

# Dynamics of X Chromosome Inactivation



Friedemann Loos

---

ISBN: 978-94-6299-028-9

Cover: Jochen Loos

Layout: Friedemann Loos; Ridderprint BV

Printing: Ridderprint BV, Ridderkerk, The Netherlands

The work described in this thesis was performed at the Department of Reproduction and Development at the Erasmus MC in Rotterdam, The Netherlands.

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

Copyright © 2015 by F. Loos. All rights reserved. No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means, without prior permission of the author.

# Dynamics of X Chromosome Inactivation

De Dynamiek van X Chromosoom-Inactivatie

Thesis

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

Prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board.

The public defense shall be held on  
Friday 6 March 2015 at 13.30 hrs

by

**Friedemann Loos**  
born in Münster, Germany



## DOCTORAL COMMITTEE

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

**Other Members:** Prof.dr. F.G. Grosveld  
Prof.dr. C.P. Verrijzer  
Prof.dr. A.B. Houtsmuller



# TABLE OF CONTENTS

List of abbreviations .....	6
<b>Chapter 1</b>	
Introduction.....	8
Aim and Scope of this Thesis.....	51
<b>Chapter 2</b>	
The <i>trans</i> -activator RNF12 and <i>cis</i> -acting elements.....	80
effectuate X chromosome inactivation independent of X-pairing	
<b>Chapter 3</b>	
<i>Xist</i> and <i>Tsix</i> transcription dynamics in mouse embryonic... stem cells is regulated by the X-to-autosome ratio and predicted by semi-stable transcriptional states	106
<b>Chapter 4</b>	
Chromatin mediated reversible silencing of sense-antisense gene pairs in ESCs is consolidated upon differentiation.....	136
<b>Chapter 5</b>	
General Discussion.....	160
<b>Appendix A</b>	
Epigenetic characterization of the <i>FMRI</i> promoter in induced pluripotent stem cells from human fibroblasts carrying an unmethylated full mutation.....	176
<b>Appendix B</b>	
Summary.....	197
Samenvatting.....	201
Zusammenfassung.....	205
Curriculum Vitae.....	209
List of Publications.....	211
PhD Portfolio.....	212
Acknowledgements.....	213

## LIST OF ABBREVIATIONS

3C	chromosome conformation capture
5-mC	5-methylcytosine
5'-UTR	5' untranslated region
BAC	bacterial artificial chromosome
BF	blocking factor
bp	base pair
ceRNA	competing endogenous RNA
CF	competence factor
CGI	CpG island
DCC	dosage compensation complex
EpiSC	epiblast derived stem cell
eRNA	enhancer-associated RNA
(h)ES cell/(h)ESC	(human) embryonic stem cell
FISH	fluorescence in situ hybridization
FXS	fragile X syndrome
HAS	high affinity site
ICM	inner cell mass
ICR	imprinting control region
(h)iPS cell/(h)iPSC	(human) induced pluripotent stem cell
iXCI	imprinted X chromosome inactivation
lincRNA	large intergenic non-coding RNA
LINE	long interspersed nuclear element
Mb	mega base
mRNA	messenger RNA
ncRNA	non-coding RNA

PcG	polycomb group
PE	primitive endoderm
PGC	primordial germ cell
PIC	preinitiation complex
Pol2	RNA polymerase II
PTGS	post-transcriptional gene silencing
RdDM	RNA-directed DNA methylation
RNAi	RNA interference
rXCI	random X chromosome inactivation
siRNA	small interfering RNA
TAD	topologically associating domains
TE	trophectoderm
TGS	transcriptional gene silencing
TrxG	trithorax group
TSS	transcription start site
X:A	X-to-autosome
Xa	active X chromosome
Xce	X controlling element
XCR	X chromosome reactivation
Xi	inactive X chromosome
Xic	X inactivation center
Xist	X inactive specific transcript
Xm	maternally derived X chromosome
Xp	paternally derived X chromosome

# CHAPTER



# 1



## Introduction

*Excerpts of this chapter adapted from:*

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

1

## Transcriptional regulation and the role of chromatin during development

### Gene regulation and chromatin structure

In the end, transcriptional regulation is responsible for nearly all biological processes. It helps adapting to and coping with different environmental conditions by changing the composition and properties of genetically identical cells. Transcriptional regulation is involved in determining which program is run in any given cell in multi-cellular organisms, thereby establishing different cell identities needed for an organism to function properly as a whole. And, in the case of malfunction, it is responsible for death and disease.

### *The core promoter and other regulatory elements*

Gene regulation is a multi-layered process in which several interwoven principles govern the final outcome. Early studies in phages, bacteria and yeast led to a simple activation by recruitment model in which transcription factor binding to specific DNA elements would recruit activators or repressors of transcription to the core promoter (reviewed in Ptashne, 2005). At the core promoter in metazoans, RNA polymerase II (Pol2), a 12 subunit complex being the catalytic entity generating RNAs using DNA as a template, and general transcription factors that form the preinitiation complex (PIC) are recruited to initiate transcription (reviewed in Roeder, 1998; Green, 2005; Grunberg and Hahn, 2013). Considering the unexpectedly low number of protein coding genes in mammals, which is in the same range as

for other less complex eukaryotes, and the vast amount of non-coding DNA present in mammalian genomes, it has been proposed that organismal complexity scales with the complexity of gene regulatory mechanisms (reviewed in Levine and Tjian, 2003). In support of this, mammals exploit an extensive number of *trans*-acting transcription factors, accounting for more than 5% of all coding genes (Zhang et al., 2014), whose combinatorial binding specificities to DNA elements provide a platform for temporally and spatially diversified gene regulation (reviewed in Levo and Segal, 2014). To orchestrate the complex networks of gene regulation employed during life and development of multi-cellular organisms, additional diversity comes from variability in core promoter architecture and usage of basal transcription machinery (reviewed in Juven-Gershon and Kadonaga, 2010). In addition to the promoter itself, distal *cis*-regulatory elements termed enhancers, which can be more than 1 Mb away from their target genes (Lettice et al., 2003), are essential for proper gene regulation in higher eukaryotes. The first enhancer was described in the SV40 virus (Banerji et al., 1981) and by now novel genome-wide approaches have identified ca. 400000 putative enhancer elements in the human genome (Consortium, 2012). These regulatory elements, in a combinatorial fashion, allow complex cell type-specific, developmental and signal-dependent transcription programs to be carried out (Andersson et al., 2014; Nord et al., 2013; Visel et al., 2009). Even though it is not entirely clear how enhancers and promoters interact, it becomes increasingly clear that the 3-dimensional organization of chromatin in the nucleus is essential for the formation of contacts between promoters and distal regulatory elements and proper gene regulation (reviewed in Wendt and Grosveld, 2014). Evidence originally obtained from the classical example of chromatin looping at the

