in its epigenetic silencing. Epigenetics & Chromatin 9, 16.

Keohane, A. M., Barlow, A. L., Waters, J., Bourn, D., and Turner, B. M. (1999). H4 acetylation, XIST RNA and replication timing are coincident and define x;autosome boundaries in two abnormal X chromosomes. Human Molecular Genetics 8, 377-383.

Keohane, A. M., O'neill, L. P., Belyaev, N. D., Lavender, J. S., and Turner, B. M. (1996). X-Inactivation and histone H4 acetylation in embryonic stem cells. Dev Biol 180, 618-630.

Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., and Wutz, A. (2004). A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. PLoS Biol 2, E171.

Kuroda, K., Han, H., Tani, S., Tanigaki, K., Tun, T., Furukawa, T., Taniguchi, Y., Kurooka, H., Hamada, Y., Toyokuni, S., et al. (2003). Regulation of marginal zone B cell development by MINT, a suppressor of Notch/RBP-J signaling pathway. Immunity 18, 301-312.

Lagarkova, M. A., Volchkov, P. Y., Lyakisheva, A. V., Philonenko, E. S., and Kiselev, S. L. (2006). Diverse epigenetic profile of novel human embryonic stem cell lines. Cell Cycle 5, 416-420.

Larschan, E., Alekseyenko, A. A., Gortchakov, A. A., Peng, S., Li, B., Yang, P., Workman, J. L., Park, P. J., and Kuroda, M. I. (2007). MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. Mol Cell 28, 121-133.

Larschan, E., Bishop, E. P., Kharchenko, P. V., Core, L. J., Lis, J. T., Park, P. J., and Kuroda, M. I. (2011). X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. Nature 471, 115-118.

Le Dily, F., Baù, D., Pohl, A., Vicent, G. P., Serra, F., Soronellas, D., Castellano, G., Wright, R. H. G., Ballare, C., Filion, G., et al. (2014). Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. Genes & Development 28, 2151-2162.

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

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

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

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

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

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

Leeb, M., and Wutz, A. (2007). Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. The Journal of Cell Biology 178, 219-229.

Lengner, C. J., Gimelbrant, A. A., Erwin, J. A., Cheng, A. W., Guenther, M. G., Welstead, G. G., Alagappan, R., Frampton, G. M., Xu, P., Muffat, J., et al. (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. CELL 141, 872-883.

Li, N., and Carrel, L. (2008). Escape from X chromosome inactivation is an intrinsic property of the Jarid1c locus. Proceedings of the National Academy of Sciences 105, 17055-17060.

Li, X., Cui, X., Wang, J., Wang, Y., Li, Y., Wang, L., Wan, H., Li, T., Feng, G., Shuai, L., et al. (2016). Generation and Application of Mouse-Rat Allodiploid Embryonic Stem Cells. 1-32.

Lingenfelter, P. A., Adler, D. A., Poslinski, D., Thomas, S., Elliott, R. W., Chapman, V. M., and Disteche, C. M. (1998). Escape from X inactivation of Smcx is preceded by silencing during mouse development. Nature Genetics 18, 212-213.

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

Lock, L. F., Takagi, N., and Martin, G. R. (1987). Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. CELL 48, 39-46.

Lv, Q., Yuan, L., Song, Y., Sui, T., Li, Z., and Lai, L. (2016). D-repeat in the XIST gene is required for X chromosome inactivation. RNA biology 13, 172-176.

Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature 190, 372-373.

Lyon, M. F. (1962). Sex chromatin and gene action in the mammalian X-chromosome. Am J Hum Genet 14, 135-148.

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

Lyon, M. F., Searle, A. G., Ford, C. E., and Ohno, S. (1964). A Mouse Translocation Suppressing Sex-Linked Variegation. Cytogenetics 3, 306-323.

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

Maclary, E., Buttigieg, E., Hinten, M., Gayen, S., Harris, C., Sarkar, M. K., Purushothaman, S., and Kalantry, S. (2014). Differentiation-dependent requirement of Tsix long non-coding RNA in imprinted X-chromosome inactivation. Nature communications 5, 4209.

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

Maenner, S., Muller, M., and Becker, P. B. (2012). Roles of long, non-coding RNA in chromosome-wide transcription regulation: lessons from two dosage compensation systems. Biochimie 94, 1490-1498.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., et al. (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell 1, 55-70.

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

Mak, W., Nesterova, T. B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A. P., and Brockdorff, N. (2004). Reactivation of the paternal X chromosome in early mouse embryos. Science 303, 666-669.

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

Marahrens, Y., Loring, J., and Jaenisch, R. (1998). Role of the Xist gene in X chromosome choosing. CELL 92, 657-664.

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

Marks, H., Kerstens, H. H. D., Barakat, T. S., Splinter, E., Dirks, R. A. M., van Mierlo, G., Joshi, O., Wang, S., Babak, T., Albers, C. A., et al. (2015). Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. Genome Biology, 1-20.

Martin GR, E. C. T. B. T. G. Y. S. M. D. J. C. S. C. S. (1978). X-chromosome inactivation during differentiation of female teratocarcinoma stem cells in vitro. Nature (5643), 329-33.

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

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

Mekhoubad, S., Bock, C., de Boer, A. S., Kiskinis, E., Meissner, A., and Eggan, K. (2012). Erosion of dosage compensation impacts human iPSC disease modeling. Cell Stem Cell 10, 595-609.

Melamed, E., and Arnold, A. P. (2007). Regional differences in dosage compensation on the chicken Z chromosome. Genome Biol 8, R202.

Meller, V. H., Wu, K. H., Roman, G., Kuroda, M. I., and Davis, R. L. (1997). roX1 RNA paints the X chromosome of male Drosophila and is regulated by the dosage compensation system. CELL 88, 445-457.

Meyer, B. J. (2010). Targeting X chromosomes for repression. Curr Opin Genet Dev 20, 179-189.

Mietton, F., Sengupta, A. K., Molla, A., Picchi, G., Barral, S., Heliot, L., Grange, T., Wutz, A., and Dimitrov, S. (2009). Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome. Mol Cell Biol 29, 150-156.

Migeon, B. R., Chowdhury, A. K., Dunston, J. A., and McIntosh, I. (2001). Identification of TSIX, encoding an RNA antisense to human XIST, reveals differences from its murine counterpart: implications for X inactivation. Am J Hum Genet 69, 951-960.

Mikami, S., Kanaba, T., Ito, Y., and Mishima, M. (2013). NMR assignments of SPOC domain of the human transcriptional corepressor SHARP in complex with a C-terminal SMRT peptide. Biomolecular NMR assignments 7, 267-270.

Miller, L. M., Plenefisch, J. D., Casson, L. P., and Meyer, B. J. (1988). xol-1: a gene that controls the male modes of both sex determination and X chromosome dosage compensation in C. elegans. CELL 55, 167-183.

Min, J., Zhang, Y., and Xu, R. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes & Development 17, 1823-1828.

Minajigi, A., Froberg, J. E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., Lessing, D., Payer, B., Boukhali, M., Haas, W., et al. (2015). Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. Science 349.

Minkovsky, A., Sahakyan, A., Rankin-Gee, E., Bonora, G., Patel, S., and Plath, K. (2014). The Mbd1-Atf7ip-Setdb1 pathway contributes to the maintenance of X chromosome inactivation. Epigenetics & Chromatin 7, 12.

Minkovsky, A., Barakat, T. S., Sellami, N., Chin, M. H., Gunhanlar, N., Gribnau, J., and 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.

Minks, J., Baldry, S. E., Yang, C., Cotton, A. M., and Brown, C. J. (2013). XIST-induced silencing of flanking genes is achieved by additive action of repeat a monomers in human somatic cells. Epigenetics & Chromatin 6, 23.

Mira-Bontenbal, H., and Gribnau, J. (2016). New Xist-Interacting Proteins in X-Chromosome Inactivation. Current Biology 26, R338-R342.

Miyake, N., Mizuno, S., Okamoto, N., Ohashi, H., Shiina, M., Ogata, K., Tsurusaki, Y., Nakashima, M., Saitsu, H., Niikawa, N., et al. (2013). KDM6A point mutations cause Kabuki syndrome. Human mutation 34, 108-110.

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

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

Moindrot, B., and Brockdorff, N. (2016). RNA binding proteins implicated in Xist-mediated chromosome silencing. Semin Cell Dev Biol 56, 58-70.

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

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

Murakami, K., Ohhira, T., Oshiro, E., Qi, D., Oshimura, M., and Kugoh, H. (2009). Identification of the chromatin regions coated by non-coding Xist RNA. Cytogenet Genome Res 125, 19-25.

Nakaki, F., Hayashi, K., Ohta, H., Kurimoto, K., Yabuta, Y., and Saitou, M. (2013). Induction of mouse germ-cell fate by transcription factors in vitro. Nature 501, 222-226.

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

Nanda, I., Shan, Z., Schartl, M., Burt, D. W., Koehler, M., Nothwang, H., Grutzner, F., Paton, I. R., Windsor, D., Dunn, I., et al. (1999). 300 million years of conserved synteny between chicken Z and human chromosome 9. Nature Genetics 21, 258-259.

Nanda, I., Haaf, T., Schartl, M., Schmid, M., and Burt, D. W. (2002). Comparative mapping of Z-orthologous genes in vertebrates: implications for the evolution of avian sex chromosomes. Cytogenet Genome Res 99, 178-184.

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

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

Navarro, P., Moffat, M., Mullin, N. P., and Chambers, I. (2011). The X-inactivation trans-activator Rnf12 is negatively regu-

lated by pluripotency factors in embryonic stem cells. Hum Genet 130, 255-264.

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

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

Nesbitt, M. N., and Gartler, S. M. (1971). The applications of genetic mosaicism to developmental problems. Annu Rev Genet 5, 143-162.

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

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

Nesterova, T. B., Barton, S. C., Surani, M. A., and Brockdorff, N. (2001b). Loss of Xist imprinting in diploid parthenogenetic preimplantation embryos. Dev Biol 235, 343-350.

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

Ng, K., Daigle, N., Bancaud, A., Ohhata, T., Humphreys, P., Walker, R., Ellenberg, J., and Wutz, A. (2011). A system for imaging the regulatory noncoding Xist RNA in living mouse embryonic stem cells. Molecular biology of the cell 22, 2634-2645.

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

Nichols, J., and Smith, A. (2009). Naive and Primed Pluripotent States. Stem Cell 4, 487-492.

Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N. L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485, 381-385.

Nora, E. P., and Heard, E. (2010). Chromatin structure and nuclear organization dynamics during X-chromosome inactivation. Cold Spring Harb Symp Quant Biol 75, 333-344.

O'Neill, L. P., Spotswood, H. T., Fernando, M., and Turner, B. M. Differential loss of histone H3 isoforms mono-, di- and tri-methylated at lysine 4 during X-inactivation in female embryonic stem cells. Biological Chemistry 389.

Ogawa, Y., Sun, B. K., and Lee, J. T. (2008). Intersection of the RNA interference and X-inactivation pathways. Science 320, 1336-1341.

Ogawa, Y., and Lee, J. T. (2003). Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. Mol Cell 11, 731-743.

Ohhata, T., Hoki, Y., Sasaki, H., and Sado, T. (2007). Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. Development 135, 227-235.

Ohhata, T., Senner, C. E., Hemberger, M., and Wutz, A. (2011). Lineage-specific function of the noncoding Tsix RNA for Xist repression and Xi reactivation in mice. Genes Dev 25, 1702-1715.

Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S. C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A., et al. (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. Nature 436, 207-213.

Ohno, S. (1967). Sex Chromosomes and-Sex-Linked Genes. Berlin: Springer Verlag.

Okamoto, I., Otte, A. P., Allis, C. D., Reinberg, D., and Heard, E. (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. Science 303, 644-649.

Okamoto, I., Patrat, C., Thepot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J. P., Renard, J. P., Duranthon, V., et al. (2011). Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature 472, 370-374.

Okamoto, I., Arnaud, D., Le Baccon, P., Otte, A. P., Disteche, C. M., Avner, P., and Heard, E. (2005). Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. Nature 438, 369-373.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature 448, 313-317.

Pasque, V., Tchieu, J., Karnik, R., Uyeda, M., Sadhu Dimashkie, A., Case, D., Papp, B., Bonora, G., Patel, S., Ho, R., et al. (2014). X chromosome reactivation dynamics reveal stages of reprogramming to pluripotency. CELL 159, 1681-1697.

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

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature 379, 131-137.

Percec, I., Plenge, R. M., Nadeau, J. H., Bartolomei, M. S., and Willard, H. F. (2002). Autosomal dominant mutations affecting X inactivation choice in the mouse. Science 296, 1136-1139.

Perche, P. Y., Vourc'h, C., Konecny, L., Souchier, C., Robert-Nicoud, M., Dimitrov, S., and Khochbin, S. (2000). Higher concentrations of histone macroH2A in the Barr body are correlated with higher nucleosome density. Current Biology 10, 1531-1534.

Pessia, E., Makino, T., Bailly-Bechet, M., McLysaght, A., and Marais, G. A. (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, 5346-5351.

Pessia, E., Engelstadter, J., and Marais, G. A. (2013). The evolution of X chromosome inactivation in mammals: the demise of Ohno's hypothesis? Cell Mol Life Sci.

Peters, A. H. F. M., Mermoud, J. E., O'Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N., and Jenuwein, T. (2002). Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. Nature Genetics 30, 77-80.

Pinter, S. F., Sadreyev, R. I., Yildirim, E., Jeon, Y., Ohsumi, T. K., Borowsky, M., and Lee, J. T. (2012). Spreading of X chromosome inactivation via a hierarchy of defined Polycomb stations. Genome Res 22, 1864-1876.

Plath, K., Talbot, D., Hamer, K. M., Otte, A. P., Yang, T. P., Jaenisch, R., and Panning, B. (2004). Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. The Journal of Cell Biology 167, 1025-1035.

Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., La Cruz, de, C. C., Otte, A. P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. Science 300, 131-135.

Popova, B. C., Tada, T., Takagi, N., Brockdorff, N., and Nesterova, T. B. (2006). Attenuated spread of X-inactivation in an X;autosome translocation. 103, 7706-7711.

Powell, J. R., Jow, M. M., and Meyer, B. J. (2005). The T-box transcription factor SEA-1 is an autosomal element of the X:A signal that determines C. elegans sex. Developmental Cell 9, 339-349.

Pullirsch, D., Hartel, R., Kishimoto, H., Leeb, M., Steiner, G., and Wutz, A. (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, 935-943.

Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. CELL 159, 1665-1680.

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

Rastan, S., and Robertson, E. J. (1985).

X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. J Embryol Exp Morphol 90, 379-388.

Rego, A., Sinclair, P. B., Tao, W., Kireev, I., and Belmont, A. S. (2008). The facultative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure. J Cell Sci 121, 1119-1127.

Reijnders, M. R. F., Zachariadis, V., Latour, B., Jolly, L., Mancini, G. M., Pfundt, R., Wu, K. M., van Ravenswaaij-Arts, C. M. A., Veenstra-Knol, H. E., Anderlid, B. M., et al. (2016). De Novo Loss-of-Function Mutations in USP9X Cause a Female-Specific Recognizable Syndrome with Developmental Delay and Congenital Malformations. Am J Hum Genet 98, 373-381.

Riggs, A. D., and Xiong, Z. (2004). Methylation and epigenetic fidelity. Proc Natl Acad Sci U S A 101, 4-5.

Ross, M. T., Grafham, D. V., Coffey, A. J., Scherer, S., McLay, K., Muzny, D., Platzer, M., Howell, G. R., Burrows, C., Bird, C. P., et al. (2005). The DNA sequence of the human X chromosome. Nature 434, 325-337.

Rossant, J. (2008). Stem cells and early lineage development. CELL 132, 527-531.

Russell, L. B. (1963). Mammalian X-chromosome action: inactivation limited in spread and region of origin. Science 140, 976-978.

Russell, L. B., and Cacheiro, N. L. (1978). The use of mouse X-autosome translocations in the study of X-inactivation pathways and nonrandomness. Basic Life Sci 12, 393-416.

Russell, L. B., and Montgomery, C. S. (1970). Comparative studies on X-autosome translocations in the mouse. II. Inactivation of autosomal loci, segregation, and mapping of autosomal breakpoints in five T (X;1) S. Genetics 64, 281-312.

Russell, W. L., Russell, L. B., and Gower, J. S. (1959). Exceptional Inheritance of a Sex-Linked Gene in the Mouse Explained on the Basis That the X/O Sex-Chromosome Constitution Is Female. Proc Natl Acad Sci U S A 45, 554-560.

Sado, T., Wang, Z., Sasaki, H., and Li, E. (2001). Regulation of imprinted X-chromosome inactivation in mice by Tsix. Development 128, 1275-1286.

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

Sado, T., Fenner, M. H., Tan, S. S., Tam, P., Shioda, T., and Li, E. (2000). X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. Dev Biol 225, 294-303.

Sadreyev, R. I., Yildirim, E., Pinter, S. F., and Lee, J. T. (2013). Bimodal quantitative relationships between histone modifications for X-linked and autosomal loci. Proceedings of the National Academy of Sciences 110, 6949-6954.

Saitou, M., and Yamaji, M. (2012). Primordial germ cells in mice. Cold Spring Harb Perspect Biol 4.

Salz, H. K., Maine, E. M., Keyes, L. N., Samuels, M. E., Cline, T. W., and Schedl, P. (1989). The Drosophila female-specific sex-determination gene, Sex-lethal, has stage-, tissue-, and sex-specific RNAs suggesting multiple modes of regulation. Genes Dev 3, 708-719.

Sarma, K., Cifuentes-Rojas, C., Ergun, A., del Rosario, A., Jeon, Y., White, F., Sadreyev, R., and Lee, J. T. (2014). ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. CELL 159, 869-883.

Sarma, K., Levasseur, P., Aristarkhov, A., and Lee, J. T. (2010). Locked nucleic acids (LNAs) reveal sequence requirements and kinetics of Xist RNA localization to the X chromosome. Proceedings of the National Academy of Sciences 107, 22196-22201.

Sasaki, K., Yokobayashi, S., Nakamura, T., Okamoto, I., Yabuta, Y., Kurimoto, K., Ohta, H., Moritoki, Y., Iwatani, C., Tsuchiya, H., et al. (2015). Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells. Stem Cell, 1-18.

Schoeftner, S., Sengupta, A. K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., and Wutz, A. (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. The EMBO Journal 25, 3110-3122.

Schoeftner, S., Blanco, R., Lopez de Silanes, I., Muñoz, P., Gómez-López, G., Flores, J. M., and Blasco, M. A. (2009). Telomere

shortening relaxes X chromosome inactivation and forces global transcriptome alterations. Proceedings of the National Academy of Sciences 106, 19393-19398.

Schulz, E. G., Meisig, J., Nakamura, T., Okamoto, I., Sieber, A., Picard, C., Borensztein, M., Saitou, M., Blüthgen, N., and Heard, E. (2014). The Two Active X Chromosomes in Female ESCs Block Exit from the Pluripotent State by Modulating the ESC Signaling Network. Stem Cell 14, 203-216.

Sharman, G. B. (1971). Late DNA replication in the paternally derived X chromosome of female kangaroos. Nature 230, 231-232.

Sharp, A. J., Spotswood, H. T., Robinson, D. O., Turner, B. M., and Jacobs, P. A. (2002). Molecular and cytogenetic analysis of the spreading of X inactivation in X;autosome translocations.

Shen, Y., Yue, F., McCleary, D. F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V. V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature 488, 116-120.

Shen, Y., Matsuno, Y., Fouse, S. D., Rao, N., Root, S., Xu, R., Pellegrini, M., Riggs, A. D., and Fan, G. (2008). X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. Proc Natl Acad Sci U S A 105, 4709-4714.

Shi, Y., Downes, M., Xie, W., Kao, H. Y., Ordentlich, P., Tsai, C. C., Hon, M., and Evans, R. M. (2001). Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. Genes & Development 15, 1140-1151.

Shibata, S., and Lee, J. T. (2004). Tsix transcription- versus RNA-based mechanisms in Xist repression and epigenetic choice. Curr Biol 14, 1747-1754.

Shin, J., Bossenz, M., Chung, Y., Ma, H., Byron, M., Taniguchi-Ishigaki, N., Zhu, X., Jiao, B., Hall, L. L., Green, M. R., et al. (2010). Maternal Rnf12/RLIM is required for imprinted X-chromosome inactivation in mice. Nature 467, 977-981.

Shin, J., Wallingford, M. C., Gallant, J., Marcho, C., Jiao, B., Byron, M., Bossenz, M., Lawrence, J. B., Jones, S. N., Mager, J., et al. (2014). RLIM is dispensable for X-chromosome inactivation in the mouse embryonic epiblast. Nature 511, 86-89.

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T. B., Webster, Z., Peters, A. H., Jenuwein, T., Otte, A. P., and Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. Dev Cell 4, 481-495.

Silva, J., Nichols, J., Theunissen, T. W., Guo, G., van Oosten, A. L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. CELL 138, 722-737.

Silva, S. S., Rowntree, R. K., Mekhoubad, S., and Lee, J. T. (2008). X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc Natl Acad Sci U S A 105, 4820-4825.

Simmler, M. C., Cattanach, B. M., Rasbery, C., Rougeulle, C., and Avner, P. (1993). Mapping the murine Xce locus with (CA) n repeats. Mamm Genome 4, 523-530.

Simon, M. D., Pinter, S. F., Fang, R., Sarma, K., Rutenberg-Schoenberg, M., Bowman, S. K., Kesner, B. A., Maier, V. K., Kingston, R. E., and Lee, J. T. (2013). High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. Nature 504, 465-469.

Singh, V., Sharma, P., and Capalash, N. (2013). DNA methyltransferase-1 inhibitors as epigenetic therapy for cancer. Current cancer drug targets 13, 379-399.

Smallwood, A., and Ren, B. (2013). Genome organization and long-range regulation of gene expression by enhancers. Current Opinion in Cell Biology 25, 387-394.

Smeets, D., Markaki, Y., Schmid, V. J., Kraus, F., Tattermusch, A., Cerase, A., Sterr, M., Fiedler, S., Demmerle, J., Popken, J., et al. (2014). Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. Epigenetics & Chromatin 7, 8.

Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M. R. F., Venselaar, H., Helsmoortel, C., Cho, M. T., Hoischen, A., et al. (2015). Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on

Wnt Signaling. Am J Hum Genet 97, 343-352.

Spencer, R. J., del Rosario, B. C., Pinter, S. F., Lessing, D., Sadreyev, R. I., and Lee, J. T. (2011). A boundary element between Tsix and Xist binds the chromatin insulator Ctcf and contributes to initiation of X-chromosome inactivation. Genetics 189, 441-454.

Splinter, E., de Wit, E., Nora, E. P., Klous, P., van de Werken, H. J. G., Zhu, Y., Kaaij, L. J. T., van IJcken, W., Gribnau, J., Heard, E., et al. (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes & Development 25, 1371-1383.

Stadtfeld, M., Maherali, N., Borkent, M., and Hochedlinger, K. (2010). A reprogrammable mouse strain from gene-targeted embryonic stem cells. Nat Methods 7, 53-55.

Stadtfeld, M., Maherali, N., Breault, D. T., and Hochedlinger, K. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell 2, 230-240.

Stavropoulos, N., Rowntree, R. K., and Lee, J. T. (2005). Identification of developmentally specific enhancers for Tsix in the regulation of X chromosome inactivation. Mol Cell Biol 25, 2757-2769.

Straub, T., Grimaud, C., Gilfillan, G. D., Mitterweger, A., and Becker, P. B. (2008). The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. PLoS Genet 4, e1000302.

Sugawa, F., Araúzo-Bravo, M. J., Yoon, J., Kim, K., Aramaki, S., Wu, G., Stehling, M., Psathaki, O. E., Hübner, K., and Schöler, H. R. (2015). Human primordial germ cell commitment in vitro associates with a unique PRDM14 expression profile. The EMBO Journal 34, 1009-1024.

Sugimoto, M., and Abe, K. (2007). X chromosome reactivation initiates in nascent primordial germ cells in mice. PLoS Genet 3, e116.

Sun, B. K., Deaton, A. M., and Lee, J. T. (2006). A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. Mol Cell 21, 617-628.

Sun, S., Fukue, Y., Nolen, L., Sadreyev, R., and Lee, J. T. (2010). Characterization of Xpr (Xpct) reveals instability but no effects on X-chromosome pairing or Xist expression. Transcription 1, 46-56.

Sun, S., del Rosario, B. C., Szanto, A., Ogawa, Y., Jeon, Y., and Lee, J. T. (2013). Jpx RNA Activates Xist by Evicting CTCF. CELL 153, 1537-1551.

Sun, S., Payer, B., Namekawa, S., An, J. Y., Press, W., Catalan-Dibene, J., Sunwoo, H., and Lee, J. T. (2015). Xist imprinting is promoted by the hemizygous (unpaired) state in the male germ line. Proceedings of the National Academy of Sciences 112, 14415-14422.

Sural, T. H., Peng, S., Li, B., Workman, J. L., Park, P. J., and Kuroda, M. I. (2008). The MSL3 chromodomain directs a key targeting step for dosage compensation of the Drosophila melanogaster X chromosome. Nat Struct Mol Biol 15, 1318-1325.

Tachibana, M., Ma, H., Sparman, M. L., Lee, H. S., Ramsey, C. M., Woodward, J. S., Sritanaudomchai, H., Masterson, K. R., Wolff, E. E., Jia, Y., et al. (2012). X-chromosome inactivation in monkey embryos and pluripotent stem cells. Dev Biol 371, 146-155.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr Biol 11, 1553-1558.

Tada, T., Obata, Y., Tada, M., Goto, Y., Nakatsuji, N., Tan, S., Kono, T., and Takagi, N. (2000). Imprint switching for non-random X-chromosome inactivation during mouse oocyte growth. Development 127, 3101-3105.

Takagi, N. (1980). Primary and secondary non-random X-chromosome inactivation in early female mouse embryos carrying Searle's translocation T(X;16)16H. Chromosoma 81, 439-459.

Takagi, N. (1993). Variable X chromosome inactivation patterns in near-tetraploid murine EC x somatic cell hybrid cells differentiated in vitro. Genetica 88, 107-117.

Takagi, N., Yoshida, M. A., Sugawara, O., and Sasaki, M. (1983). Reversal of X-inactivation in female mouse somatic cells hybridized with murine teratocarcinoma stem cells in vitro. CELL 34, 1053-1062.

Takagi, N., and Sasaki, M. (1975). Preferential inactivation of the paternally derived X

chromosome in the extraembryonic membranes of the mouse. Nature 256, 640-642.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. CELL 131, 861-872.

Tang, Y. A., Huntley, D., Montana, G., Cerase, A., Nesterova, T. B., and Brockdorff, N. (2010). Efficiency of Xist-mediated silencing on autosomes is linked to chromosomal domain organisation. Epigenetics & Chromatin 3, 1.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., et al. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. CELL 148, 664-678.

Terranova, R., Yokobayashi, S., Stadler, M. B., Otte, A. P., van Lohuizen, M., Orkin, S. H., and Peters, A. H. F. M. (2008). Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. Dev Cell 15, 668-679.

Thomson, J. A., and Marshall, V. S. (1998). Primate embryonic stem cells. Curr Top Dev Biol 38, 133-165.

Thorvaldsen, J. L., Krapp, C., Willard, H. F., and Bartolomei, M. S. (2012). Nonrandom X chromosome inactivation is influenced by multiple regions on the murine X chromosome. Genetics 192, 1095-1107.

Tian, D., Sun, S., and Lee, J. T. (2010). The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. CELL 143, 390-403.

Vallier, L., Alexander, M., and Pedersen, R. A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J Cell Sci 118, 4495-4509.

van Kruijsbergen, I., Hontelez, S., and Veenstra, G. J. C. (2015). Recruiting polycomb to chromatin. International Journal of Biochemistry and Cell Biology 67, 177-187.

Wake, N., Takagi, N., and Sasaki, M. (1976). Non-random inactivation of X chromosome in the rat yolk sac. Nature 262, 580-581.

Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J. C., Nagy, A., and Magnuson, T. (2001). Imprinted X inactivation maintained by a mouse Polycomb group gene. Nature Genetics 28, 371-375.

Wang, L., Brown, J. L., Cao, R., Zhang, Y., Kassis, J. A., and Jones, R. S. (2004). Hierarchical recruitment of polycomb group silencing complexes. Molecular Cell 14, 637-646.

Wang, X., Miller, D. C., Clark, A. G., and Antczak, D. F. (2012). Random X inactivation in the mule and horse placenta. Genome Res 22, 1855-1863.

Ware, C. B., Wang, L., Mecham, B. H., Shen, L., Nelson, A. M., Bar, M., Lamba, D. A., Dauphin, D. S., Buckingham, B., Askari, B., et al. (2009). Histone deacetylase inhibition elicits an evolutionarily conserved self-renewal program in embryonic stem cells. Cell Stem Cell 4, 359-369.

Webb, S., de Vries, T. J., and Kaufman, M. H. (1992). The differential staining pattern of the X chromosome in the embryonic and extraembryonic tissues of post-implantation homozygous tetraploid mouse embryos. Genet Res 59, 205-214.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B. E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 448, 318-324.

Williams, L. H., Kalantry, S., Starmer, J., and Magnuson, T. (2011). Transcription precedes loss of Xist coating and depletion of H3K27me3 during X-chromosome reprogramming in the mouse inner cell mass. Development 138, 2049-2057.

Wolf, J. B., and Bryk, J. (2011). General lack of global dosage compensation in ZZ/ZW systems? Broadening the perspective with RNA-seq. BMC Genomics 12, 91.

Wutz, A., Rasmussen, T. P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nature Genetics 30, 167-174.

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

Xiong, Y., Chen, X., Chen, Z., Wang, X., Shi, S., Wang, X., Zhang, J., and He, X. (2010). RNA sequencing shows no dosage compensation of the active X-chromosome. Nature Genetics 42, 1043-1047.

Xu, J., Deng, X., Watkins, R., and Disteche, C. M. (2008). Sex-specific differences in expression of histone demethylases Utx and Uty in mouse brain and neurons. The Journal of neuroscience : the official journal of the Society for Neuroscience 28, 4521-4527.

Xu, J., and Disteche, C. M. (2006). Sex differences in brain expression of X- and Y-linked genes. Brain research 1126, 50-55.

Xu, N., Tsai, C. L., and Lee, J. T. (2006). Transient homologous chromosome pairing marks the onset of X inactivation. Science 311, 1149-1152.

Xue, F., Tian, X. C., Du, F., Kubota, C., Taneja, M., Dinnyes, A., Dai, Y., Levine, H., Pereira, L. V., and Yang, X. (2002). Aberrant patterns of X chromosome inactivation in bovine clones. Nature Genetics 31, 216-220.

Yamada, N., Hasegawa, Y., Yue, M., Hamada, T., Nakagawa, S., and Ogawa, Y. (2015). Xist Exon 7 Contributes to the Stable Localization of Xist RNA on the Inactive X-Chromosome. PLoS Genet 11, e1005430.

Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y., and Saitou, M. (2008). Critical function of Prdm14 for the establishment of the germ cell lineage in mice. Nature Genetics 40, 1016-1022.

Yang, F., Babak, T., Shendure, J., and Disteche, C. M. (2010). Global survey of escape from X inactivation by RNA-sequencing in mouse. Genome Research 20, 614-622.

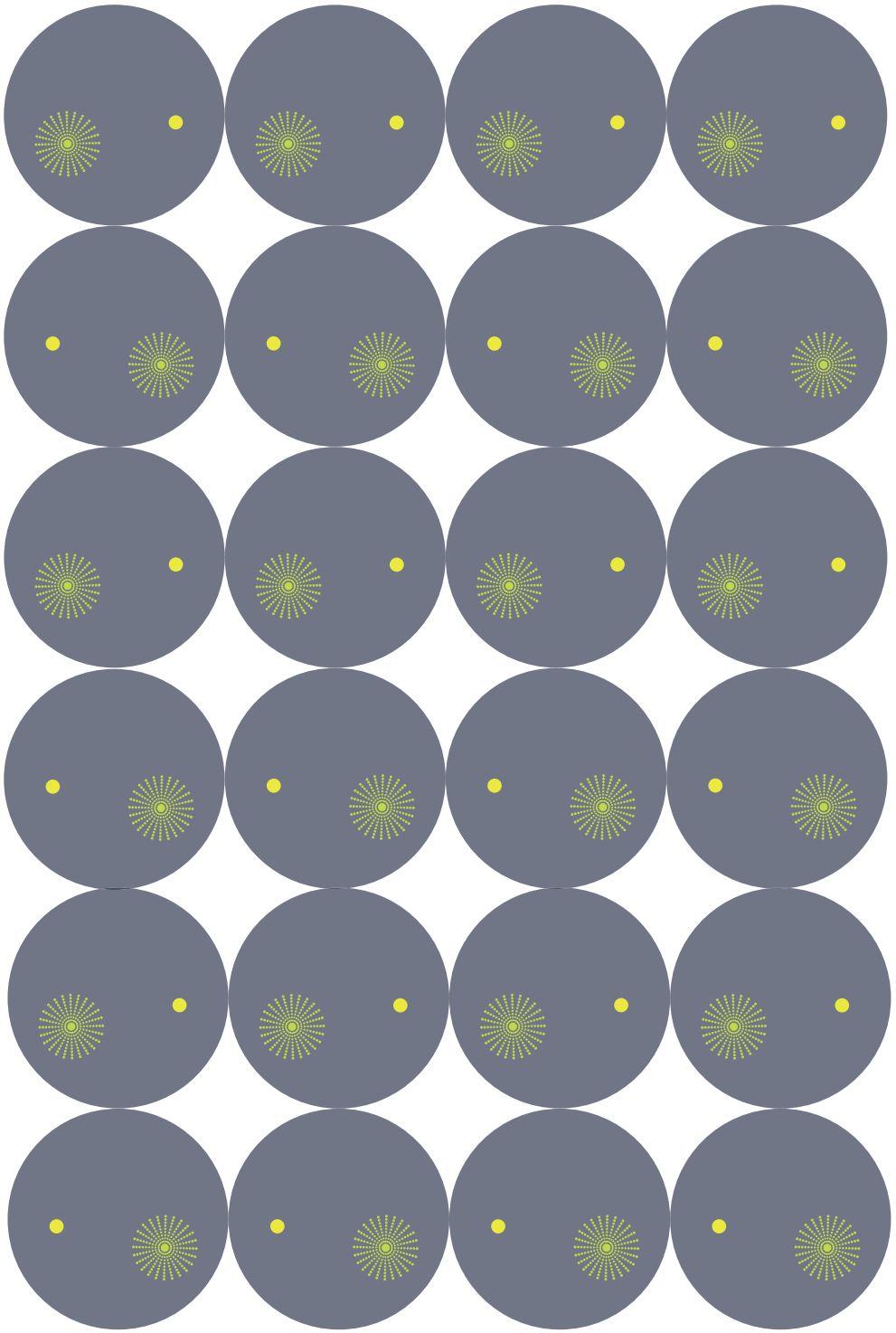
Ying, Q. L., Nichols, J., Evans, E. P., and Smith, A. G. (2002). Changing potency by spontaneous fusion. Nature 416, 545-548.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science 318, 1917-1920.

Zhang, L. F., Huynh, K. D., and Lee, J. T. (2007). Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing. CELL 129, 693-706.

Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J., and Lee, J. T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 322, 750-756.

Zhao, X. Y., Li, W., Lv, Z., Liu, L., Tong, M., Hai, T., Hao, J., Guo, C. L., Ma, Q. W., Wang, L., et al. (2009). iPS cells produce viable mice through tetraploid complementation. Nature 461, 86-90.



CHAPTER .2

Xist AND Tsix TRANSCRIPTION DYNAMICS IS REGULATED BY THE X-TO-AUTOSOME RATIO AND SEMI-STABLE TRANSCRIPTIONAL STATES

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

Accepted for publication (Mol Cell Biol.)

Xist AND *Tsix* TRANSCRIPTION DYNAMICS IS REGULATED BY THE X-TO-AUTOSOME RATIO AND SEMI-STABLE TRANSCRIPTIONAL STATES

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

Author Affiliations:
[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

Accepted for publication (Mol Cell Biol.)

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) (Borsani et al. 1991,

Brockdorff et al., 1991), and *cis*-spreading of *Xist* leads to recruitment of chromatin remodeling complexes that render the X inactive (Dixon-McDougall et al., 2016, Moindrot et al., 2016). *Tsix* is transcribed anti-sense to *Xist* and fully overlaps with *Xist* (Lee et al., 1999a). *Tsix* transcription and/or the produced Tsix RNA are involved in repression of *Xist* which includes *Tsix* mediated chromatin changes at the *Xist* promoter (Lee et al., 1999b, Navarro et al., 2006, Ohhata et al., 2008, Sado et al., 2005).

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 (Ma et al., 2011, Navarro et al., 2008, Navarro et al. 2010, Payer et al., 2013). These factors, and other ubiquitously expressed XCI-inhibitors including CTCF (Donohoe et al., 2007, Sun et al., 2013), 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 (Barakat et al., 2011, Minkovsky et al., 2013, Nesterova et al., 2011).

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) (Jonkers et al., 2009). The encoded protein RNF12 is an E3 ubiquitin ligase, which targets the XCI-inhibitor REX1 for degradation (Gontan et al., 2012). 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 (Makhlouf et al., 2014).

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 (Barakat et al., 2011). 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 (Shin et al., 2014). *Jpx* and *Ftx* have been described as putative XCI activators (Sun et et al., 2013, Chureau et al., 2011, Tian et al., 2010). Both genes are located in a region 10-100 kb 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* (Barakat et al., 2014).

Interestingly, examination of the higher order chromatin structure revealed *Xist* and *Tsix* to be located in two distinct neighboring topological associated domains (TADs) (Dixon et al., 2012, Nora et al., 2012). 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 activators and inhibitors (Nora et al., 2012, Anguera et al., 2011, Ogawa et al., 2003). 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 (Schulz et al., 2014). 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 (Marahrens et al., 1997, Penny et al. 1996), while *Tsix* is a negative regulator of XCI (Lee et al., 1999b, Stravropoulos et al., 2001). 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 (Barakat et al. 2011b), 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).

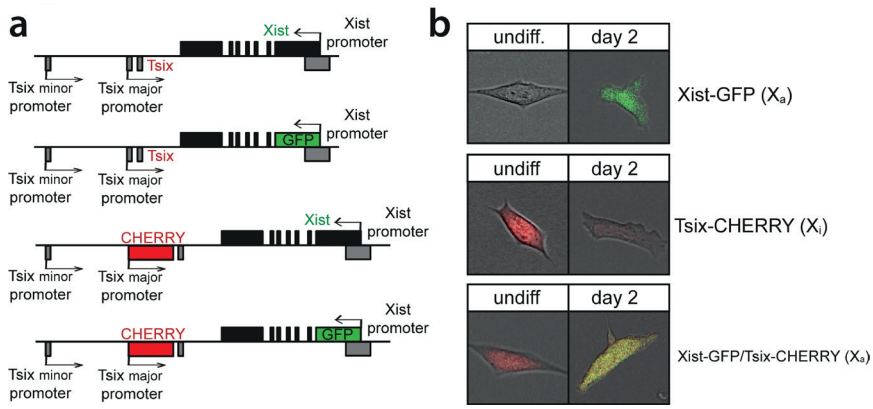


Figure 1. Generation of the reporter alleles.

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

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 (Csankovszki et al., 1999, 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.

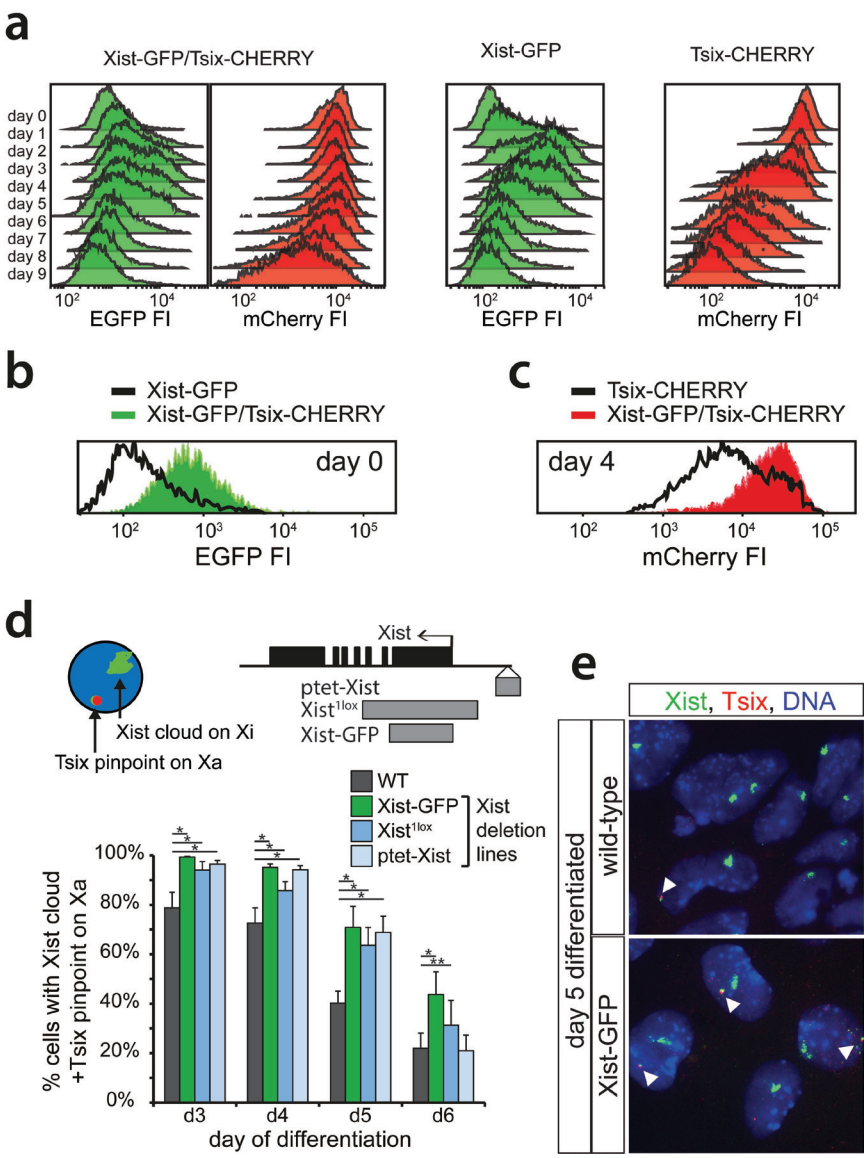


Figure 2. *Xist* and *Tsix* Reporter Lines Reveal Antagonistic Roles for *Xist* and *Tsix*.

(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.

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).

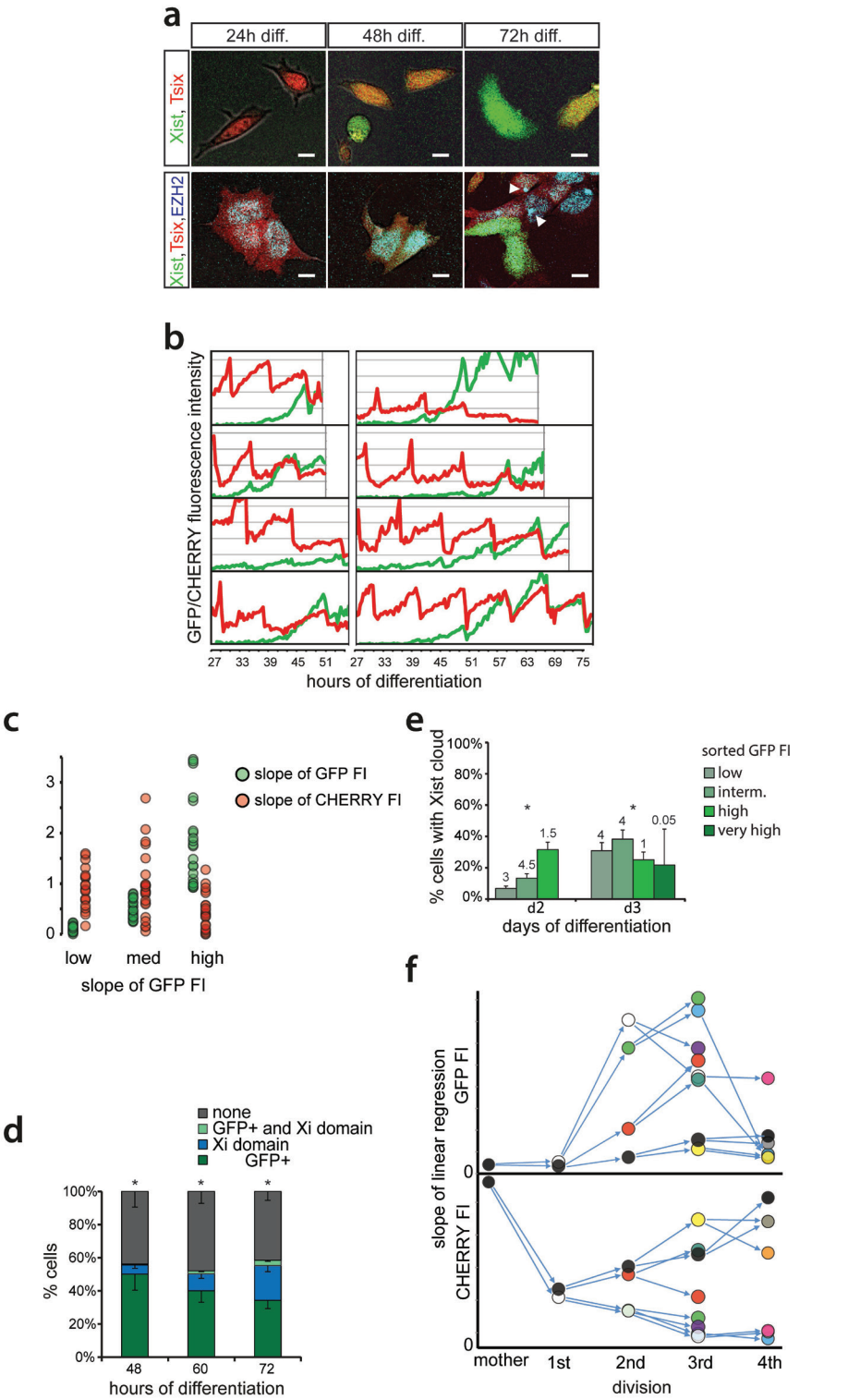
To unravel the relationship between activation of *Xist* and *Tsix* and establishment of the Xi we introduced a mTagBFP2-Ezh2 fusion gene into Xist-GFP/Tsix-CHERRY ES cells (Supplementary Fig 4b-c). Since we were not able to continually follow high numbers of cells until an Xi domain appeared, we instead 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 EGFPhigh and EGFPveryhigh FACS sorted fractions of cells contained less Xist clouds than the EGFPintermediate fraction. This suggests that EGFPhigh and EGFPveryhigh 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.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.

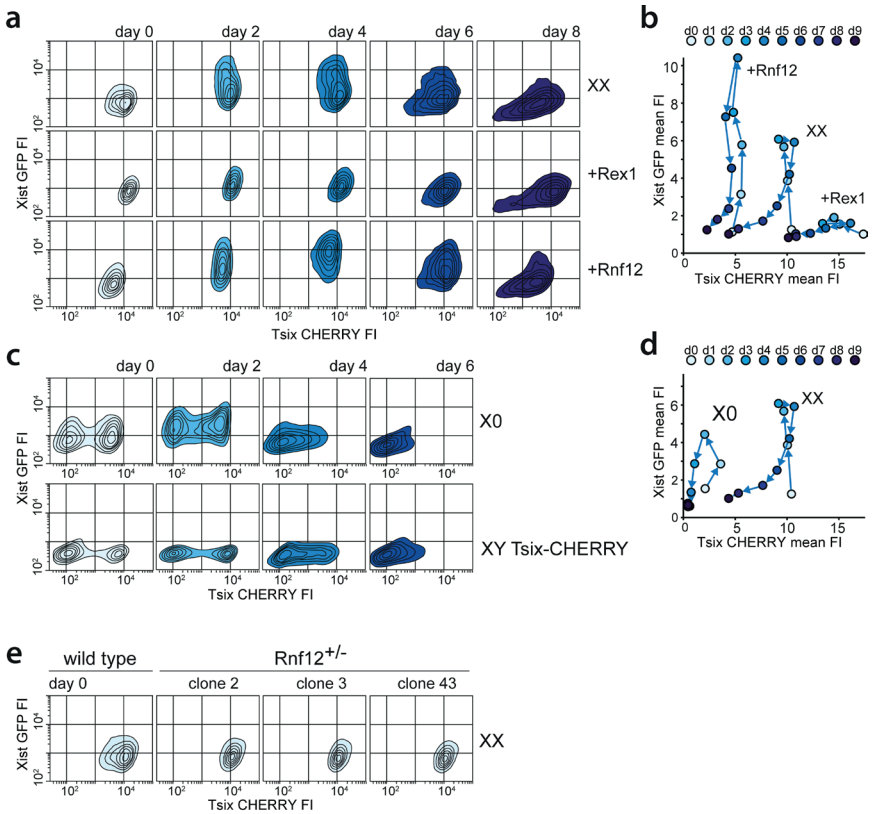

a
24h diff.
48h diff.
72h diff.
Xist, Tsix
Xist, Tsix, EZH2
b
GFP/CHERRY fluorescence intensity
27 33 39 45 51 27 33 39 45 51 57 63 69 75
hours of differentiation
c
slope of GFP FI
slope of CHERRY FI
low
med
high
slope of GFP FI
d
none
GFP+ and Xi domain
Xi domain
GFP+
% cells
100%
80%
60%
40%
20%
0%
48
60
72
hours of differentiation
e
% cells with Xist cloud
100%
80%
60%
40%
20%
0%
3
4.5
1.5
4
4
1
0.05
d2
d3
days of differentiation
sorted GFP FI
low
interm.
high
very high
f
slope of linear regression
GFP FI
CHERRY FI
0
0
mother
1st
2nd
3rd
4th
division

Effects of activators and inhibitors on XCI

RNF12 functions as a trans-activator of XCI (Jonkers et al., 2009) by targeting REX1, a repressor of XCI, for proteasomal degradation (Gontan et al., 2012). Previous work has indicated that REX1 might have a dual role in the activation of XCI by activating *Tsix* and repressing *Xist* (Navarro et al., 2010, Barakat et al., 2011a, Gontan et al., 2012). 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.

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).



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.

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.

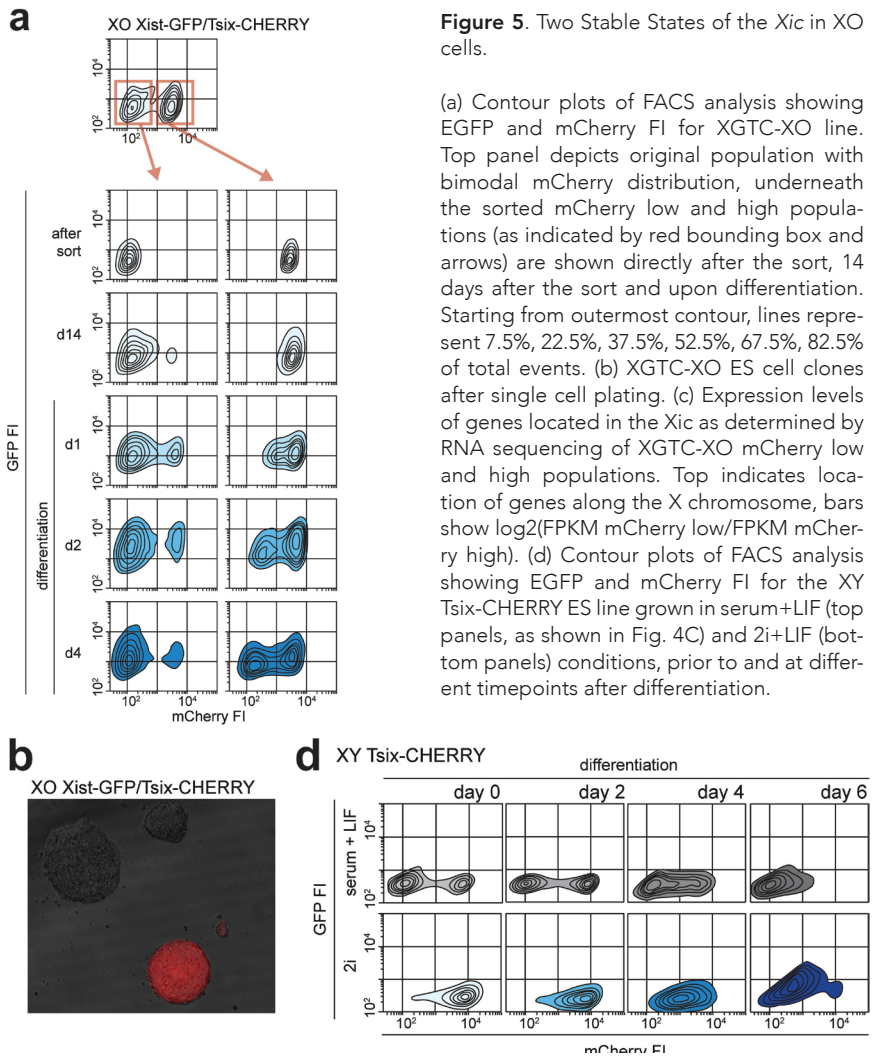
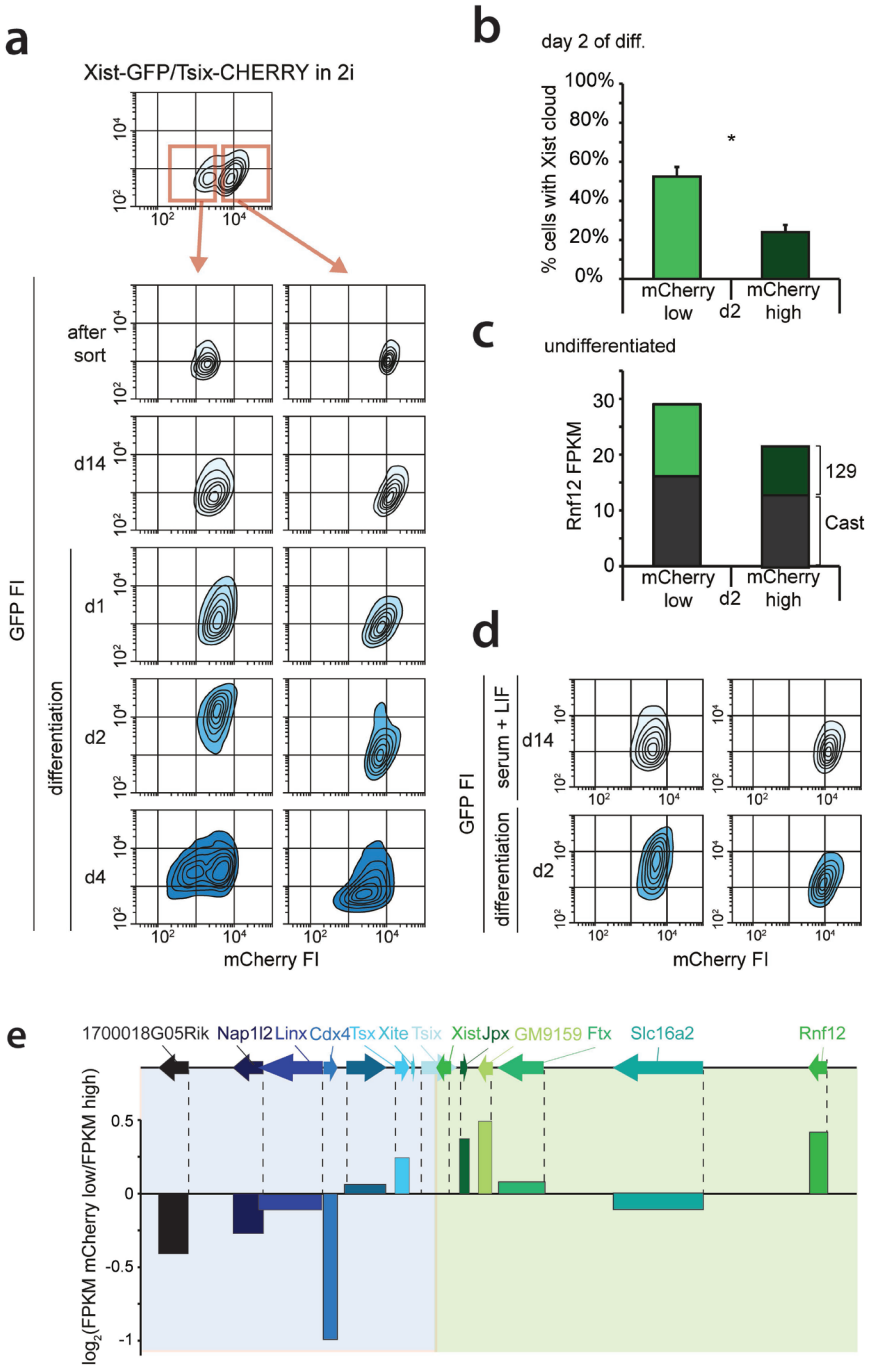


Figure 5. Two Stable States of the *Xic* in XO cells.

(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 log2(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.


a
Xist-GFP/Tsix-CHERRY in 2i
after sort
d14
d1
d2
d4
differentiation
GFP FI
mCherry FI
b
day 2 of diff.
% cells with Xist cloud
100%
80%
60%
40%
20%
0%
mCherry low
d2
mCherry high
c
undifferentiated
Rnf12 FPKM
129
Cast
d
serum + LIF
differentiation
e
1700018G05Rik
Nap1l2
Linx
Cdx4
Tsx
Xite
Tsix
XistJpx
GM9159
Ftx
Slc16a2
Rnf12
log2(FPKM mCherry low/FPKM high)

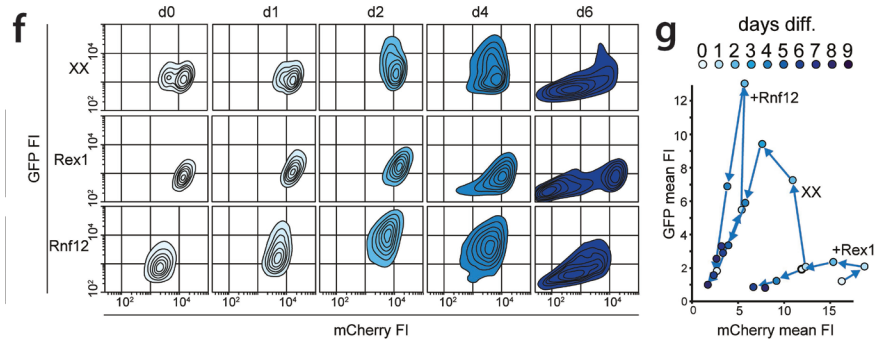


Figure 6. Two Stable States of the *Xic* Predict XCI Potential in XX cells.

(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 log2(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.

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 strong-

ly 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 (Marks et al., 2012). 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.

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 (Stavropoulos et al., 2001, Masui et al., 2011). 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 sufficientto 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 (Mlynarczyk-Evans et al., 2006). A recently developed polymer model predicted such different states of higher order chromatin structure (Giorgetti et al., 2014). 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 (Habibi et al., 2013). 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 (Masui et al., 2011), 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 (Masui et al., 2011). 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* (Payer et al., 2013, Navarro et al., 2011), 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 (Guyochin et al., 2014), 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, 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 (Makhlouf et al., 2014). 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 (Barakat et al., 2011a, Gontan et al., 2012).

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.

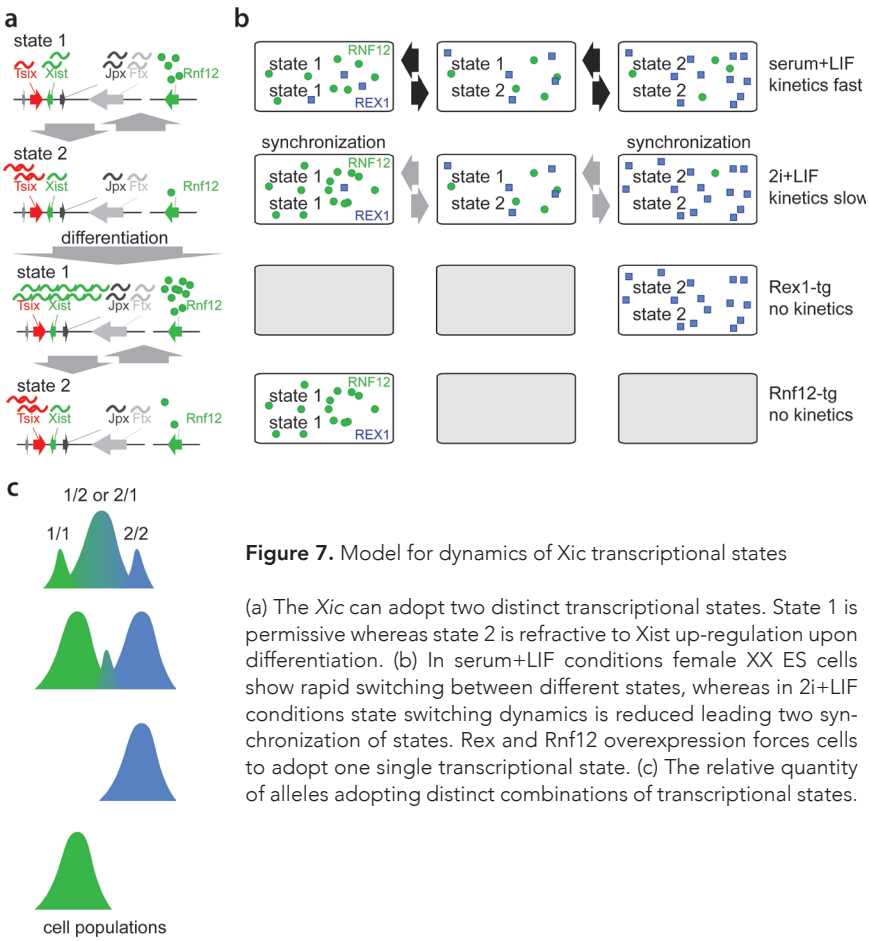


Figure 7. Model for dynamics of Xic transcriptional states

(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.

2

METHODS

Plasmids and Antibodies

Plasmids used for generation of transgenic cell lines were pCAG-Rex1-Flag, pCAG-Rnf12-Flag (Gontan et al., 2012) 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., 2011b). 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 3 μM GSK3 inhibitor CH99021 (Stemgent). Transgenic ES cell lines were generated using wild-type female line F12-1 (129/Sv-Cast/Ei) and wild-type male line J1 (129/Sv). The BAC targeting strategy was used as has been described (Barakat et al., 2011b). 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 (Barakat et al., 2011b). 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 37C. 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 ΔCT 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 (104 cells/well) in a 6-well glass bottom dish (MatTek P06G-418 1.5-20-F) coated with human plasma fibronectin (Millipore). Imaging was performed on a Leica SP5 AOBS at 37 C and 5% CO2 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, prehybridized 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:

$$p - \left[1.96\sqrt{\frac{p(1-p)}{n}}\right] \text{ to } p + \left[1.96\sqrt{\frac{p(1-p)}{n}}\right],$$

with n for the number of cells counted and p for the percentage of Xist clouds scored.

Standard deviation was calculated as: $\sqrt{\frac{\sum(x-\bar{x})^2}{n}}$,
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: $r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}}$, 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:

$$\chi^2 = \sum_i \sum_j \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

with Oij being the observed and Eij the expected frequencies.

The two-proportion z-test was calculated by:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

with n for the number of cells analyzed, and $\hat{p}$ and $\bar{p}$ corresponding to the proportion and average proportion, respectively.

2

REFERENCES

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.

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

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.

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

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.

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.

Chureau C, Chantalat S, Romito A, 655 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.

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.

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

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.

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.

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.

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.

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.

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.

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.

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

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

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.

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.

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.

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.

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.

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.

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.

Moindrot B, Brockdorff N. 2016. RNA binding proteins implicated in Xist-mediated chromosome silencing. Semin Cell Dev Biol doi:S1084-9521(16)30029-510.1016/j.semcdb.2016.01.029.

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.

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.

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 627 468:457-460.

Navarro P, Page DR, Avner P, Rougeulle C. 2006. Tsix-mediated epigenetic switch of a CTCF614 flanked region of the Xist promoter determines the Xist transcription program. Genes Dev 615 20:2787-2792.

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.

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.

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

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.

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.

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

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.

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

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.

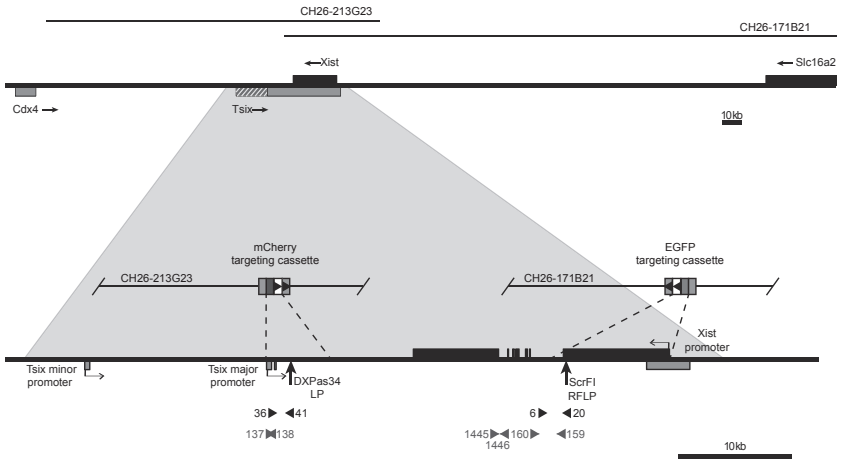
Shin J, Wallingford MC, Gallant J, Marcho C, Jiao B, Byron M, Bossenz M, Lawrence JB, Jones SN, Mager J, Bach I. 2014. RLIM is dispens-

able for X-chromosome inactivation in the mouse embryonic epiblast. Nature 511:86-89.

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.

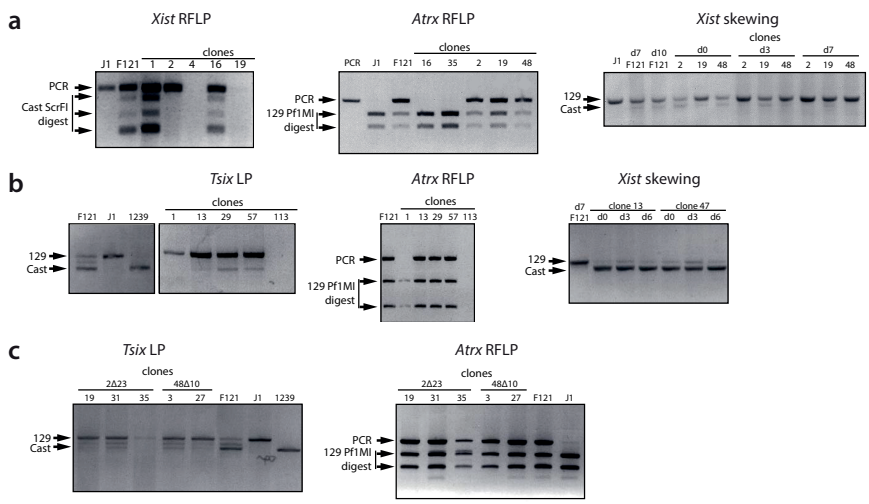
Sun S, Del Rosario BC, Szanto A, Ogawa Y, Jeon Y, Lee JT. 2013. Jpx RNA activates Xist by evicting CTCF. Cell 153:1537-1551.

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



Supplementary Figure 1. Targeting strategy.

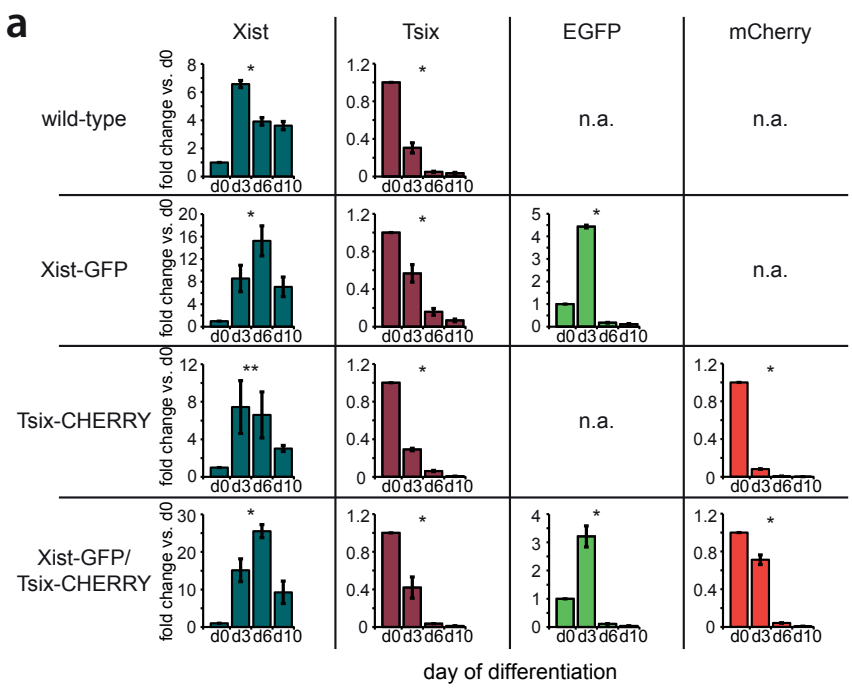
Targeting scheme for generation of Xist-GFP, Tsix-CHERRY and Xist-GFP/Tsix-CHERRY murine ES cell lines. BAC name, location and relative size are indicated on top of panel. Lower part of panel depicts mCherry and EGFP targeting cassettes, exon-intron structure of *Tsix* (grey) and *Xist* (black), and position of genotyping (black arrowheads) and phenotyping (red arrowheads) primers. Primer numbers are given as 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.

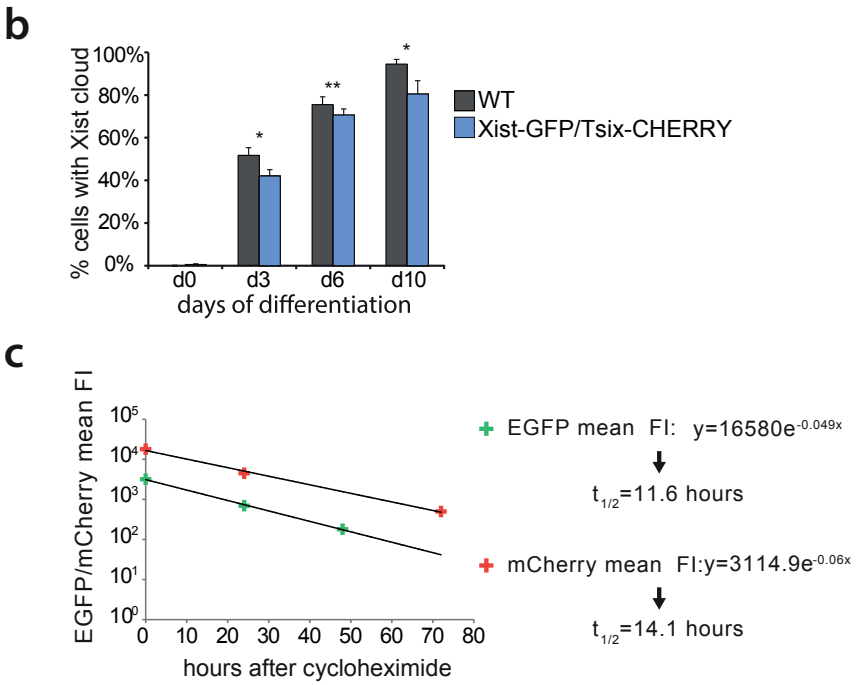


Supplementary Figure 2. Targeting of Cell Lines.

(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/

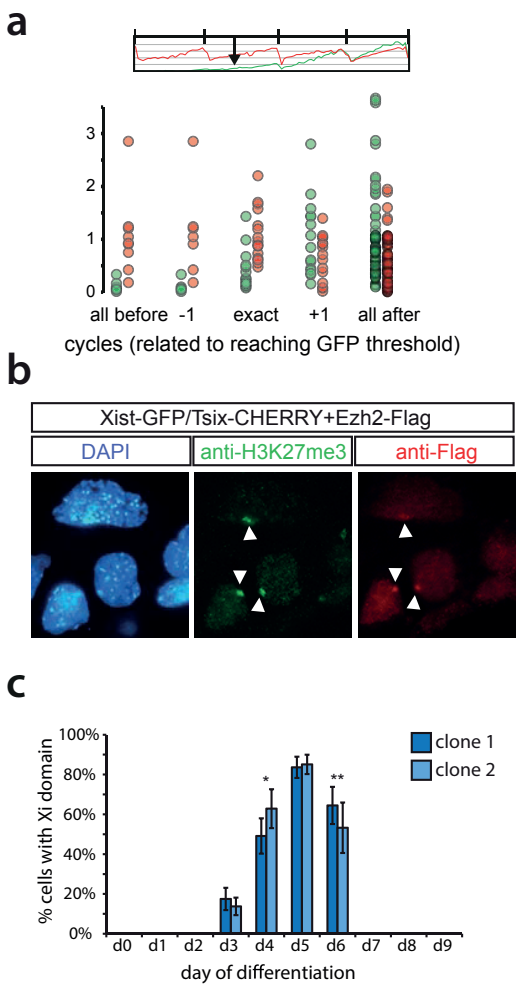
Ei *Xist* allele. Correct targeting of EGFP-cassette to Cast/Ei allele results in loss of Cast/Ei-specific restriction products, as shown for clone 2. Arrows on left indicate size of PCR product and size of ScrFI restriction fragments. J1 is a 129/Sv control, F121 the polymorphic 129/Sv-Cast/Ei mother 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 2-23.





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 mCherry 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 100mg/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).

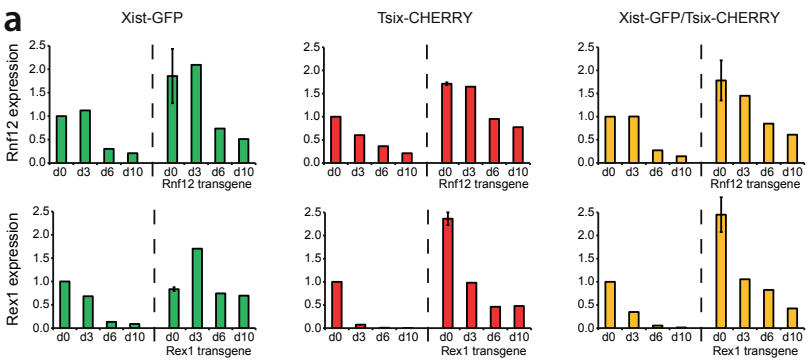


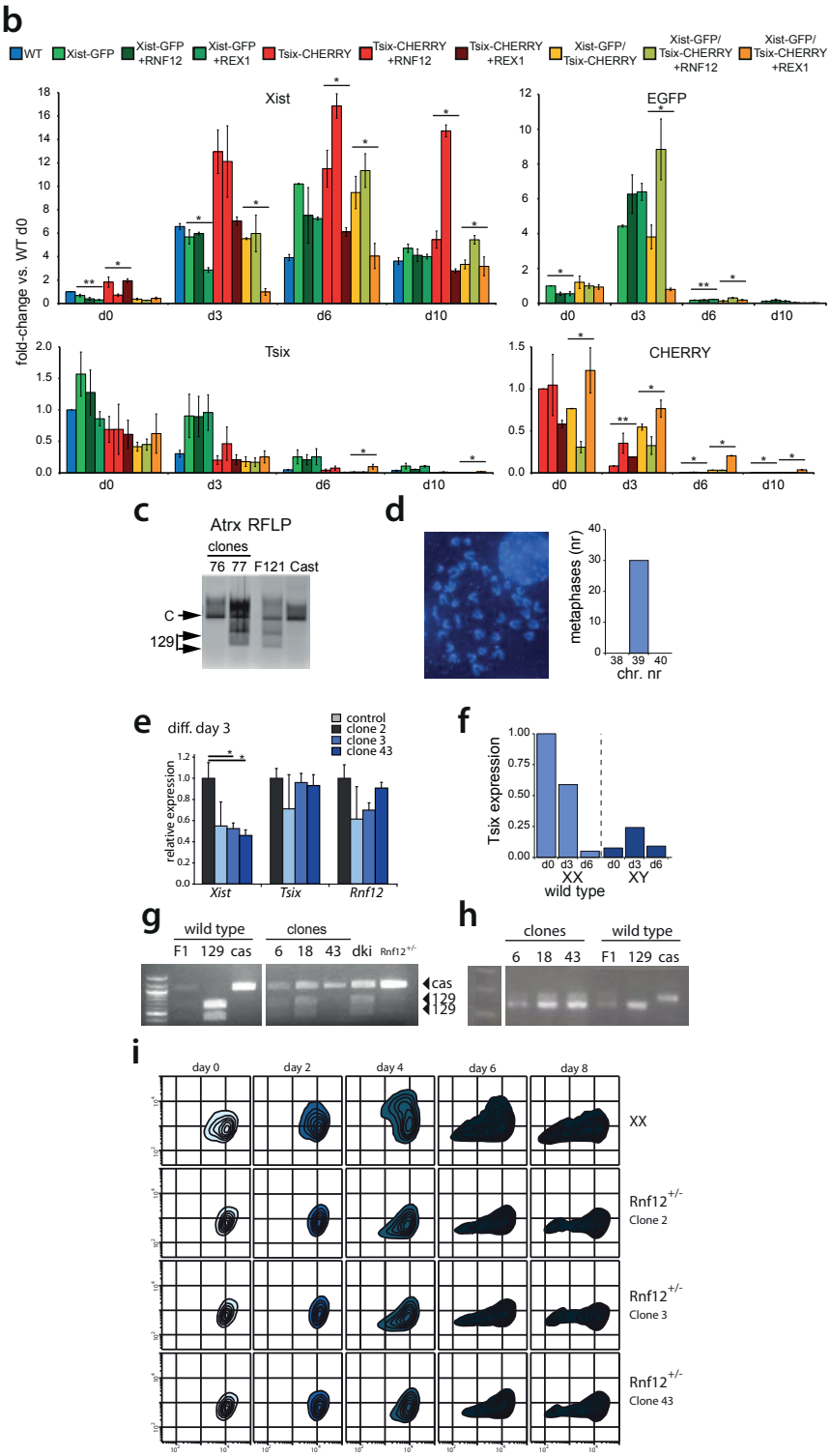
Supplementary Figure 4. Life cell imaging of reporter lines.

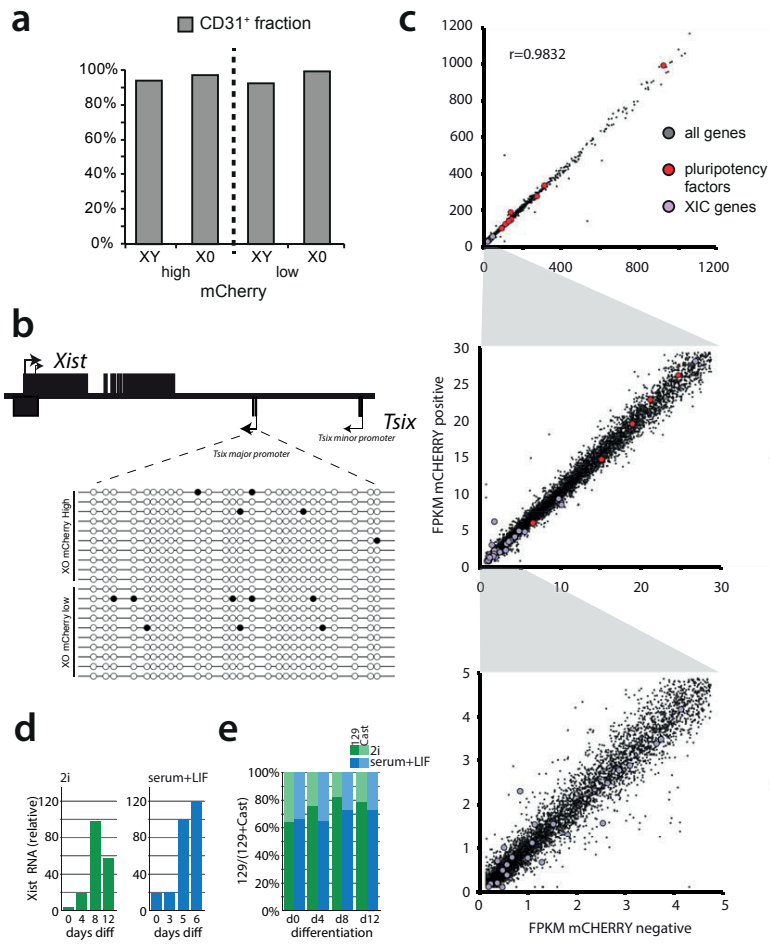
(a) Linear regression of FI over time for each cell cycle was performed. Slope of linear regression as a proxy for promoter activity is plotted. Bins are chosen according to time point of *Xist* promoter activation. Threshold for *Xist* activation was set at 3.29 SDs (corresponding to 99.9% within confidence 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.




b
WT
Xist-GFP
Xist-GFP +RNF12
Xist-GFP +REX1
Tsix-CHERRY
Tsix-CHERRY +RNF12
Tsix-CHERRY +REX1
Xist-GFP/ Tsix-CHERRY
Xist-GFP/ Tsix-CHERRY +RNF12
Xist-GFP/ Tsix-CHERRY +REX1
Xist
EGFP
Tsix
CHERRY
fold-change vs. WT d0
d0
d3
d6
d10
c
Atrx RFLP
clones
76 77 F121 Cast
C
129
d
metaphases (nr)
chr. nr
e
diff. day 3
control
clone 2
clone 3
clone 43
relative expression
Xist
Tsix
Rnf12
f
Tsix expression
XX
XY
wild type
g
wild type
F1 129 cas
clones
6 18 43 dki Rnf12+/-
cas
129
129
h
clones
6 18 43
wild type
F1 129 cas
i
day 0
day 2
day 4
day 6
day 8
XX
Rnf12+/-
Clone 2
Rnf12+/-
Clone 3
Rnf12+/-
Clone 43



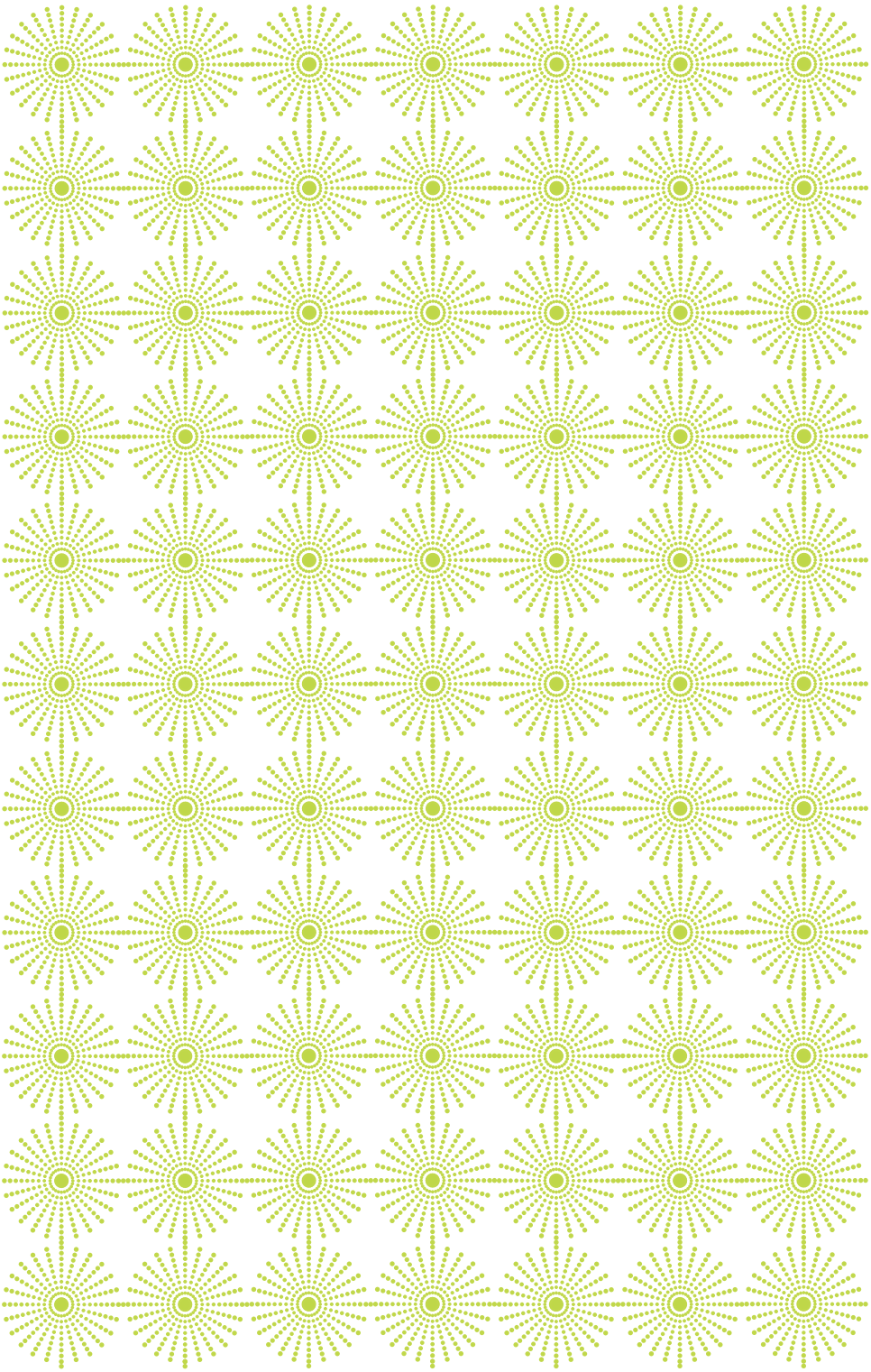
Supplementary Figure 6. RNA expression analysis of XGTC-XO, mCherry low and high subpopulations.

(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.

#	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

Supplementary Table 1.

Primers used in this study as listed in the Materials and Methods section.



CHAPTER .3

THE EFFICIENCY OF Xist-MEDIATED SILENCING OF X-LINKED AND AUTOSOMAL GENES IS DETERMINED BY THE GENOMIC ENVIRONMENT

Agnese Loda[1], Johannes H. Brandsma[2], Ivaylo Vassilev[3,4], Nicolas Servant[3,4], Friedemann Loos[5], Azadeh Amirnasr[1], Erik Splinter[6], Raymond A. Poot[2], Edith Heard[3*] and Joost Gribnau[1*]

Manuscript in preparation

THE EFFICIENCY OF Xist-MEDIATED SILENCING OF X-LINKED AND AUTOSOMAL GENES IS DETERMINED BY THE GENOMIC ENVIRONMENT.

Agnese Loda[1], Johannes H. Brandsma[2], Ivaylo Vassilev[3,4], Nicolas Servant[3,4], Friedemann Loos[5], Azadeh Amirnasr[1], Erik Splinter[6], Raymond A. Poot[2], Edith Heard[3*] and Joost Gribnau[1*]

[1]Department of Developmental Biology, Erasmus University Medical Center, Wytemaweg 80, 3015 CN Rotterdam, The Netherlands
[2]Department of Cell Biology, Erasmus University Medical Center, Wytemaweg 80, 3015 CN Rotterdam, The Netherlands
[3]Mammalian Developmental Epigenetics group, Institut Curie, CNRS UMR 3215, INSERM, U934, Paris, France
[4]Bioinformatics and Computational Systems Biology of Cancer, INSERM, Paris, France
[5]Equipe 11 labellisée par la Ligue Nationale contre le Cancer, Centre de Recherche des Cordeliers, Paris, France
[6] Cergentis B.V., Padualaan 8, 3584 CH Utrecht, The Netherlands.
*To whom correspondence should be addressed. Email: j.gribnau@erasmusmc.nl, e.heard@curie.fr

ABSTRACT

Xist is indispensable for X chromosome inactivation (XCI) in female mammalian cells. However, how Xist RNA directs chromosome-wide transcriptional inactivation of the X chromosome is largely unknown. Here, to study chromosome inactivation by *Xist*, we generated a system where ectopic *Xist* expression can be induced from several genomic contexts in aneuploid mouse ES cells. We found that ectopic *Xist* expression from any location on the X chromosome faithfully recapitulated endogenous XCI, showing the potency of *Xist* to initiate XCI. Genes that escape XCI remain consistently transcriptionally active upon ectopic XCI, regardless of their position relative to *Xist* transgenes, and we implicated the enrichment of CTCF at their promoters in directing XCI escape. *Xist* expression from autosomes facilitates their transcriptional silencing to different degrees, and gene density in proximity of the *Xist* transcription locus is instructive in determining the efficiency of gene inactivation. Enrichment of LINE elements together with a specific chromatin environment correlates with more efficient Xist-mediated silencing of both X-linked and autosomal genes. These findings provide new insights into the epigenetic mechanisms that mediate XCI and identify genomic features that promote Xist-mediated chromosome-wide gene inactivation.

3

INTRODUCTION

In mammals, dosage compensation of sex chromosomal genes between females (XX) and males (XY) is achieved through X chromosome inactivation (XCI). XCI starts with the monoallelic upregualtion of the X-linked non-coding *Xist* gene from one of the two X chromosomes in female cells and culminates in the conversion of one X chromosome into a silent heterochromatic entity known as the Barr body (Xi) (Gendrel and Heard, 2014; Mira-Bontenbal and Gribnau, 2016; Wutz, 2011). During this process, Xist RNA spreads *in cis* on the chromosome from which it is transcribed and recruits chromatin modifying complexes involved in trancriptional inactivation. Initially upon XCI, the Xi is depleted of euchromatic histone modifications such as H3K4me2/me3 and H3/H4 acetylation. Subsequently, Polycomb Repressive Complexes 1 and 2 (PRC1 and PRC2) are recruited, and Xi is enriched for repressive marks such as H3K27me3 and H2AK119ub (Chaumeil et al., 2006; Heard et al., 2001; Keohane et al., 1996; O'Neill et al.). After XCI has taken place, the silent state of the inactive X chromosome is clonally propagated through cell division.

Although *Xist* is the major player of the XCI process (Marahrens et al., 1997; Penny et al., 1996; Borsani et al., 1991; Brockdorff et al., 1991), the molecular mechanisms by which Xist RNA spreads along the entire length of the X chromosome and triggers gene silencing remain intriguing open questions. Many X to autosome translocation studies showed either inefficient Xist spreading or incomplete inactivation of the autosomal translocated material (Russell, 1963; Cattanach, 1974; White et al., 1998; Popova et al., 2006), suggesting a sequence specific model for Xist spreading *in cis*. In this context, long interspersed elements (LINE or L1) have been proposed to work as "way stations" for X-linked specific Xist spreading (Lyon, 1998). LINE elements are enriched on both human and mouse X chromosomes relative to autosomes, take part in the formation of the Xi silent compartment and were shown to facilitate silencing of X-linked genes that are more prone to escape XCI (Chaumeil et al., 2006; Chow et al., 2010). However, LINE-rich areas of the genome often correspond to gene-poor areas, whereas gene-rich areas are depleted of repetitive elements. Thus, the preferential spreading of Xist RNA in LINE-rich regions of autosomes may reflect negative selection against cells in which Xist-mediated silencing leads to functional aneuploidy (Sharp et al., 2002; Popova et al., 2006; Chow et al., 2010; Tang et al., 2010; Bala Tannan et al., 2014; Cotton et al., 2014). In addition, specific regions of the X chromosome that are initially targeted by Xist upon XCI are not enriched for LINE elements, thus questioning the role of LINE elements in conferring X chromosome specificity to Xist spreading (Simon et al., 2013; Engreitz et al., 2013).

Although XCI leads to chromosome-wide silencing of one entire X chromosome, around 12-20% and 3-7% of human and mouse X-linked genes escape from *Xist* silencing and remain expressed from both the active and inactive X chromosomes within the same nucleus (Balaton and Brown, 2016). Genes that escape XCI lack the silent epigenetic marks typical of inactivated genes and retain active histone marks such as H3K4me2 and -me3, H3K9ac and H3K27ac

(Goto and Kimura, 2009; Sadreyev et al., 2013). To date, how these genes can maintain their active transcriptional state within the silent Xi is largely unknown. Escaping genes are located outside the Xist RNA domain (Chaumeil et al., 2006; Chow et al., 2010; Splinter et al., 2011) were suggested to be intrinsically competent to resist XCI (Mugford et al., 2014; Li and Carrel, 2008) and to be flanked by *in cis* acting elements that protect neighbouring genes from escape (Horvath et al., 2013). The chromatin insulator protein CTCF has been proposed to play a role in XCI escape both as a boundary element between active and inactive X-linked loci (Filippova et al., 2005) or as an anchor that allows looping out of specific active domains from the condensed heterochromatic Xi territory (Berletch et al., 2015; Heard and Bickmore, 2007). Unravelling Xist's functions is critical to a complete understanding of XCI. Here, to address the mechanism(s) directing Xist-mediated silencing, we set up a doxycycline responsive *Xist* expression system in mouse ESC lines. By inducing ectopic XCI from several genomic regions in both karyotypically normal and abnormal ESC lines, we discovered that: (I) Xist's silencing efficiency is locus dependent, (II) specific X-linked but not autosomal loci are intrinsically prone to become inactivated or to escape XCI, (III) LINE-1 elements facilitate gene silencing but are unlikely to work as X-specific way stations, and (IV) CTCF plays a X-specific role in directing XCI escape.

RESULTS

Generation of an inducible *Xist* expression system in mouse ES cells.

To assess the efficiency of Xist-mediated silencing from several X-linked and autosomal contexts, we set up a doxycycline-responsive expression system in polymorphic F1 2-1 hybrid ESC lines (129/Sv-Cast/Ei). First, we generated an Xist-inducible transgene using a Cast/Ei BAC covering 300 kb of the X chromosome including the *Xist* endogenous locus. Through homologous recombination in bacteria (Barakat et al., 2011) a 1 kb region upstream of *Xist* TSS was replaced with a targeting cassette carrying a Ptight bidirectional doxycycline-responsive promoter driving *Xist* and a DsRed reporter gene (Figure 1). Next, the *Xist* inducible transgene was transfected into F1 2-1 (129/Sv-Cast/Ei) female ESC lines in which the reverse tetracycline transactivator M2rtTA had been targeted at the ubiquitously expressed ROSA26 locus (Figure S1). Neomycin-resistant ES colonies were screened by either RFLP-PCR or DNA FISH, and four sets of transgenic female ESC lines were selected: (1) Tg-E clones, in which the *Xist* transgene was targeted to the Xist endogenous locus of the Cast/Ei X chromosome of a wild type ESC line (40,XX) (Figure 1A). (2) Tg-X;8 and (3) Tg-X clones, in which the *Xist* transgene was randomly integrated on either X chromosomes of a karyotypically abnormal ESC line carrying two intact copies of chromosome 8 and a duplication of the distal two-thirds of the Cast/Ei chromosome 8 which is fused to the 129/Sv X chromosome in the resulting X;8 translocation (40,XX,t(X;8)). In Tg-X;8 clones the *Xist* transgene is integrated at different loci on the X to 8 translocation product, whereas Tg-X clones carry the *Xist* transgene integrated at several po

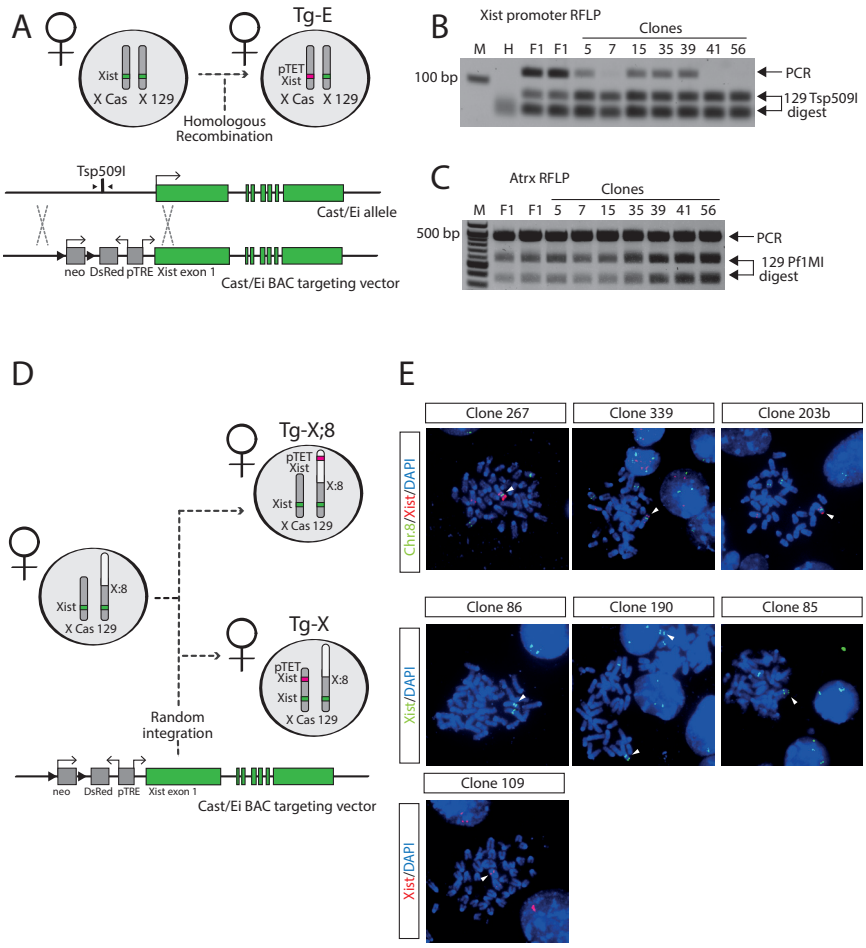
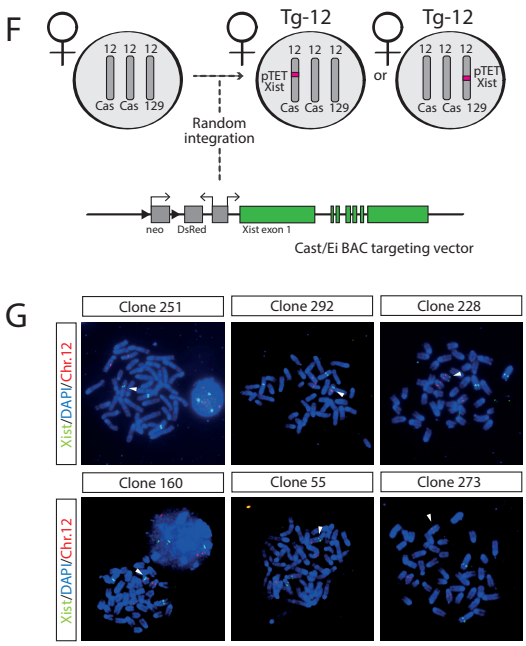


Figure 1. Generation of a tetracycline-responsive *Xist* expression system in ES cells.

(A) Targeting strategy to generate Tg-E clones. 1 kb upstream of *Xist* TSS on the Cast/Ei allele of hybrid F1 2-1 ES cells was replaced with a bidirectional tetracycline-responsive Ptight promoter, a DsRed reporter gene and a neomycin/kanamycin resistance cassette flanked by lox sites. The neo cassette was looped out after transient expression of Cre recombinase.(B) PCR amplification with primers indicated in Figure 1A and Tsp509I RFLP digest of PCR product to identify clones with a correctly targeted Cast/Ei allele. Correct targeting results in loss of Cast/Ei-specific band, as shown for clones 7, 41 and 56. Arrows on right indicate size of PCR product and Tsp509I restriction fragments. F1, F1 2-1 polymorphic 129/Sv-Cast/Ei mother cell line; M, marker; H, water. (C) PCR amplification of a fragment of the X-linked gene *Atrx* and Pf1MI digest of the PCR product to verify the presence of two X chromosomes in the targeted clones. Arrows on right indicate size of PCR product and sixe of Pf1MI restriction fragments. (D and F) Schematic representation of the strategy used to generate Tg-X, Tg-X;8 (D) and Tg-12 clones (F). The Xist transgene was randomly integrated into F1 2-1 40,XX,t(X;8) or 41,XX,dup12 ESC lines and neomycin resistant clones were screened by DNA FISH. (E and G) DNA FISH screening of Tg-X Tg-X;8 (E) and Tg-12 (G) clones. Arrows indicate the chromosome carrying *Xist* transgene. Positive Tg-X clones show two DNA FISH signals on the wild type Cast/Ei X corresponding to the endogenous *Xist* locus and to the ectopic transgene, respectively. Four independent ES clones were generated (85, 86, 190, 109). Positive Tg-X;8 clones show



co-localization of three DNA FISH signals on the translocated X;8 chromosome corresponding to the *Xist* endogenous locus, the ectopic transgene, and the chromosome 8 portion of the X;8 translocation. Three independent ES clones were generated (339, 203b, 267). Positive Tg-12 clones show co-localization of two DNA FISH signals on one of the three copies of chromosome 12, corresponding to the *Xist* inducible transgene and to chromosome 12, respectively. Six independent ES clones have been generated (251, 160, 273, 292, 55, 228).

sitions along the wild type Cast/Ei X chromosome (Figure 1B-1C). Contrarily to Tg-E clones, in these clones both the endogenous *Xist* alleles are intact. Finally, Tg-12 clones (4), in which the *Xist* transgene was randomly integrated at different loci of one copy of chromosome 12 in a trisomic ESC line (41,XX,dup12)(Figure 1F-1G).

Xist RNA induction from different genomic contexts leads to *in cis* spreading on chromosomes X and autosomes.

Next, we asked whether ectopic *Xist* could be efficiently expressed upon doxycycline induction and whether the induced Xist RNA could spread in cis on chromosomes X, 12 and 8. To this end, all ESC clones were grown in ESC medium supplemented with doxycycline for five days. By inducing ectopic XCI in undifferentiated ESCs we were able to uncouple Xist function from cell differentiation, thus allowing the efficiency of Xist spreading to be assessed independently of any selection on cell viability. Xist RNA could be ectopically expressed from both chromosomes X and autosomes, with the inducible system showing no leakiness of *Xist* expression in the absence of doxycycline (Figure 2A-2D). Tg-X;8 clone 267 and Tg-12 clone

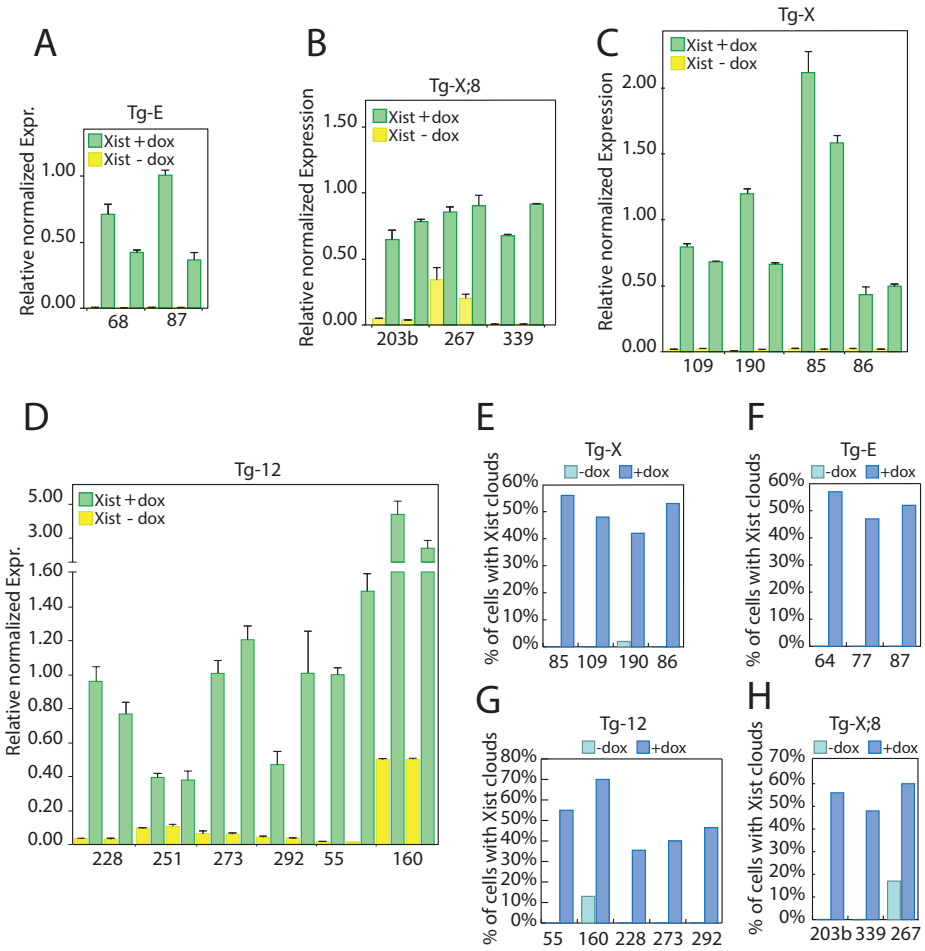
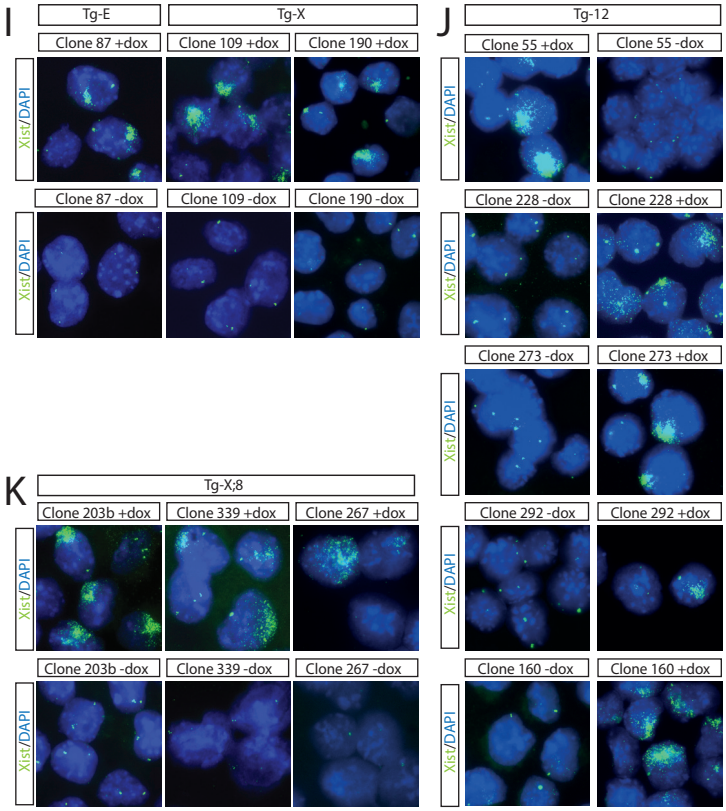


Figure 2. Ectopic Xist RNA induction in undifferentiated ES cells.

qPCR quantification of Xist RNA in clones (A) Tg-E (68, 87), (B) Tg-X;8 (203b, 267, 339), (C) Tg-X (109, 190, 86, 85), and (D) Tg-12 (228, 251, 273, 292, 55-38, 160) after five days of doxycycline treatment. (E-H). Quantification of cells showing an Xist-coated chromosome in (E) Tg-X (109, 190, 85, 86), (F) Tg-E (64, 77, 87), (G) Tg-12 (55, 228, 273, 292, 251), and (H) Tg-X;8 (203b, 339, 267) clones after five days of doxycycline induction. n>100 nuclei counted per ES clones. (I-K) Representative images of Xist RNA-FISH analysis of Tg-E (87), Tg-X(109), Tg-X;8 (203b, 339) and Tg-12 (55, 228, 273, 292) clones after five days of doxycycline treatment. Xist, FITC; DNA is stained with DAPI (blue).