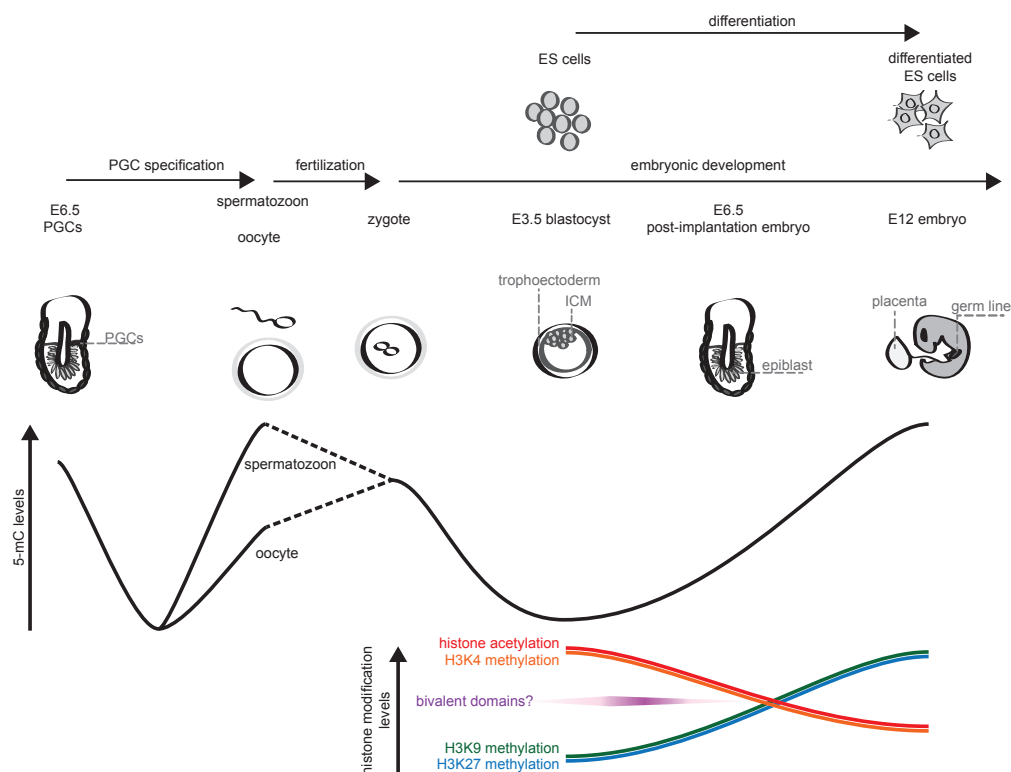
globin locus (Hanscombe et al., 1991) has been greatly extended by the advent of 3C technology (Dekker et al., 2002). Recent studies demonstrate that the entire genome is organized into topologically associating domains (TADs) of about 1Mb size, that represent domains with preferred 3 dimensional interactions (Dixon et al., 2012; Lieberman-Aiden et al., 2009). As shown for the locus responsible for X chromosome inactivation (XCI), perturbation of the boundary between two TADs results in aberrant gene regulation in the affected neighboring regions (Nora et al., 2012).

### ***Histones and their posttranslational modifications modulate chromatin structure***

Another level of gene regulation is imposed by the structure of the chromatin template itself. The basal unit of DNA packaging is the nucleosome, a histone octamer typically consisting of two H2A-H2B dimers and an H3-H4 hetero-tetramer (Kornberg, 1974; Kornberg and Thomas, 1974). Approximately 147 bp of DNA is wrapped around one nucleosome. It is instantly and intuitively conceivable that this chromatin template, constituting a barrier to DNA-related functions, affects all aspects of transcription, including transcription factor recruitment, transcriptional initiation, elongation and memory of gene regulation (reviewed in Li et al., 2007). Posttranslational modifications on histone tails received particular attention as these modifications associate with gene expression levels, and provide context-dependent binding platforms for other chromatin factors. In addition, histone modifications provide a platform for transcriptional memory inherited through cell division, forming the basis for epigenetic gene regulation. Acetylation of H3 and H4 is transient and associated with active

chromatin or regulatory regions. Possibly, this relationship stems from the ability of acetyl groups to directly influence histone-DNA interactions, for example by repelling negative charges (Hong et al., 1993; Shogren-Knaak et al., 2006). Methylation on lysine 9 of histone H3 (H3K9me2/3) occurs in heterochromatic regions and is catalyzed by G9A (Tachibana et al., 2002) and SETDB1 (Schultz et al., 2002). Also H3K27me3 is associated with gene repression. This modification is catalyzed by the Polycomb-group (PcG) protein EZH2, which is a subunit of the Polycomb Repressive Complex PRC2, and is often found to be localized to promoters (Cao et al., 2002). Both the H3K9me2/3 and H3K27me3 modifications are enriched on the inactive X chromosome (Xi) in female mammals, but recruitment seems to be restricted to different regions on the Xi (Heard et al., 2001; Plath et al., 2003). On the contrary, the H3K4me3 modification appears primarily in promoters of active genes, with H3K4me1/2 tapering off towards the 3'-end of the gene. Due to its association with elongating Pol2, H3K36me3 is also mostly found in gene bodies of active genes. Many additional histone modifications have been described in the past years and their presence at specific positions in the genome correlates with many nuclear processes (reviewed in Bannister and Kouzarides, 2011). The discovery of chromo- and bromo-domains as protein domain structures that specifically bind methylated and acetylated lysines on histone tails respectively (Dhalluin et al., 1999) provided an attractive model of how specific histone modifications or combinations thereof might confer function by recruiting "readers" with differing specificities (Bartke et al., 2010; Vermeulen et al., 2010). For example, H3K4 methylation, which is catalyzed by recruitment of Set1/MLL-like complexes (Ng et al., 2003), recruits several factors that facilitate transcription initiation and maintenance of an