160 carry a previously described Tsix Stop allele on the 129/Sv X chromosome, which explains the higher level of basal *Xist* expression in untreated conditions (Luikenhuis et al., 2001). Overall, Xist RNA enrichment in doxycycline treated cells versus untreated cells varies from 10- to 250-fold in between different ESC lines (Figure 2A-2D). In spite of this variability, the enrich-



ment of ectopic Xist in ES clones is either comparable or higher than the one reached by endogenous *Xist* upon neuronal differentiation of untreated ES cells (Figure S2A). To assess whether ectopic Xist RNA is stable and can spread *in cis* from different X and autosomal loci, we performed Xist RNA FISH on all Tg-E, Tg-X, Tg-X;8, and Tg-12 ES clones after 5 days of doxycycline induction in undifferentiated ESC cells (Figure 2E-2K). Overall, 50-70% of induced cells showed an Xist-coated chromosome (Figure 2E-2H) which was stable over time (Figure S2B-S2C). Interestingly, ectopic Xist clouds of several Tg-12 and Tg-X:8 clones seem to appear morphologically less compact and more dispersed throughout the nucleus compare to Tg-E Xist clouds (unpublished observations). Taken together, this data demonstrate that we set up a robust *Xist* expression system to separate Xist function from genomic context and cell differentiation and without aneuploidy-related phenotype.

Xist-mediated gene inactivation efficiency is locus dependent.

To assess whether ectopic Xist RNA could trigger gene silencing independently of its genomic position, we performed RNA-seq analysis of Tg-E, Tg-X, Tg-X;8 and Tg-12 clones after five days of doxycycline treatment in undifferentiated ESC cells. Amongst Tg-12 clones, clone 251 showed low levels of *Xist* induction and was left out of the analysis (Figure 2D-2F and S2C). We generated 51 RNA-seq libraries including at least two biological replicates per clone in both doxycycline- treated and untreated conditions. Total RNA-seq reads were aligned to both the 129/Sv and the Cast/Ei parental genomes of F1 hybrid ESC and the abundance of allele-specific reads was estimated as previously described (Gendrel, et al., 2014). For each gene in our dataset, we used the total counts of 129/Sv (N_{129}) and Cast/Ei (N_{Cast}) allele-specific reads to obtain the ratio of Cast-specific gene expression (N_{Cast})/(N_{Cast} + N_{129}). Our analysis was restricted to 12-, 8- and X- linked genes. In Tg-E clones, in which the inducible transgene is targeted at the *Xist* endogenous locus of the Cast/Ei X chromosome, the overall X-linked gene expression changes from biallelic expression (ratio=0.5) in untreated cells to a more 129/Sv-monoallelic gene expression (ratio < 0.5) in doxycycline- treated cells, showing efficient ectopic XCI in ES cells when *Xist* is induced from its endogenous locus on the X chromosome (Figure 3A). Similarly, in Tg-X and Tg-X;8 clones, X-linked gene expression upon doxycycline induction shifts from biallellic (ratio=0.5) to either more 129- or Cast- monoallelic expression according to which of the two X chromosomes carries the inducible *Xist* transgene. Thus, Tg-X clones 85, 86, 109, 190, show inactivation of the wild type Cast/Ei X chromosome (ratio < 0.5), whereas in Tg-X;8 clones 203b, 267 and 339 the 129/Sv X chromosome portion of the X;8 translocation product is inactivated upon doxycycline *Xist* induction (ratio > 0.5) (Figure 3B-3C). Notably, clone 267 is already biased towards Cast-monoallelic gene expression in doxycycline- untreated conditions. This bias is due to the lack of repression of endogenous *Xist* from the 129 X chromosome carrying the Tsix-Stop allele present in this line and in Tg-12 160 (Figure S3A) (Luikenhuis et al., 2001). Next, we tested whether ectopic Xist RNA could silence autosomal genes, focusing on chromosomes 12 and 8 gene expression upon *Xist* induction in Tg-12 and Tg-X;8 clones, respectively. Tg-12 clones carry three chromosomes 12, one of 129/Sv and two of Cast/Ei origin. Therefore, the overall chromosome 12 gene expression ratio of Cast-specific expression is close to 0.66 in doxycycline untreated Tg-12 clones. This allele-specific ratio shifts in either one or the other direction according to which of the three chromosome 12 carries the *Xist* inducible transgene. Notably, Xist-mediated silencing of chromosome 12 genes is more heterogeneous and in several lines less efficient compared to what we observed for chromosome X. Clones 160 and 55, carrying an *Xist* transgene on one of the two Cast/Ei chromosome 12 and on the single 129/Sv chromosome 12, respectively, show the highest efficiency of gene silencing upon doxycycline treatment (Figure 3D), whereas for clones 228, 273 and 292 induced gene silencing is poor (Figure 3D). In contrast, all Tg-X;8 clones show silencing of autosomal genes (Figure 3E). Therefore, gene expression of the trisomic portion

of chromosome 8 in clones 203b, 339 and 267 shift from a Cast-specific expression ratio of 0.66 in doxycycline- untreated cells to more biallelic gene expression upon *Xist* induction. Further, chromosome 8 gene expression does not show any change upon *Xist* induction in Tg-X clones, thus excluding any impact of doxycycline treatment on gene expression (Figure S3B). However, in all three Tg-X;8 clones the overall inactivation of the autosomal portion of the X to 8 translocation product is less robust compared to the X chromosome counterpart (compare Figure 3C with 3E). The RNA-seq results were validated by allele-specific RFLP RT-PCR analysis of both autosomal and X-linked genes (Figure S3E-3H). Taken together, this data suggests that ectopic Xist RNA can inactivate X chromosomal genes independently of the locus from which it is forced to spread, whereas in an autosomal context Xist's silencing ability differs and is dependent on the transgene integration site. For the autosomal portion of the X;8 translocation product we found that this is consistently inactivated independently of the integration site of the *Xist* transgene, whereas chromosome 12 gene silencing is overall poor, with only two Tg-12 clones out of five showing efficient autosomal gene inactivation.

In all experiments performed so far, ESC lines were grown in ESC culture media supplemented with leukemia inhibitory factor (LIF) and inhibitors of the MAPK and Gsk3ß pathways ("2i" culture conditions). ESC grown in 2i conditions are stabilized in an homogenous pluripotent ground state (Ying et al., 2008) whereas the lack of inhibitors shifts ESC to a more heterogeneous population (Graf and Stadtfeld, 2008; Wray et al., 2010). Furthermore, in undifferentiated ESCs ectopic XCI relies on *Xist* expression and gene inactivation becomes irreversible only upon ESC differentiation (Wutz and Jaenisch, 2000). Based on these observations, we asked whether 2i culture conditions and ESC differentiation might have an impact on Xist RNA silencing efficiency, especially in those Tg-12 clones that did not show robust gene inactivation. First, we set up a time course experiment in which Tg-E and Tg-12 ESC clones were grown in conventional serum+LIF culture conditions supplemented with doxycycline for six days. Allele-specific RFLP RT-PCR for X-linked and chromosome 12 genes was performed at different time points upon doxycycline induction (Figure 3F-3G and Figure S3I-S3J). In Tg-E clones, the X-linked genes *G6pdx* and *Mecp2* were consistently silenced at each tested time point (Figure 3F), whereas chromosome 12 genes inactivation was observed only in Tg-12 clone 55 although not for all tested genes. Second, we differentiated all ESC clones into neurons and EBs (Figure S4A-S4B). Upon neuronal differentiation of Tg-E clones, the X-linked genes *Mecp2* and *G6pdx* showed increased skewing toward monoallelic- 129/Sv expression in doxycycline- treated cells compared to ectopic XCI triggered in undifferentiated ESCs (day 0) (Figure 3H). In contrast, when we followed the allele-specific expression of chromosome 12 genes throughout neuronal differentiation of Tg-12 clones we found that only clones 55 and 160 show consistent gene inactivation upon doxycycline treatment (Figure 3I). Allele-specific Xist qPCR confirmed that Xist-inducible transgenes remained expressed throughout ESC differentiation in all Tg-12 clones (Figure 3J). Similar results to the neuronal differentiation exper

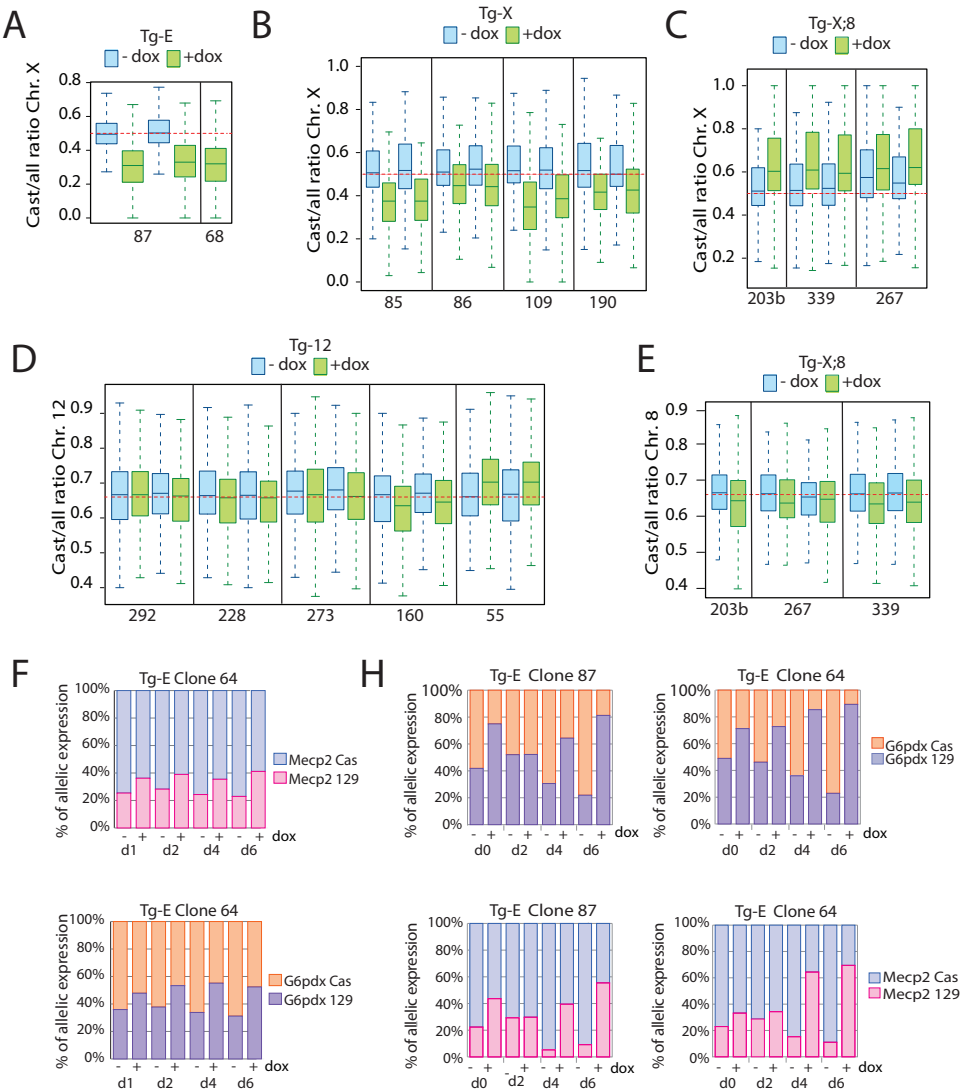
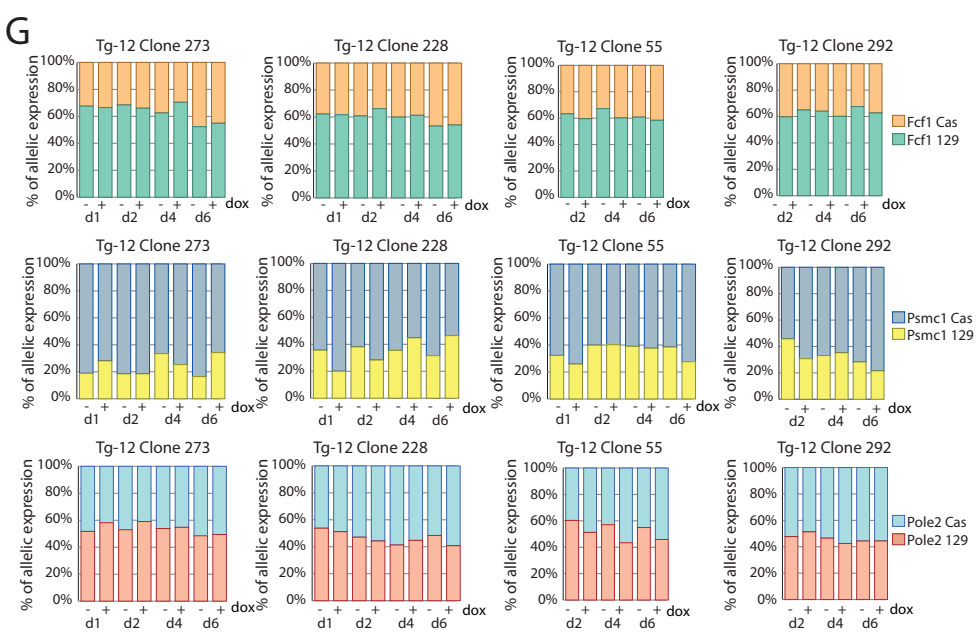
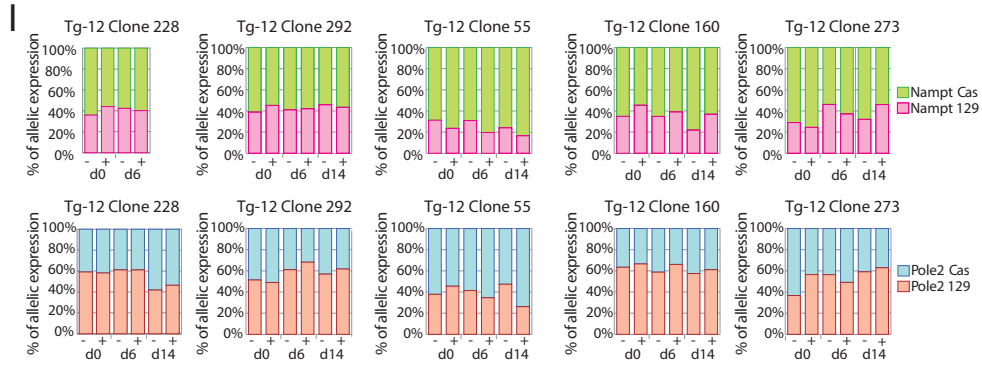
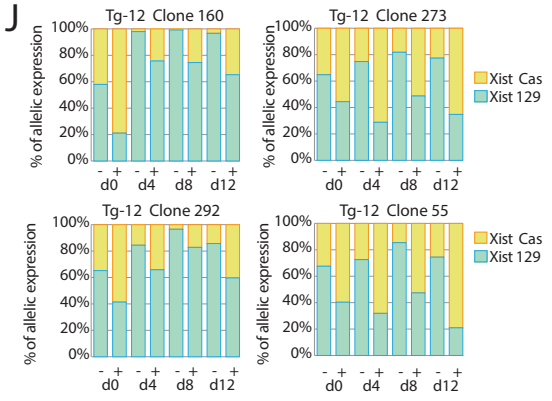


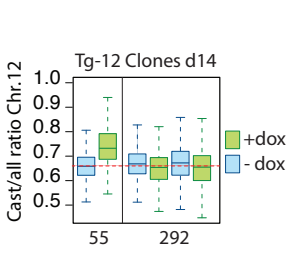
Figure 3. Xist mediated silencing at different genomic loci.

Box plot showing the Cast- specific gene expression ratio of X-linked genes for Tg-E (A) clones and Tg-X (B) and Tg-X;8 (C) clones. (D) Box plot showing the Cast- specific gene expression ratio of chromosome 12 genes in Tg-12 clones. (E) Box plot showing the Cast- specific gene expression ratio of chromosome 8 genes in Tg-X;8 clones. (F and G) Allele-specific RFLP RT-PCR analysis at different time points after doxycycline induction in undifferentiated ESC clones. Data related to *Fcf1*, *Psmc1*, *Pole2*, (Chr.12); *G6pdx* and *Mecp2* (Chr. X) are shown. (H and I) Allele-specific RFLP RT-PCR analysis upon neuronal differentiation of Tg-12 and Tg-E clones. Data for *Nampt*, *Pole2* (Chr.12); *G6pdx* and *Mecp2* (Chr. X) are shown. The percentage of allele-specific expression was determined by measuring relative band intensities of restriction fragments upon PCR products enzymatic digestion using a Typhoon image scanner and Image Quant software. (J) Xist allele-specific qPCR analysis of Tg-12 clones at different time points upon neuronal differentiation. (K) Box plot showing Cast- specific gene expression ratio of chromosome 12 genes for Tg-12 clones 55 and 292 at day 14 of neuronal differentiation.


G
Tg-12 Clone 273
Tg-12 Clone 228
Tg-12 Clone 55
Tg-12 Clone 292
% of allelic expression
100%
80%
60%
40%
20%
0%
dox
d1
d2
d4
d6
Fcf1 Cas
Fcf1 129
Psmc1 Cas
Psmc1 129
Pole2 Cas
Pole2 129


I
Tg-12 Clone 228
Tg-12 Clone 292
Tg-12 Clone 55
Tg-12 Clone 160
Tg-12 Clone 273
% of allelic expression
100%
80%
60%
40%
20%
0%
d0
d6
d14
Nampt Cas
Nampt 129
Pole2 Cas
Pole2 129


J
Tg-12 Clone 160
Tg-12 Clone 273
Tg-12 Clone 292
Tg-12 Clone 55
% of allelic expression
100%
80%
60%
40%
20%
0%
d0
d4
d8
d12
Xist Cas
Xist 129


K
Tg-12 Clones d14
Cast/all ratio Chr.12
1.0
0.9
0.8
0.7
0.6
0.5
55
292
+dox
- dox

iments were obtained for *Psmc1, Fcf1, Smc6* and *Bzw2* upon differentiation of Tg-12 clones in embryo bodies (EB) culture conditions (Figure S3K). However, we cannot exclude that the poor gene silencing observed for clones 292, 228 and 273 relies on the integration sites of *Xist* transgenes along chromosome 12 relative to the few loci we tested in our single-gene silencing assay. Therefore, we performed RNA-seq analysis at day 14 of neuronal differentiation of clones 292 and 55, two clones that showed poor and robust Xist-mediated silencing of chromosome 12, respectively (Figure 3K). Chromosome-wide analysis of clone 55 confirmed efficient gene inactivation of chromosome 12 upon *Xist* induction, showing a higher degree of skewing toward Cast/Ei- monoallelic expression in neurons (Figure 3K) compared to what we observed in undifferentiated ESCs (Figure 3D). In contrast, overall inactivation of chromosome 12 in clone 292 remains poor in fully differentiated cells (Figure 3K). Thus, we conclude that ESC differentiation stabilizes Xist-mediated silencing but does not affect the variability of Xist's silencing efficiency that we observed in between different clones.

Xist-mediated inactivation of the X;8 translocation product efficiently rescues ESC clones from lethal aneuploidy.

Next, we asked whether Xist-mediated correction of autosomal aneuploidies would be beneficial for ESC cell survival upon ESC differentiation. Clones Tg-12 carry three copies of chromosome 12, whereas in Tg-X;8 and Tg-X clones the fusion between chromosomes X and 8 results in partial trisomy of chromosome 8 genes and partial monosomy of X-linked genes (Figure 4A). When we differentiated Tg-X, Tg-X;8, Tg-E and Tg-12 clones into neurons (Figure S4A-S4B), we noticed strikingly cell death of all doxycycline- treated Tg-X clones starting between day 3 and day 4 of cell differentiation, resulting in almost no cell survival at day 8. Contrarily, Tg-X;8, Tg-E, and Tg-12 clones did not show any viability phenotype between doxycycline-treated and untreated cells (Figure 4B and data not shown). These observations suggest that silencing of the wild type X chromosome in Tg-X clones led to cell death upon ESC differentiation, whereas inactivation of the X;8 translocation product would rescue these ESC clones from lethal aneuploidy. To confirm our hypothesis, we assessed the degree of XCI skewing in Tg-E, Tg-X and Tg-X;8 clones by *Xist* allele-specific qPCR analysis at different time points throughout neuronal differentiation (Figure 4C). Although in undifferentiated ESC there is high variability of Xist RNA induction (Figure 5SA), skewing of XCI is consistently unchanged in all clone categories, suggesting that relative quantity of Xist RNA does not affect Xist function. In Tg-E clones the *Xist* endogenous promoter on the Cast/Ei X chromosome is replaced by the Ptight promoter. As expected, 100% of untreated cells up-regulate the wild type Xist 129/Sv allele upon differentiation, whereas ectopic expression of the Cast/Ei Xist allele in doxycycline- treated ESC leads to robust skewing of XCI to the Cast/Ei allele, efficiently uncoupling *Xist* expression from its endogenous regulators (Figure 4C). In Tg-X clones, Xist RNA is mostly of Cast/Ei transgenic origin between d0 and d2 of differentiation whereas from day 2 onward

the relative expression of the 129/Sv allele increases (Figure 4C). This switch from transgenic Cast/Ei to endogenous 129/sv *Xist* expression reflects progressive death of cells that inactivate the wild type X chromosome. As predicted, all doxycycline- untreated Tg-X and Tg-X;8 clones show skewed inactivation of the 129/Sv X chromosome that is fused to chromosome 8 in the X;8 translocation product (Figure 4C). Similar results are shown by allele-specific RNA-seq analysis of fully differentiated Tg-X;8 neurons, confirming that Xist-mediated correction of the X;8 chromosomal rearrangement is crucial for cell survival (Figure 4D-4E). Contrarily, wild type F1 2-1 ESC shown random XCI upon neuronal differentiation (Figure 4F).

Ectopic expression of Xist RNA leads to preferential inactivation of specific DNA loci on chromosome X but not on autosomes.

Next, we tested whether there are specific genes on chromosomes X, 8 and 12 that are prone to get inactivated regardless of the Xist-inducible transgene position along the chromosome. To this end, we analyzed the allele-specific RNA-seq data in more detail. For each gene in our datasets, we first calculated the averaged allele-specific expression ratio for all clones of each category. Second, we ranked all chromosomes X, 8 and 12 genes according to their general degree of silencing upon doxycycline treatment. Ranked genes are divided in three categories: (I) genes that are efficiently silenced, (II) genes that are partially affected and (III) not silenced genes (Figure 5A-5C). Finally, we tested whether lying in a specific chromosomal region makes a gene more prone to either get inactivated or to escape ectopic Xist inactivation (Figure 5D-5F). On chromosome 12, the distribution of silenced, partially affected and not silenced genes was invariant along the entire chromosome length and reflected general gene density (Figure 5D). Similar results were obtained for chromosome 8, with the exception of a slight tendency of strongly silenced genes to be located in a discrete region of chromosome 8 (Figure 5E). On the contrary, differentially silenced genes are not homogenously distributed along chromosome X (Figure 5F). Rather, X-linked genes are organized in chromosomal blocks that behave differently in terms of gene inactivation efficiency. Thus, centromeric genes are more prone to escape ectopic XCI compared to genes located in the sub-centromeric region of the X chromosome, independently of where the *Xist* transgene integrated on chromosome X and on the X;8 translocation product (Figure 5F and Figure S6). These results suggest that there appears to be a fundamental difference between ectopic inactivation of genes of the X chromosome and autosomal genes. While all genes on chromosomes 12 and 8 have the same likelihood to get silenced, ectopic XCI follows a specific path of inactivation on chromosome X. Based on these observations, we hypothesized that if XCI is artificially induced in undifferentiated ESC, it always recapitulates endogenous XCI, independently of the locus on the X chromosome from which Xist RNA is forced to spread. Indeed, by comparing the ranked X-linked gene list with clusters of genes that show different dynamics of silencing upon ESC differentiation (Marks et al., 2015) , we found that 74% of early silenced genes correspond to

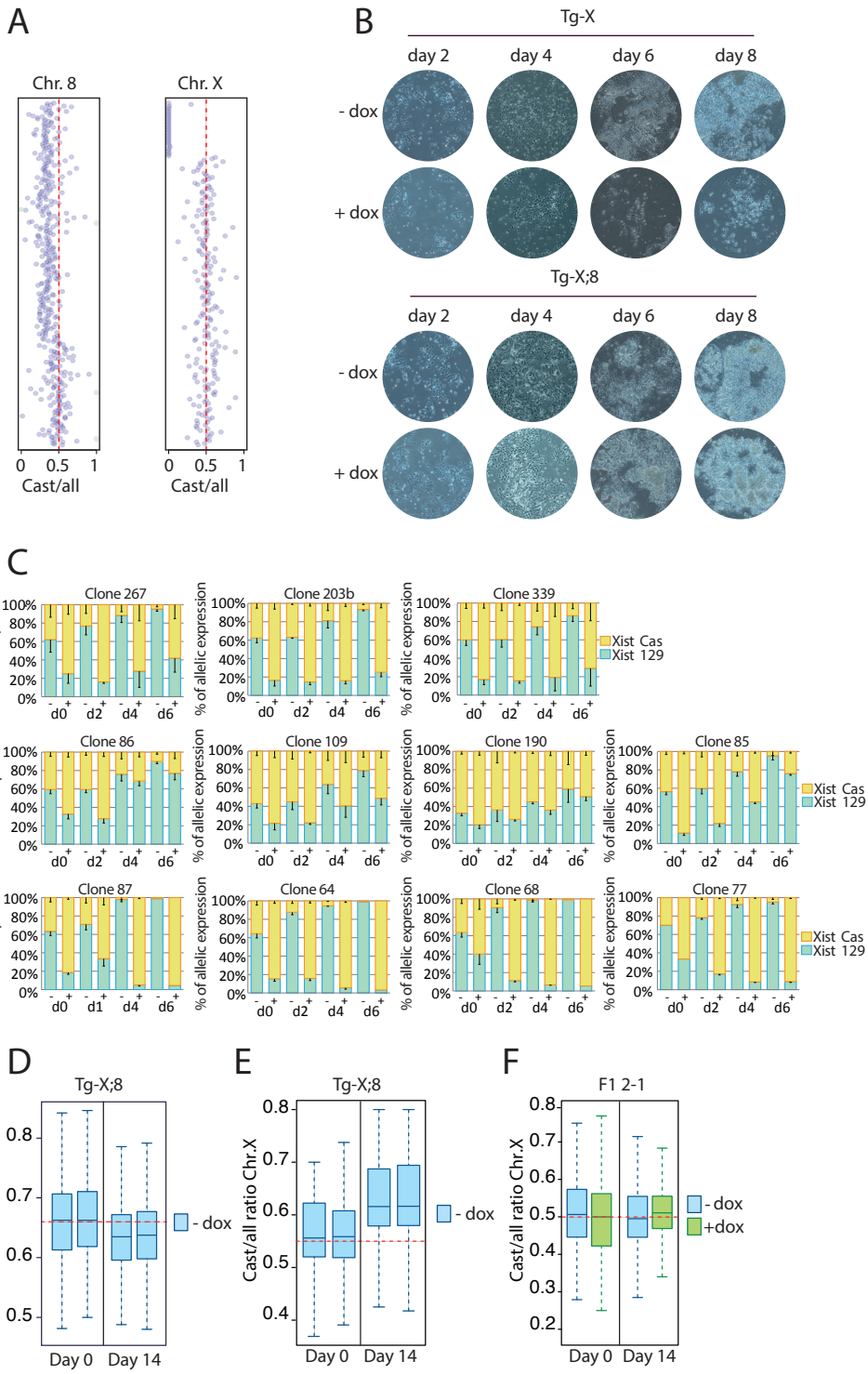
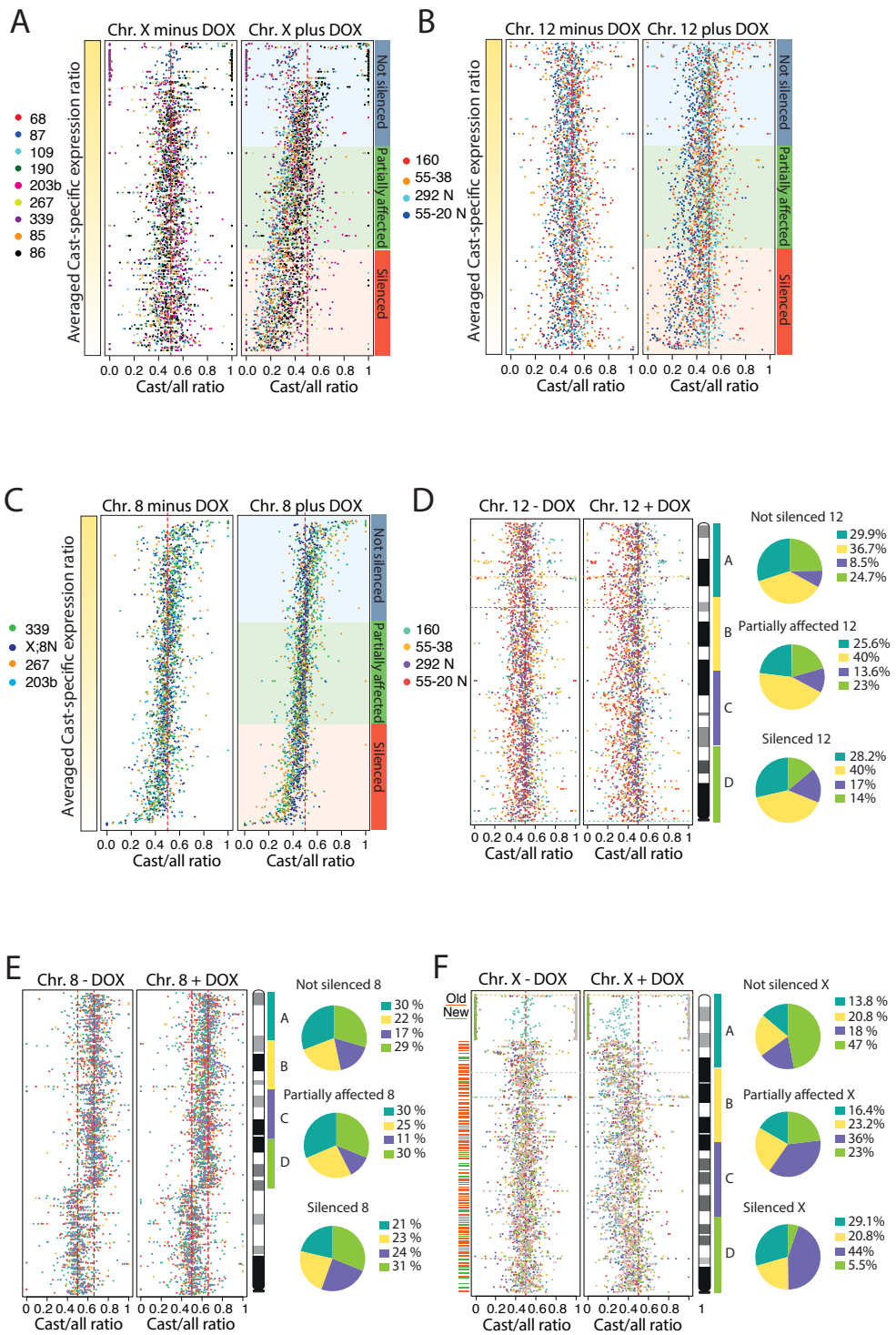

A
Chr. 8
Chr. X
0 0.5 1
Cast/all
B
Tg-X
day 2
day 4
day 6
day 8
- dox
+ dox
Tg-X;8
C
Clone 267
Clone 203b
Clone 339
Clone 86
Clone 109
Clone 190
Clone 85
Clone 87
Clone 64
Clone 68
Clone 77
Tg-X;8
Tg-X
Tg-E
% of allelic expression
Xist Cas
Xist 129
D
Tg-X;8
Cast/all ratio Chr.8
Day 0
Day 14
E
Tg-X;8
Cast/all ratio Chr.X
F
F1 2-1
+dox

Figure 4. Xist-mediated rescue of ESC clones from lethal aneuploidy.

(A) Mapping of X;8 translocation breakpoint by allele-specific RNA-seq analysis. The distal two-thirds of chromosome 8 are duplicated in the X;8 translocation product whereas 15 Mb of the telomeric end of the 129/Sv X chromosome are lost. (B) Neuronal differentiation of Tg-X;8 and Tg-X clones. Doxycycline-treated Tg-X clones show strikingly cell death starting from day 4 of differentiation whereas Tg-X;8 clones do not show any difference in cell viability between doxycycline- treated and untreated conditions. (C) *Xist* allele-specific qPCR analysis of Tg-X;8 (top), Tg-X (middle) and Tg-E (bottom) clones at different time points upon neuronal differentiation. Day 0 allele specific analysis refers to total *Xist* expression levels shown in Figure S5. The mean and SD of three to four independent experiments are shown. (D-E-F) Allele specific RNA-seq analysis. Box plots showing Cast- specific gene expression ratio of chromosome 8 (D) and X-linked genes (E) in Tg-X;8 clones at day 0 and 14 of neuronal differentiation. (F) Box plot of X-linked genes expression at day 0 and 14 of neuronal differentiation showing random XCI in F1 2-1 ESC clones.

strongly silenced genes in our dataset and 72% of the genes that escape XCI overlap with our "not silenced" genes category (Figure 5G). Importantly, 95% of the early silenced genes resulted to be at least partially affected in out dataset and only 5% of the reported escaping genes are becoming inactivated upon ectopic XCI (Fugure 5G).

X-linked genes can be classified based on the X chromosome evolutionary history: "Old" genes are those found on chicken orthologous autosomes 1 and 4 whereas "New" genes were exclusively added to the mammalian X chromosome (Bellott et al., 2010). On the human X chromosome, old genes are found in both the X conserved region (XCR) and the X added region (XAR) (Ross et al., 2005). On the mouse X chromosome, the XCR and XAR regions are rearranged leading to a different distribution of old and new genes along the chromosome (Deng et al., 2014). Here, we mapped old and new X-linked genes on the mouse X chromosome and we show that evolutionary new X-linked genes are enriched in the centromeric region of the X chromosome that is more prone to escape inactivation (Figure 5F and Figure 5H). Thus, our data confirms that the evolutionary history of the X chromosome also play a role in establishing the path of X-linked gene inactivation in XCI, as previously observed for the human X chromosome (Carrel and Willard, 2005; Ross et al., 2005).


A
Chr. X minus DOX
Chr. X plus DOX
Averaged Cast-specific expression ratio
68
87
109
190
203b
267
339
85
86
Not silenced
Partially affected
Silenced
0.0 0.2 0.4 0.6 0.8 1
Cast/all ratio
B
Chr. 12 minus DOX
Chr. 12 plus DOX
160
55-38
292 N
55-20 N
C
Chr. 8 minus DOX
Chr. 8 plus DOX
339
X;8N
267
203b
D
Chr. 12 - DOX
Chr. 12 + DOX
A
B
C
D
Not silenced 12
29.9%
36.7%
8.5%
24.7%
Partially affected 12
25.6%
40%
13.6%
23%
Silenced 12
28.2%
40%
17%
14%
E
Chr. 8 - DOX
Chr. 8 + DOX
Not silenced 8
30 %
22 %
17 %
29 %
Partially affected 8
30 %
25 %
11 %
30 %
Silenced 8
21 %
23 %
24 %
31 %
F
Chr. X - DOX
Chr. X + DOX
Old
New
Not silenced X
13.8 %
20.8 %
18 %
47 %
Partially affected X
16.4%
23.2%
36%
23%
Silenced X
29.1%
20.8%
44%
5.5%

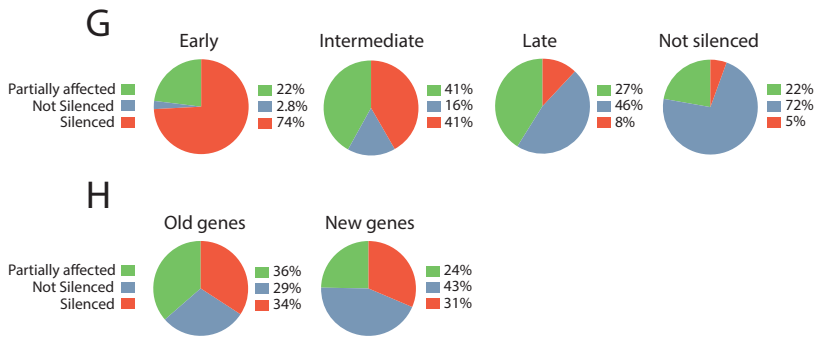


Figure 5. Preferential silencing of specific X-linked loci by ectopic Xist RNA.

Gene silencing ranking plots for X-linked (A), chromosome 12 (B) and chromosome 8 (C) genes. Every dot represents the Cast/all expression ratio for a specific gene. A total of 242 genes are shown in (A), 351 in (B) and 336 in (C). Genes are ranked based on the averaged Cast/all ratio amongst all clones in each group of clones (Tg-E, Tg-X, Tg-X;8 and Tg-12). Ranked genes are divided in three categories (I) efficiently silenced, (II) partially affected and (III) not silenced genes. To simplify data interpretation, the Cast/all expression ratios were transformed as follows: (1) For Tg-X;8 clones 339, 267 and 203b that carry the Xist- inducible transgene on the 129/Sv X chromosome, we used the reciprocal of the Cast-specific expression ratio. Thus, all Tg-X and Tg-E clones behave as *Xist* transgene was integrated on the Cast/Ei X chromosome. (2) For Tg-12 and Tg-X;8 clones, to convert the data from a trisomic to disomic we calculated a new Cast-specific gene expression ratio that takes into account the double dosage of Cast/Ei expression in these clones ($N_{Cast}/2$)/($N_{Cast}/2 + N_{129}$). (3) We transformed the Cast-specific ratio of clone 55 so that Xist-transgene results to be integrated on the Cast/Ei chromosome 12 in all Tg-12 clones. As a result of the X;8 translocation in Tg-X and Tg-X;8 clones, 15 Mb at the telomeric end of the 129/Sv X chromosome are deleted and 20 genes show either 129/Sv or Cast monoalleic expression. These genes are excluded from our analysis. (D-E-F) Genes are ordered by genomic position on chromosomes X, 12 and 8. Pie graphs show the amount of overlap between the gene categories defined in (A), (B) and (C) and chromosomal regions A, B, C, and D on chromosomes X, 12 and 8. In (F), green and red lines represent evolutionary old and new X-linked genes. Old genes are those found on chicken orthologous autosomes 1 and 4, new genes have been acquired during the evolution of the mammalian X chromosome. (G and H) Pie graphs showing the proportion of efficiently silenced, partially affected and not silenced X-linked genes overlapping with gene clusters that show different inactivation dynamics upon cell differentiation (Marks et al., 2015) and evolutionary old and new X-linked genes (Bellot et al., 2005).

Gene density in proximity of the transgene integration site is instructive for Xist spreading along the chromosome.

The variable Xist-mediated silencing efficiency that we observed at different genomic loci might rely on the *Xist* transgene integration sites along the chromosome. To assess the locality of Xist-mediated silencing, we precisely mapped the integration sites of several *Xist* transgenes on chromosomes X, 12, and 8 by targeted locus amplification (TLA) (de Vree et al., 2014). X-linked and autosomal gene expression relative to the transgene integration sites along chromosomes X and 12 show a higher degree of gene inactivation in linear proximity of *Xist* transgene (Figure 6A-6B). This proximity effect on gene silencing is independent of the overall gene inactivation that is achieved in different clones upon *Xist* induction. Clones 55, 86 and 87 efficiently inactivate chromosomes 12 and X, respectively, but genes in close proximity to the transgene integration site even showed stronger inactivation (Figure 6A-6B). Although clone 292 shows poor chromosome-wide inactivation of chromosome 12 genes in both ESC and fully differentiated neurons, expression analysis of genes lying in 4 Mb region around the Xist transgene confirmed that ectopic Xist RNA is capable of inducing local gene silencing (Figure 6B). When we compared the *Xist* integration sites in clones 55 and 292, we found that the genomic environment strongly differs in terms of gene density (Figure 6E). In clone 292 the *Xist* transgene landed in a gene desert of 2,1 Mb, whereas in clone 55 Xist transgene is located in a gene-rich region of chromosome 12. Similarly, in all Tg-X ang Tg-X;8 clones *Xist* transgene was integrated in gene-dense areas of chromosomes X and 8 (Figure 6E-6F). Interestingly, all Tg-X and Tg-X;8 clones showed comparable levels of X-linked and chromosome 8 gene inactivation whereas amongst five different Tg-12 clones only clone 55 shows robust chromosome-wide silencing of autosomal genes. Thus, we hypothesized that the distribution of genes along the entire length of the chromosome makes chromosomes X and 8 more prone to become inactivated compared to chromosome 12. Indeed, when we looked at gene distribution along chromosomes X, 8 and 12, we found that on chromosome 8 and X genes are organized in discrete blocks of gene-rich areas, whereas genes on chromosome 12 show a more homogenous distribution with the exception of a few gene-rich blocks. In particular, the proximal 70 Mb of chromosome 12 that correspond to more than half of the entire chromosome length, carries less than 10 genes per 0.5 Mb. These gene-poor regions only made up no more than 20 Mb of chromosome X and 8 (Figure 6C-6E-6G).

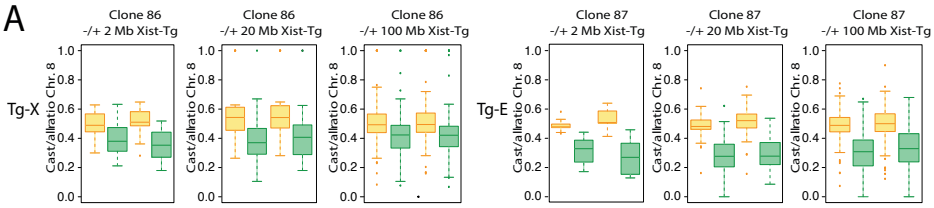
LINE-1 elements facilitate Xist-mediated silencing of both X-linked and autosomal genes.

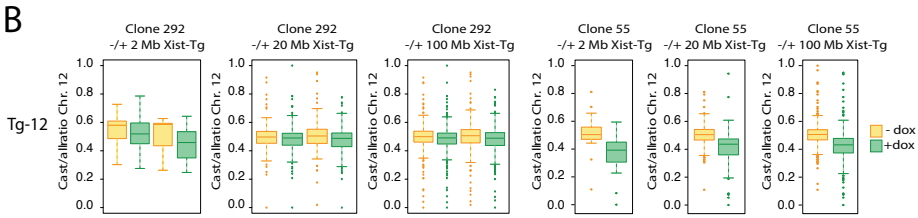
To assess whether LINE-1 elements facilitate Xist-mediated silencing, we looked at the distribution of "efficiently silenced" and "not silenced" X-linked and autosomal genes along chromosomes X, 12 and 8. We found a positive correlation between genes that are efficiently silenced and LINE density for all tested chromosomes (Figure 6I). Contrarily, not silenced genes cluster in LINE-poor regions of chromosome X, 12 and 8 (Figure 6I). To exclude that

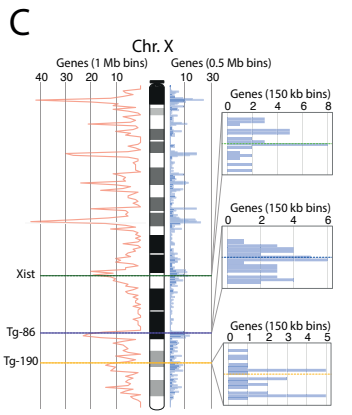
this correlation relies on the relative distribution of genes and repetitive elements along the chromosome, we performed the same analysis for short interspersed elements (SINEs). We did not find any positive correlation between SINE elements and efficiently silenced genes on chromosomes X, 8, and 12 (Figure 6J). Rather, SINE elements enrichment correlates with not silenced genes on both chromosomes X and 12 (Figure 6J). Since the X chromosome is strongly enriched for LINE-1 elements, it is nearly impossible to study Xist spreading from a LINE-poor area of the X (Figure 6 and Figure 7SA). However, in both Tg-12 clone 55 and Tg-X;8 clones 267 and 203b Xist transgenes landed in LINE-poor regions of chromosome 12 and 8 (Figure 6 and Figure S7) but chromosome-wide silencing of autosomal genes is efficiently achieved in these clones. Therefore, L1 elements are enriched around genes that are efficiently inactivated, but LINE density in close proximity of the *Xist* transgene integration sites does not affect Xist's spreading efficiency.

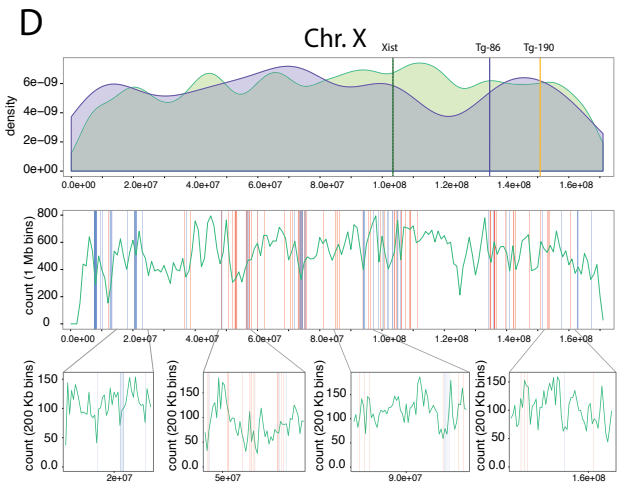
The ESC chromatin environment predisposes both X-linked and autosomal genes to efficient Xist-mediated silencing.

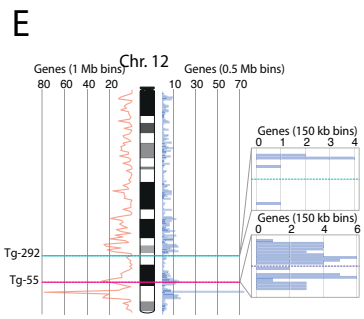
Next, we asked whether the chromatin environment of X-linked and autosomal genes in ESC prior to ectopic *Xist* induction predispose a specific gene to be either efficiently silenced or to escape ectopic inactivation. Therefore, we looked at the enrichment of both euchromatic and heterochromatic histone marks around the TSS sites of X-linked and autosomal genes after ranking them based on the degree of inactivation upon doxycycline induction (Figure 7A). To estimate the density of H3K27me3, EZH2, H3K4me3, H3K27ac and Ring1b 4 kb upstream and downstream the TSS of (I) efficiently silenced, (II) partially affected and (III) not silenced genes we used published ChIP-seq data obtained in undifferentiated male ESC lines (Ku et al., 2008; Creyghton et al., 2010; Mikkelsen et al., 2007; Das et al., 2014; Blackledge et al., 2014). Strongly inactivated genes on both chromosomes X and 12 show enrichment of H3K27me3 and Ring1b, and are depleted of active marks such as H3K4me3 and H3K27ac around their TSS prior to inactivation. In addition, the promoters of 27% and 32% of efficiently silenced X-linked and chromosome 12 genes, respectively, showed a poised chromatin state, marked by both H3K4me3 and H3K27me3 and not exclusively by H3K27me3 (data not shown). These findings indicate that the chromatin landscape guides Xist-mediated silencing.