**Figure 1. Chromatin Structure and Modifications in the Mammalian Life Cycle**

Different stages of the mammalian life cycle are shown. Reprogramming of CpG methylation is depicted by black line, showing global levels of 5-mC. During primordial germ cell (PGC) specification, genome-wide DNA demethylation takes place including at imprinting control regions. Starting at E13, sex-specific DNA methylation is restored. Upon fertilization, the genome undergoes global DNA demethylation. The inner cell mass (ICM) of the blastocyst and ES cells are strongly hypomethylated and starting with implantation or differentiation adult 5-mC patterns are established. Histone modifications are depicted by colored lines, in general ES cells show a more open chromatin configuration characterized by histone acetylation and H3K4me3. Upon differentiation, coverage of these marks decreases and H3K9/H3K27 methylation is gained over broad domains. In purple, the possible presence in pluripotent cells of bivalent domains, that is, nucleosomes with overlapping H3K4me3 and H3K27me3, is indicated.

active chromatin structure. These include general transcription factors (Lauberth et al., 2013), the splicing machinery (Sims et al., 2007), chromatin remodelers (Pray-Grant et al., 2005; Ruthenburg et al., 2011; Wysocka et al., 2006), H3K9/27-specific histone demethylases (Lin et al., 2010) and histone acetyltransferases (Taverna et al., 2006). Contrarily, only unmethylated H3K4 has been shown to have an affinity for DNA methyltransferase DN-

MT3A/B/L (Ooi et al., 2007; Otani et al., 2009). In a similar fashion, H3K36me3-dependent recruitment of DNMT3A/B via their PWWP domain (Dhayalan et al., 2010), of histone deacetylases and of H3K4me3-specific demethylases results in suppression of intragenic transcription (Carrozza et al., 2005; Houseley et al., 2008; Xie et al., 2011). Based on these findings a 'histone code' hypothesis has been introduced which proposes that

different histone modifications and combinations of these modifications are in principle instructive for certain downstream events like gene activation (Strahl and Allis, 2000). Although this hypothesis is appealing, cause and consequence are not clearly separated and the crucial question of how these histone modifications are targeted and propagated remains to be solved (Henikoff and Shilatifard, 2011). In addition to histone modifications, chromatin structure and gene regulation can be modulated by the incorporation of histone variants (reviewed in Weber and Henikoff, 2014) such as macroH2A which is enriched on the Xi and has therefore been implicated in transcriptional repression (Costanzi and Pehrson, 1998).

### ***DNA methylation is involved in gene repression***

DNA methylation represents another epigenetic mechanism involved in gene regulation of mammalian genomes. It occurs primarily as symmetric 5-methylcytosine (5-mC) at CpG dinucleotides and is often involved in development of disease like Fragile X Syndrome or cancer (reviewed in Guibert and Weber, 2013; Cedar and Bergman, 2012). Moreover, DNA methylation is essential for genomic imprinting (Li et al., 1993; Bourc'his et al., 2001; Kaneda et al., 2004; reviewed in Barlow, 2011) and plays an important role in the control of transposable elements (Walsh et al., 1998) and maintenance of XCI (Sado et al., 2000). Early studies demonstrated that most DNA is generally methylated but not in tissues in which the tested genes are expressed (Bird et al., 1985; Naveh-Many and Cedar, 1981). Thus, DNA methylation correlates with gene repression and a causal role was proven by using pre-methylated transgenes which, in contrast to their un-

methylated counterparts, were repressed in both cell culture and *in vivo* systems (Siegfried et al., 1999; Stein et al., 1982b). Of note, DNA methylation is generally believed to be a late event in the process of silencing, because it emerges only after initial gene silencing in XCI (Gendrel et al., 2012; Lock et al., 1987). Importantly, 5-mC is propagated through cell division (Stein et al., 1982a) because hemi-methylated DNA is the substrate for the maintenance DNA methyltransferase DNMT1 which recurrently catalyzes full methylation of the CpG dinucleotide (Gruenbaum et al., 1982; Leonhardt et al., 1992; Li et al., 1992). The *de novo* methyltransferases DNMT3A and DNMT3B catalyze methylation of completely unmethylated CpG's (Okano et al., 1999) and are mainly expressed in the early embryo, demarcating the major developmental window of time in which DNA methylation is established and remodeled (Borgel et al., 2010). It is not entirely clear how DNA methylation translates into actual gene repression, but one possibility is that it directly modulates chromatin structure by influencing nucleosome positioning (Chodavarapu et al., 2010). Interestingly, methylated DNA only confers a repressive effect after incorporation into chromatin (Buschhausen et al., 1987) and methylated and unmethylated DNA are assembled into DNase I-insensitive and -sensitive chromatin, respectively (Keshet et al., 1986). These findings suggest that chromatinization of DNA is essential for any 5-mC relayed gene repression. In addition, a family of methyl-binding proteins including MeCP2 and MBD3 has been identified (Lewis, 1992; reviewed in Klose and Bird, 2006), whose members associate with methylated DNA and repress transcription specifically from methylated templates (Nan et al., 1997). In support of a pivotal role for chromatin-mediated silencing, this mechanism likely depends on recruitment of histone remodeling complexes (Hashimshony

et al., 2003; Nan et al., 1998). Similarly, ZFP57 and KAP1 have been shown to specifically bind a methylated DNA motif in an imprinting control region. This binding results in tethering of additional histone modifiers and DNA methyltransferase to the locus and thereby stabilizes the local chromatin state (Quenneville et al., 2011). As unmethylated H3K4 displays enhanced affinity towards DNA methyltransferases (Ooi et al., 2007), this might create a feed-forward loop partially explaining the stability of DNA methylation-mediated silencing. CTCF, a DNA-binding protein involved in the global organization of higher order chromatin structure (reviewed in Merkenschlager and Odom, 2013), also preferentially binds to hypomethylated DNA involved in differential expression at the imprinted *IGF2/H19* locus (Bell and Felsenfeld, 2000; Hark et al., 2000). Interestingly, CTCF also demarcates the boundary between two adjacent TADs involved in the regulation of XCI (Spencer et al., 2011) and deletion of the boundary leads to aberrant gene expression from the X-inactivation center (*Xic*). Alternatively, repressive effects on transcription could be mediated by differential affinities of transcription factors to methylated and unmethylated CpGs within their binding motifs. YY1 and E2F, for example, display methylation-sensitive binding to their cognate sites (Campanero et al., 2000; Kim et al., 2003). Recently, binding of YY1 to un-methylated regulatory elements has been implicated in mono-allelic up-regulation of *Xist* at the onset of XCI (Makhlouf et al., 2014). However, since many transcription factor binding sites do not contain CpGs and several transcription factors such as Sp1 are unaffected by methylation (Holler et al., 1988), this mechanism appears to be restricted to a subset of factors.

### ***CpG islands are regions of transcriptional initiation***

In mammalian genomes, CpG density is not uniformly distributed. This is a consequence of the inherent mutagenicity of 5-mC due to deamination and implies that CpG's that are unmethylated in the germline are conserved while methylated cytosines are subject to a continuous loss over evolutionary time scales (Cohen et al., 2011). While inter- and intragenic regions are depleted for CpG's, regions of higher CpG density, so-called CpG islands (CGI's; reviewed in Deaton and Bird, 2011), map to more than 60% of annotated promoters, including almost all housekeeping and many tissue-specific genes (Saxonov et al., 2006). CGI's generally correlate with transcription start sites (TSS's), are also found at intragenic sites (Illingworth et al., 2010), and Pol2 binds to these CG-rich regions regardless of their transcriptional status (Core et al., 2008; Guenther et al., 2007). CGI promoters rarely contain TATA-boxes and usually show dispersed patterns of transcription initiation (reviewed in Juven-Gershon et al., 2008). In addition, other factors contribute to the correlation of CGI's with transcription initiation. For example, CG-rich sequences appear to destabilize nucleosomes, since SWI/SNF complexes are not required for transcriptional activation from promoters embedded in these sequences (Ramirez-Carrozzi et al., 2009) and transcription factor binding motifs are generally rich in CpG's (Landolin et al., 2010). Methylation of CpG's is bimodally distributed (Straussman et al., 2009 2009). Gene bodies are typically methylated and this methylation correlates with gene expression (Ball et al., 2009; Hellman and Chess, 2007) which is explained by the Pol2-H3K36me3-DNMT3 axis discussed before and might serve to inhibit intragenic transcription. Promoter CGI's, however, are mostly devoid of methyla-

tion. This raises the question how they are protected from CpG methylation. In general, it is likely that typical characteristics of CGI's like Pol2/transcription factor binding and low nucleosome occupancy prevent DNA methylation by steric hindrance or lack of DNA methyltransferase recruitment (Brandeis et al., 1994; Jeong et al., 2009). Moreover, proteins binding to unmethylated DNA via CXXC domains are involved in setting up particular chromatin states. Cfp1, for example, associates with H3K4 methyltransferase Set1 and creates regions of H3K4me3 in promoters devoid of 5-mC (Thomson et al., 2010), probably because methylated H3K4 is refractory to DNMT3A/B binding (Ooi et al., 2007). Other interesting proteins containing CXXC domains are KDM2A which removes H3K36me2 (Blackledge et al., 2010) thereby preventing DNMT3A/B recruitment (Dhayalan et al., 2010) and TET1/3, which catalyze the conversion of 5-mC to 5-hydroxymethylcytosine (Tahiliani et al.). This process has been implicated in the active removal of 5-mC and could therefore protect CGI's from methylation (reviewed in Williams et al., 2012).

### Changes in chromatin structure and gene regulation patterns during early embryonic development

Mouse embryos only take four days to develop from the totipotent zygote into a blastocyst with three specified lineages. The trophectoderm (TE), which gives rise to the placenta and its membranes, and the inner cell mass (ICM), containing both the primitive endoderm (PE) that will develop into the yolk sac and the epiblast, of which the latter will form the embryo proper. After formation of these three lineages the embryo implants into the uterus and undergoes gastrulation which results in the formation of the primary germ lay-

ers (reviewed in Stephenson et al., 2012). During these major transitions, which are governed by an intricate interplay of transcription factor networks and signaling pathways, the developmental potential becomes increasingly restricted while lineages are being specified (reviewed in Arnold and Robertson, 2009; Ng et al., 2012). OCT4, SOX2 and NANOG are well-characterized transcription factors that form part of the network maintaining pluripotency in the ICM and the ICM-derived embryonic stem cells (ES cells). For example, OCT4 is essential for ICM versus TE specification as *Oct4* deletion results in failed induction of the ICM, resulting in embryo's that solely develop the TE compartment (Nichols et al., 1998). In these processes heterogeneity in expression levels of transcription factors might help to reach stochastic cell fate decisions by creating "windows of opportunity" (reviewed in Torres-Padilla and Chambers, 2014).

### Chromatin structure in embryonic stem cells

Lineage commitment can be seen as an increasing yet plastic epigenetic restriction of developmental or differentiation potential starting from the totipotent zygote (reviewed in Pera and Tam, 2010). To orchestrate the temporal changes in gene expression, instructive cues and their associated signaling pathways act on chromatin structure (Figure 1; reviewed in Badeaux and Shi, 2013). Mouse ES cells are the best-studied model for epigenetic processes during early embryonic development. In these cells, the JAK-STAT pathway is important for maintaining *Nanog* and *Sox2* expression (Griffiths et al., 2011), and inhibition of the MAPK and Gsk3 $\beta$  pathways in 2i conditions results in loss of DNA methylation because increased PRDM14 levels repress

DNA methyltransferases (Ficz et al., 2013; Grabole et al., 2013; Yamaji et al., 2013). In addition, slightly increased expression of Tet1/2 tips the balance to a state of global DNA demethylation, reminiscent of cells in the ICM just after specification (Habibi et al., 2013). ES cells grown in 2i conditions are therefore considered more naïve than ES cells grown under normal conditions. Naïve ES cells also show reduced levels of H3K27me3 at promoters (Marks et al., 2012). In general ES cells harbor less heterochromatin specific epigenetic modifications, and chromatin appears to be more flexible and less condensed than in differentiated cells (Meissner et al., 2008; Meshorer et al., 2006). This is consistent with high levels of transcription across all genomic regions (Efroni et al., 2008). Another epigenetic feature specifically described for ES cells are bivalent domains that contain both H3K4me3 and H3K27me3 histone modifications, which are correlated with gene activation and repression, respectively (Bernstein et al., 2006). Intriguingly, bivalent domains mostly reside in the promoters of key developmental genes which are repressed in the pluripotent state. Upon differentiation, however, they are resolved into domains carrying either H3K4me3 or H3K27me3, depending on which developmental path they follow (Mikkelsen et al., 2007). Thus, it has been proposed that bivalent domains mark genes that are silent but “poised” for activation upon differentiation. Although this hypothesis is appealing, prevalence and function of bivalent domains are still a topic of intense debate.