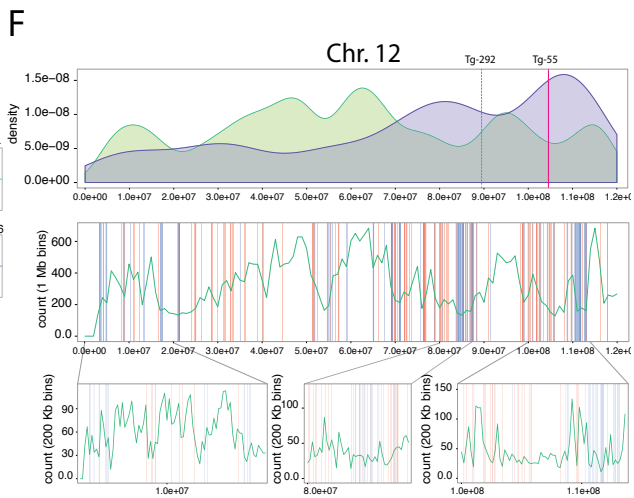

A
Clone 86 -/+ 2 Mb Xist-Tg
Clone 86 -/+ 20 Mb Xist-Tg
Clone 86 -/+ 100 Mb Xist-Tg
Clone 87 -/+ 2 Mb Xist-Tg
Clone 87 -/+ 20 Mb Xist-Tg
Clone 87 -/+ 100 Mb Xist-Tg
Tg-X
Tg-E
Cast/allratio Chr. 8


B
Clone 292 -/+ 2 Mb Xist-Tg
Clone 292 -/+ 20 Mb Xist-Tg
Clone 292 -/+ 100 Mb Xist-Tg
Clone 55 -/+ 2 Mb Xist-Tg
Clone 55 -/+ 20 Mb Xist-Tg
Clone 55 -/+ 100 Mb Xist-Tg
Tg-12
Cast/allratio Chr. 12
- dox
+dox


C
Chr. X
Genes (1 Mb bins)
Genes (0.5 Mb bins)
Genes (150 kb bins)
Xist
Tg-86
Tg-190


D
Chr. X
Xist
Tg-86
Tg-190
density
count (1 Mb bins)
count (200 Kb bins)


E
Chr. 12
Genes (1 Mb bins)
Genes (0.5 Mb bins)
Genes (150 kb bins)
Tg-292
Tg-55


F
Chr. 12
Tg-292
Tg-55
density
count (1 Mb bins)
count (200 Kb bins)

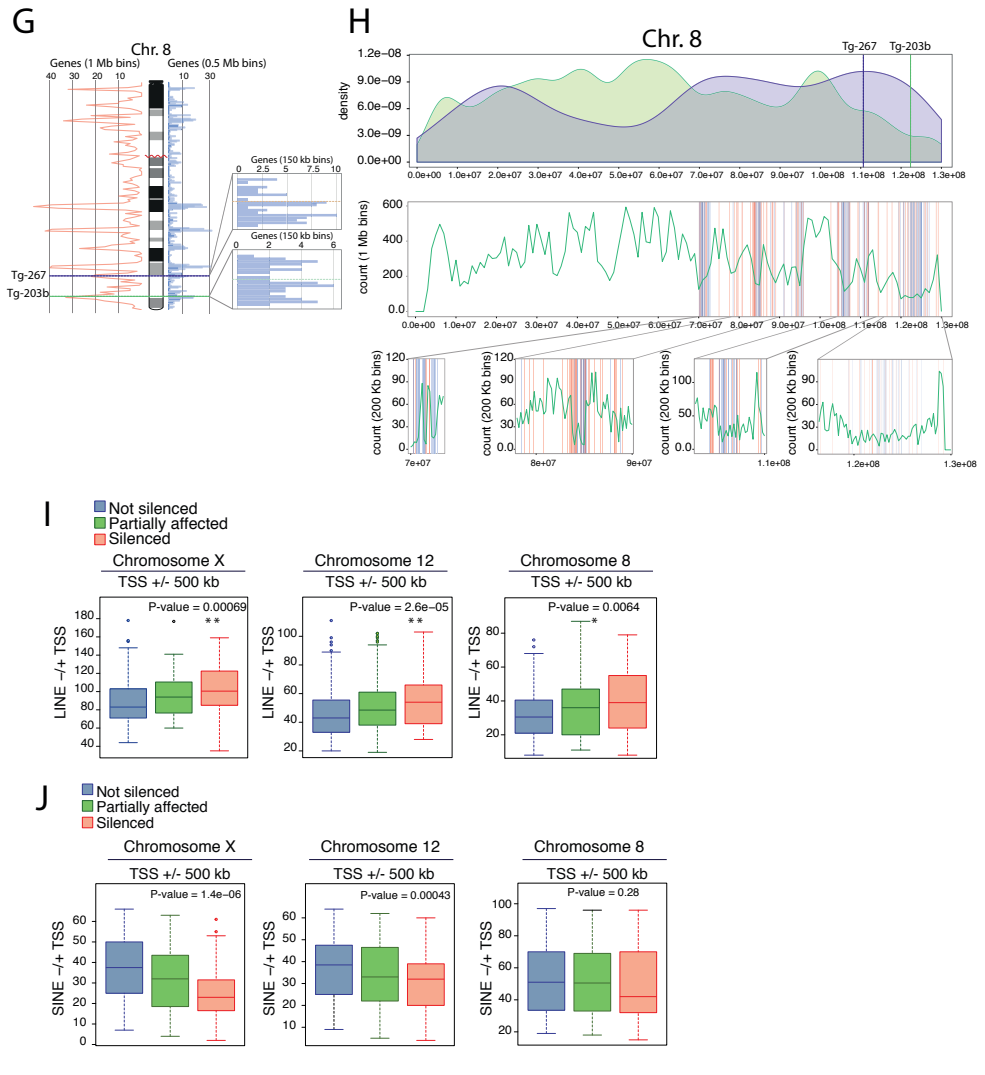


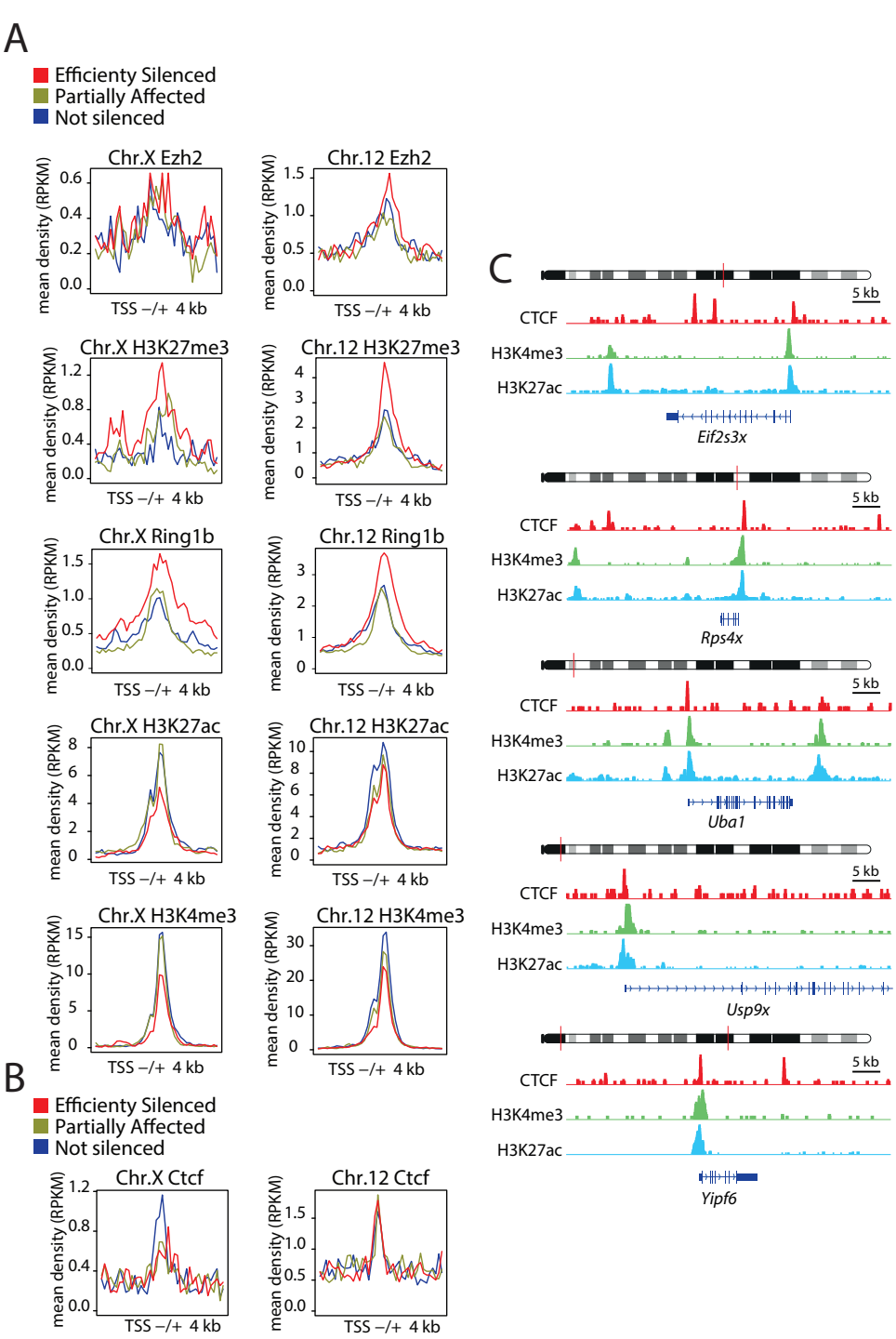
Figure 6. Xist transgenes position along the chromosome**,** gene density and LINE enrichment.

(A and B) Box plots showing the Cast/all ratios of X-linked (A) and chromosome 12 (B) genes in 2, 20 and 100 Mb bins around the Xist transgene integration sites. (C, E, G) Gene density along the entire length of chromosomes X (C), 12 (E), and 8 (G) is shown. Blue histogram bars represent 0.5 Mb bins, red frequency lines correspond to gene distribution in 1 Mb bins along the chromosomes. The integration sites of *Xist* transgenes are indicated and zoom in of the integration loci is shown in small boxes. Blue histogram bars represent 150 kb bins. (D, F, H) Top: LINEs density relative to gene density on chromosome X (D), 12 (F) and 8 (H) is shown. Gene density is shown in blue, LINE density is shown in green. Middle: LINEs distribution along chromosome X (D), 12 (F) and 8 (H) is shown, green frequency lines correspond to LINEs distribution in 1 Mb bins along the chromosomes. Blue and red lines indicate "not silenced" and "efficiently silenced" genes defined in Figure 5, respectively. Bottom: zoom in of specific loci. Green frequency lines correspond to LINE distribution in 200 kb bins. (I) Box plot showing correlation between LINE density and gene silencing. Data for chromosomes 12, 8 and X are shown. *p < 0.05 and **p < 0.005 Wilcoxon rank-sum test.

Using published data on ESCs CTCF profiles (Sleutels et al., 2012), we also assessed the enrichment of CTCF around the TSS of both X-linked and autosomal genes ranked by silencing efficiency upon *Xist* induction. Our analysis showed enrichment of CTCF at the TSS of "not silenced" X-linked genes relative to partially affected and fully silenced genes. This enrichment was absent for autosomal genes escaping XCI, pointing to an X chromosome specific role for CTCF in mediating escape from XCI (Figure 7B). Indeed, if active loci were causative of CTCF binding, CTCF would be enriched at the TSS of both X-linked and autosomal not silenced genes. Moreover, these observations also suggest that specific loss of CTCF binding around the TSS of X-linked might facilitate Xist-mediated silencing. Amongst the X-linked genes that show enrichment of CTCF at their TSS we could detect eight genes that were previously reported to escape XCI. Five of them, *Eif2s3x, Yipf6, Uba1, Rps4x* and *Usp9x* were defined as escaping genes in at least two independent studies (Berletch et al., 2015; Marks et al., 2015; Wu et al., 2014; Li et al., 2016; Carrel and Willard, 2005; Calabrese et al., 2012; Yang et al., 2010), and three of them, *Usp11, Haus7* and *Apoo* were reported to escape in one study (Li et al., 2016). Since the category of X-linked "not silenced" genes is defined by ranking the average allele-specific ratio of each gene amongst all ESC clones (Figure 5A), not silenced genes correspond to those genes than always tend to escape XCI independently of where Xist RNA is induced to spread from: (I) its endogenous locus (Tg-E clones), (II) multiple loci along the wild type Cast/Ei X chromosome (Tg-X clones), and (III) multiple loci on the autosomal portion of the X;8 translocation product (Tg-X;8). Thus, finding previously described escaping genes in this category further confirms their resistance to get inactivated in spite of Xist position along the chromosome.

Figure 7. Chromatin environment and efficiency of Xist-mediated silencing.

Average density plots for chromatin features (A) and CTCF (B) in 8 kb bins around the TSS of (I) efficiently silenced genes (red lines), (II) partially silenced genes (green lines) and (III) not silenced genes (blue lines). Data for EZH2, H3K27me3, Ring1b, H3K4me3 and H3K27ac are shown for both chromosomes X and 12. (C) ChIP-seq showing CTCF, H3K4me3 and H3K27ac enrichment at the TSS of five not silenced genes: *Eif2s3x, Rps4x, Uspx9, Uba1 and Yipf6.*


A
Efficienty Silenced
Partially Affected
Not silenced
Chr.X Ezh2
Chr.12 Ezh2
Chr.X H3K27me3
Chr.12 H3K27me3
Chr.X Ring1b
Chr.12 Ring1b
Chr.X H3K27ac
Chr.12 H3K27ac
Chr.X H3K4me3
Chr.12 H3K4me3
mean density (RPKM)
TSS –/+ 4 kb
B
Efficienty Silenced
Partially Affected
Not silenced
Chr.X Ctcf
Chr.12 Ctcf
C
5 kb
CTCF
H3K4me3
H3K27ac
Eif2s3x
Rps4x
Uba1
Usp9x
Yipf6

DISCUSSION

In this study, we developed an inducible *Xist* expression system that overcomes many of ¬the limitations that have previously hampered the study of Xist's function. First, using female ESC lines we can ectopically induce XCI for several days in ESC lines without triggering cell death. Second, by controlling ectopic *Xist* expression from different genomic locations in isogenic clones we can directly compare the efficiency of Xist-mediated silencing between sets of ESC clones that differ only in terms of *Xist* transgene integration site on chromosomes X, 12 and 8. Third, to study Xist function in an autosomal context, we exploit ESC clones that carry a complete trisomy of chromosome 12 and an unbalanced X;8 translocation. Silencing of one chromosome 12 in Tg-12 clones as well as inactivation of the X;8 translocation product in Tg-X;8 clones will not be lethal, similar to silencing of one X chromosome in female cells. This strategy allowed us to exclude that lack of autosomal inactivation reflects negative selection against those cells in which Xist-mediated silencing was robust enough to lead to lethal functional aneuploidy. Rather, inactivation of autosomal genes in our inducible system will rescue both aneuploidies of Tg-12 and Tg-X;8 clones. Finally, the high density of SNPs between 129/Sv and Cast/Ei mouse strains (Keane et al., 2011) allowed us to follow gene expression prior to and after Xist induction by allele-specific next generation sequencing of messenger RNA (RNA-seq) (Gendrel et al., 2014). We demonstrated that Xist RNA can be efficiently induced to spread *in cis* from different genomic locations, but found that its ability to trigger gene silencing is locus dependent. On the X chromosome, ectopic XCI can be triggered from different loci and always leads to efficient chromosome-wide gene inactivation. Importantly, ectopic *Xist* expression in undifferentiated ESC is sufficient to faithfully recapitulate the endogenous XCI process that only occurs upon differentiation, thus confirming that Xist is able trigger gene silencing independently of ESC differentiation (Wutz et al., 2000). Xist's robust ability to silence X-linked genes is highlighted by the strong degree of silencing found in the Tg-X;8 clones, in which Xist is forced to spread from the distal autosomal portion of the X;8 translocation product. In addition, Xist-mediated silencing of X-linked genes follows a gene-specific path of inactivation: (I) although our RNA-seq analysis was performed after five days of doxycycline induction in ESC lines rather than at different time points throughout the experiment, the X-linked "efficiently silenced" genes of our dataset highly overlap with the genes that become inactivated first upon differentiation of F1 2-1 female ESC lines (Marks et al., 2015). Again a high overlap was observed between X-linked genes that were reported to escape endogenous XCI (Marks et al., 2015). (II) The distribution of genes along chromosome X is not random but is the consequence of the mammalian sex chromosome evolution (Ross et al., 2005). Evolution of chromosome X has been heavily influenced by XCI, as evolutionary younger X-linked genes of the human X have been shown to be prone to escape XCI compared to the oldest genes (Carrel and Willard, 2005). Here, we found the same tendency that evolutionary younger genes are more likely to escape XCI when *Xist* is induced from different

locations on the mouse X chromosome. In conclusion, our data show that even when XCI is artificially induced in ESC, Xist's silencing function is perfectly recapitulated, independent of the transgene integration site.

The specific path of silencing that we described for X-linked genes strongly differs from the autosomal genes, in which we did not observe any preferential inactivation of specific genes or chromosomal regions. Overall, we found that ectopic Xist RNA cannot silence autosomal genes with the same versatility as shown for chromosome X. Three out of five Tg-12 clones showed poor chromosome-wide silencing of chromosome 12 genes in both ESC conditions and upon neuronal differentiation. These results are in line with many X;autosome translocation studies in which autosomal genes have been reported to be inactivated inefficiently in somatic cells (Russell, 1963; Disteche et al., 1979; Cattanach, 1974; White et al., 1998; Searle et al., 1983). However, all tested clones of our *Xist* inducible system show local gene silencing in close proximity of the transgene integration site, and efficient silencing of autosomal material has been shown in other *Xist* inducible systems as well (Wutz and Jaenisch, 2000; Tang et al., 2010; Chow et al., 2010). Therefore, we conclude that Xist RNA has the ability to potentially spread and inactivate any chromosome *in cis* but whether the process is efficient relies on the genomic environment in the vicinity of *Xist* transcription starting locus. Based on the Xist RNA localization pattern along the X chromosome during XCI, Xist's spreading was proposed to follow a two step mechanism, initially targeting gene-dense areas (Simon et al., 2013) with a preference for genes located in spatial proximity to the *Xist* transcription locus (Engreitz et al., 2013). However, none of these studies addressed X-linked gene silencing upon XCI, and mapping of Xist RNA to a discrete region of the X does not necessarily mean that gene inactivation has taken place. Here, we found that Xist RNA does silence genes located in close proximity to start spreading *in cis*, but can efficiently induce chromosome-wide silencing only when the initially targeted regions provide a favorable environment. In this context, our data suggests that gene density plays an important role: if Xist RNA starts spreading in close proximity of a gene-poor region, chromosome-wide gene silencing will be poor. This dependency between transgene integration sites and silencing efficiency may explain why chromosome 12 is not equally inactivated when Xist RNA is induced from different loci. According to the distribution of genes on chromosomes X, 12 and 8, random integration of *Xist* transgenes on chromosomes X and 8 will result in a higher probability to integrate in gene-rich areas compared to chromosome 12, which in turn leads to higher silencing efficiency of X-linked and chromosome 8 genes. Gene-dense areas might be more prone to become coated by Xist RNA because of their transcriptional state. Xist RNA becomes highly expressed from a chromosome that is being inactivated, and targeting discrete blocks of genes might facilitate Xist's ability to ultimately compacting all genes into the newly formed repressive compartment. In line with these observations, the robust inactivation of one entire copy of chromosome 21 in human iPS derived from Down syndrome patients was achieved after induction of XIST from a

gene-rich core of chromosome 21 (Jiang et al., 2013).

According to both high-resolution maps of Xist RNA localization on chromosome X, LINE-rich domains become coated by Xist RNA only at a late stage of XCI (Simon et al., 2013; Engreitz et al., 2013). Here, contrarily to several studies performed in somatic cells in which the role of LINE was assessed in the maintenance of XCI (Cotton et al., 2014; Bala Tannan et al., 2014), we addressed the question whether repetitive elements facilitate gene silencing at the time of Xist spreading. We found a consistent positive correlation between the efficiency of gene inactivation and LINE enrichment for both X-linked and autosomal genes. Such a positive correlation was not found for SINE elements, highlighting the specific role of LINEs in facilitating Xist silencing function. Interestingly, all autosomal *Xist* transgenes that led to robust chromosome-wide inactivation of chromosome 8 and 12 were integrated in LINE-poor areas of the autosomes. Therefore, we propose that LINE elements facilitate the propagation of silencing along the chromosome rather than working as X-linked specific binding sites for Xist spreading *in cis*. Since LINE-rich areas correlate with gene-poor areas of the genome, LINE enrichment might be beneficial for the inactivation of those genes that are located in gene-poor areas and become targeted by Xist RNA at a late stage of XCI.

We also suggest that the ESC chromatin environment at the TSS of both X-linked and autosomal genes plays a role in determining *Xist's* silencing efficiency. Genes that are better inactivated by Xist RNA show enrichment of both PRC1 and PRC2 components before *Xist* induction. Previously, allele-specific ChIP-seq analysis of PRC2 components in hybrid ESC lines showed X-chromosome specific acquisition of EZH2 sites upon ESC differentiation (Pinter et al., 2012). The majority of the acquired sites correspond to bilavent domains enriched for both H3K4me3 and H3K27me3. However, about half of these acquired X-linked sites were marked exclusively by H3K27me3 in undifferentiated ESC. Therefore, the H3K27me3 enrichment observed at the TSS of efficiently silenced genes might either be involved in Xist recruitment or might facilitate additional recruitment of PRC2 by Xist's RNA. More importantly, since PRC2 and PRC1 positively influence each others recruitment (van Kruijsbergen et al., 2015), and Xist RNA has been reported to directly or indirectly interact with both repressive complexes (Da Rocha et al., 2014; Chu et al., 2015), H3K27me3 and H2AK119ub at the TSS of efficiently silenced genes may work as initial docking stations for Xist RNA on chromosomes X and 12.

Finally, by using an unbiased approach that is independent of Xist's integration site along the chromosome we define both X-linked and autosomal genes that are prone to resist XCI in our expression system. We found many of the known escapees in our "not silenced" category of X-linked genes, thus confirming that specific X-linked loci have the intrinsic ability to consistently escape XCI (Li and Carrel, 2008; Yang et al., 2010; Carrel and Willard, 2005). Importantly, we found enrichment of CTCF at the TSS of X-linked but not autosomal genes that escape ectopic inactivation. Although escaping genes have already been found to co-localize with CTCF binding clusters on the Xi (Berletch et al., 2015), it has been very difficult to address

whether CTCF binding itself triggers escape from XCI or whether the transcriptional activity of escaping genes is causative of the CTCF binding. We found that lower enrichment of CTCF at the TSS of strongly silenced genes compared to not silenced genes was observed exclusively for X-linked genes, suggesting that upon XCI evolution loss of CTCF binding from the Xi might have facilitated chromosome-wide silencing, whereas specific maintenance of CTCF binding at the TSS of escaping genes might have protected them from becoming inactivated. Indeed, transcriptionally active escaping genes were reported to establish *in cis* contacts at the periphery of the heterochromatic Xi (Splinter et al., 2011), and XCI is associated with a collapse of topologically associated domains (TAD) which boundary regions are often enriched for CTCF (Rao et al., 2014; Deng et al., 2015; Minajigi et al., 2015; Gibcus and Dekker, 2013). In line with this hypothesis, escaping genes represent the only X-linked loci that resist Xist-mediated erasure of the active X chromosome structure upon XCI (Giorgetti et al., 2016). Contrarily, autosomal genes have not been selected in favor or against their degree of inactivation upon Xist RNA spreading, thus explaining the lack of differential CTCF enrichment at the TSS of strongly inactivates and not silent autosomal genes.

In summary, our data support a model according to which gene density in proximity of Xist transcription site is crucial for Xist function whereas genomic features such as LINE elements facilitate silencing of gene-poor areas but do not confer X-specificity to Xist spreading. The chromatin landscape of genes subjected to XCI might recruit Xist RNA in proximity of genes that will become inactivated upon Xist spreading. Finally, loss of CTCF binding at the TSS of X-linked and autosomal genes might facilitate gene silencing whereas the X-specific 3D chromosomal organization most likely allows efficient silencing of X-linked genes into the topologically unorganized Xi and CTCF-mediated escape of specific loci.

EXPERIMENTAL PROCEDURES

Recombinant BACs construction.

The X-linked Cast/Ei BAC CH26-171B21 containing the *Xist* endogenous locus was modified by bacteria-mediated homologous recombination as previously described (Barakat et al., 2011). The pTRE-DsRed-3-5-NEO targeting vector was constructed starting from pTRE-Tight-BI-DsRed2 (Clontech). Homology arms and kanamycin/neomycin resistance cassette flanked by lox sites were amplified by PCR from BAC CH26-171B21 and TOPO-Xist-GFP-NEO using primers listed in Table S1. Correctly recombined pTRE-Xist-CH26-171B21 BACs were screened by PCR using primers listed in Table S1. Similarly, the M2rtTA transactivator was targeted at the ROSA26 endogenous locus of the 129/Sv BAC RP24-140O11. The M2rtTA-ROSA26-NEO targeting vector was constructed starting from the TOPO-KanaNEO plasmid. 3' and 5' homology arms were PCR amplified from BAC RP24-140O11 and ROSA26-m2rtTA-Puro-Amp plasmid using primers listed in Table S1. Correctly recombined BACs R26-M2rtTA-RP24-140O11 were screened by PCR using primers listed in Table S1.

ES cell culture and transgenic ESC lines generation.

ESCs were grown either in standard serum+LIF ESC medium as previously described (Monkhorst et al., 2008), or feeders-free in 2i+LIF conditions containing DMEM, 100 U/ml penicillin/streptomycin, 20% KnockOut Serum Replacement (Gibco), 0.1 mM NEAA, 0.1 mM 2-mercaptoethanol, 5000 U/ml LIF, 1μM MEK inhibitor PD0325901 (Stemgent) and 3 μM GSK3 inhibitor CH99021 (Stemgent). Transgenic ESC lines were generated using polymorphic F1 2-1 hybrid ESC lines (129/Sv-Cast/Ei). ESC clones were transfected with 30μg of recombinant BACs as previously described (Barakat et al., 2011). R26-M2rtTA-RP24-140O11 carrying the reverse tetracycline transactivator M2rtTA was targeted to the 129/Sv chromosome 6 of the F1 2-1 ESC lines. Loss of a MnlI RFLP upon homologous recombination was used to screen drug-resistant clones for correct targeting events with primers listed in Table S1. Similarly, BAC pTRE-Xist-CH26-171B21 was used to target the endogenous locus of the Cast/Ei X chromosome thus generating the Tg-E clones. Loss of a Tsp509I RFLP was used to screen drug-resistant clones for correct targeting events with primers listed in Table S1. Correctly targeted clones were screened by PCR for loss of one X chromosome by a Pf1MI RFLP located in the X-linked gene Atrx using primers listed in Table S1. To generate clones Tg-12, Tg-X and Tg-X;8 F1 2-1 ESC lines 40,XX,t(X;8) and 41,XX,dup12 were transfected with BAC pTRE-Xist-CH26-171B21 and neomycin-resistant clones were screened by DNA-FISH. To induce Xist expression, ESC medium was supplemented with 2μg/ml doxycycline.

Fluorescent In Situ Hybridization

For DNA FISH, methanol-acetic acid fixed cells were dropped on glass slides and incubated at 37°C 24 hours. Slides were washed 10 minutes in 2X SSC buffer at 55°C (1XSSC: 0.15 M NaCl, 0.015 M sodium citrate), and 5 minute in 2X SSC buffer at room temperature before being dehydrate in a gradient of 70%, 90% and 100% EtOH. Nick-labeled DNA probes (DIG or BIO Nick-translation kit, Roche) were dissolved in hybridization mixture (50% formamide, 10% dextrane, 2X SSC, pH=7.5) and 100 ng/μl mouse Cot-1 DNA (Thermo Fisher Scientific) to a final concentration of 1 ng/μl. The probe mixture was applied to the cells, covered with a glass coverslip, incubated 3' at 75°C and let cooling down for 30 minutes on the heating plate after having turned it off. Slides were then incubated overnight at 37°C in a humid chamber filled with 50% formamide in 2X SSC buffer. After hybridization, slides were washed 10 minutes in 2X SSC buffer at room temperature, 2 times 10 minutes in 0.1X SSC buffer at 55°C and 10 minutes in low salt buffer at room temperature (100 mM Tris, 150 mM NaCl, 0.05% Tween). Detection was done by incuba-tion with FITC-labeled anti-digoxigenin antibody (Roche, 11207741910) and Alexa594-labeled Streptavidin (Thermo Fisher Scientific, S11227) in low salt buffer containing 1% of low fat milk for 60 minutes at 37°C. Slides were washed 10 minutes in low salt buffer and mounted with ProLong Gold Antifade with DAPI (Molecular Probes). The following BACs were used as probes: CH26171B21 (Chr. X), RP23477B14 (Chr. 8) and RP24112A14 (Chr.

12).

Xist RNA FISH was performed as previously described with minor modifications (Monkhorst et al., 2008). ESCs 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. Slides were washed twice with PBS and dehydrated in a gradient of 70%, 90%, and 100% EtOH. The Xist probe was a 5.5 kb BglII cDNA fragment covering Xist exon 3-7. The probe was dig-labeled (DIG Nick-translation kit, Roche) and dissolved in hybridization mixture (50% formamide, 2XSSC, 50 mM phosphate buffer (pH 7.0), 10% dextran sulfate) and 100 ng/µl mouse Cot-1 DNA (Thermo Fisher Scientific) to a final concentration of 1 ng/µl. After 5 minutes of denaturation, the probe was pre-hybridized for 45 min at 37°C, and slides were incubated in a humid chamber filled with 50% formamide in 2X SSC buffer at 37°C overnight. After hybridization, slides were washed once in 2X SSC, 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 incuba¬tion with anti-digoxigenin (Boehringer) and two FITC-labeled antibodies (Roche) in blocking buffer for 30 min at room temperature. Coverslips were washed twice with TST between de¬tection steps and once finally with TS (0.1 M Tris, 0.15 M NaCl). Dehydrated cover¬slips were mounted with ProLong Gold Antifade with DAPI (Molecular Probes).

ESCs differentiation

ESCs grown in conventional serum+LIF conditions were pre-plate on cell culture dishes for 40 minutes and then seeded on feeders-free gelatin-coated culture dishes containing EB differentiation medium (IMDM-glutamax, 15% fetal calf serum, 100 U/ml penicillin/streptomy¬cin, 0.1 mM NEAA and 50 µg/ml ascorbic acid). During differentiation, the culture medium was refreshed daily. To induce neuronal differentiation, ESCs grown in feeders-free 2i conditions were seeded at a density of 1*104 cells/cm2 on 10 µg/ml laminin-coated (Sigma Aldrich L2020) dishes in neuronal differentiation medium (50% Neurobasal (Gibco), 50% DMEM:F12 (Gibco), 100 U/ml penicillin/streptomy¬cin, 1% N-2 supplement (Gibco 17502-048), 2% B27 supplement (Gibco 17504-044), 2 mM L-Glutamine (Gibco 25030081), and 0.05 mM 2-mer¬captoethanol (Gibco 31350010). Cells were refreshed daily and neuronal medium was supplemented with 250 nM retinoic acid between day 4 and day 8 of differentiation. After eight days, cells were split at low density and further cultured in neuronal medium to obtain fully differentiated neurons around day 12-14 of differentiation.

Expression analysis

Cells were lysed by direct addiction of 500 µg of TRIZOL and total RNA was extracted according to the manufacturer's instructions (Invitrogen). To remove genomic DNA contamination,

samples were treated 15 minutes at 37°C with DNaseI (Invitrogen). Next, 1 µg of RNA was reverse transcribed by Superscript II reverse transcriptase with random hexamers (Invitrogen). For quantitative PCR (qPCR) and allele-specific qPCR, gene expression levels were quantified using 2x SYBR Green PCR Master Mix (Applied Biosystems) in a CFX384 Real-Time machine (Bio-Rad) with primers listed is Table S2. Expression levels were normalized to actin b using the ΔCT method. For allele-specific RT-PCR, amplicons containing restriction fragments length polymorphisms (RFLPs) were obtained with primers available upon request using Taq DNA Polymerase (Invitrogen). PCR products were digested with the indicated restriction enzymes and analyzed on 1% or 2% agarose gels stained with ethidium bromide. Allele-specific expression was determined by measuring relative band intensities using a Typhoon image scanner and ImageQuant software.

RNA sequencing

RNA samples were prepared with the Truseq RNA kit, sequenced according to the Illumina TruSeq v3 protocol on the HiSeq2000 with a single 43 bp read and 7 bp index. Allele-specific analysis was performed as previously described with minor modifications (Gendrel et al., 2014). Genes containing a single polymorphic site were included in the analysis only when they show a coverage higher than eight reads whereas for genes carrying multiple informative SNPs the coverage threshold was of five reads per polymorphic site.

REFERENCES

Bala Tannan, N., Brahmachary, M., Garg, P., Borel, C., Alnefaie, R., Watson, C. T., Thomas, N. S., and Sharp, A. J. (2014). DNA methylation profiling in X;autosome translocations supports a role for L1 repeats in the spread of X chromosome inactivation. Human Molecular Genetics 23, 1224-1236.

Balaton, B. P., and Brown, C. J. (2016). Escape Artists of the X Chromosome. 1-12.

Barakat, T. S., Rentmeester, E., Sleutels, F., Grootegoed, J. A., and Gribnau, J. (2011). Precise BAC targeting of genetically polymorphic mouse ES cells. Nucleic Acids Research 39, e121.

Bellott, D. W., Skaletsky, H., Pyntikova, T., Mardis, E. R., Graves, T., Kremitzki, C., Brown, L. G., Rozen, S., Warren, W. C., Wilson, R. K., et al. (2010). Convergent evolution of chicken Z and human X chromosomes by expansion and gene acquisition. Nature 466, 612-616.

Berletch, J. B., Ma, W., Yang, F., Shendure, J., Noble, W. S., Disteche, C. M., and Deng, X. (2015). Escape from X Inactivation Varies in Mouse Tissues. PLoS Genet 11, e1005079.

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

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

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

Calabrese, J. M., Sun, W., Song, L., Mugford, J. W., Williams, L., Yee, della, Starmer, J., Mieczkowski, P., Crawford, G. E., and Magnuson, T. (2012). Site-Specific Silencing of Regulatory Elements as a Mechanism of X Inactivation. CELL 151, 951-963.

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

Cattanach, B. M. (1974). Position effect variegation in the mouse. Genet Res 23, 291-306.

Chaumeil, J., Le Baccon, P., Wutz, A., and Heard, E. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev 20, 2223-2237.

Chow, J. C., Ciaudo, C., Fazzari, M. J., Mise, N., Servant, N., Glass, J. L., Attreed, M., Avner, P., Wutz, A., Barillot, E., et al. (2010). LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. CELL 141, 956-969.

Chu, C., Zhang, Q. C., Da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., Magnuson, T., Heard, E., and Chang, H. Y. (2015). Systematic discovery of Xist RNA binding proteins. CELL 161, 404-416.

Cotton, A. M., Chen, C. Y., Lam, L. L., Wasserman, W. W., Kobor, M. S., and Brown, C. J. (2014). Spread of X-chromosome inactivation into autosomal sequences: role for DNA elements, chromatin features and chromosomal domains. Human Molecular Genetics 23, 1211-1223.

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences 107, 21931-21936.

Da Rocha, S. T., Boeva, V., Escamilla-Del-Arenal, M., Ancelin, K., Granier, C., Matias, N. R., Sanulli, S., Chow, J., Schulz, E., Picard, C., et al. (2014). Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. Molecular Cell 53, 301-316.

Das, P. P., Shao, Z., Beyaz, S., Apostolou, E., Pinello, L., de Los Angeles, A., O'Brien, K., Atsma, J. M., Fujiwara, Y., Nguyen,

M., et al. (2014). Distinct and combinatorial functions of Jmjd2b/Kdm4b and Jmjd2c/Kdm4c in mouse embryonic stem cell identity. Molecular Cell 53, 32-48.

de Vree, P. J. P., de Wit, E., Yilmaz, M., van de Heijning, M., Klous, P., Verstegen, M. J. A. M., Wan, Y., Teunissen, H., Krijger, P. H. L., Geeven, G., et al. (2014). Targeted sequencing by proximity ligation for comprehensive variant detection and local haplotyping. Nature biotechnology 32, 1019-1025.

Deng, X., Ma, W., Ramani, V., Hill, A., Yang, F., Ay, F., Berletch, J. B., Blau, C. A., Shendure, J., Duan, Z., et al. (2015). Bipartite structure of the inactive mouse X chromosome. Genome Biology 16, 67.

Deng, X., Berletch, J. B., Di K Nguyen, and Disteche, C. M. (2014). X chromosome regulation: diverse patterns in development, tissues and disease. Nat Rev Genet 15, 367-378.

Disteche, C. M., Eicher, E. M., and Latt, S. A. (1979). Late replication in an X-autosome translocation in the mouse: correlation with genetic inactivation and evidence for selective effects during embryogenesis. Proc Natl Acad Sci U S A 76, 5234-5238.

Engreitz, J. M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E. S., et al. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science 341, 1237973.

Filippova, G. N., Cheng, M. K., Moore, J. M., Truong, J., Hu, Y. J., Di Kim Nguyen, Tsuchiya, K. D., and Disteche, C. M. (2005). Boundaries between Chromosomal Domains of X Inactivation and Escape Bind CTCF and Lack CpG Methylation during Early Development. Developmental Cell 8, 31-42.

Gendrel, A., Attia, M., Chen, C., Diabangouaya, P., Servant, N., Barillot, E., and Heard, E. (2014). Developmental Dynamics and Disease Potential of Random Monoallelic Gene Expression. Developmental Cell 28, 366-380.

Gendrel, A., and Heard, E. (2014). Noncoding RNAs and Epigenetic Mechanisms During X-Chromosome Inactivation. Annual Review of Cell and Developmental Biology 30, 561-580.

Gibcus, J. H., and Dekker, J. (2013). The Hierarchy of the 3D Genome. Molecular Cell 49, 773-782.

Giorgetti, L., Lajoie, B. R., Carter, A. C., Attia, M., Zhan, Y., Xu, J., Chen, C., Kaplan, N., Chang, H. Y., Heard, E., et al. (2016). Structural organization of the inactive X chromosome in the mouse. Nature 535, 575-579.

Goto, Y., and Kimura, H. (2009). Inactive X chromosome-specific histone H3 modifications and CpG hypomethylation flank a chromatin boundary between an X-inactivated and an escape gene. Nucleic Acids Research 37, 7416-7428.

Graf, T., and Stadtfeld, M. (2008). Heterogeneity of Embryonic and Adult Stem Cells. Stem Cell 3, 480-483.

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

Heard, E., and Bickmore, W. (2007). The ins and outs of gene regulation and chromosome territory organisation. Current Opinion in Cell Biology 19, 311-316.

Horvath, L. M., Li, N., and Carrel, L. (2013). Deletion of an X-inactivation boundary disrupts adjacent gene silencing. PLoS Genet 9, e1003952.

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

Joel B Berletch, W. M. F. Y. J. S. W. S. N. C. M. D. X. D. (2015a). Escape from X Inactivation Varies in Mouse Tissues. 1-26.

Keane, T. M., Goodstadt, L., Danecek, P., White, M. A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M., et al. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. Nature 477, 289-294.

Keohane, A. M., O'neill, L. P., Belyaev, N. D., Lavender, J. S., and Turner, B. M. (1996). X-Inactivation and histone H4 acetylation in embryonic stem cells. Dev Biol 180, 618-630.

Ku, M., Koche, R. P., Rheinbay, E., Mendenhall, E. M., Endoh, M., Mikkelsen, T. S., Presser,

A., Nusbaum, C., Xie, X., Chi, A. S., et al. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet 4, e1000242.

Li, N., and Carrel, L. (2008). Escape from X chromosome inactivation is an intrinsic property of the Jarid1c locus. Proceedings of the National Academy of Sciences 105, 17055-17060.

Li, X., Cui, X., Wang, J., Wang, Y., Li, Y., Wang, L., Wan, H., Li, T., Feng, G., Shuai, L., et al. (2016). Generation and Application of Mouse-Rat Allodiploid Embryonic Stem Cells. 1-32.

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

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

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

Marks, H., Kerstens, H. H. D., Barakat, T. S., Splinter, E., Dirks, R. A. M., van Mierlo, G., Joshi, O., Wang, S., Babak, T., Albers, C. A., et al. (2015). Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. Genome Biology, 1-20.

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T., Koche, R. P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448, 553-560.

Minajigi, A., Froberg, J. E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., Lessing, D., Payer, B., Boukhali, M., Haas, W., et al. (2015). Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. Science 349.

Mira-Bontenbal, H., and Gribnau, J. (2016). New Xist-Interacting Proteins in X-Chromosome Inactivation. Current Biology 26, R338-R342.

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

Mugford, J. W., Starmer, J., Williams, R. L., Calabrese, J. M., Mieczkowski, P., Yee, D., and Magnuson, T. (2014). Evidence for local regulatory control of escape from imprinted X chromosome inactivation. Genetics 197, 715-723.

O'Neill, L. P., Spotswood, H. T., Fernando, M., and Turner, B. M. Differential loss of histone H3 isoforms mono-, di- and tri-methylated at lysine 4 during X-inactivation in female embryonic stem cells. Biological Chemistry 389.

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature 379, 131-137.

Pinter, S. F., Sadreyev, R. I., Yildirim, E., Jeon, Y., Ohsumi, T. K., Borowsky, M., and Lee, J. T. (2012). Spreading of X chromosome inactivation via a hierarchy of defined Polycomb stations. Genome Res 22, 1864-1876.

Popova, B. C., Tada, T., Takagi, N., Brockdorff, N., and Nesterova, T. B. (2006). Attenuated spread of X-inactivation in an X;autosome translocation. 103, 7706-7711.

Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. CELL 159, 1665-1680.

Ross, M. T., Grafham, D. V., Coffey, A. J., Scherer, S., McLay, K., Muzny, D., Platzer, M., Howell, G. R., Burrows, C., Bird, C. P., et al. (2005). The DNA sequence of the human X chromosome. Nature 434, 325-337.

Russell, L. B. (1963). Mammalian X-chromosome action: inactivation limited in spread and region of origin. Science 140, 976-978.

Sadreyev, R. I., Yildirim, E., Pinter, S. F., and Lee, J. T. (2013). Bimodal quantitative relationships between histone modifications for X-linked and autosomal loci. Proceedings of the National Academy of Sciences 110, 6949-6954.

Searle, A. G., Beechey, C. V., Evans, E. P., and Kirk, M. (1983). Two new X-autosome translocations in the mouse. Cytogenet Cell Genet 35, 279-292.

Sharp, A. J., Spotswood, H. T., Robinson, D. O., Turner, B. M., and Jacobs, P. A. (2002). Molecular and cytogenetic analysis of the spreading of X inactivation in X;autosome translocations.

Simon, M. D., Pinter, S. F., Fang, R., Sarma, K., Rutenberg-Schoenberg, M., Bowman, S. K., Kesner, B. A., Maier, V. K., Kingston, R. E., and Lee, J. T. (2013). High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. Nature 504, 465-469.

Sleutels, F., Soochit, W., Bartkuhn, M., Heath, H., Dienstbach, S., Bergmaier, P., Franke, V., Rosa-Garrido, M., van de Nobelen, S., Caesar, L., et al. (2012). The male germ cell gene regulator CTCFL is functionally different from CTCF and binds CTCF-like consensus sites in a nucleosome composition-dependent manner. Epigenetics & Chromatin 5, 8.

Splinter, E., de Wit, E., Nora, E. P., Klous, P., van de Werken, H. J. G., Zhu, Y., Kaaij, L. J. T., van IJcken, W., Gribnau, J., Heard, E., et al. (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes & Development 25, 1371-1383.

Tang, Y. A., Huntley, D., Montana, G., Cerase, A., Nesterova, T. B., and Brockdorff, N. (2010). Efficiency of Xist-mediated silencing on autosomes is linked to chromosomal domain organisation. Epigenetics & Chromatin 3, 1.

van Kruijsbergen, I., Hontelez, S., and Veenstra, G. J. C. (2015). Recruiting polycomb to chromatin. International Journal of Biochemistry and Cell Biology 67, 177-187.

White, W. M., Willard, H. F., van Dyke, D. L., and Wolff, D. J. (1998). The spreading of X inactivation into autosomal material of an x;autosome translocation: evidence for a difference between autosomal and X-chromosomal DNA. Am J Hum Genet 63, 20-28.

Wray, J., Kalkan, T., and Smith, A. G. (2010). The ground state of pluripotency. Biochemical Society Transactions 38, 1027-1032.

Wu, H., Luo, J., Yu, H., Rattner, A., Mo, A., Wang, Y., Smallwood, P. M., Erlanger, B., Wheelan, S. J., and Nathans, J. (2014). Cellular Resolution Maps of X Chromosome Inactivation: Implications for Neural Development, Function, and Disease. Neuron 81, 103-119.

Wutz, A. (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat Rev Genet 12, 542-553.

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

Yang, F., Babak, T., Shendure, J., and Disteche, C. M. (2010). Global survey of escape from X inactivation by RNA-sequencing in mouse. Genome Research 20, 614-622.

Ying, Q., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature 453, 519-523.

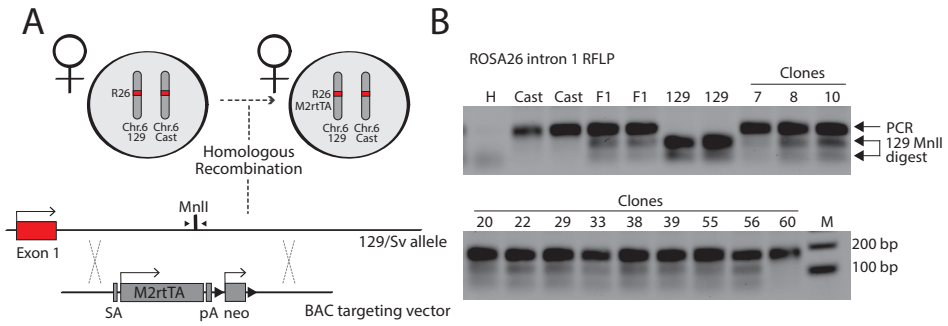


Figure S1, Related to Figure 1. Targeting of M2rtTA at the *ROSA26* endogenous locus of F1 2-1 ESC lines.

(A) Targeting strategy to generate hybrid F1 2-1 ESC lines constitutively expressing the reverse tetracycline transactivator M2rtTA. M2rtTA has been targeted at the *ROSA26* locus together with a neomycin/kanamycin resistance cassette flanked by lox sites. SA, slice acceptor; pA, polyadenylation sequence. (B) PCR-RFLP analysis with primers indicated in Figure S1A spanning a MnlI RFLP discriminating between the Cast/Ei (no MnlI site) and the 129/Sv (MnlI site) alleles, which was used to target the M2rtTA cassette. Correct targeting of the M2rtTA cassette results in loss of 129/Sv specific restriction products, as shown for clone 60. Arrows on right indicate size of PCR product and size of MnlI restriction fragments. F1, F1 2-1 polymorphic 129/Sv-Cast/Ei mother cell line; M, marker; H, water.

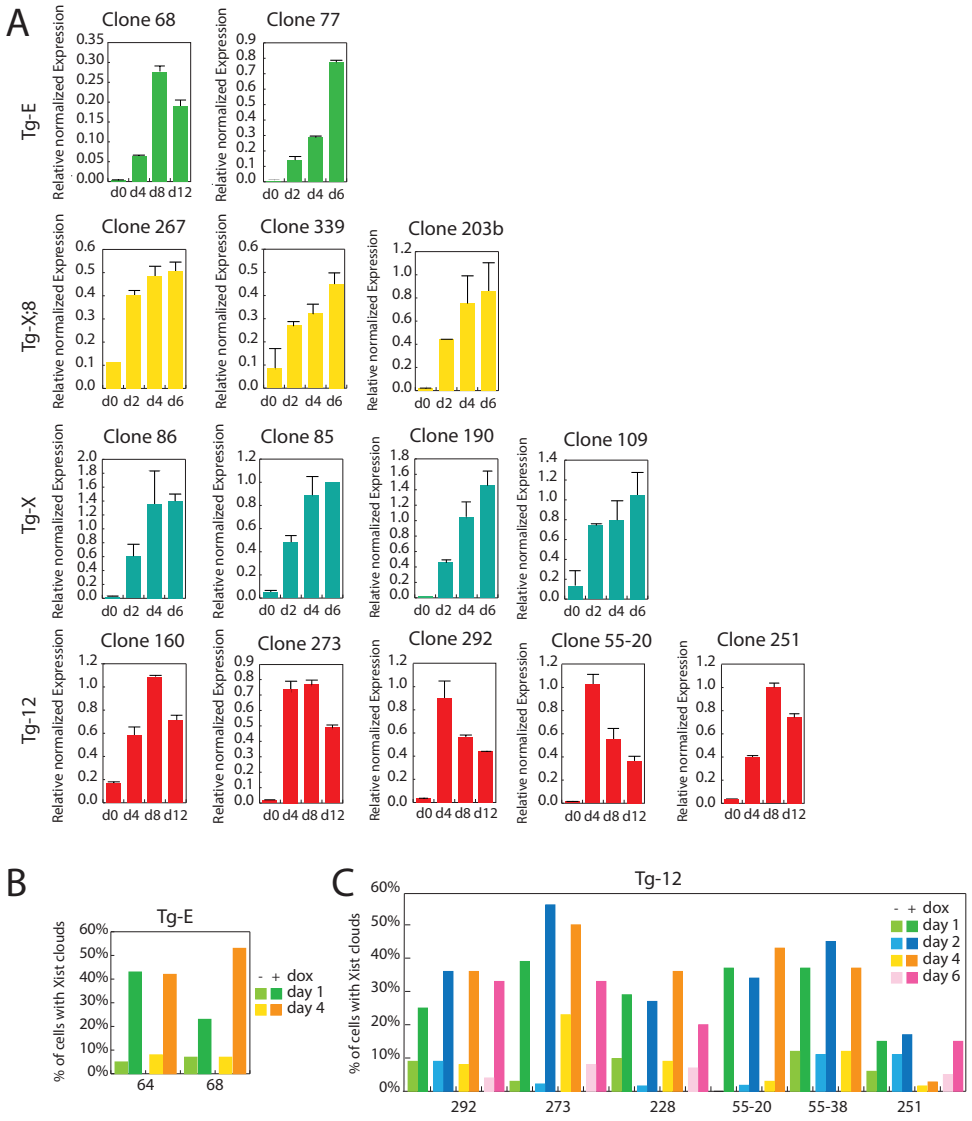
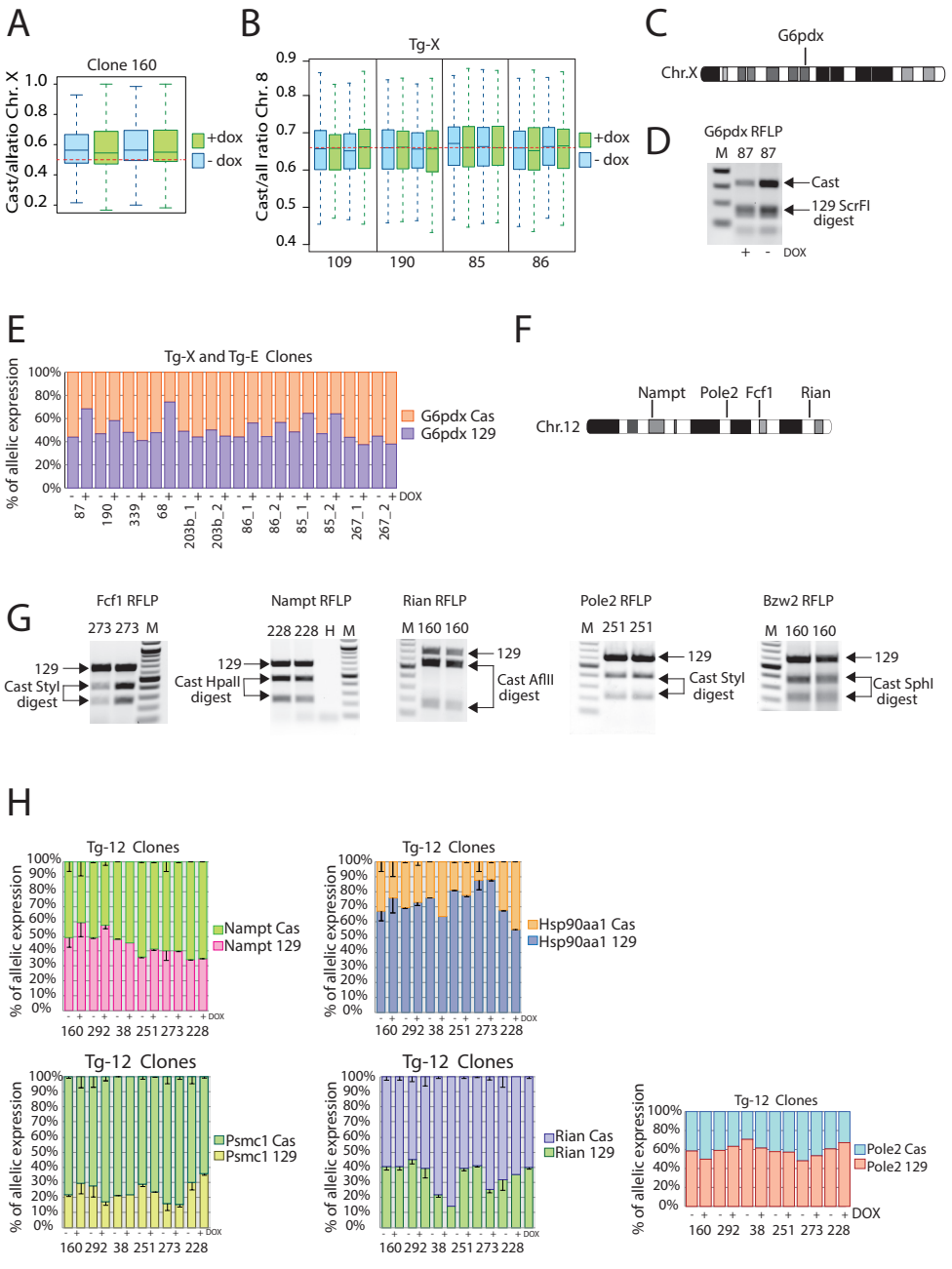


Figure S2, Related to Figure 2.

(A) qPCR quantification of Xist RNA expression level in doxycycline untreated cells at different time points upon neuronal differentiation. Mean and SD of two independent differentiation experiments are shown for clones Tg-E (68, 77), Tg-X (86, 85, 190, 109), Tg-X;8 (203B, 267, 339) and Tg-12 (160, 273, 292, 55-20, 251). (B and C) Quantification of cells showing an Xist-coated chromosome at different time points after doxycycline induction. Data related to Tg-E (64, 87) and Tg-12 (292, 273, 228, 55-20, 55-38, 251) clones is shown. n>150 nuclei counted per time point. Higher levels of Xist-coated chromosomes in doxycycline untreated clones compare to Figure 2F is due to serum versus 2i (Ying et al., 2008) ESC culturing conditions.