### ***Chromatin dynamics in embryos and differentiating embryonic stem cells***

Upon ES cell differentiation, general chromatin condensation is observed as shown by increased levels of H3K9me2

(Wen et al., 2009) and H3K27me3 (Zhu et al., 2013). Also DNA methylation, coinciding with the expression of DNMT's, is gained particularly at the promoters of pluripotency and germ line-specific factors (Borgel et al., 2010; Meissner et al., 2008; Mohn et al., 2008). This is consistent with the described cycles of methylation and demethylation of CpG's during *in vivo* development and contrasts with stable 5-mC patterns in somatic cells (Kafri et al., 1992; Monk et al., 1987; Sanford et al., 1987; reviewed in Smith and Meissner, 2013). Using whole genome bisulfite-sequencing and methylated DNA immunoprecipitation at different stages of pre- and post-implantation embryos, the *in vivo* dynamics of DNA methylation have been described in greater detail. Just after fertilization the relatively hypomethylated female and the hypermethylated male pronucleus undergo further demethylation until the ICM stage of the blastocyst is reached, which is essentially devoid of DNA methylation. This demethylation relies on passive dilution and, in the case of the male pronucleus, active demethylation by TET3 (Gu et al., 2011). During subsequent implantation, DNA methylation is regained particularly in regions of low CpG density and at a subset of promoters belonging to germ line-specific genes, pluripotency factors and some lineage-specific genes (Borgel et al., 2010; Smith et al., 2012). Also at later developmental stages, for example during specification of the hematopoietic lineages, small changes in DNA methylation patterns occur, and depending on cellular identity promoters of lineage-specific genes are re- or demethylated (Bock et al., 2012; Borgel et al., 2010). In human ES cells, differentiation towards cell types of different germ layers results in a lineage-specific switch from high DNA methylation to either H3K4me1 or H3K27me3 at regulatory sequences (Gifford et al., 2013). Since this switch did not automatically correlate



with altered expression levels of target genes, a model of epigenetic priming similar to “poised” bivalent domains was proposed. An epigenetic pre-disposition for cell fate decisions has also been demonstrated in liver and pancreas progenitors (Xu et al., 2011a) and might possibly involve the higher order chromatin structure as shown by pre-existing chromatin contacts between enhancers and target promoters (Jin et al., 2013). Pioneer transcription factors that initially bind to regulatory elements and are thought to allow quick induction by modulating chromatin structure might also play a role in this process (reviewed in Zaret and Carroll, 2011). Another interesting finding from human and mouse ES cell differentiation experiments suggests that CpG-rich promoters of developmental regulators that are active early during differentiation are repressed by H3K27me3-related mechanisms while tissue-specific promoters, which display a lower CG content and are expressed at later stages, gain DNA methylation upon silencing (Meissner et al., 2008; Xie et al., 2013). Genetic studies have shed further light on the role of chromatin modifying machinery in early development. It was shown that DNMT1, DNMT3A and DNMT3B are essential for mouse development (Li et al., 1992; Okano et al., 1999), while *Dnmt3L*<sup>-/-</sup> mice fail to properly establish genomic imprints in male and female gametes (Bourc’his et al., 2001). Interestingly, ES cells carrying homozygous deletions for DNMT’s are viable and only show marked effects upon differentiation (Chen et al., 2003; Jackson et al., 2004; Panning and Jaenisch, 1996; Tsumura et al., 2006). Not surprisingly, deletion of many other chromatin remodeling and modifying complexes show broad phenotypes observed in undifferentiated and differentiated ES cells and different stages of embryonic development (reviewed in Chen and Dent, 2014). The following non-exhaustive list describes several inter-

esting cases especially demonstrating that complexes associated with gene silencing have important functions in differentiation and development. During early lineage specification, OCT4 recruits SETDB1 to repress a set of target genes responsible for differentiation of the TE. Therefore, deletion of *Setdb1* resembles phenotypes of OCT4 ablation in that it restricts ES cells to a TE fate (Bilodeau et al., 2009; Yuan et al., 2009). Also the NuRD complex, which has histone remodeling and histone deacetylase activities, is important for lineage commitment of pluripotent cells, as targeted loss of NuRD subunit MBD3 prevents ES cells from exiting the pluripotent state (Kaji et al., 2006; Reynolds et al., 2012). The repressive PcG complexes PRC1 and PRC2 play an important role in the repression of developmental genes. ES cells deficient for subunits of these complexes display de-repression and preferential activation of PRC target genes upon differentiation (Boyer et al., 2006) and homozygous knockout mouse embryos arrest at the latest at the gastrulation stage (O’Carroll et al., 2001; Voncken et al., 2003). Analysis of deletions and point mutations in the histone methyltransferase G9A showed its involvement in a stepwise process of *Oct4* silencing during ES cell differentiation. G9A mediates initial heterochromatinization via H3K9 methylation and HP1 recruitment. Subsequently, recruitment of DNMT3A/B results in *de novo* DNA methylation as a last nearly irreversible event, supporting a putative role for DNA methylation in terminal silencing (Epsztejn-Litman et al., 2008; Feldman et al., 2006). In addition to chromatin factors involved in transcriptional repression, activating complexes like SAGA have also been implicated in embryonic development. For example, deletion of the histone acetyltransferase GCN5, a subunit of SAGA, results in embryonic lethality and reduced developmental potential (Lin et al., 2007; Xu et al., 2000).

## Conclusions

Technological advances in recent years have emphasized the crucial role of chromatin in gene regulation and development. However, the function of many additional epigenetic factors during embryonic development remains to be elucidated and it will be of special interest to further investigate the cause-consequence relationship of epigenetic phenomena. New technological developments will also help to examine the intricate interplay of environmental cues, chromatin structure, genome topology and gene regulatory networks. Finally, understanding the role of chromatin biology in health and disease will open the door for therapeutic intervention by targeting the epigenetic landscape of the genome.

## Functional non-coding RNAs and antisense transcription

### Pervasive transcription in the mammalian genome

Xist and H19, two of the best-studied non-coding RNAs (ncRNAs), were the first to be discovered. Especially Xist provided the first evidence for a role of ncRNAs in gene regulation (Brannan et al., 1990; Brockdorff et al.; Brown et al., 1992). Initially, ncRNAs were seen as an exotic exception to the rule of protein-coding genes. However, the arrival of novel technology that made genome-wide analysis of transcription and protein occupancy feasible, resulted in an exponential rate of discovery of novel ncRNAs. An effort to characterize the murine transcriptome revealed over 180000 transcripts, far exceeding the number of protein-coding genes (Carninci et al., 2005; Carninci et al., 2003), while oligonucleotide-array based transcript profiling (Bertone et al., 2004) and a pilot study by the ENCODE project (Consortium et al., 2007) showed that as much as 93% of the human genome might be transcribed. However, due to very moderate levels of conservation (Marques and Ponting, 2009; Ponjavic et al., 2007; Wang et al., 2004), unspliced variants, low expression levels (Cabili et al., 2011; Derrien et al., 2012) and the expected inherent infidelity of Pol2 transcriptional initiation (Struhl, 2007) most of this transcription was regarded as “transcriptional noise” and the extent and function of this pervasive transcription is still under debate (Clark et al., 2011; van Bakel et al., 2010). In addition, it is not entirely clear in how far presumed ncRNA might encode small peptides which would not be detected as ORFs due to their short size as shown for *tarsal-less* in *Drosophila* (Galindo et al., 2007; Kondo et al., 2010). Ribosome

profiling indicates that many long ncRNAs are bound to ribosomes (Ingolia et al., 2011), but not necessarily give rise to peptides (Guttman et al., 2013). Different strategies to verify bona fide ncRNAs (Derrien et al., 2012; Hangauer et al., 2013; Numata et al., 2003; Ravasi et al., 2006), including approaches exploiting chromatin signature of transcribed genes (Guttman et al., 2009; Khalil et al., 2009), have been employed subsequently, indicating that at least a portion of these transcripts might be regulated and functional. Divergent transcription at promoters showing co-regulation of ncRNAs and protein-coding RNAs on the other hand suggests that a substantial part might indeed be a mere byproduct of expression of protein-coding genes (Sigova et al., 2013).

### Functional non-coding RNAs

#### *Non-coding RNAs and the regulation of genomic imprinting*

The functional relevance of ncRNAs is highlighted by imprinted gene clusters in mammals. Imprinted genes display parent-specific expression which is controlled by a differentially methylated imprinting control region (ICR) (Peters, 2014). Interestingly, all imprinted gene clusters appear to harbor ncRNA genes and genetic studies have shed light on their function. While conditional deletion of *H19*, a maternally expressed, non-overlapping ncRNA gene in the *Insulin-Igf2* cluster, does not result in loss of imprinted silencing of *Igf2* (Schmidt et al., 1999), deletions of five other imprinted ncRNAs showed marked phenotypes indicating their involvement in silencing of genes in *cis*. A paternal truncation of *Airn*, removing over 95% of its transcript, showed loss of silencing of the entire gene cluster, including *Igf2r*, which overlaps in the opposite direction with



*Airn* (Sleutels et al., 2002). Similarly, a paternally inherited truncation of *Kcnq1ot1*, introduced by insertion of a premature polyA signal results in aberrant activation of the overlapping *Kcnq1* and other genes in this cluster in *cis* (Mancini-Dinardo et al., 2006; Thakur et al., 2004). Promoter deletion or premature termination of transcription of paternally inherited *Ube3a-ATS* gene, the non-coding antisense partner of *Ube3a*, also abrogated repression of *Ube3a* on this same chromosome (Meng et al., 2012). Also, a paternally transmitted *Nespas* hypomorph showing reduced transcription levels resulted in loss of DNA methylation and gain of H3K4 methylation at the *Nesp* promoter in *cis*, concomitant with de-repression of the overlapping gene *Nesp*. The same hypomorphic allele displayed low levels of ectopic *Nespas* activation when maternally transmitted, which correlated with reduced *Nesp* expression in *cis* (Williamson et al., 2011). In a less well characterized example, deletion of the ncRNA *Gtl2* in the *Dlk1-Gtl2-Dio3* locus causes a lethal phenotype in mice (Takahashi et al., 2009; Zhou et al., 2010). Taken together, these studies suggest that ncRNAs are essential for the proper regulation of a wide range of imprinted genes, particularly their direct antisense partners, during development and adult life.

### ***Small non-coding RNAs repress transcription transcriptionally and post-transcriptionally***

A large group of RNAs involved in gene regulation are small ncRNAs. First hints of a post-transcriptional gene silencing (PTGS) mechanism by RNAs came forth from transgene studies in plants suggesting “co-suppression” of an introduced transgene and a homologous endogenous gene (Napoli et al., 1990; van der Krol et al., 1990). These and further studies led

to the seminal discovery of sequence-specific, double-stranded RNA-dependent pathways that are ubiquitously used to degrade endo- and exogenous RNAs, pooled under the term RNA interference (RNAi) (Chang et al.; Fire et al., 1998; Grishok, 2013; Hammond et al., 2000; Hannon, 2002; Lee and Ambros, 2001; Malone et al., 2009; Voinnet, 2009; Wightman et al., 1993). Apart from involvement in direct degradation of RNAs, several classes of small ncRNAs have also been implicated in heterochromatin formation and transcriptional gene silencing (TGS). In the fission yeast *Schizosaccharomyces pombe*, components of the RNAi pathway are necessary for formation of pericentromeric heterochromatin (Hall et al., 2002; Volpe et al., 2002). Both strands of these genomic regions are transcribed giving rise to complementary siRNAs. These siRNAs supposedly guide a complex containing RNAi effector proteins, the RITS complex, to homologous stretches of DNA (Verdel et al., 2004), which then recruits heterochromatin-inducing factors (Buhler et al., 2006; Cam et al., 2005; Noma et al., 2004; Zhang et al., 2008). Pol2 subunits important for transcriptional elongation in heterochromatin and RDRC, a complex with an RNA-directed RNA polymerase activity, are responsible for a positive feedback loop which enhances the RNAi response (Kato et al., 2005; Motamedi et al., 2004). In addition to classical heterochromatin formation, RNA degradation by RNAi ribonucleases and the exosome has also been implicated in heterochromatic gene silencing. This effect is mediated by a specialized poly(A) polymerase that has been proposed to mark transcripts emerging from heterochromatin for degradation (Buhler et al., 2007). In plants, specialized RNA polymerases POLIV and POLV, RNAi components and RNA-directed RNA polymerases are part of chromatin modification pathways that methylate CpGs in transposons and re-

peats resulting in transcriptional repression. This process is called RNA-directed DNA methylation (RdDM) and shares many similarities with siRNA-dependent gene silencing in *S. pombe*, e.g. target recognition by siRNAs and positive feedback loops to enhance silencing (Bologna and Voinnet, 2014; Matzke and Mosher, 2014). Roles for TGS in metazoans are less well defined, but synthetic siRNAs can induce epigenetic silencing of endogenous loci under certain circumstances (Janowski et al., 2006; Kim et al., 2006; Morris et al., 2004), and disturbing Dicer or Argonaute protein levels results in defects in heterochromatin formation (Deshpande et al., 2005; Fukagawa et al., 2004). In addition, several proteins have been identified that are essential for nuclear RNAi, suggesting a role in TGS (Guang et al., 2010).