A
Clone 160
Cast/allratio Chr. X
+dox
- dox
B
Tg-X
Cast/all ratio Chr. 8
109
190
85
86
C
G6pdx
Chr.X
D
G6pdx RFLP
M 87 87
Cast
129 ScrFI digest
+ - DOX
E
Tg-X and Tg-E Clones
% of allelic expression
G6pdx Cas
G6pdx 129
DOX
F
Nampt Pole2 Fcf1 Rian
Chr.12
G
Fcf1 RFLP
273 273 M
129
Cast StyI digest
Nampt RFLP
228 228 H M
Cast HpaII digest
Rian RFLP
M 160 160
Cast AflII digest
Pole2 RFLP
M 251 251
Cast StyI digest
Bzw2 RFLP
M 160 160
Cast SphI digest
H
Tg-12 Clones
Nampt Cas
Nampt 129
Hsp90aa1 Cas
Hsp90aa1 129
Psmc1 Cas
Psmc1 129
Rian Cas
Rian 129
Pole2 Cas
Pole2 129
160 292 38 251 273 228

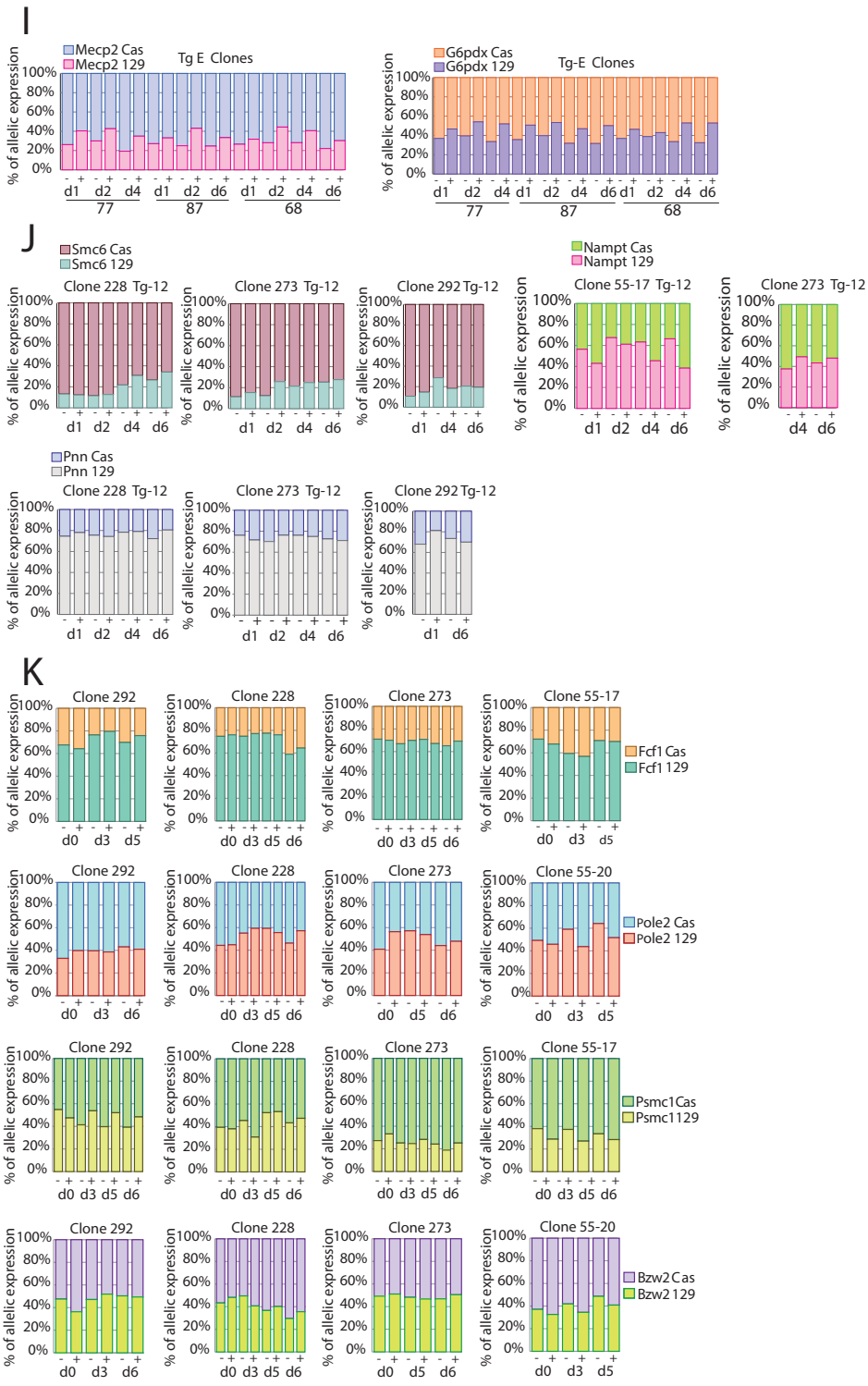

I
Mecp2 Cas
Mecp2 129
Tg E Clones
% of allelic expression
100%
80%
60%
40%
20%
0%
d1 d2 d4 d1 d2 d6 d1 d2 d4 d6
77
87
68
G6pdx Cas
G6pdx 129
Tg-E Clones
d1 d2 d4 d1 d2 d4 d6 d1 d2 d4 d6
77
87
68
J
Smc6 Cas
Smc6 129
Clone 228 Tg-12
Clone 273 Tg-12
Clone 292 Tg-12
Nampt Cas
Nampt 129
Clone 55-17 Tg-12
Clone 273 Tg-12
Pnn Cas
Pnn 129
Clone 228 Tg-12
Clone 273 Tg-12
Clone 292 Tg-12
K
Clone 292
Clone 228
Clone 273
Clone 55-17
Fcf1 Cas
Fcf1 129
d0 d3 d5
Clone 292
Clone 228
Clone 273
Clone 55-20
Pole2 Cas
Pole2 129
d0 d3 d6
Clone 292
Clone 228
Clone 273
Clone 55-17
Psmc1Cas
Psmc1129
d0 d3 d5 d6
Clone 292
Clone 228
Clone 273
Clone 55-20
Bzw2 Cas
Bzw2 129
d0 d3 d6

Figure S3, Related to Figure 3.

Box plot showing the Cast- specific gene expression ratio of X-linked genes in Tg-12 clone 160 (A) and of chromosome 8 genes in Tg-X clones (109, 190, 86, 85) (B). (C and F) Schemes showing the position on chromosomes 12 and X of the genes used for allele-specifc RFLP RT-PCR. (D and G) Examples of allele-specific RFLP RT-PCR analysis of chromosome 12 genes *Rian, Fcf1, Pole2, Nampt, Smc6, Pnn* and X-linked genes *G6pdx, Mecp2.* Arrows indicate size of RT-PCR products corresponding to either Cast or 129 alleles and size of restriction fragments upon enzymatic digestion. (E and H) Validation of RNA-seq analysis by allele-specific RFLP RT-PCR for Tg-X (86, 85, 109, 190), Tg-X;8 (203b, 267, 339), Tg-E (68, 87) and Tg-12 (160, 292, 55, 228, 273) clones. Expression data for *Rian, Fcf1, Pole2, Nampt* (Chr.12) *G6pdx* (Chr.X) and are shown. (I and J) Allele-specific RFLP RT-PCR analysis of *G6pdx, Mecp2* (Chr. X) and *Smc6, Pnn, Nampt* (Chr.12) at different time points after Xist induction in ESC clones maintained in conventional serum+Lif conditions. (K) Allele-specific RFLP RT-PCR analysis of chromosome 12 genes in Tg-12 292, 228, 273 and 55 at different time points upon EB differentiation. Expression data for *Fcf1, Psmc1, Pole2* and *Bzw2* are shown. For allele-specific RFLP RT-PCR analysis the percentage of allele-specific expression was determined by measuring relative band intensities of restriction fragments upon enzymatic digestion of PCR products using a Typhoon image scanner and ImageQuant software.

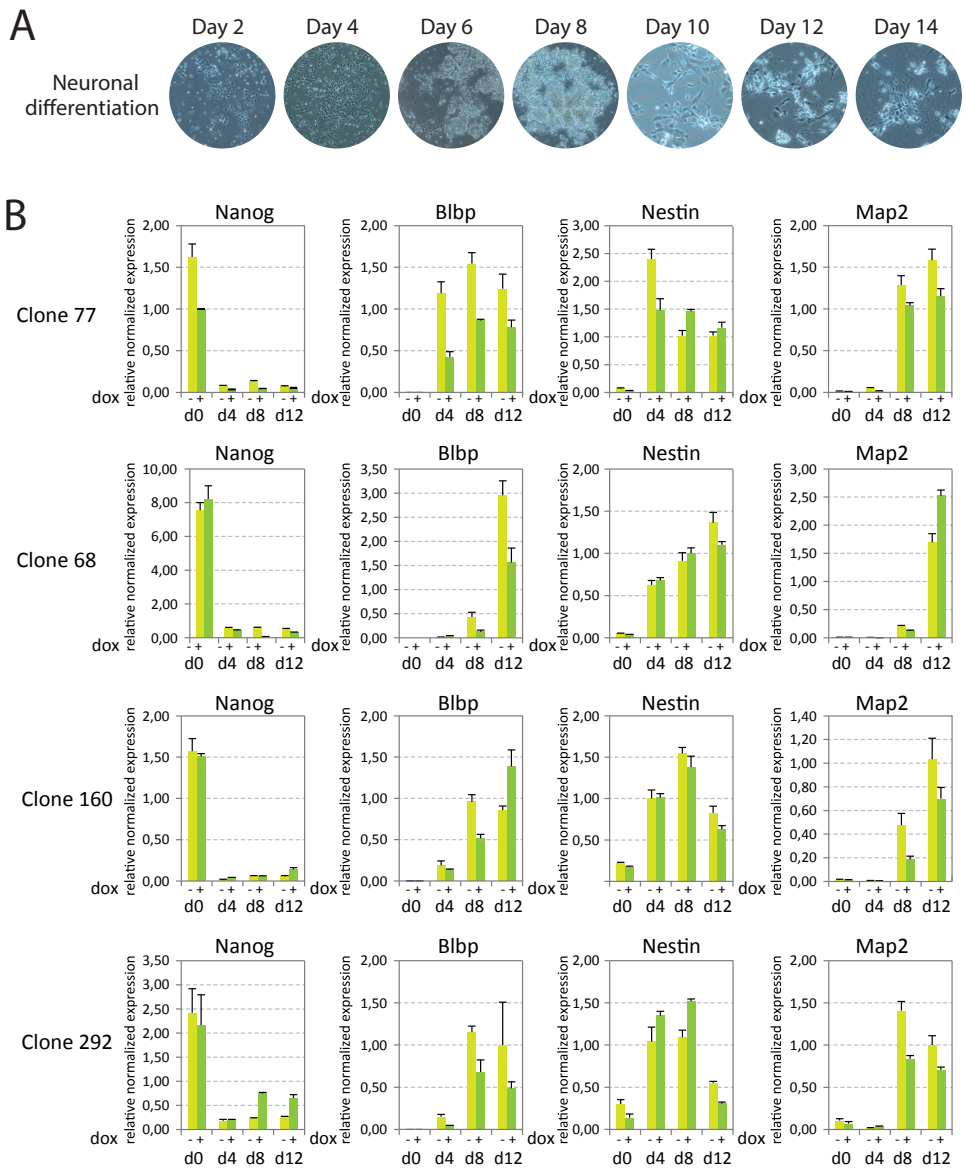


Figure S4, Related to Figure 3.

(A) Representative images of *in vitro* neuronal differentiation of F1 2-1 ESC lines. (B) qPCR analysis of Tg-E (77, 68) and Tg-12 (160, 292) at different time points upon neuronal differentiation. Data for *Nanog, Blbp, Nestin and Map2* are shown.

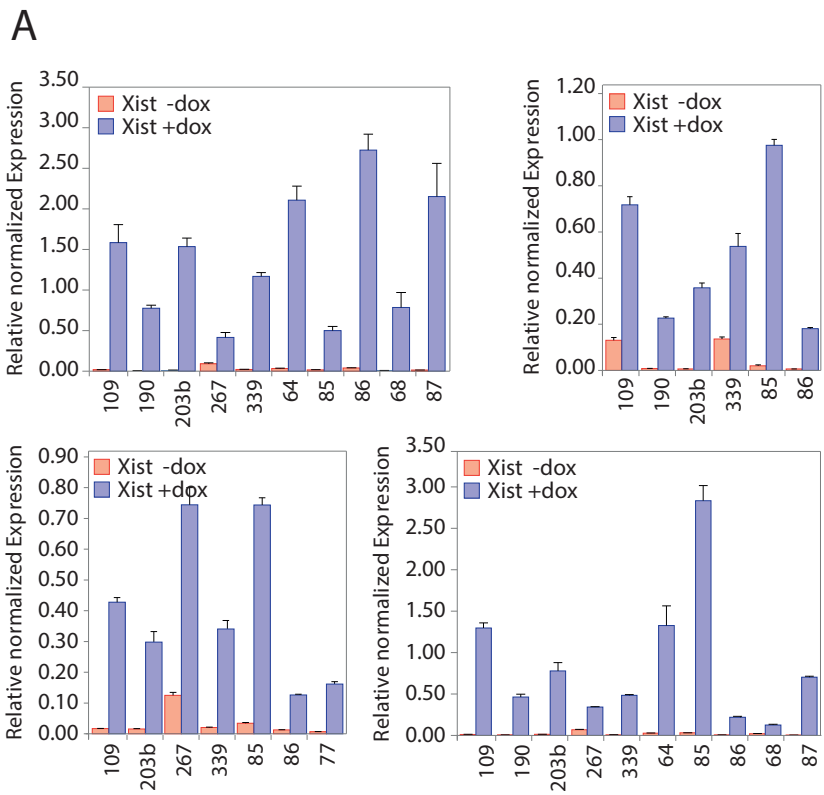


Figure S5

(A) Xist qPCR analysis of Tg-E (87,68,77), Tg-X (109, 190, 86, 85) and Tg-X;8 (203B, 339, 267) clones after five days of doxycycline treatment prior to neuronal differentiation. Data of three to four independent experiments per clone are shown.

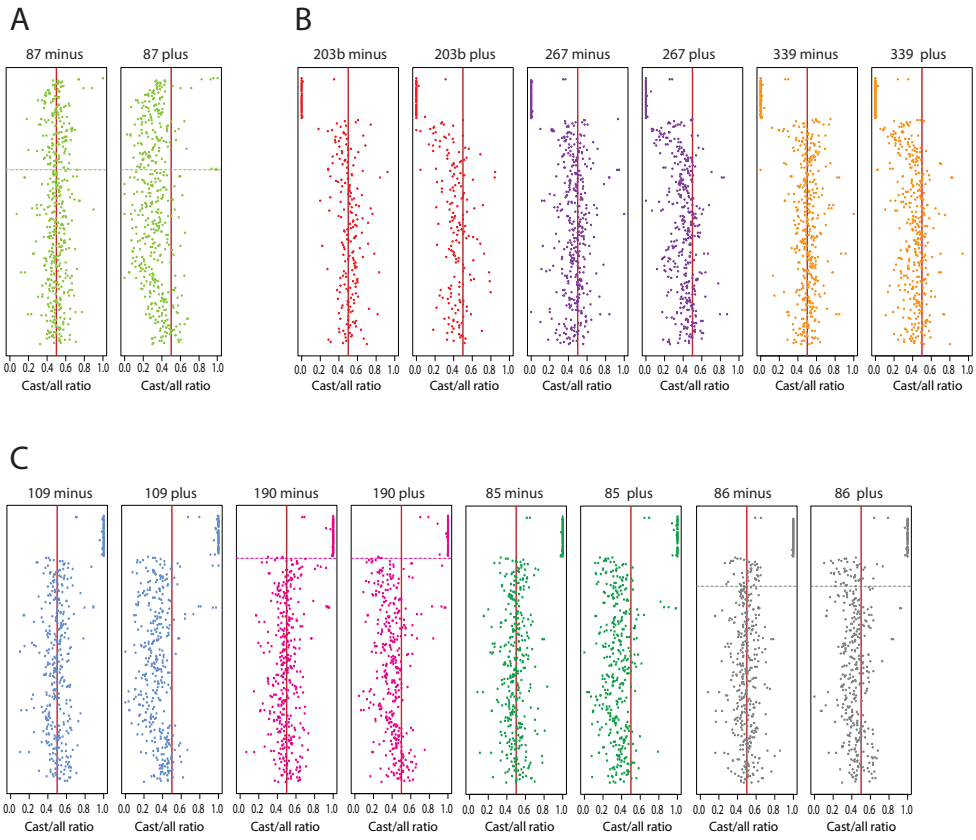


Figure S6, Related to Figure 5

Plots showing Xist mediated silencing of X-linked genes in (A) Tg-E, (B) Tg-X;8 and (C) Tg-X clones. Every dot represents a gene. Genes are ordered by genomic position on chromosomes X from telomeric (top) to centromeric (bottom) ends.

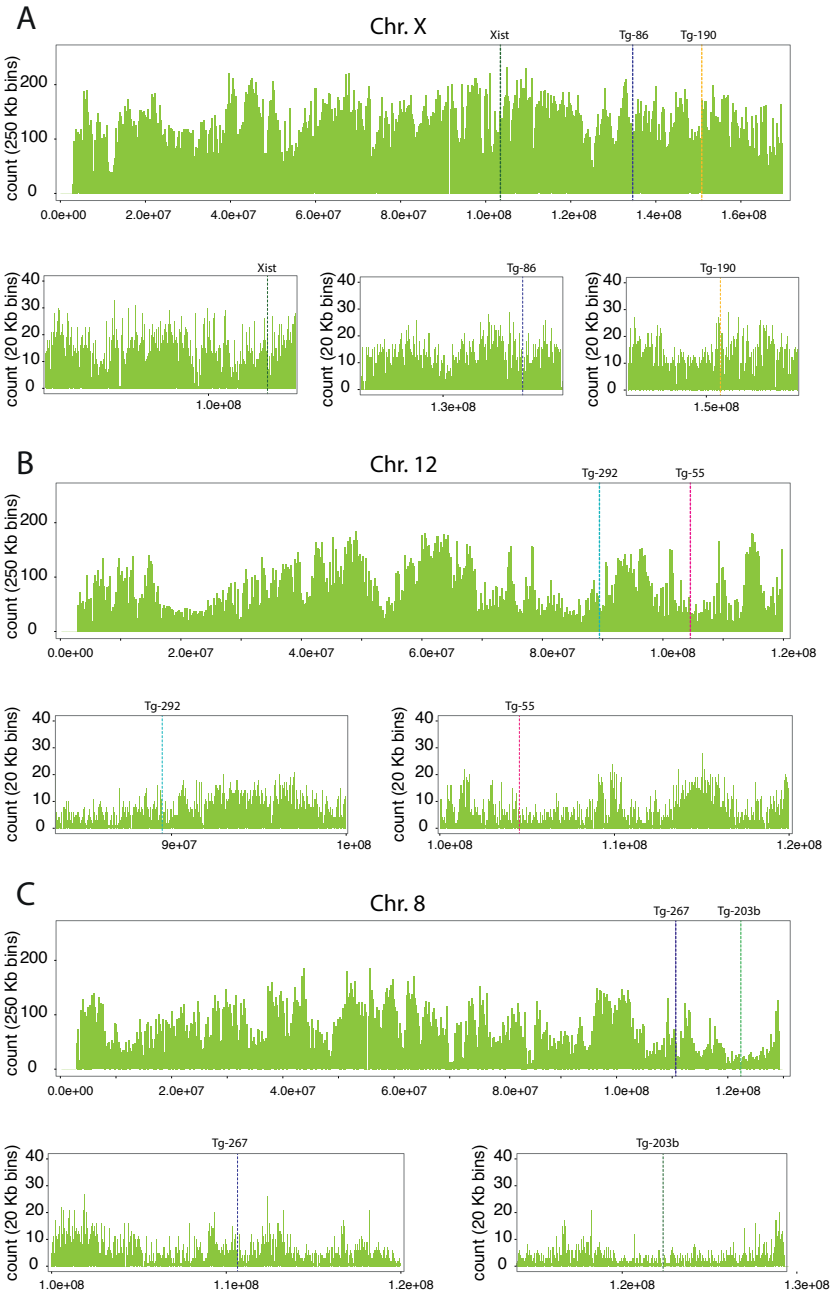


Figure S7, Related to Figure 6

LINE density along the entire length of chromosomes X (A), 12 (B), and 8 (C) is shown. Green histogram bars represent 250 kb bins. Zoom in around the integration sites of *Xist* transgene are shown, histogram bars correspond to 20 kb bins.

Table S1

OLIGO	SEQUENCE	DESCRIPTION
94	ggtaccCAGCCATGTTTGCTCGTTT	PTET XIST_3`ARM_KpnI F
95	ggatccTAAACGCAGGTATCCGAGGT	PTET XIST_3`ARM_BamHI R
96	cactgtcggtcactgttcaga	PTET XIST_5`ARM_F
97	CACACCAAAAGCATCACCAG	PTET XIST_5`ARM_R
98	ctgcagCCCCCTGAACCTGAAACATA	Floxed NEO F
99	ctgcagAAGtctgggtcaagccattg	Floxed NEO R
119	TGGGTCCTTGTTTCTTGACC	PTET XIST 5`BB F k53_BAC_REC
120	TCCCTTTAGGGTTCCGATTT	PTET XIST 5`BB R k53__BAC_REC
121	GCCATCACGAGATTTCGATT	PTET XIST 5`BB R k35__BAC_REC
127	ATCGCCTGGAGAATTCGAG	PTET XIST 3`BB F__BAC_REC
128	CGATGGGCAAAAGAAAAAGA	PTET XIST 3`BB F__BAC_REC
90	CGCGTCATGTCACTGAGCTT	Xist targeting TspeI F__ES
91	CGTTGCACGCCTTTAACTGA	Xist targeting TspeI R__ES
114	ATATAgcggccgcGGGCCTATTCTCAGTCCAG	3`ARM ROSA26 F NotI
115	CGATAGtctagaAGAATGCCATGAGTCAAGCC	3`ARM ROSA26 R XbaI
116	tagactGACGTCtctgggggagtcgttttacc	5`ARM ROSA26 AatII F
117	atctaaGAGCTCtttcgaggtcgatcgaggtc	5`ARM ROSA26 SacI R
136	TGAAAACACAAATGGCGTGT	R26 BB 5`F__BAC_REC
137	gcgaagagtttgtcctcaacc	R26 BB 5`R__BAC_REC
138	caatggcttgacccagaCTT	R26 BB 3`F__BAC_REC
139	ACACACCAGAAGAGGGCATC	R26 BB 3`R__BAC_REC
156	ggagagaggcattcatgggag	R26_targeting_22__ES_REC
157	ctttttgttgatcctttgccttgatcc	R26_targeting_23__ES_REC

Supplementary Table 1.

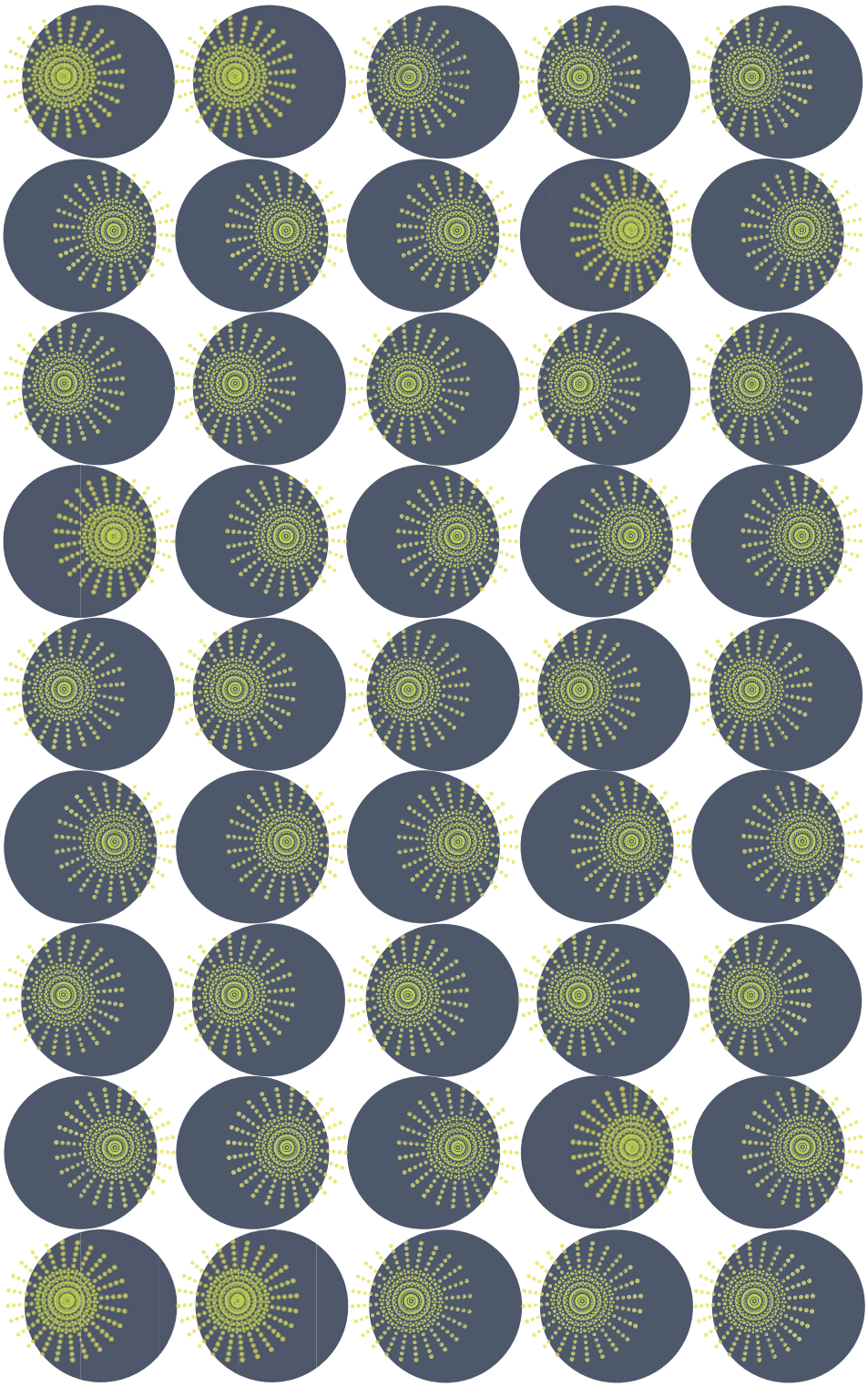
Primers used in this study to generate transgenic ES clones as described in the experimental procedures section.

Table S2

OLIGO	SEQUENCE	DESCRIPTION
160	GGATCCTGCTTGAACTACTGC	Forward primer *Xist* expression
161	CAGGCAATCCTTCTTCTTGAG	Reverse primer *Xist* expression
162	AACCCTAAGGCCAACCGTGAAAAG	Forward primer *Actin* expression
163	CATGGCTGGGGTGTTGAAGGTCTC	Reverse primer *Actin* expression
327	AACCAAAGGATGAAGTGCAAGCGG	Forward primer *Nanog* expression
328	TCCAAGTTGGGTTGGTCCAAGTCT	Reverse primer *Nanog* expression
323	AACCTGGAAGCTGACAGACAGT	Forward primer *Blbp* expression
324	TCACAGTTGGTTTGGTCACG	Reverse primer *Blbp* expression
333	AGGCGCTGGAACAGAGATT	Forward primer *Nestin* expression
334	TTCCAGGATCTGAGCGATCT	Reverse primer *Nestin* expression
345	CCTTTTAAAACCGGGAGAGG	Forward primer *Map2* expression
346	AAGGAGAGTGGGCCTGAACT	Reverse primer *Map2* expression
255	GGACCAAAATACACAGATGGCCTA	*Xist*_CAS_Forward_allele_specific
257	CATCATTCCGTCCGGTCAAG	*Xist*_129_Forward_allele_specific
256	CTTGGAAGTCACAGGTGTCCTGTA	*Xist*_Reverse_allele_specific
251	CCCTGAAACCCTACAATGAGTTTTa	*G6pdx*_Forward_129_allele_specific
252	CCTGAAACCCTACAATGAGTTTTg	*G6pdx*_Forward_CAS_allele_specific
253	GGACCGCCATTTTGTCCTAT	*G6pdx*_Reverse_allele_specific
266	CTAAGACTCAGCCTATGGTCGACA	*Mecp2*_Forward_129_ allele_specific
267	TAAGACTCAGCCTATGGTCTCCG	*Mecp2*_Forward_CAS_allele_specific
268	GGCATGGAAGATGAAACAATGTC	*Mecp2*_Reverse_allele_specific

Supplementary Table 2.

Primers used in this study for gene expression analysis as described in the experimental procedures section.



CHAPTER .4

CHROMATIN MEDIATED REVERSIBLE SILENCING OF SENSE-ANTISENSE GENE PAIRS IN ESCS IS CONSOLIDATED UPON DIFFERENTIATION

Friedemann Loos[a], **Agnese Loda[a]**, Louise van Wijk[a],
J. Anton Grootegoed[a], Joost Gribnau[a#]

Published in Mol Cell Biol
35:2436-2447 (2015)

CHROMATIN MEDIATED REVERSIBLE SILENCING OF SENSE-ANTISENSE GENE PAIRS IN ESCS IS CONSOLIDATED UPON DIFFERENTIATION

Friedemann Loos[a], **Agnese Loda[a]**, Louise van Wijk[a], J. Anton Grootegoed[a], Joost Gribnau[a#]
[a]Department of Developmental Biology, Erasmus MC, University Medical Center, Rotterdam, The Netherlands
[#]Address correspondence to Joost Gribnau, j.gribnau@erasmusmc.nl.

ABSTRACT

Genome wide gene expression studies have indicated that the eukaryotic genome contains many gene pairs showing overlapping sense and anti-sense transcription. Regulation of these coding, and/or non-coding gene pairs involves intricate regulatory mechanisms. Here, we have utilized an EGFP reporter plasmid *cis*-linked to a doxycycline inducible antisense promoter, generating antisense transcription that fully overlaps EGFP, to study the mechanism and dynamics of gene silencing after induction of non-coding antisense transcription, in undifferentiated and differentiating mouse embryonic stem cells (ESCs). We find that EGFP silencing is reversible in ESCs but is locked into a stable state upon ESC differentiation. Reversible silencing in ESCs is chromatin dependent, and is associated with accumulation of H3K36me3 at the EGFP promoter region. In differentiating ESCs, antisense transcription-induced accumulation of H3K36me3 is associated with an increase in CpG methylation at the EGFP promoter. Repression of the sense promoter is affected by small molecule inhibitors which interfere with DNA methylation and histone demethylation pathways. Our results indicate a general mechanism for silencing of fully overlapping sense-antisense gene pairs involving antisense transcription-induced accumulation of H3K36me3 at the sense promoter, resulting in reversible silencing of the sense partner, which is stabilized during ESC differentiation by CpG methylation.

INTRODUCTION

Strand specific RNA sequencing analysis of the mammalian transcriptome has indicated that more than 20% of the sequenced transcripts belong to sense-antisense gene pairs (Katayama et al., 2005). Many of these gene pairs show full overlap of at least one template, or antisense transcription through the sense promoter, and may consist of coding genes, non-coding genes, or a combination of coding and non-coding genes. Sense-antisense gene pairs are frequently found in imprinted gene clusters, involved in setting up and maintaining parent specific gene expression profiles. Imprinted gene loci are regulated by differentially methylated imprinting control regions (ICRs), which often direct parent specific transcription of non-coding RNA (ncRNA) transcripts. Studies involving knockout alleles and alleles with introduced

transcriptional stop sequences have indicated that these antisense non-coding genes play a crucial role in the regulation of the coding sense partner. The imprinted non-coding antisense genes *Kcnq1ot1*, *Ube3a-ATS*, *Nespas* and *Airn* are the master regulators of the *Kcnq1*, Prader-Willi / Angelman syndrome, *Gnas* and *Igf2r* clusters respectively, by regulating the sense protein coding partner. *Kcnq1ot1*-mediated repression of *Kcnq1* and silencing of *Igf2r* by *Airn* does not depend on dsRNA molecules, but has been attributed to the act of transcription involving transcription through the promoter of *Kcnq1* and *Igf2r* (Thakur et al., 2004; Latos et al., 2012). This repression might involve transcriptional interference mechanisms of the sense partner, but may also include recruitment of chromatin remodeling complexes leading to local accumulation of histone modifications and DNA methylation, as was found for *Nespas* and *Airn* respectively (Williamson et al., 2011; Sleutels et al., 2002). In addition, for *Kcnq1ot1* and *Airn* it has been shown that recruitment of chromatin remodeling complexes is involved in cis spreading of silencing towards non-overlapping genes, leading to parent specific inhibition of expression of flanking genes over long distances, *in cis* (Pandey et al., 2008; Nagano et al., 2008).

The *Xist/Tsix* gene pair represents one of the best studied mammalian sense-antisense gene loci. In contrast to most imprinted gene loci, both *Xist* and *Tsix* are non-coding and the respective transcriptional activities or the transcribed ncRNAs are involved in mutual repressive mechanisms. *Xist* and *Tsix*, which is fully overlapping *Xist*, are the main players in the X chromosome inactivation (XCI) process. Random XCI occurs, and can be studied, in differentiating female mouse ESCs, with two X chromosomes. Xist is up-regulated on the future inactive X chromosome, and cis spreading and ncRNA-mediated recruitment of chromatin remodeling complexes, including PRC2, leads to inactivation of that one X chromosome. *Tsix*-mediated repression of *Xist* on the active X chromosome does not involve dsRNA and RNA interference mechanisms (Nesterova et al., 2008), but is dependent on *Tsix* antisense transcription through the *Xist* promoter, which leads to Xist promoter associated changes in histone modifications and CpG methylation (Sado et al., 2005; Ohhata et al., 2008). Whether this local recruitment of chromatin remodelers is ncRNA-mediated or is dependent on a transcriptional interference mechanism is unknown.

Loss or gain of expression of a non-coding antisense partner of a sense gene has often been implicated in disease. For instance, in fragile X syndrome (FXS), a repeat expansion of a CGG repeat in the 5'UTR of the human *FMR1* gene results in induction of antisense transcription through the *FMR1* promoter (Ladd et al., 2007), initiating at the expanded repeat producing an unstable non-coding transcript. This antisense transcription results in epigenetic silencing of *FMR1*, which involves CpG methylation of the expanded repeat. This silencing of *FMR1* happens during a defined window of neuronal differentiation (Brouwer et al., 2009). One form of alpha-thalassemia has been associated with juxtaposition of *LUC7L* to *HBA2*, resulting in antisense transcription through *HBA2*. This aberrant antisense transcription leads to silencing

of *HBA2* and DNA methylation of its associated CpG island by an unknown mechanism, during a specific developmental time window (Tufarelli et al., 2003). These examples highlight the close relationship between transcription, ncRNAs and gene regulation with human disease in a developmental context.

For all these examples the exact mechanisms involved in silencing of the sense partner by antisense transcription remain elusive, as the effects of the act of transcription and biological activity of the respective ncRNA product cannot be separated. Therefore, a general question is whether transcriptional interference, e.g. collision of RNA polymerase II (pol II) complexes, torsonial strain or displacement/occlusion of transcription factors/regulatory elements, or chromatin mediated mechanisms are responsible for silencing of overlapping genes. Direct transcriptional interference has mostly been studied in prokaryotes and yeast, and shows that collision of pol II complexes and displacement/occlusion of transcription factors/regulatory elements does influence the activity of overlapping genes to some extent (reviewed in Shearwin et al., 2005). In the case of *Airn*, transcriptional overlap with the *Igf2r* promoter alone is necessary and sufficient for silencing of *Igf2r*, thus precluding a role for the RNA molecule itself (Latos et al., 2012). However, chromatin modulation based on the process of transcription per se might still be involved. This model is supported by the observed association of pol II with chromatin remodelers like *Set2* (Li et al., 2002), which catalyzes deposition of H3K36me3 in transcribed regions. H3K36me3 in turn has been implicated in recruitment of chromatin factors correlating with transcriptional repression like histone deacetylases, histone demethylases and DNA methyltransferases (Carrozza et al., 2005; Xie et al., 2011; Dhayalan et al., 2010). In contrast, *Airn, Kcnq1ot1, Xist* and other ncRNAs have been shown to recruit histone modifying complexes like G9A and PRC2 *in cis* (Pandey et al., 2008, Nagano et al., 2008, Plath et al., 2003), suggesting the requirement of the RNA molecule itself. To be able to exclusively study the effects of antisense transcription, we have utilized an artificial sense-antisense gene pair consisting of an EGFP reporter and a fully overlapping inducible antisense transcription unit. Our studies indicate that antisense mediated silencing of the EGFP gene is reversible in embryonic stem cells (ESCs), and is dependent on modifications of the chromatin environment. Interestingly, silencing is locked into a stable state upon ESC differentiation, concomitant with accumulation of EGFP promoter-associated CpG methylation. Antisense transcription-induced silencing is augmented by blocking JARID1/JMJD2 family histone demethylases, suggesting that the transcription-coupled histone modification H3K36me3 provides a repressive environment for sense transcription initiation, which is locked into a stable state by CpG methylation upon ESC differentiation.

MATERIALS AND METHODS

Plasmids, Reagents and Antibodies

Plasmids used for generation of transgenic lines and transient transfections were pTRE-Tight-

BI-DsRed2 (Clontech) and pCAG-EGFP-N1, which was generated by replacing the CMV promoter in pEGFP-N1 (Clontech) with the CAG promoter from pCAG-Rnf12-Flag (Gontan et al., 2012). Reagents used were doxycycline, 5-aza-dC, SAHA, 2,4-PDCA, 2-PCPA, curcumin, pargyline, JQ1 (all Sigma).

Cell Lines

Culture media and conditions for ESC culture and differentiation have been described (Barakat et al., 2011). Final concentration of doxycycline was 2 µg/ml. Final concentrations of small molecule inhibitors were: 20nM 5-aza-dC; 400nM SAHA; 5mM 2,4-PDCA; 200µM 2-PCPA; 10µM CUR; 1.5mM PAR; 150nM JQ1. Transgenic ESC lines were generated using polymorphic male 129/Sv-Cast/Ei line harboring an M2rtTA transcriptional activator in the ROSA26 locus (Friedrich et al., 1991) as follows: A tetracycline-responsive ptet promoter excised from pTRE-Tight-BI-DsRed2 was inserted downstream and in antisense direction to the EGFP in pCAG-EGFP-N1 (pCAG-EGFP-as-ptet). This construct was transfected into M2 ESC line by electroporation (Bio-Rad Gene Pulser Xcell) and stable clones with random integrations were obtained by one week selection in 350 µg/ml G-418 (Gibco). Clones were screened for expression of EGFP and responsiveness to doxycycline. For transient transfections, M2 ESCs were co-transfected with pCAG-EGFP-as-ptet and pTRE-Tight-BI-DsRed2 using an AMAXA nucleofector device and Mouse ESC Nucleofector Kit (Lonza). After 18 hours, EGFP-positive cells were sorted and used for experiments.

FACS Analysis

Single cell suspensions were prepared by TE treatment for 7 minutes at 37°C and 30 minutes pre-plating to remove feeder cells if necessary. Duplets were excluded by appropriate gating and dead/dying cells by Hoechst 33258 straining (1 µg/ml, Molecular Probes). Relative fluorescence intensities (FI) 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.

Strand-specific 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 SuperScriptII (Invitrogen) with strand-specific primers for target and control in the same reaction. Quantitative RT-PCR was performed on a CFX384 Real-Time PCR Detection System (Biorad) using Fast SYBR Green Master Mix (Applied Biosystems). Primers are listed in Table S2. Results were normalized to Actin, using the ΔCT method (Livak et al., 2001) and mostly represented as fold-change versus undifferentiated no doxycycline control.

Bisulfite Sequencing

Phenol-chloroform extracted DNA was converted using the EpiTect Bisulfite Kit (QIAGEN) following the manufacturer's instructions. Part of the CAG promoter was amplified from bisulfite converted DNA with Platinum Taq (Invitrogen) using primers 204+207. The PCR product was gel-purified and subcloned into pGEM T-Easy (Promega) and transformed into bacteria. Single bacterial clones were isolated and the fragment of the CAG promoter was amplified by colony PCR using primers 208+209, followed by Sanger sequencing using primer 302. Sequence reads were analyzed using QUMA (Kumaki et al., 2008).

Chromatin Immunoprecipitation

In short, approximately $5x10^6$ cells were cross-linked in dish for 10 minutes at room temperature by adding 1/10 volume 11% buffered formaldehyde solution (50 mM Hepes-KOH pH 7.6, 100 mM NaCl, 1 mM EDTA pH 8, 0.5 mM EGTA pH 8, 11% v/v formaldehyde) and quenched for 10 minutes at room temperature with 125 mM glycine. Cells were washed twice in ice-cold PBS+ protease inhibitors and resuspended in SDS lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA pH 8, 1% SDS), followed by sonication until a fragment size of ca. 500 bp was reached. Chromatin was diluted 1:10 in ChIP dilution buffer (16.7 mM Tris-HCl pH 8, 167 mM NaCl, 1.2 mM EDTA pH 8, 1.1% Triton X100, 0.01% SDS) and incubated with antibodies overnight at 4°C. Chromatin was then incubated with pre-blocked Protein G Dynabeads (Novex) for 1 hour at 4°C. Beads were washed thrice in low salt buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM EDTA pH 8, 1% Triton X100, 0.1% SDS), once in high salt buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 2 mM EDTA pH 8, 1% Triton X100, 0.1% SDS), once in LiCl buffer (10 mM Tris-HCl pH 8, 250 mM LiCl, 1 mM EDTA, 0.5% deoxycholate, 0.5% NP-40), and once in TE buffer (10 mM Tris-HCl pH 8, 50 mM NaCl, 1 mM EDTA pH 8). Complexes were eluted in elution buffer (10 mM TrisHCl pH 8, 150 mM NaCl, 1 mM EDTA pH 8, 5 mM DTT, 1% SDS) for 15 minutes at 65°C and cross-links were reversed by incubation overnight at 65°C. DNA fragments were recovered by Proteinase K treatment followed by phenol-chloroform extraction and analyzed by quantitative RT-PCR. Enrichment was estimated by determining the original amount of template in pull-down and input fractions as 2-CT(pull-down)/ 2-CT(input).

RESULTS

Inducible antisense transcription reversibly silences an EGFP reporter

To be able to study the general effect of antisense transcription on gene regulation during development, we generated transgenic mouse ESC lines containing an EGFP reporter cassette and a doxycycline-responsive promoter in antisense direction downstream of the reporter. This antisense ptet promoter was intended to initiate antisense transcription that fully overlaps with the sense EGFP reporter. To this end, the tetracycline-responsive ptet promoter was inserted downstream and in antisense direction to a CAG promoter driven (Alexopoulou et

al., 2008) EGFP (Fig.1A). This construct was randomly integrated into male ESCs generated from a cross of a male Cast/Ei mouse and a female 129/Sv mouse carrying a M2rtTA transcriptional activator in the *ROSA26* locus (Friedrich et al., 1991), allowing us to induce doxycycline(DOX)-dependent antisense transcription through the EGFP coding sequence and its promoter. EGFP positives clones were isolated and expanded, and reactivity to DOX was tested. Most clones showed a reduction in EGFP intensity upon DOX addition and six clones, denominated M2-3, M2-4, M2-5, M2-8, M2-20 and M2-29, were used for further analysis (example in Fig.1B).

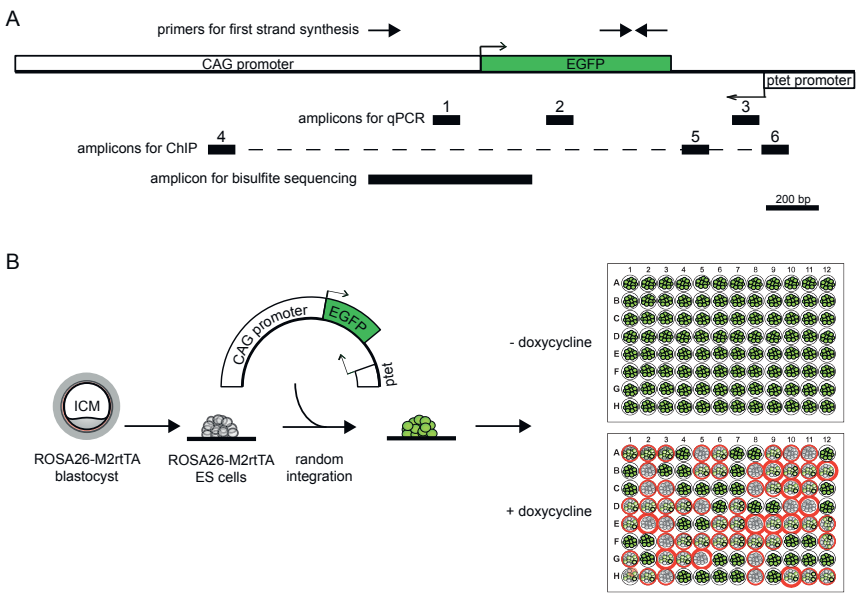


Figure 1. Map of EGFP-antisense-ptet construct and transfection strategy.

(A) Map to scale showing the construct used for generation of transgenic cell lines. Promoters and EGFP ORF are depicted as white and green boxes, respectively. Location of primers used for first strand synthesis for strand-specific expression analysis is indicated on top, and the black boxes numbered 1-6 mark the location of amplicons for expression analysis and ChIP analysis. Area amplified from bisulfite-treated DNA for bisulfite sequencing is shown on bottom. (B) Cartoon depicting transfection and screening strategy used to identify clones for further analysis. ESCs harboring a M2rtTA in the *ROSA26* locus were transfected, and stable random integration was forced by G-418 selection. EGFP-positive clones were expanded, plated in duplicate, and responsiveness to doxycycline was tested in 96-well plates (as an example, red circles denote clones with a decreased level of EGFP expression).

When undifferentiated M2-3 ESCs were grown for two days in the presence of DOX and analyzed by FACS, the EGFP relative fluorescence intensity (FI) level was reduced to 40% as

compared to control cells. This demonstrates that antisense transcription can reduce transcriptional activity of a sense partner, in this experimental context, where sense and antisense partners are biologically unrelated (Fig.2A,B; Fig. S1). This silencing appeared very dynamic as DOX removal within one day resulted in almost complete recovery of EGFP FI to levels measured without induction of antisense transcription (Fig. 2A,B; Fig. S1). Comparable results were found with several other M2 clones. Due to heterogeneity between independent experiments, statistical significance for the difference between recovered and repressed (+DOX) was not reached for all clones. However, the trend was highly similar in all independent experiments. This is also reflected by the statistical significance for the difference when all clones are pooled (Fig. 2B). To evaluate the time course of synthesis of both EGFP mRNA and the transcript originating from the ptet promoter, RNA was isolated from these pulse-chase experiments and strand-specific quantitative RT-PCR was performed. Two different sets of primers were used, one for amplification of a ptet proximal and one for a ptet distal product (Fig. 1A). As expected, quantification of the proximal transcript demonstrated reversible induction of the ptet promoter by doxycycline, which resulted in a concomitant reduction of the EGFP mRNA level. After DOX washout, the EGFP mRNA level recovered. As for fluorescence levels, heterogeneity between independent experiments was observed, but these independent experiments showed always the same trend. Therefore, not all clones show statistically significant differences, whereas all clones pooled do (Fig.2C). The same pattern was observed for the distal ptet amplicon, confirming that ptet induced transcription fully overlaps with the EGFP reporter and runs through the CAG promoter (Fig. 2D). These results show that, in the present system using undifferentiated mouse ESCs, repression of a coding gene by antisense transcription is completely reversible, so that the repression is dependent on continuous antisense transcription.

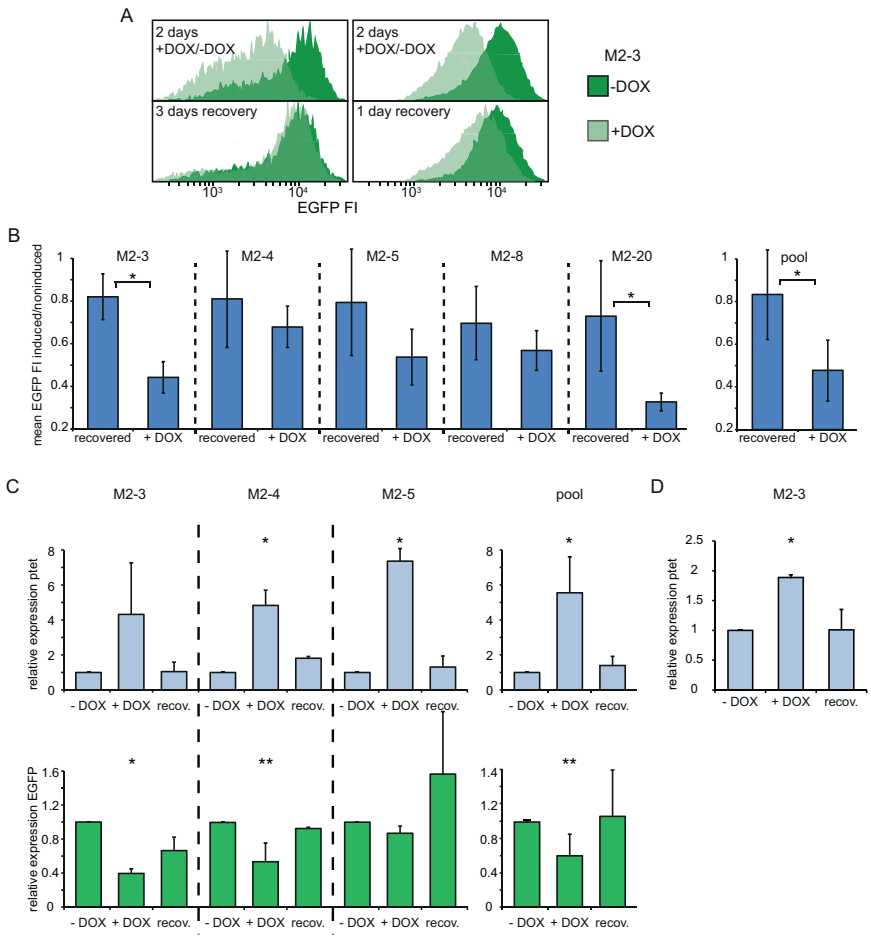


Figure 2. Antisense transcription and EGFP reporter expression in undifferentiated ESCs.

(A) Histograms of EGFP fluorescence intensity (FI) distribution in M2-3 as determined by FACS analysis. Dark green represents uninduced cells, and light green the cells treated with doxycycline. Left upper panel shows repression after two days of doxycycline treatment as compared to untreated control, and the left lower panel shows the respective situation three days later, after doxycycline has been washed out. Right panels depict another experiment with the same experimental setup, but with only one day recovery from doxycycline induction. (B) Mean EGFP FI of induced/noninduced cells is shown after two days of doxycycline treatment (+DOX) and after two days of doxycycline treatment followed by one day of recovery (recovered). (C) Strand-specific expression analysis in noninduced (-DOX), induced (+DOX), and recovered (recov.) cells as outlined in (B) and main text. Top panels (light blue bars) show the results for the ptet proximal amplicon (amplicon 3 in Figure 1A), and bottom panels (green bars) show the results for the EGFP amplicon (amplicon 2). Quantification is depicted as fold change (relative expression) compared to noninduced cells. (D) Same as in top panel of (C) but results for the ptet distal amplicon are shown (amplicon 1). The mean and SD of at least two independent experiments for each clone, and of all clones pooled, are shown in (B), (C) and (D); * $p < 0.05$, ** $p < 0.1$, two-sample Student's t test (B) or single factor ANOVA (C), (D).

Antisense transcription changes the chromatin structure of CAG promoter

For several specific examples of sense-antisense gene pairs, it has been described that silencing of the sense partner is accompanied by changes in promoter chromatin structure. In some cases, transcriptional overlap was found to be sufficient for silencing (Latos et al., 2012), but for other such gene pairs a requirement for antisense ncRNA to recruit chromatin modifying complexes has been reported (Pandey et al., 2008). Thus, mechanisms of regulation appear to vary and it is not fully understood how antisense transcription is converted into a repressive chromatin environment. Our ptet-EGFP system provides an experimental tool to study the effect of antisense transcription on chromatin structure, for specific genes and RNA sequences which are not taking part, in a biological context, in a sense-antisense regulation system. Hence, we anticipated that the present experiments would provide information regarding the more general aspects of such regulatory mechanisms.

Elongating pol II itself interacts with a plethora of histone modifying proteins, and we hypothesized that transcriptional read-through per se might be sufficient to create a specific repression-instructive chromatin signature in promoters. We were particularly interested in methylations of histone H3 at residues K4 and K36, which are catalyzed by two pol II associated proteins, SET1 (Ng et al., 2003) and SET2 (Li et al., 2002), respectively. ChIP analysis of the CAG promoter in undifferentiated, not induced M2-3 cells showed strong enrichment of H3K4me3, while H3K36me3 levels were close to background (Fig. 3A). The H3K36me3 signal between the EGFP cassette and the ptet promoter (amplicon 5 in Fig.1A) found in uninduced cells is most likely caused by read-through transcription derived from the EGFP cassette. Upon DOX addition, however, H3K36me3 signal at the CAG promoter increased approximately 3-fold with a concomitant decrease in H3K4me3 (Fig. 3A; Fig. 5C), thereby creating a specific chromatin environment that corresponded with repressed expression of the EGFP reporter. Gain of H3K36me3 just downstream of the ptet promoter also demonstrates effective transcriptional elongation, while enrichment at the ptet promoter itself most likely reflects a lack of resolution or initiation of transcription slightly upstream of the amplicon tested by qPCR. This might also explain lack of induction of H3K4me3 signal at the ptet promoter (amplicon 6 in Fig.1A) itself. Analogous to fluorescence and mRNA abundance measurements, this effect was almost completely reversible after DOX washout (Fig. 3A).

CpG islands (CGIs) are CG rich genomic regions which frequently initiate transcription and constitute more than 50% of all annotated promoters in vertebrates (Saxonov et al., 2006). Most CGIs remain unmethylated, but DNA methylation of CpG residues is correlated with stable repression of transcription. Several promoters of developmentally regulated genes acquire DNA methylation during development (Meissner et al., 2008; Mohn et al., 2008). To test if DOX-induced repression of the CAG promoter that drives EGFP expression and contains a CGI involves DNA methylation, we performed bisulfite sequencing on undifferentiated ESCs grown in absence and presence of DOX. In most clones the CAG promoter was found to be

completely devoid of DNA methylation, regardless of induction of antisense transcription (Fig.3B). Only clone M2-5 showed higher levels of DNA methylation, but this was unresponsive to induction of antisense transcription, meaning that this higher level most likely is related to a position-effect. Thus, antisense transcription generates a special chromatin state at an unrelated promoter that is located nearby and transcribed in sense direction. This sense partner is reversibly repressed in cis. Importantly, this effect is not dependent on any specific RNA sequences or locus requirements.