A special case of small ncRNA-directed TGS is the germ line-specific silencing of transposons by piRNAs in animals. During meiosis, transcription from piRNA clusters which are derived from defective transposons leads via self-amplifying cycles of piRNA production (“ping-pong” cycle) to incorporation of piRNAs into Argonaute/PIWI complexes and silencing of transposons (Aravin et al., 2007; Brennecke et al., 2007; Malone et al., 2009). Importantly, apart from degradation of transposon-derived transcripts, these complexes direct histone H3K9 methylation (Wang and Elgin, 2011), HP1 recruitment (Brower-Toland et al.) and DNA methylation (Watanabe et al., 2011) to the genomic loci of transposons. Several RNAi pathways thus directly confer epigenetic silencing of their target genes by sequence-specific recruitment of heterochromatin-inducing protein complexes.

## Emerging new classes of non-coding RNAs

Apart from the well-studied cases in XCI, genomic imprinting and RNAi many other examples and themes of ncRNAs and their function in gene regulation have emerged (Figure 2). A few of these cases are of special interest regarding their biology and will be briefly introduced. Detection of transcriptional activity at enhancers (De Santa et al., 2010; Kim et al., 2010) and knockdown studies of these transcripts (Orom et al., 2010) established enhancer-associated RNAs (eRNAs) as a new class of functional RNAs. Even though some evidence based on interaction studies and chromosome conformation capture technology suggests that eRNAs are directly transducing an activating effect by chromatin looping and association with Mediator (Lai et al., 2013; Li et al., 2013), the majority of this heterogeneous species of RNA appears to mostly act in *cis*. Moreover, the presumed activating effects of eRNAs might be conferred by the DNA elements they are produced from (reviewed in Darrow and Chadwick, 2013; Orom and Shiekhattar, 2013).

Another class of RNAs, termed competing endogenous RNAs (ceRNAs), functions as “sponges” (Ebert et al., 2007) that compete for a limited pool of miRNAs and therefore establish a paradigm for RNA crosstalk (Tay et al., 2014). Diverse types of RNAs are able to titrate out endogenous miRNAs, amongst others pseudogene-derived RNAs and circular RNAs (Hansen et al., 2013). The pseudogene *PTENP1*, for example, contains conserved seed matches for several *PTEN*-targeting miRNA families, thereby protecting the tumor suppressor gene *PTEN* from repression by miRNA-mediated degradation or translational repression (Poliseno et al., 2010).