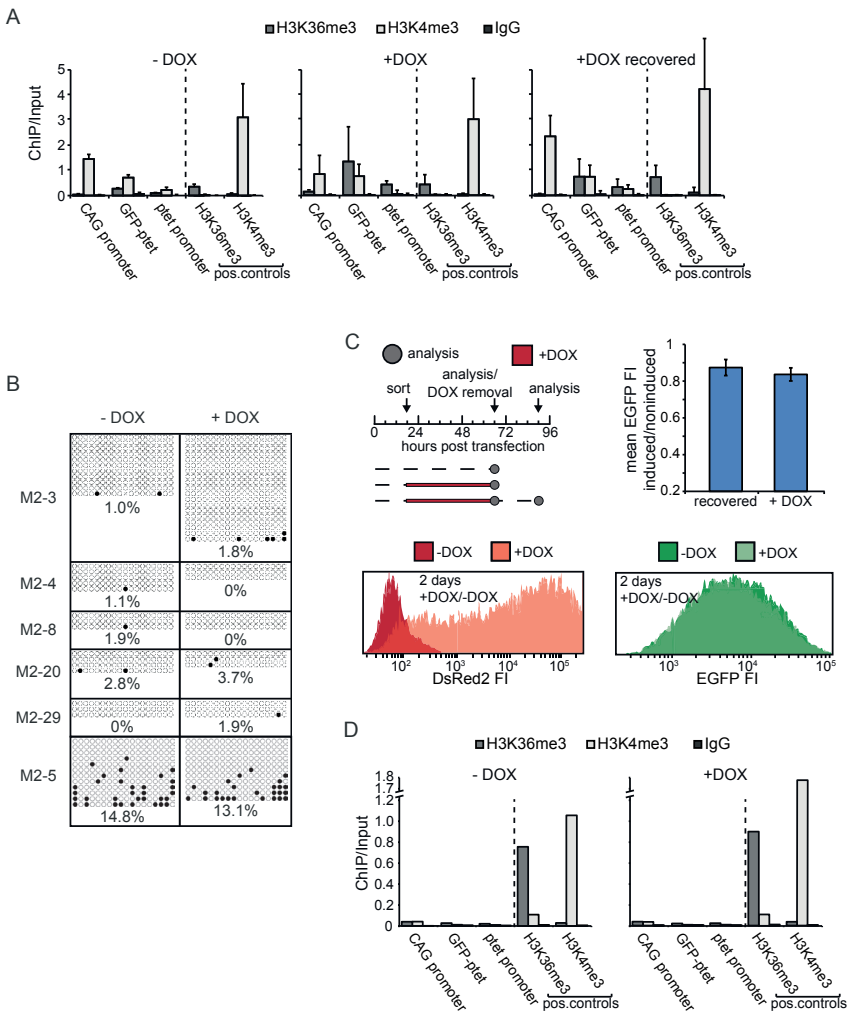


Figure 3. Chromatin structure in the CAG and ptet promoter regions in undifferentiated ESCs.

(A) ChIP-qPCR analysis of the region encompassing CAG and ptet promoters (CAG promoter: amplicon 4 in Fig.1A; GFP-ptet: amplicon 5; ptet promoter: amplicon 6) in the M2-3 clone for noninduced (-DOX), induced (+DOX), and recovered (+DOX recovered) conditions as outlined in main text. Antibodies against H3K36me and H3K4me3, and whole IgG as mock control, were used as indicated. Loci analyzed, including H3K36me3 and H3K4me3 positive controls, are indicated. Values are plotted as the ratio of original amount of template DNA, for pull-down and input fractions. The mean and SD of three independent experiments are shown. (B) Bisulfite sequencing of the CAG promoter in several clones treated (+DOX) or not treated with doxycycline (-DOX) for two days. For independent clones represented in the separate panels, filled circles are methylated, empty circles unmethylated CpGs. Average percentages for each condition and clone are indicated. (C) Upper left shows timing of transient transfection experiments, arrows indicate time point of sorting, analysis and doxycycline removal; dashed lines represent time cells were grown in absence, red line in presence of doxycycline. Lower panels display FACS histograms of DsRed2 (red) and EGFP (green) fluorescence intensity (FI) distribution, for the conditions without doxycycline (shown in dark color) and with doxycycline (in lighter color), respectively. Upper right gives mean EGFP FI of induced/noninduced cells after two days of doxycycline treatment (+DOX), and after two days of doxycycline treatment followed by one day of recovery (recovered). Error bars represent SD of three independent experiments. (D) Exactly as in (A), but for transient transfections.

Antisense transcription mediated repression requires an intact chromatin template

ChIP analysis of the CAG promoter suggested that antisense transcription induces a specific chromatin signature over promoters. Next, we asked whether chromatin modifications are important for silencing by antisense transcription. We therefore exploited transient transfections as a system in which the regular chromatin structure is perturbed (Jeong et al., 1994; Hebbar et al., 2008). The same sense-EGFP-antisense-ptet construct that was used for generation of M2 cell clones was transiently transfected into M2rtTA-ROSA26 male ESCs, and EGFP positive cells were sorted after 18hrs into medium containing DOX or no DOX. As a control for DOX induction, a ptet-DsRed construct was co-transfected. After two days of either DOX or no DOX treatment, cells were analyzed by FACS. For another set of cells DOX was removed after 48hrs and cells were left to recover for an additional 24 hours before FACS analysis. Surprisingly, addition of DOX almost completely failed to repress EGFP transcription from the transiently transfected plasmid, even though DOX induction of ptet transcription per se was functional as demonstrated by expression of DsRed (Fig. 3C). Thus, recovery from DOX treatment did not significantly increase EGFP FI levels as compared to the induced condition (Fig. 3C). To study if a perturbed chromatin arrangement on a transiently transfected template carries chromatin modifications as they are laid down by the transcription machinery and thus are involved in the specific chromatin state induced by antisense transcription, ChIP was performed on transiently transfected cells. Intriguingly, neither H3K4me3 nor H3K36me3 were found to reside on the transiently transfected plasmid (Fig. 3D). Thus, even though EGFP expression and ptet induction are not hampered on a transiently transfected template devoid of the histone modifications H3K4me3 and H3K36me3, repression of EGFP mediated by antisense transcription does not occur in that situation.

Antisense transcription-mediated repression during ESC differentiation

To investigate the effect of antisense transcription on expression of the EGFP reporter during differentiation, we performed pulse-chase type time-course experiments. Cells were differentiated by removal of feeders and LIF, and kept in culture until day three of differentiation. We opted for this window of time because, even in the absence of DOX, the EGFP reporter was increasingly silenced at later time-points in all clones analyzed. Moreover, in the present system, 50% of female ESCs initiate XCI during the first three days of differentiation, demonstrating that during this developmental time window major epigenetic and gene regulatory changes occur. The different conditions for day one, two, and three of differentiation were: i) no DOX, ii) with two days DOX followed by washout and one day recovery, and iii) addition of DOX for two days before analysis (Fig. 4A). FACS analysis showed that, while DOX treatment until the start of differentiation did not interfere with recovery of FI levels, the DOX treatment exerted a stronger inhibitory effect on recovery if DOX was administered during differentiation (Fig. 4B). Plotting the difference of EGFP FI ratios between recovered and repressed cells for all clones reveals significant differences over time of differentiation (Fig. 4D). This difference trends towards zero demonstrating that at later time points of differentiation cells are not able to recover from antisense transcription induced EGFP repression (Fig. 4D). In addition, when ptet antisense transcription through the EGFP cassette was induced from day one to three of differentiation, repression of the EGFP was attenuated. This suggests either that antisense transcription at the onset of differentiation is important for proper silencing of the antisense partner, or that at later time-points during differentiation any kind of forced expression helps to maintain the locus in an open conformation. To test whether the loss of reversible silencing during differentiation is caused by general repression of the EGFP reporter and to verify that induction of transcription from the ptet promoter is working under differentiation conditions, RNA was isolated from the same time-course experiments and transcripts emanating from the ptet and CAG promoters were quantified by strand-specific qPCR. Abundance of ptet transcripts increased 1.5- to 3-fold upon DOX addition at all time points analyzed, and this expression reverted to levels similar to the noninduced condition after DOX removal (Fig. 4C), demonstrating that the inducible system functions normally in differentiating cells. Of note, the ptet promoter appeared to become increasingly de-repressed while differentiation lasted. Upon addition or wash-out of DOX, EGFP transcripts displayed the expected anti-correlation with ptet-derived transcription, mirroring the data obtained by FACS analysis (Fig. 4C). Plotting the differences in expression levels for all clones pooled, we found that ptet levels do not differ significantly over time of differentiation. The difference between uninduced and recovered cells cluster around zero, showing that ptet expression is effectively terminated upon DOX washout, whereas the difference between recovered and repressed cells is negative, showing that ptet expression can be induced at all time points during differentiation (Fig. 4E, left panel).

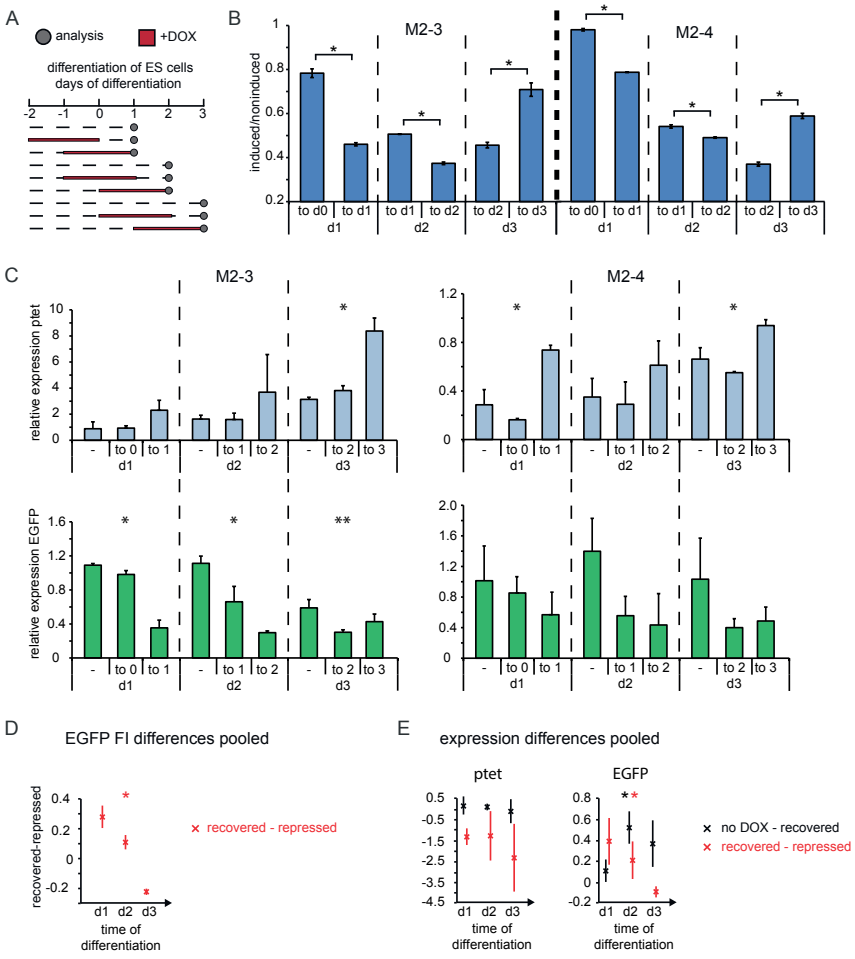


Figure 4. Stable repression of CAG promoter by antisense transcription in differentiating ESCs.

(A) Time schedule of induction experiments in differentiating ESCs. Dashed lines represent time that cells were grown in absence of doxycycline, red lines represent time that cells were grown in presence of doxycycline. Grey dots indicate time point of analysis. (B) Mean EGFP FI of induced/noninduced cells is shown for cells treated for two days with doxycycline until day of analysis and for cells treated for two days followed by one day of recovery as outlined in (A). Upper label on x-axis gives time period of doxycycline treatment and lower label indicates day of analysis. (C) Strand-specific expression analysis in noninduced, recovered, and induced cells as outlined in (A) and main text. Upper label on x-axes gives time period of doxycycline treatment (- is noninduced) and lower label indicates day of analysis. Top panels show ptet proximal amplicon (amplicon 3 in Figure 1a), and bottom panels show EGFP amplicon (amplicon 2). Quantification is depicted as fold change compared to noninduced, undifferentiated cells. (D) Difference in EGFP FI ratios from (B) between cells after doxycycline washout (recovered) and under doxycycline treatment (repressed). (E) Difference in relative expression of ptet and EGFP from (C) between cells with no doxycycline (no DOX) and recovered cells (black), and between recovered and repressed cells (red). The mean and SD of two to three independent experiments for each clone are shown in (B) and (C), and for all clones pooled in (D) and (E); * $p < 0.05$, ** $p < 0.1$, two-sample Student's t test (B) or single factor ANOVA (C), (D), (E).

In contrast, expression levels of EGFP change significantly over time of differentiation. The difference between uninduced and recovered cells increases, whereas the difference between recovered and repressed cells decreases, demonstrating that EGFP expression levels do not recover from DOX-induced repression at later time points of differentiation (Fig. 4E, right panel). Importantly, EGFP mRNA abundance in the absence of DOX did not substantially decrease during differentiation, but stayed at levels comparable to those in undifferentiated cells. Taken together, these data indicate that antisense transcription during differentiation, in contrast to the reversible silencing observed in undifferentiated cells, might lead to stable repression of a gene on the opposite strand.

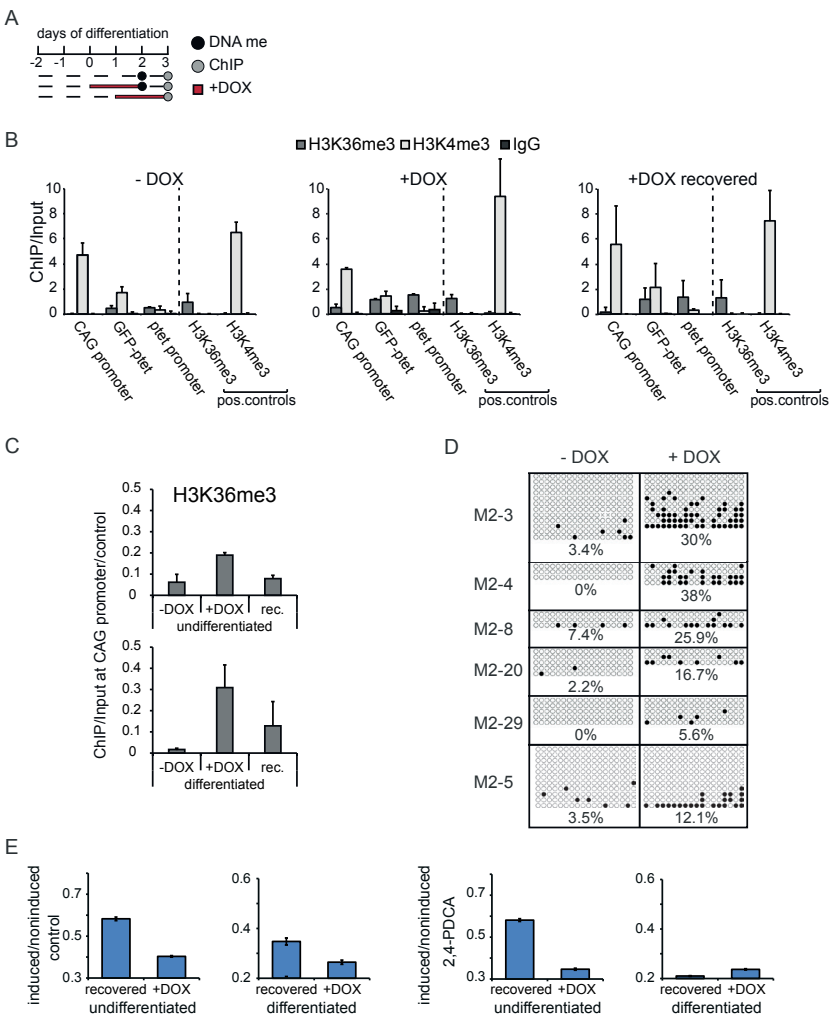


Figure 5. Epigenetic modifications of the CAG promoter in differentiating ESCs.

(A) Time schedule of induction experiments in differentiating ESCs. Dashed lines represent time that cells were grown in absence of doxycycline, red lines represent time that cells were grown in presence of doxycycline. Grey dots indicate time point of ChIP analysis, black dots indicate time point of bisulfite sequencing. (B) ChIP-qPCR analysis of region encompassing CAG and ptet promoters in differentiating M2-3 clone for noninduced (-DOX), induced (+DOX), and recovered (+DOX recovered) conditions as outlined in (A) and main text. Antibodies against H3K36me and H3K4me3, and whole IgG as mock control, were used as indicated on top. Loci analyzed, including H3K36me3 and H3K4me3 controls, are indicated. Values are plotted as the ratio of original amount of template DNA, for the pull-down and input fractions. The mean and SD of two independent experiments are shown. (C) Direct comparison of H3K36me3 levels on the CAG promoter between undifferentiated and differentiating cells. Values for undifferentiated (Fig. 3A) and differentiating cells (Fig. 5A) were divided by corresponding values of H3K36me3 positive controls for normalization. The mean and SD of two (differentiated) or three (undifferentiated) independent experiments are shown. (D) Bisulfite sequencing of CAG promoter in several differentiating clones treated (+DOX) or not treated with doxycycline (-DOX) for two days as outlined in (A). For independent clones represented in the separate panels, filled circles are methylated, empty circles unmethylated CpGs. Average percentages for each condition and clone are indicated. (E) Mean EGFP FI of induced/noninduced M2-3 cells after two days of doxycycline treatment (+DOX) and after two days of doxycycline treatment followed by one day of recovery (recovered). Left panels show undifferentiated, right panels differentiating cells, all were treated with 2,4-PDCA. Error bars represent SD of three independent experiments.

Chromatin structure induced by antisense transcription during differentiation

Since we observed that transcriptional antisense read-through resulted in a specific chromatin signature at the CAG promoter without altering DNA methylation levels in undifferentiated ESCs, we next asked which effect antisense transcription would have on chromatin structure during differentiation. Therefore, differentiating cells were grown in absence or presence of DOX, or were allowed to recover for 1 day after DOX removal, and ChIP analysis of H3K4me3 and H3K36me3 was performed at day three of differentiation (Fig. 5A,B,C). Similar to undifferentiated cells, without DOX, in the absence of ptet transcription, H3K4me3 was found to be strongly enriched at the CAG promoter, while H3K36me3 levels were close to background. In DOX-induced cells, H3K36me3 becomes enriched at the CAG promoter, while H3K4me3 levels decrease slightly, suggesting that antisense transcription through the CAG promoter creates a specific chromatin environment also during differentiation. However, one day after DOX wash-out, reversal of chromatin marks to the noninduced state was less complete than in undifferentiated cells, even though the absolute drop was equally prominent (Fig. 5C). This might be attributed to either the enhanced levels of DOX-induced ptet transcription during later phases of differentiation (Fig. 4C) or to a more stable silencing of EGFP. DNA methylation, which is believed to be important for stable repression of the inactivated X chromosome (Sado et al., 2000) and several other genes (Mohn et al., 2008), is highly dynamic during and essential for embryonic development (Borgel et al., 2010; Okano et al., 1999). We therefore tested if DNA methylation is involved in the repression of the EGFP reporter, by bisulfite sequencing of cells differentiated for 2 days. Strikingly, all clones analyzed displayed a marked increase in DNA methylation in the CAG promoter which was strictly dependent on antisense transcription (Fig. 5D). Values ranged from 0% to 8% methylated CpGs without DOX addition

up to 40% CpG methylation after two days of DOX treatment. These findings indicate that antisense transcription generates a particular chromatin signature and, contrary to the situation in undifferentiated ESCs, is capable of inducing promoter-associated CGIs DNA methylation only in the specific context of differentiation. These events might lead, in turn, to stable gene repression.

To follow up on the observations that properly assembled chromatin and specific combinations of chromatin modifications may have a role in the antisense transcription-mediated repression of the EGFP reporter, we used several small molecule inhibitors (Table S1) to interfere with enzymes that catalyze DNA methlation or the addition or removal of histone modifications. As before, pulse-chase type experiments on undifferentiated and differentiating M2-3 cells were performed, but this time in presence of these inhibitors. Most inhibitors had only minimal effects on repression and recovery in undifferentiated and differentiating cells (Fig. S2). However, in contrast to all other small molecule inhibitors used in this study, 2,4-PDCA, a histone demethylase inhibitor with high specificity for JARID1 and JMJD2 family demethylases which are responsible for H3K4me3 and H3K36me3 demethylation respectively (Kristensen et al., 2012), strongly enhanced both the direct repressive effect and stable silencing by antisense transcription during differentiation (Fig. 5E). Taken together, these data point to a general role for H3K4me3 and/or H3K36me3 in antisense transcription-mediated repression and the establishment of silent chromatin at the CAG promoter driving EGFP. In particular, maintenance of the silent state, which is only put into place during differentiation, appears to be influenced by the H3K4me3 and H3K36me3 histone modifications.

DISCUSSION

Transcriptional interference mechanisms have been thoroughly studied in prokaryotes and yeast (reviewed in Shearwin et al., 2005). These studies indicate inhibition of transcription of the sense gene of a sense-antisense gene pair, which might involve transcription invoked torsional effects or transcriptional collision. In addition, studies of the SRG1/SER3 system in yeast have implicated SRG1 transcription dependent nucleosome occupancy in the SER3 promoter in repression of SER3 (Hainer et al., 2011; Thebault et al., 2011).Torsional or topological effects have also been implicated in transcriptional interference in higher eukaryotes (Eszterhas et al., 2002), and for the *Xist/Tsix* locus chromatin remodelling of the *Xist* promoter by overlapping Tsix transcription has been shown to be involved in Xist silencing (Sado et al., 2005; Ohhata et al., 2008). Our findings indicate that antisense-mediated repression of a sense gene is absent on transiently transfected templates, precluding an important role for direct transcriptional interference in silencing of the reporter. Our ChIP experiments revealed the absence of chromatin modifications that are normally found on templates randomly integrated at different positions in the genome, suggesting that histone modifications play a key role in antisense mediated repression of fully overlapping sense-antisense gene pairs.

We found that silencing of a stably integrated reporter plasmid is reversible in ESCs and is accompanied by an increase in H3K36me3 and a reduction of H3K4me3 in the CAG promoter region driving EGFP transcription. Interestingly, repression of EGFP is stabilized during the ESC differentiation process, concomitant with partial maintenance of accumulated H3K36me3 and loss of H3K4me3, and a significant increase in CpG methylation at the EGFP promoter. Addition of inhibitors interfering with specific epigenetic pathways had little effect on GFP expression during and after recovery of antisense transcription, for most compounds tested. However, addition of 2,4-PCDA had a pronounced effect on EGFP expression in differentiating ESCs, both during doxycyclin-induced antisense transcription and after recovery from this antisense transcription. The compound 2,4-PCDA inhibits H3K4me3 and H3K36me3 demethylases, so that the results suggest that accumulation of these modifications leads to silencing of the CAG promoter, which might be the case also for any other gene promoter with overlapping antisense transcription. The observed effect of 2,4-PCDA was also present, but less pronounced, in undifferentiated ESCs, possibly related to the reversibility of the silencing process. Interestingly, treatment of differentiating ESCs with the DNA methylation inhibitor 5-aza-dC revealed the opposite effect, resulting in an increase in EGFP expression, pointing at a specific role for DNA methylation in silencing of antisense promoters in a developmental context.

Although we cannot formally exclude a role for the non-coding antisense RNA in this process, the synthetic nature of our reporter construct favors a model were the act of transcription and pol II associated chromatin remodelers play a crucial role in the regulation of sense-antisense gene pairs where at least one of the genes initiates transcription through the promoter of the other gene. In yeast, the methyltransferase *Set2* associates with pol II, and the resulting accumulation of H3K36me3 in the gene body is important for recruitment of the histone deacytelase *Rpd3* (Li et al., 2002; Carrozza et al., 2005; Houseley et al., 2008). In mammals, H3K36me3 is involved in recruitment of the H3K4me3 demethylase KDM5B (Xie et al., 2011), and the DNA *de novo* methyltransferases DNMT3A (Dhayalan et al., 2010), implicating a role for H3K36me3 in repression of transcription from cryptic intragenic promoters. Our data are consistent with these findings and suggest that promoter-associated H3K36me3 reversibly represses transcription initiation in ESCs, which might be dependent on the recruitment of KDM5B whose function is associated with ESC self-renewal. Upon ESC differentiation, H3K36me3 enriched gene bodies, including our antisense transcribed EGFP reporter, might be targeted specifically by DNMT3A which is upregulated during this differentiation process. Whether silencing of the EGFP promoter only requires H3K36me3 or is also dependent on H3K4me3 needs further investigation. Similar observations have been made for *Xist* and *Tsix*, two endogenous overlapping gene loci. Loss of *Tsix* antisense transcription through the *Xist* promoter has been shown to result in promoter-associated chromatin changes, allowing aberrant initiation of *Xist* transcription (Sado et al., 2005; Ohhata et al., 2008; Navarro et al.,

2006). In addition, forced *Tsix* expression during development results in *Xist* promoter methylation (Ohhata et al., 2011). *Xist* promoter methylation is required to stably repress *Xist* at later stages of development, also in the absence of ongoing *Tsix* transcription, which is shut down in differentiated cells (Beard et al., 1995). Also for the imprinted *Igf2r/Airn* locus, Airn antisense transcription through the *Igf2r* promoter is required for silencing of *Igf2r* (Sleutels et al., 2002). Studies with an inducible *Airn* promoter indicate that antisense *Airn* transcription leads to CpG methylation of the *Igf2r* promoter, stabilizing the silent state, which can then be maintained in the absence of *Airn* transcriptional readthrough (Santoro et al., 2013). The present findings obtained for an experimental sense-antisense gene pair in undifferentiated and differentiating ESCs, taken together with the above-described findings on physiological gene pairs, clearly demarcate a developmental time window where irreversible silencing is established. Our experimental system provides a powerful tool to study the respective regulatory mechanisms.

In prokaryotes and yeast, it has been found that genes with a clear 'on-off' switch show an enrichment for antisense expression from a neighbouring locus (Duhring et al., 2006). This antisense transcription has been implicated in providing thresholds that need to be overcome for sense genes to be expressed (Xu et al., 2011). Also in higher eukaryotes, the best studied sense-antisense fully overlapping gene pairs, including *Xist/Tsix* and *Igf2r/Airn*, show such a binary switch pattern in gene expression during development, where the antisense partner provides a threshold for transcription initiation of the sense gene. Our findings with an engineered reporter indicate that fully overlapping sense-antisense, and possibly sense-sense, gene pairs might be subject to a general silencing mechanism, which does not involve transcriptional interference, but at least partially relies on transcription mediated accumulation of histone modifications in promoters leading to gene silencing (Fig.6). Our model does not exclude alternative models like nucleosome occupancy mediated repression of overlapping promoters as observed for the SRG1/SER3 locus in yeast (Hainer et al., 2011, Thebault et al., 2011) or ncRNA-mediated recruitment of chromatin modifiers at specific loci as proposed for *Airn*, *Kcnq1ot1* and *Xist* (Pandey et al., 2008; Nagano et al., 2008; Plath et al., 2003). Rather, several complementary mechanisms might act cooperatively to ensure faithful regulation of overlapping gene pairs. Genome-wide strand specific RNA-seq and ChIP-seq studies will be required to determine whether such a general surveillance mechanism is indeed active in mammalian systems.

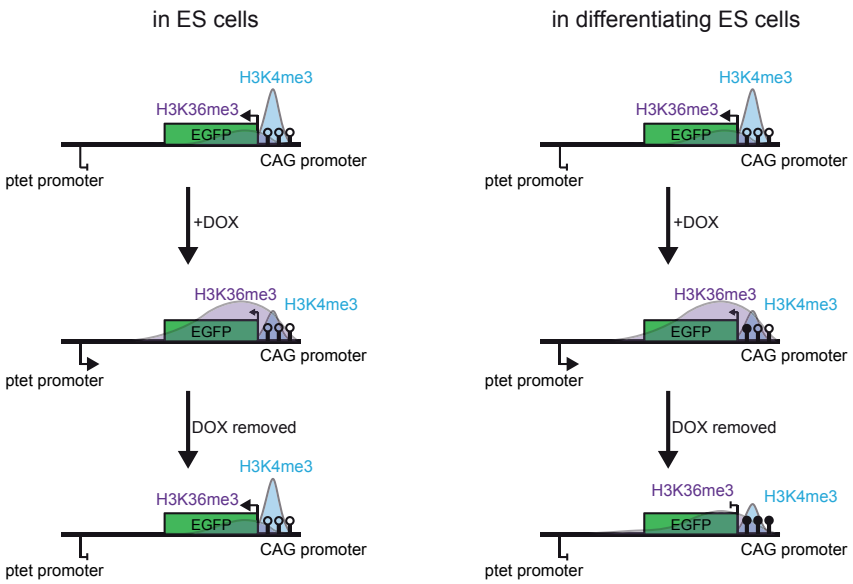


Figure 6. Model of chromatin mediated repression of sense-antisense gene pairs.

In undifferentiated ES cells (left panel), transcription through the CAG promoter driving EGFP expression results in a specific chromatin signature at and repression of this promoter. This effect is completely dependent on antisense transcription and thus reversible upon doxycycline washout. In contrast, antisense transcription in differentiating ES cells (right panel) leads to accumulation of DNA methylation (filled circles) in the CAG promoter and stable silencing of EGFP even if antisense transcription is stopped at later stages of differentiation. Only histone modifications resulting from ptet transcription are displayed.

ACKNOWLEDGEMENT

This work was supported by NWO VICI (projectnr: 865.10.003) and ERC grants (project nr:260587).

We would like to thank all department members for helpful discussions.

4

REFERENCES

Alexopoulou AN, Couchman JR, Whiteford JR. 2008. The CMV early enhancer/chicken beta actin (CAG) promoter can be used to drive transgene expression during the differentiation of murine embryonic stem cells into vascular progenitors. BMC Cell Biol 9:2.

Barakat TS, Gunhanlar N, Gontan Pardo C, 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.

Beard C, Li E, Jaenisch R. 1995. Loss of methylation activates Xist in somatic but not in embryonic cells. Genes Dev 9:2325-2334.

Borgel J, Guibert S, Li Y, Chiba H, Schubeler D, Sasaki H, Forne T, Weber M. 2010. Targets and dynamics of promoter DNA methylation during early mouse development. Nat Genet 42:1093-1100.

Brouwer JR, Willemsen R, Oostra BA. 2009. The FMR1 gene and fragile X-associated tremor/ataxia syndrome. Am J Med Genet B Neuropsychiatr Genet 150B:782-798.

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP, Workman JL. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell 123:581-592.

Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A. 2010. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. J Biol Chem 285:26114-26120.

Duhring U, Axmann IM, Hess WR, Wilde A. 2006. An internal antisense RNA regulates expression of the photosynthesis gene isiA. Proc Natl Acad Sci U S A 103:7054-7058.

Eszterhas SK, Bouhassira EE, Martin DI, Fiering S. 2002. Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. Mol Cell Biol 22:469-479.

Friedrich G, Soriano P. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev 5:1513-1523.

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.

Hainer SJ, Pruneski JA, Mitchell RD, Monteverde RM, Martens JA. 2011. Intergenic transcription causes repression by directing nucleosome assembly. Genes Dev 25:29-40.

Hebbar PB, Archer TK. 2008. Altered histone H1 stoichiometry and an absence of nucleosome positioning on transfected DNA. J Biol Chem 283:4595-4601.

Houseley J, Rubbi L, Grunstein M, Tollervey D, Vogelauer M. 2008. A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. Mol Cell 32:685-695.

Jeong S, Stein A. 1994. Micrococcal nuclease digestion of nuclei reveals extended nucleosome ladders having anomalous DNA lengths for chromatin assembled on non-replicating plasmids in transfected cells. Nucleic Acids Res 22:370-375.

Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J, Suzuki H, Carninci P, Hayashizaki Y, Wells C, Frith M, Ravasi T, Pang KC, Hallinan J, Mattick J, Hume DA, Lipovich L, Batalov S, Engstrom PG, Mizuno Y, Faghihi MA, Sandelin A, Chalk AM, Mottagui-Tabar S, Liang Z, Lenhard B, Wahlestedt C, Group RGER, Genome Science G, Consortium F. 2005. Antisense transcription in the mammalian transcriptome. Science 309:1564-1566.

Kristensen LH, Nielsen AL, Helgstrand C, Lees M, Cloos P, Kastrup JS, Helin K, Olsen L, Gajhede M. 2012. Studies of H3K4me3 demethylation by KDM5B/Jarid1B/PLU1 reveals strong substrate recognition in vitro and identifies 2,4-pyridine-dicarboxylic acid as an in vitro and in cell inhibitor. FEBS J 279:1905-1914.

Kumaki Y, Oda M, Okano M. 2008. QUMA: quantification tool for methylation analysis. Nucleic Acids Res 36:W170-W175.

Ladd PD, Smith LE, Rabaia NA, Moore JM, Georges SA, Hansen RS, Hagerman RJ, Tassone F, Tapscott SJ, Filippova GN. 2007. An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. Hum Mol Genet 16:3174-3187.

Latos PA, Pauler FM, Koerner MV, Senergin HB, Hudson QJ, Stocsits RR, Allhoff W, Stricker SH, Klement RM, Warczok KE, Aumayr K, Pasierbek P, Barlow DP. 2012. Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. Science 338:1469-1472.

Li J, Moazed D, Gygi SP. 2002. Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. J Biol Chem 277:49383-49388.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402-408.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454:766-770.

Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D. 2008. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol Cell 30:755-766.

Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, Fraser P. 2008. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science 322:1717-1720.

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.

Nesterova TB, Popova BC, Cobb BS, Norton S, Senner CE, Tang YA, Spruce T, Rodriguez TA, Sado T, Merkenschlager M, Brockdorff N. 2008. Dicer regulates Xist promoter methylation in ES cells indirectly through transcriptional control of Dnmt3a. Epigenetics Chromatin 1:2.

Ng HH, Robert F, Young RA, Struhl K. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell 11:709-719.

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.

Ohhata T, Senner CE, Hemberger M, Wutz A. 2011. Lineage-specific function of the noncoding Tsix RNA for Xist repression and Xi reactivation in mice. Genes Dev 25:1702-1715.

Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99:247-257.

Pandey R, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C. 2008. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell 32:232-246.

Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, de la Cruz CC, Otte AP, Panning B, Zhang Y. 2003. Role of histone H3 lysine 27 methylation in X inactivation. Science 300:131-135.

Sado T, Fenner MH, Tan SS, Tam P, Shioda T, Li E. 2000. X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. Dev Biol 225:294-303.

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

Santoro F, Mayer D, Klement RM, Warczok KE, Stukalov A, Barlow DP, Pauler FM. 2013. Imprinted Igf2r silencing depends on continuous Airn lncRNA expression and is not restricted to a developmental window. Development 140:1184-1195.

Saxonov S, Berg P, Brutlag DL. 2006. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A 103:1412-1417.

Shearwin KE, Callen BP, Egan JB. 2005. Transcriptional interference--a crash course. Trends Genet 21:339-345.

Sleutels F, Zwart R, Barlow DP. 2002. The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature 415:810-813.

Thakur N, Tiwari VK, Thomassin H, Pandey RR, Kanduri M, Gondor A, Grange T, Ohlsson R, Kanduri C. 2004. An antisense RNA regulates the bidirectional silencing property of the Kcnq1 imprinting control region. Mol Cell Biol 24:7855-7862.

Thebault P, Boutin G, Bhat W, Rufiange A, Martens JA, Nourani A. 2011. Transcription regulation by the noncoding RNA SRG1 requires Spt2-dependent chromatin deposition in the wake of RNA polymerase II. Mol Cell Biol 31:1288-1300.

Tufarelli C, Stanley JA, Garrick D, Sharpe JA, Ayyub H, Wood WG, Higgs DR. 2003. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. Nat Genet 34:157-165.

Williamson CM, Ball ST, Dawson C, Mehta S, Beechey CV, Fray M, Teboul L, Dear TN, Kelsey G, Peters J. 2011. Uncoupling antisense-mediated silencing and DNA methylation in the imprinted gnas cluster. PLoS Genet 7:e1001347.

Xie L, Pelz C, Wang W, Bashar A, Varlamova O, Shadle S, Impey S. 2011. KDM5B regulates embryonic stem cell self-renewal and represses cryptic intragenic transcription. EMBO J 30:1473-1484.

Xu Z, Wei W, Gagneur J, Clauder-Munster S, Smolik M, Huber W, Steinmetz LM. 2011. Antisense expression increases gene expression variability and locus interdependency. Mol Syst Biol 7:468.

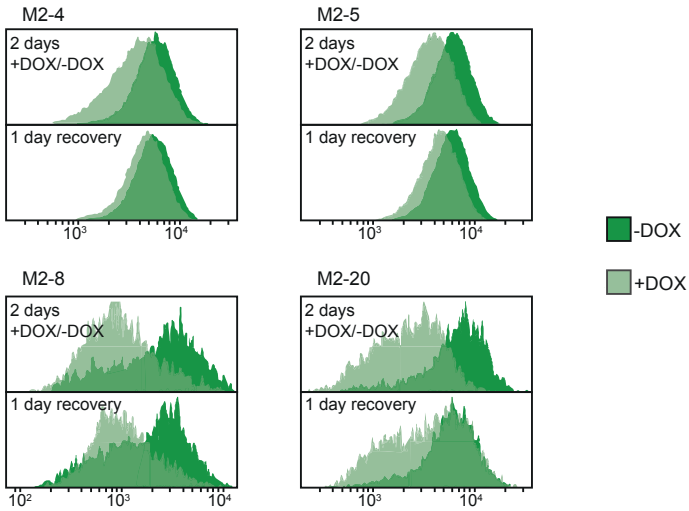


Figure S1.

Representative histograms of EGFP fluorescence intensity (FI) distribution in clones M2-4, M2-5, M2-8 and M2-20 as determined by FACS analysis. Dark green represents uninduced cells, and light green the cells treated with doxycycline. Upper panels shows repression after two days of doxycycline treatment as compared to untreated control, and lower panels shows the respective situation one day later, after doxycycline has been washed out.

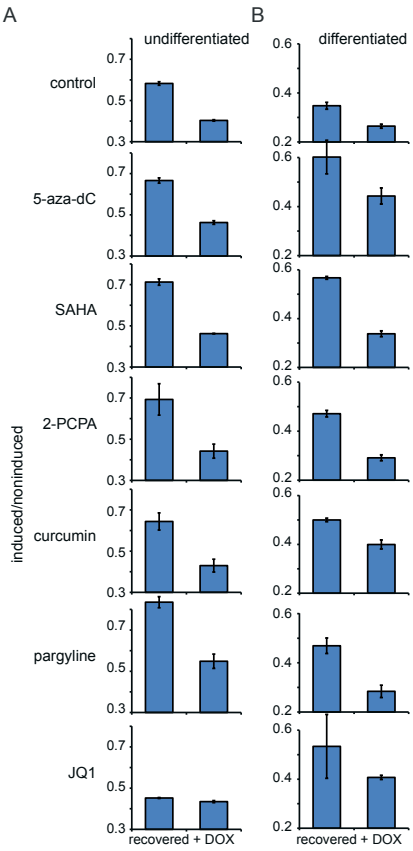


Figure S2. Interference of small molecule inhibitors with chromatin modifiers.

(A) Mean EGFP FI of induced/noninduced M2-3 cells after two days of doxycycline treatment (+DOX) and after two days of doxycycline treatment followed by one day of recovery (recovered). Small molecule inhibitors used are indicated. Error bars represent SD of three independent experiments. (B) Same as (A) but in differentiating M2-3 cells. Time schedule as in Figure 5A but all experimental steps taken one day earlier and cells analyzed at day 2 of differentiation. Error bars represent SD of three independent experiments.

name	abbreviation	targets	inhibition
5-aza-2'-deoxycytidine	5-aza-dC	DNMTs	DNA methylation
suberanilohydroxamic acid	SAHA	HDACs 1-9	histone deacetylation
2,4-pyridinedicarboxylic acid	2,4-PDCA	JARID1/JMJD2 family	H3K4me3/H3K36me3 demethylation
tranylcypromine	2-PCPA	LSD1	H3K4me1/2 demethylation
curcumin	CUR	HATs	histone acetylation
pargyline	PAR	LSD1	H3K9me1/2 demethylation
thieno-triazolo-1,4-diazepine	JQ1	BET family/BRD4	BRD4/transcriptional elongation

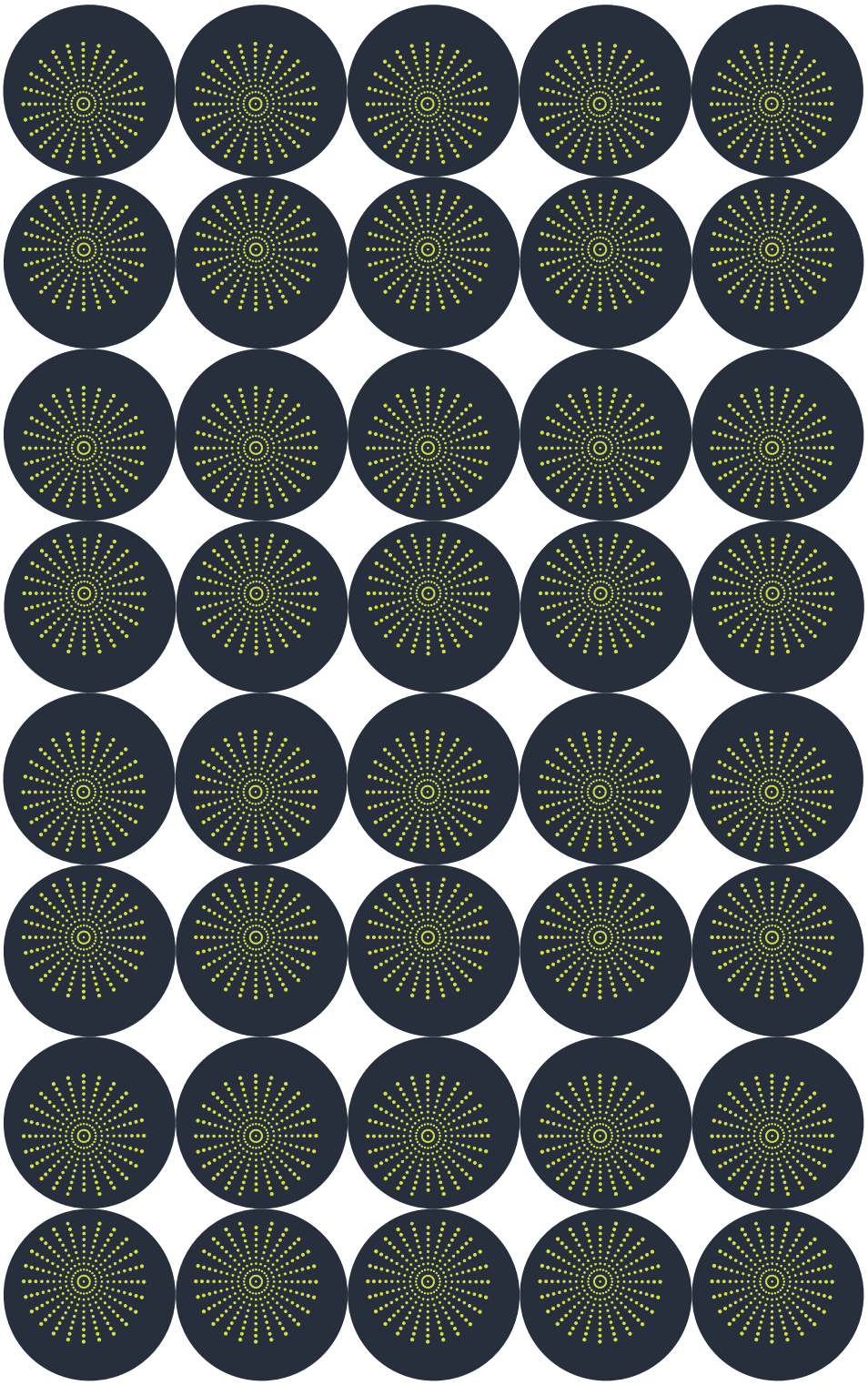
Table S1.

Small molecule inhibitors and their effects on histone modifying enzymes. First column gives name, second abbreviation, third enzymes targeted and fourth biological process inhibited.

#	SEQUENCE	DESCRIPTION
228	ACACGCAGCTCATTGTAG	First strand primer for strand-specific *Actb* expression
226	GATATCGCTGCGCTGGTCGT	FOR primer *Actb* expression
227	AGATCTTCTCCATGTCGTCC	REV primer *Actb* expression
217	CTTCTCGTTGGGGTCTTTGC	First strand primer for strand-specific EGFP expression
106	AGGGCATCGACTTCAAGGAG	FOR primer EGFP expression
107	CACCTTGATGCCGTTCTTCTG	REV primer EGFP expression
104	CAAGATCCGCCACAACATCG	First strand primer for strand-specific ptet proximal expr.
222	TTTCACTGCATTCTAGTTGTGGT	FOR primer ptet proximal expression
223	GGTACCCGGGGATCCTCTA	REV primer ptet proximal expression
220	TGGTAATCGTGCGAGAGGG	First strand primer for strand-specific ptet distal expr.
231	TCCCCTTCTCCCTCTCCAG	FOR primer ptet distal expression
232	CTGCAGAATTCTAGAGCCGC	REV primer ptet distal expression
195	CTCTGACTGACCGCGTTACT	FOR primer ChIP for CAG promoter
196	TTTCACGCAGCCACAGAAAA	REV primer ChIP for CAG promoter
323	CACATTTGTAGAGGTTTTACTTGCT	FOR primer ChIP for GFP-ptet
324	AGCTGCAATAAACAAGTTAACAACA	REV primer ChIP for GFP-ptet
317	GAGCTCGAATTCTCCAGGCG	FOR primer ChIP for ptet promoter
318	GTATGTCGAGGTAGGCGTG	REV primer ChIP for ptet promoter
343	TCTCCCAGCATCCTCTACACA	FOR primer ChIP for H3K36me3 control in *Mcm2*
344	CTATGGTATGTGTGGTGGGCA	REV primer ChIP for H3K36me3 control in *Mcm2*
331	GCTAGGTTAGGAGAGCCCAGA	FOR primer ChIP for H3K4me3 control in *Mrps14*
332	AGGTCTCAATCATCCGACTCTC	REV primer ChIP for H3K4me3 control in *Mrps14*
204	GGATTTTTTTTGTTTTAAATTTGTG	FOR primer for bisulfite sequencing amplicon
207	AAATAAACTTCAAAATCAACTTACC	REV primer for bisulfite sequencing amplicon
209	GTAAAACGACGGCCAG	FOR primer for amplification from bacteria (M13 -20 FOR)
208	CAGGAAACAGCTATGAC	REV primer for amplification from bacteria (M13 REV)
302	ATTTAGGTGACACTATAG	Sequencing primer for bisulfite sequencing (Sp6)

Table S2.

Primers used in this study.



CHAPTER .5

GENERATION OF A NOVEL *IN VITRO* DIFFERENTIATION STRATEGY TO STUDY THE DYNAMICS OF X CHROMOSOME INACTIVATION IN RAT

Aristea Magaraki[#1], **Agnese Loda[#1]**, Cristina Gontan Pardo[1], Stephen Meek[2], Willy M. Baarends[1], Tom Burdon[2] and Joost Gribnau[1*]

Manuscript in preparation

GENERATION OF A NOVEL *IN VITRO* DIFFERENTIATION STRATEGY TO STUDY THE DYNAMICS OF X CHROMOSOME INACTIVATION IN RAT

Aristea Magaraki[#1], **Agnese Loda[#1]**, Cristina Gontan Pardo[1], Stephen Meek[2], Willy M. Baarends[1], Tom Burdon[2] and Joost Gribnau[1*]

[1]Department of Developmental Biology, Erasmus University Medical Center, Wytemaweg 80, 3015 CN Rotterdam, The Netherlands
[2]The Roslin Institute and R(D)VS, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, Scotland
[#]These authors contributed equally to this work
*To whom correspondence should be addressed. Email: j.gribnau@erasmusmc.nl

ABSTRACT

X chromosome inactivation (XCI) is developmentally regulated and relies on several mechanisms including antisense transcription, non-coding RNA-mediated silencing, and recruitment of chromatin remodelling complexes. *In vitro* modelling of XCI provides a powerful tool to study the dynamics of this process by genetically modifying key regulatory players. Importantly, *in vitro* strategies are based on differentiation of pluripotent stem cells into functional cell types and overcome the need to use early developing embryos, thus increasing the number of species in which XCI can be investigated. However, to date, robust XCI *in vitro* has been achieved exclusively upon differentiation of mouse pluripotent cells. Here, we established a novel monolayer differentiation protocol for rat ES cells to study XCI. We show that efficient XCI initiation can only be achieved upon complete withdrawal of MEK and GSK3 inhibitors upon differentiation. We also show that in differentiating rat female cells, Xist RNA starts accumulating *in cis* on the X chromosome around day 2 of differentiation, and is rapidly followed by H3K27me3 enrichment on the inactive X (Xi) domain. Finally, we demonstrate that the critical roles of RNF12 and REX1 in mediating XCI are well conserved in rats. Our work provides the basis to investigate the mechanisms directing the XCI process in a model organism different from the mouse.

INTRODUCTION

In mammals, X chromosome inactivation (XCI) ensures the dosage compensation of sex chromosomal genes between females (XX) and males (XY) (Gendrel and Heard, 2014; van Bemmel et al., 2016). The process of XCI occurs early upon female embryonic development and is mediated by a multitude of epigenetic mechanisms that result in the complete transcriptional inactivation of one entire X chromosome within the nucleus of every female somatic cell. In

eutherians, initiation of XCI is mediated by long non-coding RNAs, with the non-coding gene *Xist* being the major player of XCI in placental mammals (Grant et al., 2012; Marahrens et al., 1997; Penny et al., 1996; Borsani et al., 1991; Brockdorff et al., 1991). During XCI, Xist RNA spreads *in cis* along the entire length of the X chromosome and trigger chromosome-wide silencing of X-linked genes. The molecular mechanisms by which *Xist* induces transcriptional inactivation remain largely unknown. The study of XCI relies both on *in vivo* and *in vitro* models that allow genetic manipulation of the factors involved, and the vast majority of our current knowledge has been achieved by using the mouse as a model organism. *In vivo* studies have shown that XCI starts around the 4-8 cell stage of female mouse embryonic development and is initially imprinted (iXCI), resulting in exclusive inactivation of the paternal X chromosome (Xp) (Huynh and Lee, 2003; Mak et al., 2004; Okamoto et al., 2004; Patrat et al., 2009). Later on in development, at the blastocyst stage (~E4.5), the Xp becomes reactivated in the inner cell mass (ICM) of the embryo, whereas iXCI is maintained in the extra-embryonic lineages (Mak et al., 2004; Okamoto et al., 2004). Reactivation of Xp in the ICM is then followed by random inactivation (rXCI) of either the paternal or maternal X chromosome in cells of the developing epiblast. *In vitro*, mouse embryonic stem cells (mESCs) have been extensively used to model rXCI. In fact, undifferentiated mESCs carry two active X chromosomes and faithfully mimic the pluripotent environment of the ICM, whereas their differentiation results in random inactivation of one of the two X chromosomes. Mouse ESC-based *in vitro* studies have led to the discovery of the long non-coding gene *Tsix*, which is transcribed antisense to *Xist* and represents the major repressor of *Xist* up-regulation at the onset of XCI (Lee and Lu, 1999; Navarro et al., 2006; Sado et al., 2005; Ohhata et al., 2007). XCI is tightly linked to loss of the pluripotent state (Wutz and Jaenisch, 2000; Schulz et al., 2014), and several pluripotency factors including NANOG, SOX2, OCT4, REX1 and PRDM14 have been described to function as XCI-inhibitors either by directly inhibiting *Xist* expression or by enhancing *Tsix* upregulation (Ma et al., 2011; Navarro et al., 2008; 2010; Payer et al., 2013). Activation of XCI is mediated by the X-linked E3 ubiquitin ligase RNF12 involved in dose-dependent degradation of REX1 (Jonkers et al., 2009; Gontan et al., 2012). Interestingly, the study of XCI in female pre-implantation embryos from different species suggested that the epigenetic processes that mediate XCI might be more heterogeneous than expected. Indeed, iXCI occurs in the extra-embryonic lineages of rat and cow (Wake et al., 1976; Dindot et al., 2004; Xue et al., 2002), whereas in other species such as human, monkey, horse, pig and rabbit, rXCI has been exclusively observed in both embryonic and extra-embryonic tissues (Okamoto et al., 2011; Moreira de Mello et al., 2010). Comparative analysis of Xist RNA expression dynamics and X-linked gene silencing between rabbit and human pre-implantation embryos confirmed substantial diversity in the timing and regulation of XCI initiation among mammals, with cells of the human ICM showing two active X chromosomes regardless of Xist RNA coating (Okamoto et al., 2011). In addition, the overall *Xist* gene structure appears to be conserved in all placental mammals, but *Xist's* sequence evolved

rapidly and differs between species (Chureau et al., 2002; Duret et al., 2006; Nesterova et al., 2001; Elisaphenko et al., 2008). Finally, *Tsix* antisense transcription through the *Xist* promoter has not been found in human (Migeon et al., 2001; 2002) but appears to be conserved in rodents (Shevchenko et al., 2011). Interestingly, differentiation of mouse-rat allodiploid ES cells leads to specific primary inactivation of the mouse X chromosome (Li et al., 2016). This mouse allele-biased expression of *Xist* has been proposed to result from the higher expression of *Tsix* from the rat allele, interfering with expression of *Xist in cis* (Li et al., 2016).

Clearly, the development of novel *in vitro* systems derived from different species is necessary to reach a comprehensive understanding of the XCI process. However, although the induced pluripotent stem cells (iPSCs) technology has allowed the generation of several ES cell-like lines from different species (Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Ben-Nun et al., 2011), both the characterization of the X chromosomes status and the generation of *in vitro* differentiation protocols that recapitulate XCI have been proven to be challenging (Tchieu et al., 2010; Mekhoubad et al., 2012; Pasque and Plath, 2015). In this context, rat ES cells (rESCs) only recently became well characterized (Meek et al., 2014; 2010; 2013; Buehr et al., 2008; Li et al., 2008; Kawamata and Ochiya, 2010a; 2010b; Hirabayashi et al., 2010; Men et al., 2012), and the establishment of the novel CRISPR/Cas9 system for genome editing rapidly enhanced the generation of genetically modified rat models potentially facilitating genetic studies on XCI in rESCs (Shao et al., 2014; Guan et al., 2014). Therefore, we set out to generate a robust *in vitro* system that could faithfully mimic the dynamics of XCI in rat. By developing a novel monolayer differentiation protocol for rESCs similar to the one recently reported by by Vaskova and colleagues (Vaskova et al., 2015), we were able to follow some aspects of XCI regulation. Similar to mouse, we were able to observe (I) *Xist* up-regulation at an early stage of rESCs differentiation followed by (II) transcriptional inactivation of X-linked genes and (III) H3K27me3 accumulation on the inactive X chromosome (Xi). In addition, (IV) overexpression experiments in rESCs confirmed that the REX1-RNF12 axis of *Xist* regulation is conserved between rat and mouse. Thus, our data has established the technical basis to study the dynamics of XCI in a different system from the mouse and suggests that specific aspects of XCI are conserved in rodents.

RESULTS

In vitro neuronal differentiation of rESCs

In vitro differentiation of mESCs towards different functional cell types including neurons, cardiomyocytes, hepatocytes and pancreatic cells can be efficiently achieved by several established protocols (Schroeder et al., 2009). Usually, differentiation strategies are based on the formation of embryoid bodies (EB) followed by growth-factor-mediated induction of early progenitor cells to differentiate into their respective lineages. In spite of the growing list of differentiation protocols for mESCs, differentiation of rESC is extremely difficult to achieve

in vitro. To date, only two strategies have been described in which rESCs were triggered to differentiate into either cardiomyocytes or neuronal precursors (Cao et al., 2011; Peng et al., 2013). XCI is closely linked to loss of pluripotency and the presence of an inactive X chromosome is a powerful readout for cell differentiation. We initially set out to assess rat XCI after inducing rESCs differentiation according to the already established protocols. Several rESCs derived from different rat inbred strains were differentiated, including three pure Lewis lines (LEW) (A4p20, A9p20, A10p20), and two lines of a mixed background of dark agouti (DA) and Sprague-Dawley (SD) (135-7, 141-6). In all cases, we were never able to see either *Xist* up-regulation or its associated X-linked gene silencing, although both female and male rat cells appeared to be morphologically differentiated (data not shown). Interestingly, we observed massive cell death of female rESCs compared to male cells around day 3 of differentiation, suggesting that impairment of XCI initiation might have an impact on cell survival (data not shown).

We reasoned that the lack of XCI features upon differentiation could rely on our culture conditions. In fact, in both experimental strategies, rESC differentiation is never achieved without complete withdrawal of the MEK and GSK3 inhibitors ("2i" culture conditions). Inhibiting both the MAPK and Gsk3ß pathways is necessary for the maintenance of the homogeneous pluripotent ground state of rESCs (Buehr et al., 2008). However, since *Xist* regulation and function are strictly linked to cell differentiation, stabilizing the pluripotent state results in tight repression of *Xist* expression (Schulz et al., 2014; Navarro et al., 2008; 2010; Payer et al., 2013). Therefore, we hypothesized that the supplement of MEK and GSK3 inhibitors at the onset of rESCs differentiation might interfere with *Xist* up-regulation, thus preventing XCI initiation. To test our hypothesis, we implemented the neuronal differentiation protocol initially described by Peng and colleagues as follows (Peng et al., 2013): (I) we completely eliminated the presence of both 2i inhibitors starting from day 1 of neuronal differentiation, (II) we increased the concentration of ROCK (rho-associated protein kinase) inhibitor, which has been shown to prevent dissociation-induced apoptosis in cultured human ES cells (Ishizaki et al., 2000; Watanabe et al., 2007) and finally, (III) we started rESCs differentiation with a greater number of cells. Using these modified conditions, we were able to maintain viable differentiating male and female rESCs in the absence of 2i inhibitors (Figure 1A). Importantly, qPCR analysis of both pluripotency and differentiation marker expression levels at different time points upon differentiation confirmed efficient downregulation of the pluripotency factors *Esrrb, Prdm14* and *Rex1*, and parallel up-regulation of the neuronal precursor marker *Nestin* (Figure 1B).

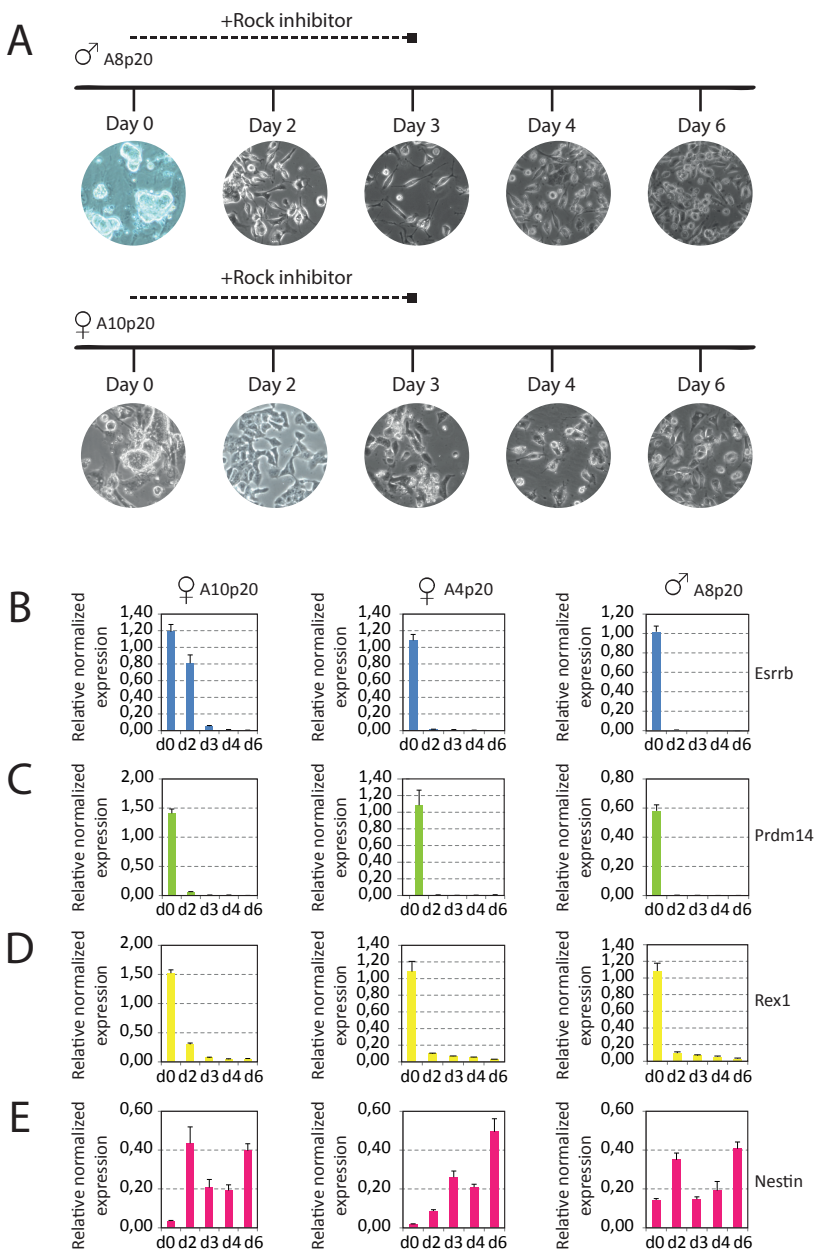


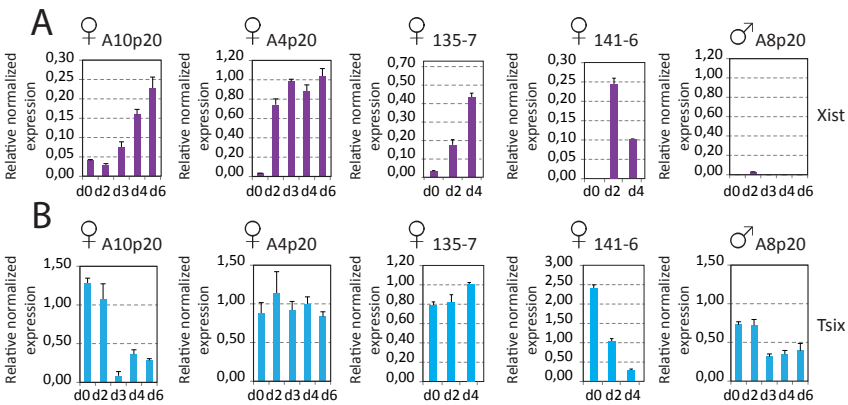
Figure 1. Neuronal differentiation of rESCs.

A. Schematic representation of our neuronal differentiation strategy. Brightfield images of male (A8p20, top) and female (A10p20, bottom) rESCs at several time points upon differentiation are shown. (B-C-D-E) qPCR analysis of (B) Esrrb, (C) Prdm14, (D) Rex1 and (E) Nestin expression levels in female (A10p20, A4p20) and male (A8p20) differentiating rESCs.

Female rESCs underdo XCI upon *in vitro* neuronal differentiation

We then addressed the question of whether differentiating rESCs without the supplement of 2i inhibitors would facilitate XCI to occur. To this end, four independent female rESC lines were differentiated and the Xist RNA expression level was assessed by qPCR analysis at different time points upon neuronal differentiation. Importantly, in order to assess the sex-specific regulation of Xist RNA, one male rESC line was also included into our analysis. As in mouse, *Xist* upregulation occurs exclusively in female rat cells between day 2 and day 4 of differentiation (Figure 2A). In parallel, we also assessed *Tsix* expression levels and contrarily to what is observed in mouse (Loos et al., 2016, chapter 2), *Tsix* appears to be efficiently downregulated only upon differentiation of two out of four female rESC lines. Moreover, male rESCs showed persistent *Tsix* expression throughout differentiation, although the expression levels decreased around day 2 (Figure 2B). These observations suggest that the interplay between *Tsix* and *Xist* regulation at the onset of XCI might slightly differ between mouse and rat.

Next, we addressed the dynamics of *Xist* expression by performing Xist RNA FISH analysis at different time points upon neuronal differentiation. In undifferentiated rESCs, Xist RNA pinpoint signals were observed within the nuclei of both female and male cells (Figure 2C-D). However, since the Xist RNA FISH probe can hybridize to either Xist or Tsix RNA, the pinpoint signal might represent *Tsix* expression instead of *Xist*. Around day 2 of neuronal differentiation, Xist RNA starts to accumulate exclusively on a single X chromosome within female nuclei, whereas Xist RNA accumulation was never observed in differentiating male rESCs (Figure 2C). Importantly, upon differentiation of A10p20 and A4p20 rESC female lines, 60% of the nuclei showed a Xist RNA-coated X chromosome at day 6 of differentiation. Taken together, these observations show that neuronal differentiation of rESCs in absence of 2i inhibitors allows Xist RNA to be upregulated and to spread in cis from a single X chromosome in female cells.



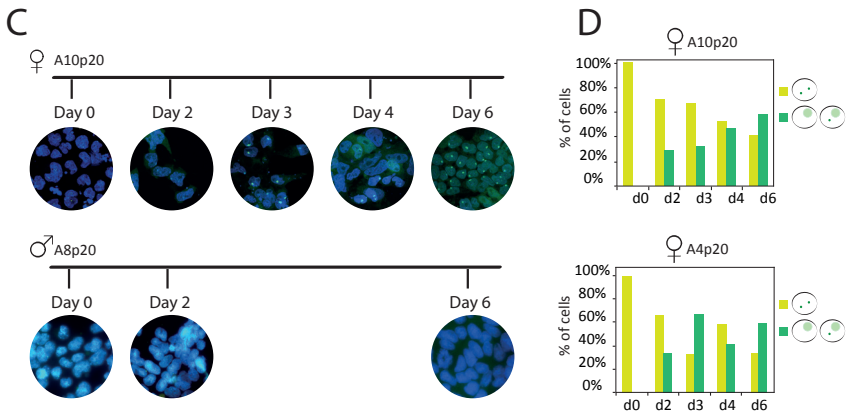


Figure 2. Monoallelic upregulation of Xist RNA upon female rESCs differentiation.

(A) *Xist* and (B) *Tsix* qPCR expression analysis in female (A10p20, A4p20, 135-7, 141-6) and male (A8p20) differentiating rESCs. Expression levels of *Xist* and *Tsix* at different time points upon neuronal differentiation are shown. (C) Representative images of Xist RNA FISH (green) analysis upon differentiation of male (A8p20) and female (A10p20). DNA is stained with DAPI (blue). (D) Quantification of relative number of Xist RNA signals (pinpoints or clouds) in A10p20 and A4p20 female rESCs at day 0, 2, 3, 4 and 6 upon neuronal differentiation.

In addition, we determined at which time point upon neuronal differentiation Xist-mediated silencing of X-linked genes is established. In mouse, the gene silencing-associated H3K27me3 histone modification represents one of the earliest histone marks that accumulate on the Xi during XCI (Chaumeil et al., 2006; Silva et al., 2003; Plath et al., 2003). Therefore, in order to resolve the overall degree of XCI, we followed enrichment of H3K27me3 by immunofluorescence analysis upon differentiation of both male and female rESCs. In undifferentiated rESCs, no H3K27me3 domains were observed in neither male nor female cells (Figure 3A-B). However, starting from day 2 of differentiation and in line with female-specific upregulation of Xist RNA, H3K27me3 starts to accumulate into specific nuclear domains within female cells. Later on, by day 6, more that 60% of the female nuclei show one H3K27me3 domain, thus confirming that XCI is efficiently initiated upon female rESCs differentiation (Figure 3B). Finally, to precisely assess the dynamics of X-linked gene silencing, we followed the Xist-mediated inactivation of the X-linked gene *Pgk1* by two-colour RNA-FISH analysis at different time points upon rESCs differentiation. While the single copy of *Pgk1* in male cells remains actively transcribed throughout differentiation, the transcriptional inactivation of one copy of *Pgk1* in female cells starts around day 2 of differentiation (Figure 3C). However, robust *Pgk1* inactivation in up to 70% of the female nuclei is only reached around day 6 of differentiation (Figure 3C).

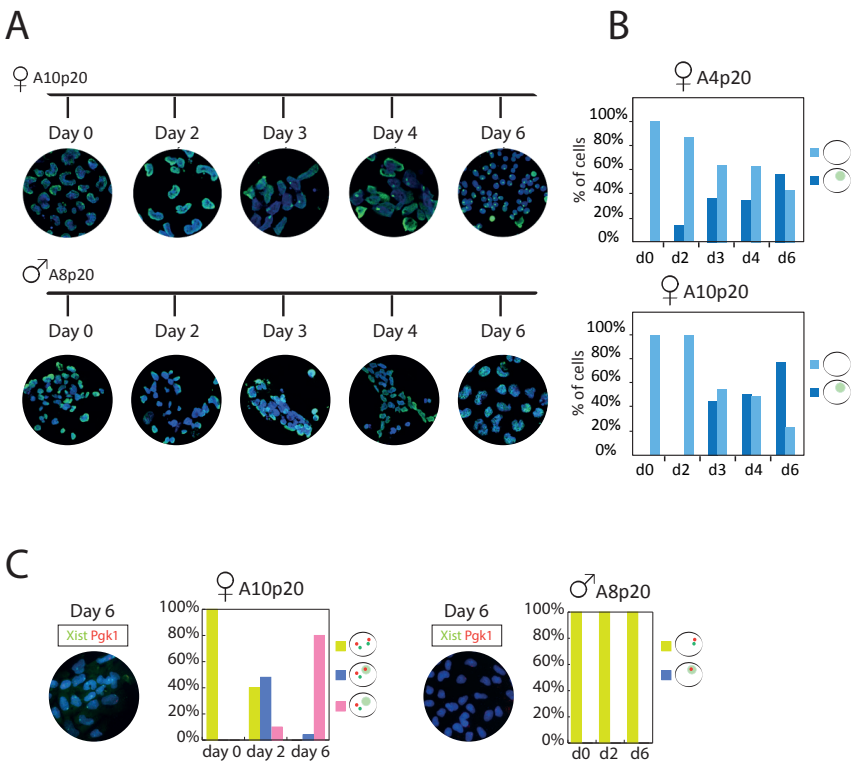
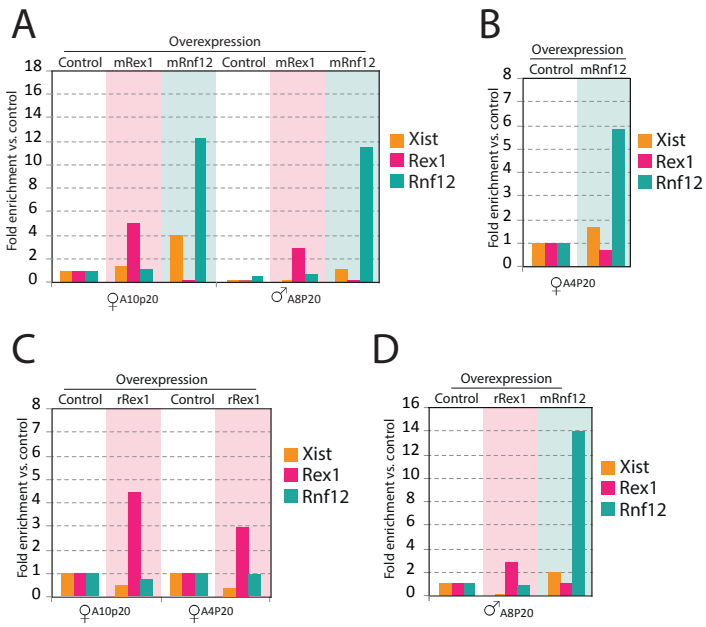


Figure 3. Xist-mediated silencing of X-linked genes.

(A) Representative images of H3K27me3 (green) immunofluorescence analysis in female (A10p20) and male (A8p20) rESC at different time points upon neuronal differentiation. DNA is stained with DAPI (blue). (B) Quantification of relative number of cells carrying a H3K27me3 domain at day 0, 2, 3, 4 and 6 of neuronal differentiation. Data of to A10p20 and A4p20 female rESC lines are shown. (C) Xist(green)/Pgk1(red) two-colour RNA-FISH analysis at different time points upon neuronal differentiation of female (A10p20) and male (A8p20) rESCs. The relative number of cells showing either biallelic or monoallelic *Pgk1* expression is quantified, together with the relative number of cells carrying Xist pinpoints or clouds signals. DNA is stained with DAPI (blue).



E

```
Mouse Rex1 MNEQMKKTAKTSGQKGPGGRALDRLTLKQDEARPVQNTRVEAPRVTYTIRDESEISPETEEDGFPDGY
Rat   Rex1 MNEQMKKMAKTSGQKGPGRRALDRMTLNHDEARPVQNTRVEPPHVTYTIHEEGEISPENEEDVFPNGY

Mouse Rex1 LECIIRGEFSEPILEEDFLFKSFESLEEVEQNLSRQVLEASSLLESSLEYMTKGTKQEKREVTQETPP
Rat   Rex1 LECIIRGEFSEPILEEDLLFKSFESLEEAEQDLSRQVLEASSLLECSLEYMTKGTKQEKRA-TQEEPQ

Mouse Rex1 LRVGASS------------LLAGGPAE---------------KPEGGVYCGVLSMLECPQAGCKKKLR
Rat   Rex1 QTVGANSVPECSEYLTSRKLPVGEPSEVDLSYHEQLSGFTGEKLKGGPYCGNISMLECPQAGCKKKLR

Mouse Rex1 DKTALRKHMLVHGPRRHVCAECGKAFTESSKLKRHFLVHTGEKPHQCTFEGCGKRFSLDFNLRTHTRI
Rat   Rex1 DKTALRKHMLVHGPRQHVCAECGKAFTESSKLKRHFLVHTGERPHQCTFQGCGKRFSLDFNLRTHTRI

Mouse Rex1 HTGERRFVCPFDGCEKSFIQSNNLKIHILTHAKAGKKC
Rat   Rex1 HTGEQRFVCPFDGCEKSFIQSNNLKIHILTHAKAGKKC
```

F

```
Mouse Rnf12 MENSDSNDKGSDQSAAQRRSQMDRLDREEAFYQFVNNLSEEDYRLMRDNNLLGTPGESTEEELLRRL
Rat   Rnf12 MENSDSNDKGSDQSAAQRRSQMDRLDREEAFYQFVNNLSEEDYRLMRDNNLLGTPGESTEEELLRRL

Mouse Rnf12 QQIKEGPPPQSPDENRAGESSDDVTNSDSIIDWLNSVRQTGNTTRSGQRGNQSWRAVSRTNPNSGDF
Rat   Rnf12 QQIKEGPPPQSSDENRAGESSDDVSNSDSIIDWLNSVRQSGNTTRSGQRGNQSWRAVSRTNPNSGDF

Mouse Rnf12 RFSLEINVNRNNGSQTSENESEPSTRRLSVENMESSSQRQMENSASESASARPSRAERNSAEAVTEV
Rat   Rnf12 RFSLEINVNRNNGSQTSENEAEPSTRRLSVESMESSSQRQMESHASESTSARPSRAERNSTEAL-EV

Mouse Rnf12 PTTRAQRRARSRSPEHRRTRARAERSRSPLQPTTEIPRRA------TTLEQSSENEPEGSSRTRHHV
Rat   Rnf12 PLTRGQRRARSRSPEHRRTRARAERSRSPLHPTTEIPRRAHHNISPLTLEQPLENEPEGSSRTRHHV

Mouse Rnf12 TLRQQISGPELLGRGLFAASGSRNPSQGTSSSDTGSNSESSGSGQRPPTIVLDLQVRRVRPGEYRQR
Rat   Rnf12 TLRQQISGPELLGRGLFAASGSRNPSQGTSSSDTGSNSESSGSGQRPPTIVLDLQVRRVRPGEYRQR

Mouse Rnf12 DSIASRTRSRSQAPNNTVTYESERGGFRRTFSRSERAGVRTYVSTIRIPIRRILNTGLSETTSVAIQ
Rat   Rnf12 DSIASRTRSRSQTPNNTVTYESERGGFRRTFSRSERAGVRTYVSTIRIPIRRILNTGLSETTSVAIQ

Mouse Rnf12 TMLRQIMTGFGELSYFMYSDSDSEPSASVSSRNVERVESRNGRGSSGGGNSSGSSSSSPSPSSSGE
Rat   Rnf12 TMLRQIMTGFGELSYFMYSDSDSEPSASVSSRNTERAESRNGRGSS--GAGNSSASSSSPSPSSNGE

Mouse Rnf12 SSESSSEMFEGSSEGGSSGPSRRDGRHRAPVTFDESGSLPFLSLAQFFLLNEDDEDQPRGLTKEQID
Rat   Rnf12 SSESSSDLFEGTSEGGPSGPSRRDGRHRAPVAFDESGSLPFLSLAQFFLLNEDDEDQPRGLTKEQID

Mouse Rnf12 NLAMRSFGENDALKTCSVCITEYTEGNKLRKLPCSHEYHVHCIDRWLSENSTCPICRAVLSSGNRE
Rat   Rnf12 NLAMRSFGENDALKTCSVCITEYTEGNKLRKLPCSHEYHVHCIDRWLSENSTCPICRAVLSSGNRE
```

Figure 4. RNF12 and REX1 overexpression in rESCs.

qPCR analysis of *Xist*, *Rex1* and *Rnf12* expression levels after overexpression of mREX1 and mRNF12 proteins (A-B) or rREX1 (C-D) protein in undifferentiated male (A8p20) and female (A10p20) rESCs. (E-F) Amino acid sequences alignment of mouse and rat REX1 (E) and RNF12 (F). Highlighted in red are the not conserved amino acids. In blue, the four zinc fingers domains of REX1 protein (E) and the RING finger domain of RNF12 (F).

Overexpression of RNF12 and REX1 in undifferentiated rESCs modulates Xist expression

The X-linked E3 ubiquitin ligase RNF12 has been previously shown to activate Xist transcription at the onset of XCI (Barakat et al., 2011; Jonkers et al., 2009). Importantly, the pluripotency factor REX1 has been identified as a key target of RFN12, and dose-dependent degradation of REX1 by RNF12 has been proposed to act as a crucial mechanism directing the initiation of XCI upon differentiation of female mESCs (Gontan et al., 2012). Since the RNF12-REX1 axis represents an important pathway for XCI to occur in mouse, we asked whether these factors play similar roles in rat XCI. To this end, we transiently overexpressed the mouse RNF12 (mRNF12) and REX1 (mREX1) proteins in rESCs, and determined the impact of overexpression on Xist RNA regulation. According to the mouse data, we expected REX1 overexpression to result in the inhibition of *Xist* transcription whereas overexpressing RNF12 would lead to Xist up-regulation (Barakat et al., 2011; Gontan et al., 2012). Xist RNA expression levels were determined by qPCR analysis, and the experiment was performed in two independent female rESC lines and a single male rESC line (Figure 4). Overexpression of mRNF12 consistently resulted in upregulation of Xist RNA in both male and female rESCs, thus confirming RNF12 to act as major *trans*-acting activator of XCI in both mouse and rat (Figure 4A-B-D). Interestingly, overexpressing mREX1 in rESCs did not lead to a clear inhibition of Xist RNA transcription (Figure 4A). This observation might be explained by the fact that Xist RNA repression in undifferentiated mESCs does not exclusively rely on REX1 function but is rather mediated by the combined action of several pluripotency factors (Navarro et al., 2010; 2008; Payer et al., 2013). Therefore, the effect of mREX1 overexpression might be masked by the rESCs pluripotent state itself, with cells that are more prone to differentiate having the tendency to de-repress Xist RNA more easily. However, a second possibility is that the different impact of mRNF12 and mREX1 overexpression on Xist RNA regulation depends on the degree of conservation between the rat and mouse proteins. Indeed, the catalytic ring finger domain of RNF12 shows 100% of amino acid sequence identity between mouse and rat whereas the zinc finger domains of REX1 appeared to be less well conserved (Figure 4E-F). To confirm our hypothesis, overexpression of rat REX1 (rREX1) in undifferentiated rESCs resulted in downregulation of Xist RNA in both female and male cells (Figure 4C-D).

DISCUSSION

Our knowledge concerning the regulation of XCI in developing rat embryos is very limited and relies on conservation of the key regulators *Xist* and *Tsix* between mouse and rat and a very few studies in which, similar to mouse, iXCI has been proposed to occur in early rat embryonic development (Wake et al., 1976; Nesterova et al., 2001; Chureau et al., 2002; Elisaphenko et al., 2008; Duret et al., 2006). Studying the XCI process in rESCs offers the opportunity to explore species-specific epigenetic features and will generally help to reach a more comprehensive understanding of the XCI process in mammals. Despite that rESCs *in vitro*

differentiation protocols have been previously established (Cao et al., 2011; Peng et al., 2013), the difference between female and male differentiation has never been taken into account, and the transcriptional status of the X chromosomes upon rESCs has never been characterized. Here, we did not observe any of the XCI-related epigenetic features of rESCs during differentiation according to the previously established protocols, and we therefore set up a novel monolayer *in vitro* differentiation strategy that efficiently recapitulates the XCI process in rat cells. Importantly, the key feature that allowed us to achieve robust initiation of XCI is the complete absence of 2i inhibitors throughout the entire differentiation protocol. However, in the meantime, Vaskova and colleagues reported a similar strategy to trigger XCI upon differentiation of pluripotent rat cells (Vaskova et al., 2015). As in our protocol, withdrawing 2i inhibitors from the differentiating culture medium resulted in efficient *Xist* up-regulation, thus confirming that inhibition of the MAPK and Gsk3ß pathways upon *in vitro* differentiation of rESCs prevents XCI to occur. We next exploited our *in vitro* system to assess the dynamics of XCI in differentiating rESCs. As in mESCs, both X chromosomes are active in undifferentiated rESCs and Xist RNA monoallelic upregulation starts to occur around day 2 of neuronal differentiation. Interestingly, the downregulation of *Tsix* expression at the onset of rat XCI appeared to be heterogeneous compared to what we observed in mESCs (Loos 2016, chapter 2). However, since allele-specific analysis of *Tsix* expression levels cannot be assessed in our rESCs system, whether the observed *Tsix* expression is derived from the active or the inactive chromosome remains an open question and will need to be addressed in hybrid cell lines.
In addition, we showed that transcriptional inactivation of X-linked genes directly follows *Xist* RNA accumulation on one of the two X chromosomes. In fact, the exclusive enrichment of H3K27me3 loci in female nuclei starts around day 3 of neuronal differentiation, and Xist-mediated silencing of X-linked gene *Pgk1* occurs with similar dynamics. Finally, by overexpressing the XCI key regulators RNF12 and REX1 in undifferentiated rESCs, we have confirmed the conservation of their critical function in directing *Xist* expression. Importantly, we show that overexpression of mRNF12 protein in rESCs efficiently recapitulates RNF12 function, whereas in the case of REX1 only the overexpression of the rat REX1 homologue results in Xist RNA downregulation in rESCs. In line with this observation, the DNA-binding domains of REX1 proteins from different species show an average of 11-20 amino acid differences (Kim et al., 2007) thus confirming that the degree of protein conservation between mouse and rat REX1 homologues may explain our results. Contrarily, RNF12 is highly conserved among mammals (Bach et al., 1999), and the observed up-regulation of rat Xist upon mRNF12 overexpression confirms what we previously observed upon overexpression of human RNF12 in mESCs (Jonkers et al., 2009). In conclusion, we were able to set up a robust *in vitro* system to study the regulation of XCI in differentiating rESCs, recapitulating in addition the main steps of mouse XCI. In the future, the generation of hybrid F1 polymorphic rESCs and several applications of the recently developed CRISPR/Cas9 technology for genomic editing will most likely increase

the use of rat as a model organism in basic epigenetic and biomedical research.

EXPERIMENTAL PROCEDURES

Cell culture and DNA transfection

rESCs were derived as previously described (Meek et al., 2010) and subsequently maintained in N2B27 medium supplemented with 3μM CHIR99021 (Stemgent), 1μM PD0325901 and 1000U/ml mouse LIF on mouse feeders.

For the monolayer differentiation culture plates were coated with 100μg/ml laminin (Sigma-Aldrich) for at least 4 hours at 37o C, followed by three PBS washes. Single rESCs were plated at a density of 105/cm2 for the female cell lines and 2x104/cm2 for the male cell lines in N2B27 supplemented with 10μM of ROCK inhibitor (Sigma-Aldrich) for the first three days. Thereafter, the ROCK inhibitor was eliminated. Medium was refreshed daily.

For the overexpression experiments, the *mRex1*, *rRex1* and *mRnf12* coding sequences were subcloned into pCAG-Flag, a CAG-driven expression vector containing a Flag-tag. rESCs were transfected using lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions, followed by 48 hours of puromycin selection

Probe preparation and Fluorescent in Situ Hybridization

For preparing probes detecting *Xist* and *Pgk1* mRNAs, BACs harbouring these genes were labelled as a whole, with digoxigenine and biotin (Roche) respectively, by nick translation following the manufacturer's instructions.

For RNA-FISH at different time points of neuronal differentiation, cells were grown on glass coverslips and then fixed with 3% PFA for 10 minutes at room temperature followed by three washes in PBS. Next, cells were permeabilised with 0.5% Triton and washed again three times in PBS. Cytoplasm was removed by treating the cells with 0.025% pepsin in 0.01N HCL for 3 minutes. Subsequently, cells were dehydrated with sequential ethanol washes (70%, 85% and 100% 2 minutes each) and air-dried. Finally, probes were applied on the samples overnight at 37o C in a 50% Formamide/2XSSC humid chamber. The next day, slides were washed two times in 50% Formamide/2xSSC pH=7.4 at 37o C, followed by two washes in 2xSSC at 37° C and cells were blocked for 30 minutes at room temperature with TSBSA (2 mg/ml bovine serum albumin in 0.1 M Tris and 0.15 M NaCl) in a humified champer at room temperature. Detection was performed by incubation with anti-digoxigenine FITC (Boehringer, 1:250) and streptavidin alexa fluor 555 (Thermofisher Scientific, 1:400) in TSBSA for 30 minutes at room temperature. Slides were then washed two times 5 minutes each with TS (0.1 M Tris, 0.15 M NaCl) and mounted with ProLong Gold ProLong® Gold Antifade Mountant with Dapi (ThermoFisher Scientific). Imaging was performed on a Zeiss LSM700 microscope (Carl Zeiss, Jena).

Expression analysis

Cells were lysed by direct addiction of 500 µg of TRIZOL and total RNA was extracted according to the manufacturer's instructions (Invitrogen). To remove genomic DNA contamination, samples were treated 15 minutes at 37°C with DNaseI (Invitrogen). Next, 1 µg of RNA was reverse transcribed by Superscript II reverse transcriptase with random hexamers (Invitrogen). For quantitative PCR (qPCR) gene expression levels were quantified using 2x SYBR Green PCR Master Mix (Applied Biosystems) in a CFX384 Real-Time machine (Bio-Rad) with primers listed is Table S1. Expression levels were normalized to actin b using the ΔCT method.

Immunocytochemistry

For immunofluorescence analysis on different time points of neuronal differentiation, cells were grown on glass coverslips and then fixed with 3% PFA for 10 minutes at room temperature followed by three washes in PBS. Therafter, cells were permeabilised with 0.5% Triton, washed with PBS (3x5') and blocked with 2% BSA, 5% donkey serum in PBS (blocking solution) for 30 minutes at room temperature. This was followed by anti- H3K27me3 rabbit (Diagenode, 1:500) incubation, diluted in blocking solution, at 4o C overnight in a humid chamber. The next day, slides were washed in PBS (3x5 minutes) and blocked with donkey anti-rabbit alexa fluor 488 (ThermoFischer Scientific, 1:250) secondary antibody, diluted in blocking buffer, for 1 hour at room temperature, in a humid chamber. Slides were then washed in PBS (3x5 minutes) and mounted with ProLong® Gold Antifade Mountant with Dapi (ThermoFisher Scientific). Confocal imaging was performed on a Zeiss LSM700 microscope (Carl Zeiss, Jena).

5

REFERENCES

Bach I., Rodriguez-Esteban C., Carrière C., Bhushan A., Krones A., Rose D.W., Glass C.K., Andersen B., Izpisúa Belmonte J.C., Rosenfeld M.G. (1999). RLIM inhibits functional activity of LIM homeodomain transcription factors via recruitment of the histone deacetylase complex. Nat Genet. 1999 Aug;22(4):394-9.

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

Ben-Nun, I. F., Montague, S. C., Houck, M. L., Tran, H. T., Garitaonandia, I., Leonardo, T. R., Wang, Y., Charter, S. J., Laurent, L. C., Ryder, O. A., et al. (2011). Induced pluripotent stem cells from highly endangered species. Nat Methods 8, 829-831.

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

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

Buehr, M., Meek, S., Blair, K., Yang, J., Ure, J., Silva, J., McLay, R., Hall, J., Ying, Q., and Smith, A. (2008). Capture of authentic embryonic stem cells from rat blastocysts. CELL 135, 1287-1298.

Cao, N., Liao, J., Liu, Z., Zhu, W., Wang, J., Liu, L., Yu, L., Xu, P., Cui, C., Xiao, L., et al. (2011). In vitro differentiation of rat embryonic stem cells into functional cardiomyocytes. Cell research 21, 1316-1331.

Chaumeil, J., Le Baccon, P., Wutz, A., and Heard, E. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev 20, 2223-2237.

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

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

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

Elisaphenko, E. A., Kolesnikov, N. N., Shevchenko, A. I., Rogozin, I. B., Nesterova, T. B., Brockdorff, N., and Zakian, S. M. (2008). A dual origin of the Xist gene from a protein-coding gene and a set of transposable elements. PLoS One 3, e2521.

Gendrel, A., and Heard, E. (2014). Noncoding RNAs and Epigenetic Mechanisms During X-Chromosome Inactivation. Annual Review of Cell and Developmental Biology 30, 561-580.

Gontan, C., Achame, E. M., Demmers, J., Barakat, T. S., Rentmeester, E., van IJcken, W., Grootegoed, J. A., and Gribnau, J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature 485, 386-390.

Grant, J., Mahadevaiah, S. K., Khil, P., Sangrithi, M. N., Royo, H., Duckworth, J., McCarrey, J. R., VandeBerg, J. L., Renfree, M. B., Taylor, W., et al. (2012). Rsx is a metatherian RNA with Xist-like properties in X-chromosome inactivation. Nature 487, 254-258.

Guan, Y., Shao, Y., Li, D., and Liu, M. (2014). Generation of site-specific mutations in the rat genome via CRISPR/Cas9. Methods in enzymology 546, 297-317.

Hirabayashi, M., Kato, M., Kobayashi, T., Sanbo, M., Yagi, T., Hochi, S., and Nakauchi, H. (2010). Establishment of rat embryonic stem cell lines that can participate in germline chimerae at high efficiency. Molecular reproduction and development 77, 94.

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

Ishizaki, T., Uehata, M., Tamechika, I., Keel, J., Nonomura, K., Maekawa, M., and Narumiya, S. (2000). Pharmacological properties of Y-27632, a specific inhibitor of rho-associated kinases. Molecular pharmacology 57, 976-983.

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

Kawamata, M., and Ochiya, T. (2010a). Establishment of embryonic stem cells from rat blastocysts. Methods in molecular biology (Clifton, N.J.) 597, 169-177.

Kawamata, M., and Ochiya, T. (2010b). Generation of genetically modified rats from embryonic stem cells. Proceedings of the National Academy of Sciences 107, 14223-14228.

Kim, J. D., Faulk, C., and Kim, J. (2007). Retroposition and evolution of the DNA-binding motifs of YY1, YY2 and REX1. Nucleic Acids Research 35, 3442-3452.

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

Li, P., Tong, C., Mehrian-Shai, R., Jia, L., Wu, N., Yan, Y., Maxson, R. E., Schulze, E. N., Song, H., Hsieh, C., et al. (2008). Germline competent embryonic stem cells derived from rat blastocysts. CELL 135, 1299-1310.

Li, X., Cui, X., Wang, J., Wang, Y., Li, Y., Wang, L., Wan, H., Li, T., Feng, G., Shuai, L., et al. (2016). Generation and Application of Mouse-Rat Allodiploid Embryonic Stem Cells. 1-32.

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

Mak, W., Nesterova, T. B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A. P., and Brockdorff, N. (2004). Reactivation of the paternal X chromosome in early mouse embryos. Science 303, 666-669.

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

Meek, S., Buehr, M., Sutherland, L., Thomson, A., Mullins, J. J., Smith, A. J., and Burdon, T. (2010). Efficient gene targeting by homologous recombination in rat embryonic stem cells. PLoS One 5, e14225.

Meek, S., Sutherland, L., and Burdon, T. (2014). Methods in Molecular Biology K. Turksen, ed. (New York, NY: Springer New York).

Meek, S., Wei, J., Sutherland, L., Nilges, B., Buehr, M., Tomlinson, S. R., Thomson, A. J., and Burdon, T. (2013). Tuning of ?-catenin activity is required to stabilize self-renewal of rat embryonic stem cells. Stem cells (Dayton, Ohio) 31, 2104-2115.

Mekhoubad, S., Bock, C., de Boer, A. S., Kiskinis, E., Meissner, A., and Eggan, K. (2012). Erosion of dosage compensation impacts human iPSC disease modeling. Cell Stem Cell 10, 595-609.

Men, H., Bauer, B. A., and Bryda, E. C. (2012). Germline transmission of a novel rat embryonic stem cell line derived from transgenic rats. Stem cells and development 21, 2606-2612.

Migeon, B. R., Chowdhury, A. K., Dunston, J. A., and McIntosh, I. (2001). Identification of TSIX, encoding an RNA antisense to human XIST, reveals differences from its murine counterpart: implications for X inactivation. Am J Hum Genet 69, 951-960.

Migeon, B. R., Lee, C. H., Chowdhury, A. K., and Carpenter, H. (2002). Species differences in TSIX/Tsix reveal the roles of these genes in X-chromosome inactivation. Am J Hum Genet 71, 286-293.

Moreira de Mello, J. C., de Araújo, E. S. S., Stabellini, R., Fraga, A. M., de Souza, J. E. S., Sumita, D. R., Camargo, A. A., and Pereira, L. V. (2010). Random X inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome. PLoS One 5, e10947.

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

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

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

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

Ohhata, T., Hoki, Y., Sasaki, H., and Sado, T. (2007). Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. Development 135, 227-235.

Okamoto, I., Otte, A. P., Allis, C. D., Reinberg, D., and Heard, E. (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. Science 303, 644-649.

Okamoto, I., Patrat, C., Thepot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J. P., Renard, J. P., Duranthon, V., et al. (2011). Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature 472, 370-374.

Pasque, V., and Plath, K. (2015). X chromosome reactivation in reprogramming and in development. Current Opinion in Cell Biology 37, 75-83.

Patrat, C., Okamoto, I., Diabangouaya, P., Vialon, V., Le Baccon, P., Chow, J., and Heard, E. (2009). Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. Proceedings of the National Academy of Sciences 106, 5198-5203.

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

Peng, X., Gao, H., Wang, Y., Yang, B., Liu, T., Sun, Y., Jin, H., Jiang, L., Li, L., Wu, M., et al. (2013). Conversion of rat embryonic stem cells into neural precursors in chemical-defined medium. Biochemical and biophysical research communications 431, 783-787.

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature 379, 131-137.

Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., La Cruz, de, C. C., Otte, A. P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. Science 300, 131-135.

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

Schroeder, I. S., Wiese, C., Truong, T. T., Rolletschek, A., and Wobus, A. M. (2009). Differentiation analysis of pluripotent mouse embryonic stem (ES) cells in vitro. Methods in molecular biology (Clifton, N.J.) 530, 219-250.

Schulz, E. G., Meisig, J., Nakamura, T., Okamoto, I., Sieber, A., Picard, C., Borensztein, M., Saitou, M., Blüthgen, N., and Heard, E. (2014). The Two Active X Chromosomes in Female ESCs Block Exit from the Pluripotent State by Modulating the ESC Signaling Network. Stem Cell 14, 203-216.

Shao, Y., Guan, Y., Wang, L., Qiu, Z., Liu, M., Chen, Y., Wu, L., Li, Y., Ma, X., Liu, M., et al. (2014). CRISPR/Cas-mediated genome editing in the rat via direct injection of one-cell embryos. Nature protocols 9, 2493-2512.

Shevchenko, A. I., Malakhova, A. A., Elisaphenko, E. A., Mazurok, N. A., Nesterova, T. B., Brockdorff, N., and Zakian, S. M. (2011). Variability of sequence surrounding the Xist gene in rodents suggests taxon-specific regulation of X chromosome inactivation. PLoS One 6, e22771.

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T. B., Webster, Z., Peters, A. H., Jenuwein, T., Otte, A. P., and Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. Dev Cell 4, 481-495.

Takahashi, K., Tanabe, K., Ohnuki, M., Nari-

ta, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. CELL 131, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. CELL 126, 663-676.

Tchieu, J., Kuoy, E., Chin, M. H., Trinh, H., Patterson, M., Sherman, S. P., Aimiuwu, O., Lindgren, A., Hakimian, S., Zack, J. A., et al. (2010). Female human iPSCs retain an inactive X chromosome. Cell Stem Cell 7, 329-342.

van Bemmel, J. G., Mira-Bontenbal, H., and Gribnau, J. (2016). Cis- and trans-regulation in X inactivation. Chromosoma 125, 41-50.

Vaskova, E. A., Medvedev, S. P., Sorokina, A. E., Nemudryy, A. A., Elisaphenko, E. A., Zakharova, I. S., Shevchenko, A. I., Kizilova, E. A., Zhelezova, A. I., Evshin, I. S., et al. (2015). Transcriptome Characteristics and X-Chromosome Inactivation Status in Cultured Rat Pluripotent Stem Cells. Stem cells and development 24, 2912-2924.

Wake, N., Takagi, N., and Sasaki, M. (1976). Non-random inactivation of X chromosome in the rat yolk sac. Nature 262, 580-581.

Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J. B., Nishikawa, S., Nishikawa, S., Muguruma, K., et al. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nature biotechnology 25, 681-686.

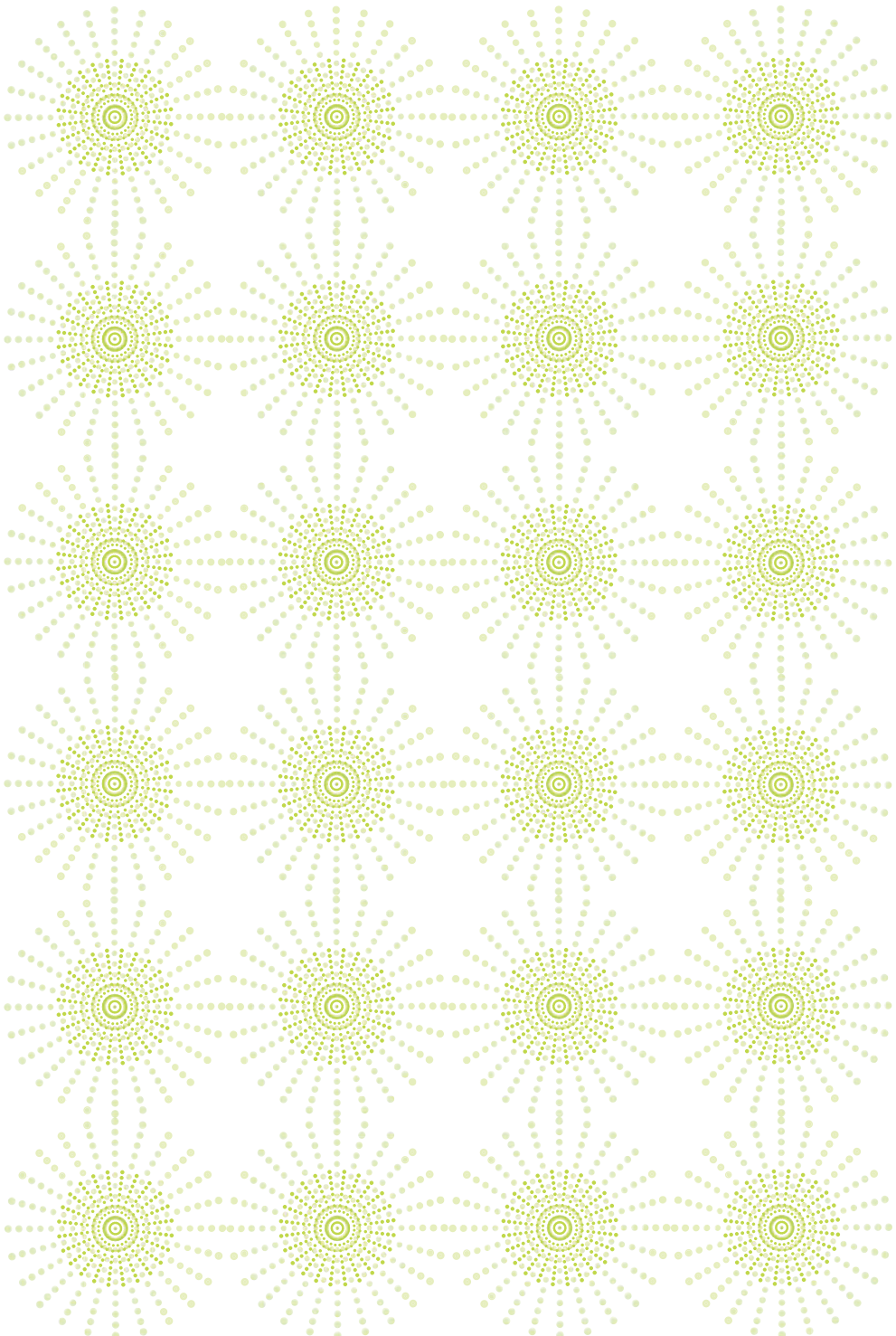
Wutz, A., and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol Cell 5, 695-705.

Xue, F., Tian, X. C., Du, F., Kubota, C., Taneja, M., Dinnyes, A., Dai, Y., Levine, H., Pereira, L. V., and Yang, X. (2002). Aberrant patterns of X chromosome inactivation in bovine clones. Nature Genetics 31, 216-220.

OLIGO	SEQUENCE	DESCRIPTION
454	TGCCTGGATTTAGAGGAG	Forward primer *Xist* expression
455	CTCCACCTAGGGATCGTCAA	Reverse primer *Xist* expression
456	GTATCCACAGCCCCGATG	Forward primer *Tsix* expression
457	ACCTCGGATACCTGCGTTT	Reverse primer *Tsix* expression
372	TAGCCCTGATTCTTCTAGCA	Forward primer *Nanog* expression
373	TTTGCTGCAACGGCACATAA	Reverse primer *Nanog* expression
400	GGCGTTCTTCAAGAGAACCA	Forward primer *Esrrb* expression
401	CCCACTTTGAGGCATTTCAT	Reverse primer *Esrrb* expression
396	AGGAACTGCGCTTCGTTCT	Forward primer *Prdm14* expression
397	GGCATCACCAAAAGCTGTCT	Reverse primer *Prdm14* expression
376	AAATCATGACGAGGCAAGGC	Forward primer *Rex1* expression
377	TGAGTTCGCTCCAACAGTCT	Reverse primer *Rex1* expression
394	CTCTGCTGGAGGCTGAGAAC	Forward primer *Nestin* expression
395	TGGTATCCCAAGGAAATTCG	Reverse primer *Nestin* expression
388	GCTGGCCTTAGAGACCACAG	Forward primer *Actin* expression
389	AAGCAATTCAGCAACACCAA	Reverse primer *Actin* expression

Supplementary table 1.

Primers used in this study as described in the experimental procedures section.



CHAPTER .6

GENERAL DISCUSSION

Agnese Loda

GENERAL DISCUSSION

The process of XCI results in a fascinating functional asymmetry between two genetically identical DNA molecules within the same nucleus. Such a striking epigenetic difference between the inactive X chromosome (Xi) and its active homologue (Xa) is necessary to ensure dosage compensation of sex chromosomal genes between females (XX) and males (XY), and allows female embryos to survive throughout development.

The idea of XCI was first put forward by Mary Lyon in 1961 (Lyon, 1961), and was based on mouse genetic studies: (I) Lyon noticed that the few X-linked mutations known at that time would result in variegated effects exclusively in heterozygous females, whereas homozygous females and hemizygous males showed the same phenotype (Falconer, 1953, Fraser, 1953, Lyon, 1960). In the same years, (II) XO female mice had been reported to survive embryonic development (Welshons and Russell, 1959), and (III) the Barr body (see **chapter 1**), had been suggested to correspond to a heterochromatic X chromosome (Ohno and Hauschka, 1960). Lyon published her hypothesis more than fifty years ago, at a time when less than 100 genes of the entire mouse genome had been mapped by following their patterns of segregation through several generations: (I) "the heteropyknotic X chromosome can be either paternal or maternal in origin", (II) "the mosaic phenotype is due to the inactivation of one or other X chromosome in embryonic development", (III) "all female mammals heterozygous for sex-linked mutant genes would be expected to show the same phenomena as those in the mouse. The coat of the tortoiseshell cat, being a mosaic of the black and orange colours of the two homozygous types, fulfils this expectation" (Lyon, 1961). Today, in spite of the fact that XCI has developed into a powerful tool to address a multitude of basic epigenetic mechanisms, our understanding of the process remains far from complete, and Lyon's observations are still the basis of our research. In this thesis, by exploiting a multitude of transgenic strategies in ES cells, we provided new insights into several aspects of the XCI process, including: (I) XCI initiation (**chapter 2**), (II) Xist RNA in cis spreading (**chapter 3**), (III) Xist-mediated chromosome-wide gene silencing (**chapter 3**), (IV) escape from XCI (**chapter 3**), and (V) the developmental features that tightly link *Xist* regulation to its functions (**chapter 4**). Finally (VI), we developed a novel *in vitro* differentiation strategy that enables us to study the dynamics of XCI in a model organism different from the mouse (**chapter 5**). The following paragraphs aim to integrate our findings into the currently existing picture of XCI regulation (reviewed in **chapter 1**), and also include considerations about how to unravel several of the remaining open questions.

The two antisense transcribed noncoding genes *Xist* and *Tsix* represent the major players at the onset of XCI and their tight reciprocal regulation is critical to ensure *Xist* monoallelic expression. However, the structural overlap of *Xist* and *Tsix* at the *Xic* makes dissecting the func-

tion of one gene from the other technically challenging. To be able to uncouple *Xist* and *Tsix* regulation in mouse ES cells, in **chapter 2**, we developed two independent knock-in alleles by replacing the first exons of *Xist* and *Tsix* with fluorescent reporter genes. Thus, in Xist-GFP and Tsix-CHERRY ES clones, the regulation of EGFP and mCherry reporters is under the control of either the *Xist* or *Tsix* endogenous promoter, respectively. In addition, a third knock-in line was generated in which both the Xist-GFP and Tsix-CHERRY alleles were targeted to the same X chromosome. The mutual antagonistic role of *Xist* and *Tsix* at the onset of XCI was faithfully recapitulated either by following the expression of the reporters or by live cell imaging. As expected, we observed increased *Xist* driven GFP expression in the double knock-in ES clone compared to the single Xist-GFP clone, confirming the repressive role of *Tsix* on *Xist* expression (reviewed in **chapter 1**). Interestingly, we found that *Xist* also acts locally to downregulate *Tsix*, not only on the Xi but on the future Xa as well, suggesting that the interplay between *Xist* and *Tsix* is a key event on both alleles at the onset of XCI. Moreover, overexpression of known activators and inhibitors of XCI seems to properly recapitulate their function exclusively in the double knock-in environment in which the antisense transcription between *Xist* and *Tsix* is abrogated. Therefore, *cis*-mediated mutual regulation of *Xist* and *Tsix* within the *Xic* seems to overrule their *trans*-acting regulation. However, since the double knock-in ES line consistently responds to both *Rnf12* and *Rex1* overexpression, we believe that this transgenic system can be exploited as a powerful strategy to discover novel *trans*-acting activators and inhibitors of XCI. GFP and mCherry expression levels could be easily assessed by FACS analysis after genome wide mutagenesis to screen for factors that directly impact *Xist* and *Tsix* expression. Finally, two semi-stable transcriptional states characterized by either high or low *Tsix* expression levels were found to be present in XO Xist-GFP/Tsix-CHERRY cells. These states were also present to a lesser extend in the double knock-in Xist-GFP/Tsix-CHERRY XX cells in 2i culture conditions. These two subpopulations are thus responsive to the X to autosome ratio and to external signalling, and can be further stabilized by overexpressing regulators of XCI. However, although switching from the serum-based primed pluripotent state (Wray et al., 2010) to the more homogeneous 2i-dependent ground state of pluripotency (Ying et al., 2008) did modulate the stability of these states, no difference in terms of expression of pluripotency factors was observed when comparing populations with high and low mCherry to each other. In addition, DNA methylation of the *Tsix* promoter also appears to be indifferent between the two transcriptional states, thus excluding this epigenetic modification to play a role in the maintenance of the semi-stable states. Rather, we suggest that the two transcriptional states might correspond to different higher order chromatin conformations of the *Xic*. In line with this hypothesis, physical modelling of the structural variations within the *Tsix* TAD has recently revealed the existence of two main classes of conformations that differ from each other in terms of chromatin compaction and long-range interaction frequency (Giorgetti et al., 2014). These two clusters of alternative configurations have been related to the *Tsix* TAD transcrip-

tional activity, and the dynamic fluctuations between different conformations has been proposed to ensure asymmetric transcription of *Tsix* between the two *Xics*, which would result in monoallelic up-regulation of *Xist* and initiation of XCI (Giorgetti et al., 2014). Therefore, the two *Tsix* semi-stable transcriptional states that we observed in our transgenic system might indeed correspond to the different clusters of conformations in XX cells, with the low mCherry population assuming a structure more favourable for *Xist* expression and the high mCherry population in which *Tsix* is more likely to be expressed. However, further functional validation by either high-resolution 3D DNA-FISH or chromatin capture experiments is needed to confirm this hypothesis.

As reviewed in **chapter 1**, X to autosome rearrangements and *Xist* transgenic strategies have been extensively exploited to understand the XCI process, from defining the *Xic* to understanding the developmental context in which XCI can take place. In **chapter 3**, we developed a novel transgenic *Xist* expression system in female mouse ESCs. Our strategy brings together several advantages that allow us to dissect the mechanism(s) directing Xist-mediated silencing. In fact, we were not only able to (I) induce ectopic *Xist* expression from different genomic loci, thus uncoupling *Xist* RNA function from its X chromosome specific context, but we could also (II) trigger Xist RNA spreading from many independent loci along the entire length of the same chromosome, either the X chromosome or autosomes. In addition, (III) by exploiting ectopic XCI to correct for unbalanced autosomal aneuploidies, we were able to exclude any confounding bias arising from functional monosomy of autosomal genes.
Although we could efficiently induce *Xist* expression from several independent regions of the mouse genome, we found that *Xist* silencing efficiency is locus dependent and strongly relies on the genomic environment from which Xist RNA is forced to spread. Thus, if Xist RNA starts spreading from gene-rich areas, transcriptional silencing of autosomal genes can be achieved with the same efficiency as for X-linked genes. In addition, we found that enrichment of LINE elements and specific chromatin features also contributes to the generation of a favourable environment for Xist-mediated silencing. Particularly, we proposed that H3K27me3 and H2AK119ub histone marks might work as docking stations for Xist RNA localization to the Xi. However, although Xist RNA might indeed exploit these chromatin marks in spatial proximity of its transcription site to recruit PRC2 and PRC1 repressive complexes to the Xi, such polycomb recruitment is most likely to occur only after gene silencing is already established, most likely in parallel with SPEN-mediated HDAC3 recruitment (McHugh et al., 2015). Indeed, as reviewed in **chapter 1**, *Xist* expression is the only event known to be critical for XCI to occur, whereas PRC2 and PRC1 repressive complexes are both dispensable for the initial phase of transcriptional inactivation. Moreover, Xist RNA deleted for the repeat A silencing domain still mediates the formation of a repressive heterochromatic compartment from which RNA Pol II and several transcription factors are excluded, suggesting that the A repeat-Spen-Smrt-Hda3

axis is not the sole mechanism directing gene silencing (**chapter 1**). In this context, a fascinating example of RNA-mediated gene silencing has been described for the X-linked FMR1 gene, whose dynamic mutations are associated with the Fragile X Syndrome (FXS) (Colak et al., 2014). FXS syndrome is caused by a trinucleotide (CGG) repeat expansion in the promoter region of the *FMR1* gene that results in epigenetic loss of *FMR1* expression. The expanded CGG repeat has been shown to form a hairpin structure at the 5' UTR of the FMR1 transcript that binds to its own DNA while it is being transcribed, thus forming an RNA-DNA duplex that triggers gene silencing. Accordingly, treatment with small molecules that alter the hairpin structure of the FMR1 transcript results in impaired *FMR1* transcriptional repression (Colak et al., 2014). Several secondary structures into which Xist RNA is likely to be organized have been extensively modelled (Duszczyk et al., 2008; 2011; Maenner et al., 2010; Caparros et al., 2002; Fang et al., 2015), and the existence of different evolutionary conserved RNA structural domains within the Xist RNA has been recently proven (Lu et al., 2016). Therefore, it is temping to speculate that, similarly to FMR1, Xist RNA itself could directly trigger gene silencing by physically interfering with ongoing transcription. However, this hypothesis implies the recognition of specific DNA sequence by Xist RNA, and such specific interaction is unlikely to occur for each gene along the X chromosome. Nonetheless, the repetitive nature of Xist RNA, together with several putative secondary structures into which Xist RNA can fold, might allow the formation of DNA-RNA duplexes between Xist RNA and repetitive elements which are located either in introns or *in cis* acting regulatory elements of X-linked genes. Such multiple interactions between Xist RNA and X-linked transcribed genes might in turn limit accessibility of X-linked gene promoters, thus leading to exclusion of the transcriptional machinery from the Xist RNA-dependent Xi domain.

Interestingly, Xist RNA-FISH analysis after induction of ectopic XCI in undifferentiated ES cells resulted in a clear morphological difference between the X- and the autosomal- Xist RNA coated chromosomes. In fact, independently of the degree of silencing efficiency that is achieved upon *Xist* induction, autosomal clones often showed dispersed Xist RNA clouds, a phenotype that partially resembles the one observed upon knock-down of the matrix protein hnRNP U, which was shown to ensure Xist RNA localization to the Xi (Hasegawa et al., 2010). These observations can be differently interpreted: (I) Xist RNA might not properly localize to the autosome from which it is transcribed because autosomal DNA lacks X chromosome specific features, or (II) less efficient Xist-mediated silencing of autosomal genes compared to X-linked genes might impair the formation of a "Xi-like" silent territory within the nucleus. However, robust chromosome-wide inactivation of autosomal genes could be achieved in several clones in which Xist RNA domains appeared to be less compacted than the endogenous Xi domain. Moreover, although Xist-mediated silencing becomes robust and irreversible upon ES cells differentiation (Wutz and Jaenisch, 2000, **chapter 3**), the difference in terms of Xist RNA clouds compaction between the Xi and inactive autosomes appeared to be main-

tained in fully differentiated neurons. Therefore, different levels of Xist silencing efficiency are unlikely to explain the observed morphological discrepancies. In line with these observations, Xist silencing and localization functions have been shown to be mediated by different domains of Xist RNA (Wutz et al., 2002), and interfering with Xist RNA localization to the Xi partially impairs but does not completely abrogate Xist-mediated silencing (Hasegawa et al., 2010; Beletskii et al., 2001; Chu et al., 2015; McHugh et al., 2015).

Another possibility is that autosomal chromatin might tend to retain its TAD-based 3D spatial organization upon ectopic inactivation, whereas the Xi might be intrinsically more prone to acquire a topologically unorganized chromatin structure (see below) (Rao et al., 2014; Deng et al., 2015; Giorgetti et al., 2016). Such a difference in structural organization might in turn lead to a more or less dispersed Xist RNA domain within the nucleus as observed by Xist RNA FISH. However, since our morphological observations are qualitative, we cannot exclude that there is a correlation between Xist RNA cloud compaction and silencing efficiency in between different autosomal clones, and further experiments involving quantitative imaging analysis will be needed to clarify this point.

In general, understanding the link between higher order chromatin organization of the Xi and its transcriptional activity represents one of the major open questions of the XCI field. Minajigi and colleagues have recently reported that deleting Xist from the Xi in somatic cells results in a restored the Xa-like conformation of the Xi (Minajigi et al., 2015). However, although the transcriptional activity of the Xist-deleted Xi was not assessed, it is not likely to result in Xi reactivation, as previously observed (Csankovszki et al., 1999) . Indeed, a similar study in which the 3D conformation of the Xi was assessed by allele-specific 4C analysis has shown that deleting Xist in neuronal precursor cells (NPCs) results in partial refolding of the Xi into a structure reminiscent of the Xa, but such a structural shift is never associated with X-linked gene reactivation (Splinter et al., 2011) . However, both the establishment and the maintenance of Xi chromosome-wide silencing rely on several redundant layers of repressive epigenetic features that accumulate on the Xi upon differentiation. Therefore, the lack of Xi reactivation following the structural changes induced by *Xist* deletion in fully differentiated cells does not exclude that Xist-mediated reorganization of the Xi conformation at an early stage of XCI is a key feature of Xist RNA function (see below).

By using an unbiased approach based on induction of Xist RNA from several loci along the X, in **chapter 3**, we were able to show that ectopic inactivation of X-linked genes always recapitulates endogenous XCI. Thus, the large majority of X-linked genes that were not affected by Xist-mediated silencing correspond to previously described escapees. This result confirms that escaping genes are intrinsically able to resist XCI, regardless of their relative position to the Xist transcription site (reviewed in **chapter 1**). Escape from XCI plays a major role in proper female mammalian development and susceptibility to X-linked diseases. Indeed, different

from XO female mice that do not develop a phenotype and are fertile, women with a single X are affected by Turner syndrome (45,XO), and the majority of the XO embryos die in utero (Hook and Warburton, 1983; Bondy, 2009). The different phenotype associated with a single copy of the X chromosome between mice and women might rely on the different percentage of X-linked genes that escape XCI, 12-20% in human and 3-7% in mice. In spite of the fact that lack of biallelic expression of several X-linked genes leads to harmful effects on female development (reviewed in **chapter 1**), very little is known about the molecular mechanism(s) directing XCI escape. How can escaping genes remain actively transcribed from an otherwise completely inactive chromosome? Which *cis*- or *trans*- acting features render escaping genes intrinsically able to resist XCI? How would *cis* acting elements mediate escape? By blocking spreading of heterochromatin into active euchromatin? Or by preventing inactivated genes from being reactivated? Do escaping genes exploit a common strategy to resist XCI? Or is escape regulation directed at a gene-specific level?

In **chapter 3**, we found exclusive CTCF enrichment at the TSS of X-linked escaping genes prior to XCI, whereas the autosomal genes that resulted to be unaffected by Xist-mediated silencing do not show such an enrichment. We propose that, during XCI, the maintenance of CTCF binding at X-linked escaping loci protects them from becoming inactivated, whereas lack of CTCF enrichment at the TSS of X-linked genes facilitates gene silencing. Recently, genome-wide chromosome capture (Hi-C) analysis of mouse NPCs has shown that escaping genes correspond to the only loci that maintain a TAD-like structure within the otherwise TAD-depleted Xi. This observation, together with the role of CTCF in directing TAD formation (Ong and Corces, 2014; Zuin et al., 2014; Sofueva et al., 2013; Rudan et al., 2015), and the fact that CTCF binding sites in proximity of escapees retain DNA accessibility in NPCs (Giorgetti er al., 2016), strongly support our finding that CTCF binding is directly involved in XCI escape. Our findings also led us to hypothesize that, upon evolution, loss of CTCF binding following Xist RNA spreading might have conferred to the X chromosome the advantage of becoming easily reorganized into its silenced compartment. In line with this idea, contrarily to the X chromosome, autosomal chromatin would be refractory to the loss of CTCF binding upon ectopic Xist induction. Thus, the fact that X-linked genes are generally more efficiently inactivated compared to autosomal genes (reviewed in **chapter 1**, **chapter 3**) might be at least partially explained by the tendency of autosomal chromatin to resist Xist-mediated structural reorganization. However, if this would be the case, we would expect strongly silenced autosomal genes to be preferentially depleted of CTCF at their TSS. Instead, we did not find differential enrichment of CTCF binding at the TSS of autosomal genes that are differentially affected by Xist silencing. Moreover, this hypothesis cannot explain our finding that exactly the same autosome can become either efficiently or poorly inactivated upon Xist RNA induction from different integration sites along the autosome. Therefore, although loss of CTCF-binding upon XCI might facilitate the establishment of Xist-mediated silencing, the Xist-mediated

reorganization of the Xi structure is unlikely to be the sole critical event ensuring proper XCI. Interestingly, two recent studies have shown that deleting the DXZ4 macrosatellite boundary region that mediates the partitioning of the Xi into two large mega-domains in both human and mouse does not affect initiation XCI nor silencing establishment (Giorgetti et al., 2016; Darrow et al., 2016). However, fusion of the two Xi mega-domains upon deletion of the DXZ4 boundary in mouse ES cells results in at least partial mis-regulation of facultative escaping genes upon XCI (Giorgetti et al., 2016). Thus, although only escaping genes that resist XCI in specific tissues or developmental contexts are affected by DXZ4 deletion, whereas constitutive escaping genes remain properly expressed, this study suggests that the structural organization of the Xi is not completely dispensable to ensure accurate XCI regulation (Giorgetti et al., 2016).

The developmental regulation of *Xist* function is one of the most fascinating aspects of the XCI process. The ability of inducible *Xist* transgenes to induce gene silencing and faithfully recapitulate XCI is restricted to a short time window at the onset of ES cells differentiation (Wutz and Jaenisch, 2000). In the same special regulatory phase during early development, the interplay between *Xist* and *Tsix* directs the initiation of XCI (reviewed in **chapter 1**). Importantly, developmentally regulated *Tsix* repression of *Xist* relies on antisense transcription through the *Xist* promoter (Sado et al., 2005; Ohhata et al., 2007), and understanding the epigenetic features of such antisense regulation might provide us with new insights into the events that direct Xist RNA function at the onset of XCI. Therefore, in **chapter 4**, we developed a controllable expression system that allowed us to mimic the effects of antisense transcription on gene regulation in the context of development, independently of a specific genomic locus. We generated an artificial sense-antisense gene pair by overlapping an inducible antisense transcription unit to a GFP reporter gene. Induction of antisense transcription in either undifferentiated ES cells or upon ES cell differentiation led to completely different outcomes in terms of stability of GFP repression: (I) in ES cells, gene repression is reversible and associated with enrichment of H3K36me3 at the GFP promoter, contrarily, (II) upon ES cells differentiation, antisense-mediated silencing of the reporter gene is stabilized, and CpG methylation of the GFP promoter seems to be responsible for the irreversible nature of gene silencing. These observations suggest the existence of differentiation-specific chromatin modifying complexes that mediate epigenetic silencing, but further studies are clearly needed to discover their identity. In this context, dissecting the developmentally regulated events that make Xist RNA able or unable to induce gene silencing as well as the epigenetic mechanisms that determine the difference between reversible and irreversible Xist-mediated silencing is critical for a complete understanding of XCI.

Finally, in **chapter 5**, we established a novel monolayer differentiation protocol for rat pluriP-

otent stem cells (rESCs). We demonstrated that, at the onset of XCI, monoallelic *Xist* up-regulation can be achieved only when female rESCs are triggered to differentiate in complete absence of 2i inhibitors. This finding explains why we were not able to track XCI upon differentiation of rESCs according to the previously established protocols (Peng et al., 2013; Cao et al., 2011), and underlines the strict link between Xist RNA function and pluripotency network regulation (Schulz et al., Navarro et al., 2008, Navarro et al., 2010). Furthermore, we also showed that rat XCI occurs with similar dynamics as the one observed upon differentiation of mouse pluripotent cells, and that the critical role of the XCI regulators RNF12 and REX1 is conserved in rat (**chapter 1**). In general, developing novel *in vitro* strategies to explore the dynamics of XCI in different species is of great importance for a comprehensive understanding of the XCI process. In fact, although several *in vivo* observations in early developing embryos suggested XCI to be quite variable amongst different mammals, (Moreira de Mello et al., 2010; Okamoto et al., 2011; Xue et al., 2002) the majority of our knowledge concerning XCI regulation is based on mouse ES cell-based studies. In this context, the induced pluripotent stem cell (iPSCs) technology allows us to derive pluripotent stem cells from potentially any species of interest. However, although exploiting *in vitro* strategies would prevent the use of embryos in XCI research, thus extending the possibility to explore XCI in many additional mammalian species, the lack of *in vivo* studies is a critical limitation for the establishment of robust differentiation protocols that faithfully recapitulate XCI. For example, characterizing the XCI status of human iPSCs cells has emerged to be challenging (**chapter 1**), and many efforts have been made to establish specific culture conditions that would maintain both X chromosomes active in human pluripotent cells (**chapter 1**). However, since the mouse system is always used as a reference, there is a risk to develop protocols that mimic mouse XCI rather than human XCI upon early embryonic development. Indeed, single-cell RNA-sequencing analysis of human preimplantation embryos has recently confirmed that, in contrast to the mouse, human *XIST* is expressed from both X chromosomes upon early development, and X-linked genes remain bi-allelically expressed while dosage compensation occurs (Petropoulos et al., 2016, Okamoto et al., 2011). Thus, only by integrating both *in vivo* and observations we will be able to faithfully address species-specific mechanisms involved in XCI regulation.

In conclusion, after more than fifty years of XCI research, understanding how two genetically identical X chromosomes become different epigenetic entities remains fascinating and challenging. In the future, new insights in XCI regulation will shed new light on several basic epigentic aspects, transcriptional regulation, genomic spatial organization and facultative heterochromatin formation, with an impact that will surely reach far beyond the field of dosage compensation. I am curious to experience our next discoveries.

REFERENCES

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

Bondy, C. A. (2009). Turner syndrome 2008. Hormone research 71 Suppl 1, 52-56.

Cao, N., Liao, J., Liu, Z., Zhu, W., Wang, J., Liu, L., Yu, L., Xu, P., Cui, C., Xiao, L., et al. (2011). In vitro differentiation of rat embryonic stem cells into functional cardiomyocytes. Cell research 21, 1316-1331.

Caparros, M., Alexiou, M., Webster, Z., and Brockdorff, N. (2002). Functional analysis of the highly conserved exon IV of XIST RNA. Cytogenet Genome Res 99, 99-105.

Chu, C., Zhang, Q. C., Da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., Magnuson, T., Heard, E., and Chang, H. Y. (2015). Systematic discovery of Xist RNA binding proteins. CELL 161, 404-416.

Colak, D., Zaninovic, N., Cohen, M. S., Rosenwaks, Z., Yang, W., Gerhardt, J., Disney, M. D., and Jaffrey, S. R. (2014). Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. Science 343, 1002-1005.

Csankovszki, G., Panning, B., Bates, B., Pehrson, J. R., and Jaenisch, R. (1999). Conditional deletion of Xist disrupts histone macro-H2A localization but not maintenance of X inactivation. Nature Genetics 22, 323-324.

Darrow, E. M., Huntley, M. H., Dudchenko, O., Stamenova, E. K., Durand, N. C., Sun, Z., Huang, S., Sanborn, A. L., Machol, I., Shamim, M., et al. (2016). Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture. Proceedings of the National Academy of Sciences.

Deng, X., Ma, W., Ramani, V., Hill, A., Yang, F., Ay, F., Berletch, J. B., Blau, C. A., Shendure, J., Duan, Z., et al. (2015). Bipartite structure of the inactive mouse X chromosome. Genome Biology 16, 67.

Duszczyk, M. M., Zanier, K., and Sattler, M. (2008). A NMR strategy to unambiguously distinguish nucleic acid hairpin and duplex conformations applied to a Xist RNA A-repeat. Nucleic Acids Research 36, 7068-7077.

Duszczyk, M. M., Wutz, A., Rybin, V., and Sattler, M. (2011). The Xist RNA A-repeat comprises a novel AUCG tetraloop fold and a platform for multimerization. RNA (New York, N.Y.) 17, 1973-1982.

Falconer, D. S. (1953). [Total sex-linkage in the house mouse]. Zeitschrift für induktive Abstammungs-und Vererbungslehre 85, 210-219.

Fang, R., Moss, W. N., Rutenberg-Schoenberg, M., and Simon, M. D. (2015). Probing Xist RNA Structure in Cells Using Targeted Structure-Seq. PLoS Genet 11, e1005668.

Fraser AS, Sobey S, Spicer CC. (1953). Mottled, a sex-modified lethal in the house mouse. J. Genet. 51:217–21

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

Giorgetti, L., Lajoie, B. R., Carter, A. C., Attia, M., Zhan, Y., Xu, J., Chen, C., Kaplan, N., Chang, H. Y., Heard, E., et al. (2016). Structural organization of the inactive X chromosome in the mouse. Nature 535, 575-579.

Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., Tsutui, K., and Nakagawa, S. (2010). The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. Dev Cell 19, 469-476.

Hook, E. B., and Warburton, D. (1983). The distribution of chromosomal genotypes associated with Turner's syndrome: livebirth prevalence rates and evidence for diminished fetal mortality and severity in genotypes associated with structural X abnormalities or mosaicism. Hum Genet 64, 24-27.

Lu, Z., Zhang, Q. C., Lee, B., Flynn, R. A., Smith, M. A., Robinson, J. T., Davidovich, C., Gooding, A. R., Goodrich, K. J., Mattick, J. S., et al. (2016). RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. CELL 165, 1267-1279.

Lyon, M. F. (1960). A further mutation of the mottled type in the house mouse. J. Hered. 51:116–2

Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature 190, 372-373.

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

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

Minajigi, A., Froberg, J. E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., Lessing, D., Payer, B., Boukhali, M., Haas, W., et al. (2015). Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. Science 349.

Moreira de Mello, J. C., de Araújo, E. S. S., Stabellini, R., Fraga, A. M., de Souza, J. E. S., Sumita, D. R., Camargo, A. A., and Pereira, L. V. (2010). Random X inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome. PLoS One 5, e10947.

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

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

Ohhata, T., Hoki, Y., Sasaki, H., and Sado, T. (2007). Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. Development 135, 227-235.

Ohno, S., and HAUSCHKA, T. S. (1960). Allocycly of the X-chromosome in tumors and normal tissues. Cancer research 20, 541-545.

Okamoto, I., Patrat, C., Thepot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J. P., Renard, J. P., Duranthon, V., et al. (2011). Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature 472, 370-374.

Ong, C., and Corces, V. G. (2014). CTCF: an architectural protein bridging genome topology and function. Nature Publishing Group 15, 234-246.

Peng, X., Gao, H., Wang, Y., Yang, B., Liu, T., Sun, Y., Jin, H., Jiang, L., Li, L., Wu, M., et al. (2013). Conversion of rat embryonic stem cells into neural precursors in chemical-defined medium. Biochemical and biophysical research communications 431, 783-787.

Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S. P., Codeluppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R., and Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. CELL 165, 1012-1026.

Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. CELL 159, 1665-1680.

Rudan, M. V., Barrington, C., Henderson, S., Ernst, C., Odom, D. T., Tanay, A., and Hadjur, S. (2015). Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. CellReports 10, 1297-1309.

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

Schulz, E. G., Meisig, J., Nakamura, T., Okamoto, I., Sieber, A., Picard, C., Borensztein, M., Saitou, M., Blüthgen, N., and Heard, E. (2014). The Two Active X Chromosomes in Female ESCs Block Exit from the Pluripotent State by Modulating the ESC Signaling Network. Stem Cell 14, 203-216.

Sofueva, S., Yaffe, E., Chan, W., Georgopoulou, D., Rudan, M. V., Mira-Bontenbal, H., Pollard, S. M., Schroth, G. P., Tanay, A., and Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. The EMBO Journal 32, 3119-3129.

Splinter, E., de Wit, E., Nora, E. P., Klous, P., van de Werken, H. J. G., Zhu, Y., Kaaij, L. J. T., van IJcken, W., Gribnau, J., Heard, E., et al. (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes & Development 25, 1371-1383.

Welshons, W. J., and Russell, L. B. (1959). THE Y-CHROMOSOME AS THE BEARER OF MALE DETERMINING FACTORS IN THE MOUSE. Proc Natl Acad Sci U S A 45, 560-566.

Wray, J., Kalkan, T., and Smith, A. G. (2010). The ground state of pluripotency. Biochemical Society Transactions 38, 1027-1032.

Wutz, A., Rasmussen, T. P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nature Genetics 30, 167-174.

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

Xue, F., Tian, X. C., Du, F., Kubota, C., Taneja, M., Dinnyes, A., Dai, Y., Levine, H., Pereira, L. V., and Yang, X. (2002). Aberrant patterns of X chromosome inactivation in bovine clones. Nature Genetics 31, 216-220.

Ying, Q., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature 453, 519-523.

Zuin, J., Dixon, J. R., van der Reijden, M. I. J. A., Ye, Z., Kolovos, P., Brouwer, R. W. W., van de Corput, M. P. C., van de Werken, H. J. G., Knoch, T. A., van IJcken, W. F. J., et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proceedings of the National Academy of Sciences 111, 996-1001.

APPENDIX

SUMMARY

In mammals, sex chromosome composition differs between males and females with males carrying a Y chromosome and a single X chromosome and females carrying two X chromosomes. Early during embryonic development, X chromosome inactivation (XCI) takes place to completely shut down one X chromosome in female somatic cells, thus achieving dosage compensation of X-linked genes between the sexes. XCI is an extraordinary epigenetic paradigm that involves several mechanisms such as antisense transcription, non-coding RNA-mediated silencing, recruitment of chromatin remodelling complexes and 3D chromatin structure reorganization. The X-linked non-coding gene Xist is indispensable for XCI to occur. During XCI, Xist RNA is monoallelically up-regulated from one of the two X chromosomes and spread *in cis* along the entire length of the X, thus leading to chromosome-wide transcriptional silencing. *Xist* is negatively regulated by *Tsix*, a long non-coding gene that is transcribed in antisense direction to and fully overlaps with *Xist*. Once XCI has taken place, the silent state of the inactive X chromosome is clonally propagated through a near infinite number of cell divisions.

This thesis unravels several aspects of *Xist* regulation and Xist RNA function in a specific developmental context, by employing genetically modified mouse ES cells as a model system. **Chapter 1** provides a general overview of the state-of-the-art of science in the field of XCI and highlights the major open questions that remain to be addressed. **Chapter 2** describes a series of experiments in which *Xist* and its antisense partner *Tsix* are dissected from each other by replacing the two non-coding genes with GFP and mCHERRY fluorescent reporters, respectively. By exploting this system, we were able to study the genetic and dynamic regulation of *Xist* and *Tsix* independently from their reciprocal antisense transcription. We found mutually antagonistic roles for *Tsix* and *Xist* and vice versa, and we observed the presence of semi-stable transctiptional states of the X inactivation center (Xic) predicting the outcome of XCI. **Chapter 3** focuses on the mechanisms underlying Xist RNA silencing function and is based on a novel *Xist* expression system that we developed. By triggering *Xist* expression from different genomic contexts, we found that gene density in proximity of the *Xist* transcriprion locus is a key feature to determine the efficiency of -*cis* inactivation. We also showed that LINE elements facilitate Xist-mediated silencing of both X-linked and autosomal genes and that endogenous XCI is faithfully recapitulated upon induction of ectopic *Xist* expression from different loci along the X chromosome. In particular, escaping genes remain consistently active regardless of their position relative to the *Xist* transgenes, and the enrichment of CTCF at their promoters is implicated in directing their escape from XCI. The results presented in **chapter 4** rely on the generation of a EGFP reporter plasmid *cis*-linked to an inducible antisense promoter that mimics the structure of sense-antisense gene pairs such as *Xist* and *Tsix*.

By exploiting this transgenic system during ES cells differentiation, we found that induced antisense transcription leads to completely reversible silencing of the reporter gene in undifferentiated ES cells whereas stable EGFP silencing can only be achieved upon differentiation. Reversible silencing is mediated by chromatin modifications rather than by direct transcriptional interference, and stable silencing is characterized by CpG methylation of the EGPF promoter. Thus, we described a specific phase during early development in which antisense transcription generates a specific chromatin signature that triggers promoter-associated DNA methylation, which in turn is responsible for stable gene repression. **Chapter 5** describes a novel *in vitro* differentiation protocol for rat ES cell that allowed us to address the dynamics of XCI in a model organism different from the mouse. By mimicking the XCI process upon female rat ES cells differentiation, we found that Xist RNA starts accumulating in cis on the X chromosome during the second day of differentiation and is rapidly followed by H3K27me3 enrichment. We also showed that the critical roles of key XCI regulators discovered in mouse-based studies are well conserved in rat. Finally, in **chapter 6**, our findings are discussed in the context of the existing knowledge and future potential perspectives in the XCI field are presented.

SAMENVATTING

Vrouwelijke lichaamscellen bevatten twee X-chromosomen, het ene geërfd van moeder en het andere van vader. Mannen bezitten in iedere cel slechts één X, altijd van moeder, met daarnaast een klein Y-chromosoom. Op het X-chromosoom bevinden zich ruim 1.000 genen, tegenover slechts 50 genen op het Y-chromosoom. Om de hoeveelheid actieve X-genen bij vrouw en man min of meer gelijk te trekken (dosis-compensatie), wordt in vrouwelijke lichaamscellen altijd één van beide X-chromosomen uitgeschakeld. Dat gebeurt al vroeg tijdens de embryonale ontwikkeling, en vanaf dat moment zijn meisjes verder opgebouwd als een mozaïek van cellen waarin óf de X van moeder óf die van vader actief is. X chromosoom inactivatie (XCI) representeert een krachtig epigenetisch model systeem, omdat er veel verschillende epigenetische processen bij betrokken zijn, zoals functionele niet coderende RNAs, tegenovergestelde transcriptie, chromatine veranderingen, en de 3D structuur van het chromatine.

Twee genen, *Xist* en *Tsix*, spelen een cruciale rol bij het XCI proces. Beide genen liggen op het X chromosoom, overlappen elkaar maar worden afgeschreven in tegenovergestelde richting, en spelen een antagonistische rol in initiatie van het X chromosoom inactivatie (XCI) proces. In tegenstelling tot de meeste andere genen coderen *Xist* en *Tsix* niet voor een eiwit maar produceren lange niet coderende RNA moleculen. Het Xist RNA is functioneel en wordt alleen aangemaakt op het X chromosoom dat wordt uitgezet waar het accumuleert en het X chromosoom inpakt. *Xist* rekruteert daarbij verschillende eiwit complexen die ervoor zorgen dat het chromatine zodanig verandert dat genexpressie onmogelijk wordt. Als de X eenmaal is uitgezet tijdens de embryonale ontwikkeling dan wordt de inactieve staat gehandhaafd, ook in dochter cellen na celdeling.

In dit proefschrift worden verschillende aspecten van *Xist* regulatie en functie onderzocht tijdens de embryonale ontwikkeling, door gebruik te maken van genetisch gemodificeerde embryonale stam (ES) cellen. In **hoofdstuk 1** wordt een overzicht gegeven van de huidige kennis met betrekking tot XCI, en worden de belangrijkste wetenschappelijke vragen in dit onderzoeksveld belicht. **Hoofdstuk 2** beschrijft een serie van experimenten waarin de regulatie van *Xist* en zijn tegenovergesteld afgeschreven partner *Tsix*, onafhankelijk van elkaar wordt onderzocht in ES cellen waarin de gen-sequenties van *Xist* en *Tsix* zijn vervangen door GFP en mCHERRY fluorescente reporters. Door middel van dit systeem hebben we de dynamiek van *Xist* en *Tsix* expressie onderzocht, en gevonden dat ook in de afwezigheid van reciproke transcriptie beide genen tegengesteld gereguleerd worden. In **hoofdstuk 3** worden de mechanismen onderzocht die *Xist* in staat stellen een heel X chromosoom inactief te maken. Om dit

A

te bewerkstelligen is een *Xist* expressie systeem ontwikkeld waarmee *Xist* op verschillende plaatsen in het genoom tot expressie wordt gebracht. Dit onderzoek laat zien dat een hoge gen dichtheid in de omgeving van het transgene *Xist* locus erg belangrijk is voor *Xist* om zich te kunnen spreiden over een heel X chromosoom of autosoom, en dat een hoge dichtheid aan repetitieve LINE elementen daarbij blijkt te helpen. Onder de juiste omstandigheden is *Xist* in staat om vanuit verschillende posities op het X chromosoom, dit hele X chromosoom efficiënt uit te schakelen. Een aantal X-gebonden genen dat normaal genomen ontsnappen aan het XCI proces doen dit ook wanneer *Xist* vanuit een andere positie tot expressie wordt gebracht, en de mate waarmee een gen ontsnapt aan XCI is geassocieerd met verrijking van het eiwit CTCF in de promotors van deze genen. In **hoofdstuk 4** wordt de regulatie onderzocht van overlappende genen in de context van cel differentiatie. Een EGFP reporter gekoppeld aan een induceerbare promotor die transcriptie initieert in tegenovergestelde richting wordt gebruikt als paradigma voor het *Xist-Tsix* locus. Door gebruik te maken van dit systeem is gevonden dat EGFP uitgezet wordt als transcriptie in tegenovergestelde richting wordt geïnitieerd. Inactiviteit van EGFP is reversibel in ongedifferentieerde ES cellen, maar wordt gefixeerd tijdens ES cel differentiatie. Reversibele inactivatie is afhankelijk van chromatine en niet van transcriptie interferentie. Stabiele inactivering van EGFP wordt daarentegen gekarakteriseerd door accumulatie van CpG methylering rondom de EGFP promotor. Dit onderzoek laat zien dat tegenovergestelde transcriptie een chromatine markering achterlaat die tijdens cel differentiatie leidt tot CpG methylering en daarmee stabiele gen repressie. In
wordt een rat ES cel differentiatie systeem beschreven dat het mogelijk maakt XCI in vitro te onderzoeken. Uit dit onderzoek blijkt dat de rol van belangrijke XCI regulerende genen in muis geconserveerd zijn in de rat. In **hoofdstuk 6** worden de bevindingen beschreven in de voorgaande hoofdstukken bediscussieerd in de context van de huidige kennis, en worden daarnaast nieuwe inzichten en ideeën met betrekking tot XCI gepresenteerd.

A

CURRICULUM VITAE

Name	Agnese Loda
Date of Birth	September 27, 1986 in Savona, Italy
Nationality	Italian
Email	agnese.loda@gmail.com

Education

Mar 2011 to present	PhD candidate, Erasmus Medical Center, Rotterdam, The Netherlands Date of promotion 28th October 2016
Oct 2008 to Feb 2011	Master in Experimental and Applied Biology, University of Pavia, Italy Final mark: 110/110 cum laude
Oct 2005 to Oct 2008	Bachelor Degree in Biological Sciences, University of Pavia, Italy Final mark: 110/110 cum laude

Research experience

Mar 2011 to present	**PhD Thesis**: "X chromosome inactivation: spreading of silencing" **Prof. Gribnau laboratory**, Department of Developmental Biology. Erasmus MC, Rotterdam, The Netherlands
Sept 2009 to Dec 2010	**Master Thesis**: "Analysis of FBXO7 proteins in families with PARK15 (parkinsonian-pyramidal syndrome) **Prof. Bonifati laboratory,** Department of Clinical Genetics. Erasmus MC, Rotterdam, The Netherlands
Oct 2007 to Oct 2008	**Bachelor Thesis**: "Site-directed mutagenesis of MUTYH to study the pathogenesis of MAP (MUTYH Associated Polyposis). **Prof. Ranzani laboratory,** Department of Genetics and Microbiology University of Pavia, Italy

A

PUBLICATIONS

1. **Agnese Loda**, Johannes H. Brandsma, Ivaylo Vassilev, Nicolas Servant, Friedemann Loos, Azadeh Amirnasr, Erik Splinter, Raymond A. Poot, Edith Heard°, Joost Gribnau°, The efficiency of Xist-mediated silencing of X-linked and autosomal genes is determined by the genomic environment (manuscript in preparation).

2. Aristea Magaraki*, **Agnese Loda***, Cristina Gontan Pardo, Stephen Meek, Willy M. Baarends, Tom Burdon and Joost Gribnau1°, Generation of a novel in vitro differentiation strategy to study the dynamics of X chromosome inactivation in rat. (manuscript in preparation).

3. Loos, F., Maduro, C.*, **Loda, A.***, Lehmann, J., Kremers, G.J., Ten Berge, D., Grootegoed, J.A., Gribnau, J.°, Xist and Tsix transcription dynamics is regulated by the X-to-autosome ratio and semi-stable transcriptional states. Mol Cell Biol, pii: MCB.00183-16, epub ahead of print, 2016 Aug 15.

4. Loos, F., **Loda, A**., van Wijk L., Grootegoed J.A., Gribnau J.°, Chromatin mediated reversible silencing of sense-antisense gene pairs in ESCs is consolidated upon differentiation. Mol Cell Biol 35(14): 2436-47, 2015 May 11.

5. **Loda, A.***, Loos, F.*, Gribnau, J.°, Chapter 9 - X Chromosome Inactivation in Stem Cells and Development, in: Charbord, P., Durand, C. (Eds.), Stem Cell Biology and Regenerative Medicine. River Publishers. ISBN: 978-87-93237-07-0 (print), 2015 Jan.

6. Zhao T, De Graaff E, Breedveld GJ, **Loda A**, Severijnen LA, Wouters CH, Verheijen FW, Dekker MC, Montagna P, Willemsen R, Oostra BA, Bonifati V. Loss of nuclear activity of the FBXO7 protein in patients with parkinsonian-pyramidal syndrome (PARK15). PLoS One. 11;6(2):e16983, 2011 Feb

*equal contribution °corresponding author

A

PHD PORTFOLIO
SUMMARY OF PHD TRAINING AND TEACHING ACTIVITIES

Name PhD student: Agnese Loda Erasmus MC Department: Developmental Biology Research School: Biomedical Sciences	PhD period: March 2011- Septempber 2016 Promotor(s): Prof.dr. Joost Gribnau

1. PhD training

	Year
General courses	
- Laboratory animal sciences (art.9/FELASA-C)	2013
- Handling of laboratory animals (IVC)	2013
General academic skills	
- Genetics	2011
- Biochemistry and Biophysics	2011
- Cell and Developmental Biology	2012
- Literature course	2012
General research skills	
- Course on R	2014
- Analysis of microarray and RNA Seq expression data using R/BioC and web tools	2014
- NGS: analysis of RNA-seq data	2015
Presentations	
- Developmental Biology/Biochemistry Work Discussion	2011-2016
- 23nd MGC Symposium, Rotterdam, The Netherlands ("*Autosomal spreading of Xist revisited*")	2013
- 20th MGC Workshop, Luxembourg (*"Epigenetic characterization of X chromosomes in human somatic cells"*)	2013
- 21st MGC Workshop, Münster, Germany (*"Xist-mediated inactivation of autosomal and X-linked genes: does genomic position matter?"*)	2014
- Paris-Rotterdam joint XCI retreats	2012-2014
- 9th Winter school of the International Graduiertenkolleg, Kleinwalsertal, Austria	2016
- X-chromosome inactivation: a tribute to Mary Lyon, London UK (*"The efficiency of Xist-mediated silencing of X-linked and autosomal genes is determined by the genomic environment"*)	
(Inter)national conferences	
- 5th International NIRM symposium, Amsterdam, The Netherlands (poster)	2011 2011
- 4th Dutch Stem Cell meeting, Leiden, The Netherlands	2011
- EMBO Workshop 50 years of X-inactivation research, Oxford, UK (poster)	2011 2013

- TRR81 International Chromatin Symposium, Giessen, Germany	2013
- 11th Annual ISSCR Meeting, Boston, USA (poster)	2014
- 24nd MGC Symposium, Rotterdam, The Netherlands	
- Max Plank Epigenetics Meeting, Freiburg, Germany (poster)	
Seminars and workshops	
- 18th MGC PhD student Workshop, Maastricht, The Netherlands	2011
- 4th Dutch Stem Cell meeting, Leiden, The Netherlands	2011
- Photoshop and Illustrator CS6 Workshop for PhD-students	2012
- 19th MGC PhD student Workshop, Düsseldorf, Germany	2012
- Joint XCI retreat, Spetses, Greece	2012
- 20th MGC Workshop, Luxembourg	2013
- Joint XCI retreat, Rotterdam, The Netherlands	2013
- 21st MGC Workshop, Münster, Germany	2014
- Joint XCI retreat, Tinos, Greece	2016
2. Teaching activities	
	Year
Supervising	
- Tutoring Human Genetics Bachelor students (University of Pavia)	2008-2009
- Azadeh Amirnasr's Master Thesis	2013
3. Awards	**Year**
- Best oral presentation at the 20th MGC Workshop, Luxembourg	2013
- EMBO Short Term Fellowship	2015

A

ACKNOWLEDGMENTS

The last five years in Rotterdam have been an interesting mix of very different feelings including pure joy and huge frustration. I am grateful for everything I have learned and I feel fortunate to have never been alone along this path. To all the beautiful people that have cheered up for me, these pages are yours.

First of all, I am thankful to **Prof. Gribnau**, who was the first one to believe I could become a scientist. Dear **Joost**, you certainly are a great mind and pursuing my PhD in your lab was an exciting and totalizing experience. The freedom of research you put in my hands was scary at the beginning, and I spent the first months in your lab wondering how you could actually think I was going to make it. Now, I am glad you let me explore my projects, you made me grow enormously and you definitely helped me to realize the direction in which I am determined to proceed. From you I learned that an exam that leaves you without open questions is a failed exam, which was a culturally completely new concept to me. I am especially glad you taught me not to always take myself too seriously, because this made a happier person out of me. More than anything else, you supported me in a very delicate phase of my personal life, and I will never forget it. Thank you.

I am thankful to the members of my inner doctoral committee, **Prof. Heard**, **Prof. Verrijzer** and **Dr. Poot** for taking the time to read my work.

Dear **Edith**, it's an honour to have you sitting in my committee. The months I spent in your lab were instrumental to me, I realized what I really want to learn and which kind of scientist I want to become. I am looking forward to the French adventure to begin.

Dear **Peter**, the feedback you provided at our Monday Morning Meetings was always very much appreciated, you really made me think about which experiment was still missing to make my point.

Dear **Raymond**, thank you for your thoughtful suggestions on how to improve my manuscript, you made this book a better one.

I am thankful to **Prof. Grosveld**, **Prof. de Laat** and **Dr. Baarends** for completing my doctoral committee and for being there for my thesis defense.

Dear **Willy**, during the years your help was very precious to me, and not only scientifically. You are a super woman and supervisor and I really like your attitude towards life. I learned a lot from you, thank you.

Dear **Anton**, **Prof. Grootegoed**, I am very glad I had the opportunity to work with you. Your dedication to science has been a great motivation to me.

Thank you to my paranymphs, for having been there since the very beginning.

Dear **Fede**, in very simple words, you are the best. You have always supported this work in all possible ways, from staying long nights at the bench optimizing the crazy-fishing-screening to

A

all the times we have laughed together during these amazing years. I know how lucky I am to have you as a friend. **You**, **Ale** and **Mattia** make me feel loved. *Grazie, Vi voglio bene.*
Dear **Isa**, the first time I stepped into a lab you were there, it was 2007 and we were in Pavia. Many things have changed since then and I am happy our friendship grew up with us. Without you, this book would never be here. To me, you have always been inspiring, both as a scientist and as a human being. I am thankful for your empathy, it makes you one of a very unique kind. *Per tutto quello che fai per me, Grazie.*
Dear **Friedo**, you are the biggest surprise of my time in The Netherlands. Your contribution to this thesis is obviously enormous, I am glad our projects overlapped, you are brilliant and I have learned a lot from you. Your determination is amazing. More than anybody else, you found the way to deal with my enthusiasm (and especially with my frustration) which is the opposite of an easy task. Your pure allegria makes my life better, and I am very proud of our European family. *Thank you, grazie, danke schön* and now *merci beaucoup* as well.
I am thankful to the all the past and present members of the **Developmental Biology department** for having created such a sparkling environment. **Cristina**, you are a great scientist, friend and mother and I am sure I am going to miss you. Thank you for everything you have taught me along the way. I really had fun, a lot of fun. **Aristea**, you are my favourite, I deeply admire you and your unbelievable strength. I am thankful for the huge amount of work you did for chapter 5, without you, this story would have never been started. You always took care of me like only good friends can do, especially when things got difficult. Thank you for never letting go. **Fabrizia**, I am glad for the time we spent together in the last year, it was precious. You are thoughtful, and I have no doubt there is a brilliant future waiting for you. *Non dubitare, non saremo mai davvero lontane dalle nostre Alice*. **Hegias**, you are smart and cheerful and I really enjoyed being around you, you are one of these people who makes the lab a better place to be. **Joke**, there is something very special about you, to me you have been a mentor and a friend, and I could not have asked for more, thank you. **Sarra**, I am happy you joined the J-lab, I have enjoyed working with you since the very beginning of our collaboration. You are sharp, trustful and skilled, and especially fun. Thank you for all the scientific input you gave me. **Eskeww**, **Helen** (and little **Artemis**), I could always count on you, thank you for having been there, I will never forget it. **Cheryl**, thousands of clones later, we made it! Good luck with the final sprint. **Ruben**, it was fun to share the office and to be paranymphs in crime. **Evelyne**, **Esther**, **Eveline**, **Teresa**, **Andrea**, **Azi**, **Annegien**, **Bas**, **Selma**, **Marja**, **Jos**, **Daphne**, **Godfried**, **Stefan** and **Catherine**, thank you for the bench-time in the lab, it was super!
Many of the cool people I met in Rotterdam happened to be scientists. To them goes a huge thank you, it still feels impossible but we really made it. **Nesrin**, one thing is sure, no matter in how many cell-culture rooms I will be working, I will be missing our culturing afternoons (and evenings, and weekends). You are strong and smart, I know our scientific paths will cross again. **Fanny**, after having tested several people I can finally say you are the best at the stat-

ue-mimic context we invented in Maastricht. You really are a good one, don't ever change. **Johan**, I could not have had a better maest-**R**-o. Thank you for having undertaken the task of teaching me how to deal with NGS data. You are gifted and generous and it was an honour to stand close to you at your PhD defense. Science is totally missing you. **Maaike**, you were right: there is still life after the writing! **Andrea**, after having tried your gluten-free lactose-free cheese cake I directly realized that Isa had found the right one. **Jessica** and **Thomas**, I am so happy you are coming all the way from Freiburg to celebrate, R'dam is really not the same without you! *Jessi, sei eccezionale, sai sempre come starmi vicina in un modo molto speciale, grazie*. **Simone** e **Flavia**, *promessi sposi, siete fortissimi*. Olgiati, I really miss our chats in the office, whether it was about writing our thesis, living abroad or Italian politics, I always enjoyed it. **Jovana** and **Enrico**, it's basically impossible not to love you, especially at Franco's. **Luca**, you really are the next one now. I am sure you will do a great job. **Francesca**, the days in Naples were great, you are doing a wonderful job there at the sea and I am thankful I could get to know your new world.

I am thankful to **Alessia** and **Chiara**, for the huge amount of work they have done to make this book look beautiful. You are the perfect example of very dedicated people who really love what they do. I like you more and more, Thank you. *Ale, mi dai una forza incredibile senza nemmeno bisogno di parole. Per esserci da sempre, grazie*.

I am thankful to **Prof. Rosati** and **Dr. Bailo** of the San Raffaele hospital (Milan), and to **Dr. Giacchello** of the Regina Montis Regalis hospital (Mondovì), for having shown me the most noble means of research, which is helping others.

To my old friends spread all over Europe and the world, you are my sunshine! Sono felice che i binari riescano sempre ad incrociarsi, grazie! **Valerie**, tra di noi le parole non servono, ma sono grata che tu riesca sempre a trovare quelle giuste. Con te sono sempre a casa, anche a distanza da quella vera. **Giulia**, sei una forza della natura. Senza di te l'ultimo inverno non si sarebbe mai trasformato in primavera. **Federico**, vederti è sempre una gioia pazzesca, ti voglio bene. **Ylenia**, la nostra amicizia è potente e bellissima e tu mi manchi sempre un po' di più. **MaFre**, non cambi mai e questa è una cosa *veramente eccezionale*, grazie. **Mazzo**, sei mitico, ti vorrei sempre vicino! **Claudia**, tu più di tutti sai come mi sono sentita in questi mesi, sei il più bello dei regali Erasmus, grazie. **Irene**, per me hai saputo fare la cosa più difficile e te ne sarò per sempre grata. Sei una meraviglia, da sempre. **Sara**, sono vent'anni che ci capiamo al volo! **Cristina**, non mi hai lasciata sola mai, grazie. **Alessandra**, il tuo lavoro è faticoso e difficile, ma non mollare mai. **Mauri**, potere micron, sono fiera di noi. **Erika**, ritrovarti è stato un regalo bellissimo, sei un'amica preziosa, grazie. **Greta**, non vedo l'ora di festeggiarvi, siete unici. **Valentina**, sei sincera, forte e coraggiosa, grazie.

Grazie a **Nuccio**, per avermi saputo indicare la strada senza dubitare, e a **Luca**, per *il carattere*.

Sono grata ad **Umberto**, **Patrizia** e **Andrea** per non essersi spaventati mai ed essermi stati accanto sempre, soprattutto nei momenti in cui abbiamo pensato di non farcela. Grazie a

nonna **Mirosa**, per la sua energia, e a **Olga**, **Silvio** e **Stefano** per essere rimasti. A **Grazia** e **Alberto**, per avere compreso quanto fosse importante per me questo traguardo, sono felice di avervi nella mia vita.
Finally, I am thankful to my parents, **Aldo** e **Rita**, for raising me to believe that *anything* is possible, to my sister **Alice**, for being the extraordinary human being she is, and to baby **Elsa**, for having brought the sun back. Mamma, questa tesi è per noi quattro, se siamo stati sempre così forti il merito è soprattutto tuo, sei una donna straordinaria, è da te che ho imparato a non mollare mai. Alice, il solo fatto che tu ci sia è pura gioia, tu sai già tutto, quello che sento io lo senti anche tu ed è questa la nostra forza. Per il modo in cui mi proteggi, grazie.
Aramambo splende.