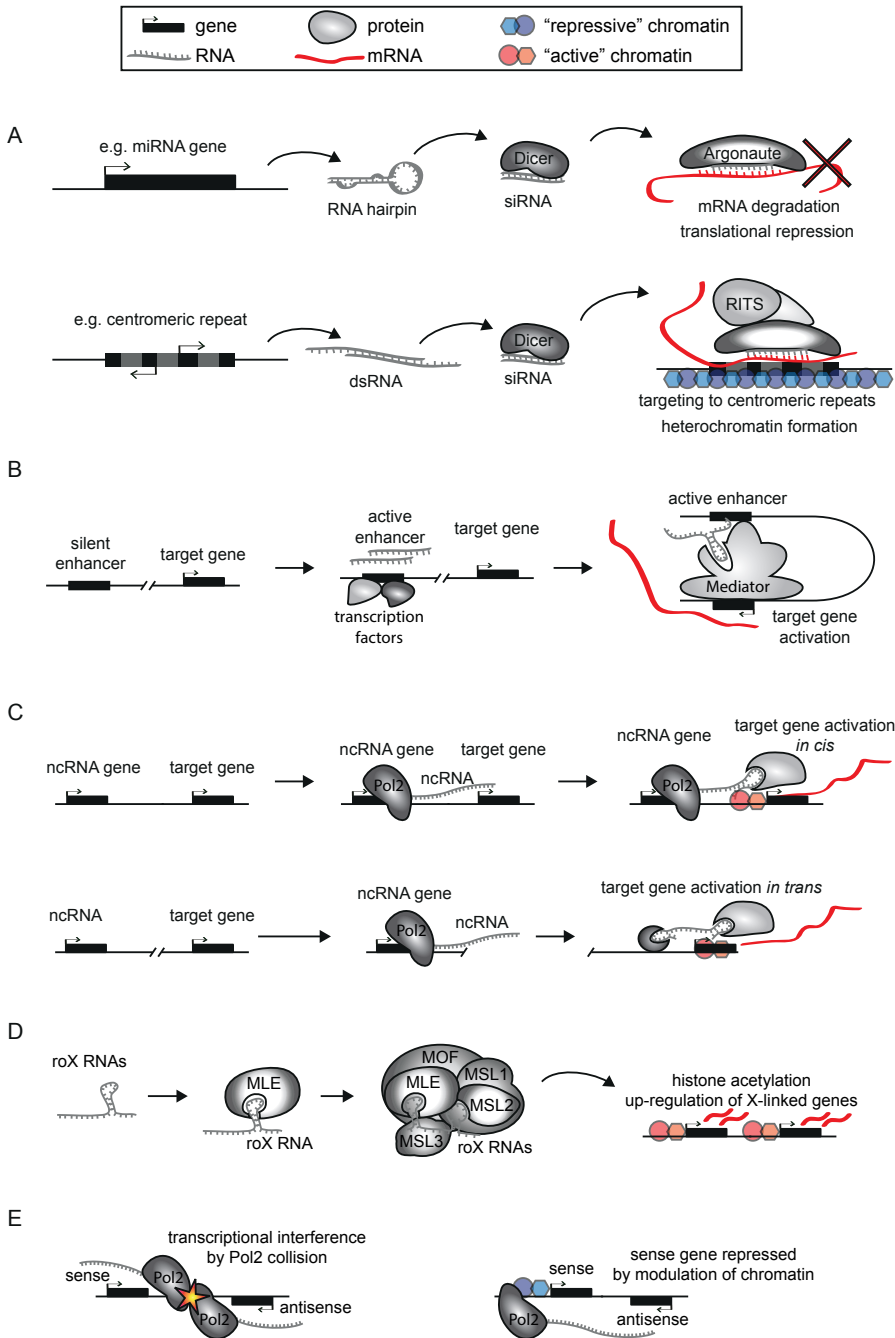
A ncRNA with a more direct role in

gene regulation is ANRIL which interacts with H3K27me3 and polycomb complexes. Disruption of this interaction leads to de-repression of the antisense transcribed tumor suppressor INK4a (Yap et al., 2010; Yu et al., 2008). Malat1 is a conserved and abundant long ncRNA that localizes to nuclear speckles and whose expression is affected in several types of cancer. These features made it an attractive candidate for a functional ncRNA and several studies indicated that Malat1 knockdown causes a variety of phenotypes on the cellular and molecular level. However, Malat1 knock-out mice are viable and phenotypes, if any at all, are mild and controversial (Gutschner et al., 2013). A different, complicated situation is presented at the *FLC* locus in *Arabidopsis thaliana*, which is responsible for control of flowering. COOLAIR is the antisense transcript of *FLC* and exists in two differentially spliced variants. Up-regulation of the shorter version which ends in the 3' end of sense *FLC* is correlated with loss of H3K4me2 in the body of *FLC* and reduction of sense and antisense transcription, while the longer splice variant, which is transcribed through the *FLC* promoter, is associated with activation of *FLC*. During vernalization, another sense ncRNA at the *FLC* locus, COLDAIR, is up-regulated after extended periods of cold. This ncRNA recruits PRC2 leading to epigenetic silencing of the entire locus (Ietswaart et al., 2012).

Finally, ncRNAs have been implicated in the regulation of *Hox* genes, which are essential for the anterior-posterior patterning of animals during embryonic development. Interestingly, the regulation of these genes by ncRNAs involves the very same PcG and Trithorax-group (TrxG) proteins that originally were discovered as repressors and activator of *Hox* genes, and subsequently became paradigmatic examples of chromatin modulators (reviewed in Krumlauf, 1994 ). Mistral in mouse

and HOTTIP in human are long ncRNAs that interact with transcriptional activators and were shown to recruit MLL1 to neighboring homeotic genes, thereby activating their expression (Bertani et al., 2011; Wang et al., 2011). Conversely, the long ncRNA HOTAIR appears to repress *Hox* genes. SiRNA-mediated depletion of this RNA, which is situated within *HOXC* locus on human chromosome 12 and interacts with PRC2, causes de-repression of a 40kb region of the *HOXD* locus on chromosome 2 in *trans* (Rinn et al., 2007). Ectopic *trans*-activation of part of this *HOXD* locus is accompanied by loss of H3K27me3 and PRC2 occupancy. In line with this result, over-expression of HOTAIR results in aberrant PRC2 localization at the *HOXD* locus (Gupta et al., 2010). Interestingly, HOTAIR appears to bind to many genomic sites on all chromosomes, but this recruitment is independent of the PRC2 component EZH2, suggesting an upstream function for HOTAIR in the recruitment of PcG proteins (Chu et al., 2011). In spite of these functional implications in human, *Hotair* knock-out mice develop normally and no effects on *Hoxd* gene expression were found (Schorderet and Duboule, 2011). This highlights the difficulties researchers face in distinguishing functional ncRNAs from ncRNAs produced as a byproduct of transcriptional activity of a specific region. The list of functional ncRNAs discussed herein is not exhaustive and new examples are discovered frequently. Nevertheless, it underscores the diversity and scope of ncRNAs in the regulation of genomes.

In contrast to the dissection of the specific cases of RNA-mediated gene regulation discussed above, large scale efforts in the wake of the discovery of widespread transcription and large numbers of ncRNAs were undertaken to verify and characterize their functionality on a more general level. RNA immunoprecip-



itation (Dekker et al., 2002) experiments showed that as much as 20% of large intergenic ncRNAs (lincRNAs) in human

cell lines might bind to PRC2 and equally high numbers have been reported for mouse ES cells (Khalil et al., 2009; Zhao

## ◀ Figure 2. Mechanisms of noncoding RNA function and antisense transcription

(A) Small ncRNAs like miRNAs transcribed from dedicated genes or siRNAs derived from centromeric repeats can modulate gene expression by post-transcriptional degradation of mRNAs or transcriptional silencing by modulation of chromatin structure. (B) Noncoding transcription from enhancers might induce looping events and target gene activation by tethering of Mediator complex to the target gene promoter. (C) Noncoding RNA can act in *cis* or in *trans*, e.g. by recruiting effector protein complexes like PRC2 to neighboring genes (in *cis*) or distant genes (in *trans*). The lower panel exemplifies the modular nature of ncRNA-protein interactions possibly involved in these processes: One domain of the ncRNA can bind to a tethering protein while another domain might bind a transcriptional activator. Effects can be activating (shown) or repressive. (D) Noncoding RNAs like roX might act as building platforms for protein complexes. Interaction of RNA and protein can induce conformational changes in both interactors, thereby revealing or creating additional interaction domains. (E) Effects of antisense transcription can be caused by transcriptional interference, which might include polymerase collision. Alternatively, antisense transcription through a sense gene can modulate chromatin structure resulting in expression modulation.

et al., 2010). Knockdown of more than 100 lincRNAs in murine ES cells resulted in genome-wide changes of gene expression and implicated lincRNAs in the maintenance of the pluripotent state of ES cells (Guttman et al., 2011). A similar strategy using siRNAs to knockdown long ncRNAs expressed in adipogenic cells showed varying degrees of effects on adipogenesis (Sun et al., 2013a). Attempts to find evolutionary conservation of ncRNAs suggest that the intron-exon and the secondary structure of many ncRNAs might be conserved (Guttman et al., 2010; Smith et al., 2013). Taken together these findings open the possibility that a larger portion of ncRNAs than previously anticipated might be functional. However, the exact mechanisms of function of many of the studied examples remain elusive and additional experimental validation is required.

### Models and mechanisms of non-coding RNA function in the nucleus

From the described ncRNAs and many additional examples common themes of their function have emerged and several models of their mechanistic action have been proposed. The diversity of modes of action of ncRNAs, however, still raises the question whether the well-studied examples are isolated cases

or follow more general mechanisms. In principle, the function of ncRNAs can be dependent on either the RNA molecule or the process of transcription per se. Moreover, it is important to distinguish between effects in *cis*, that is in the proximity of the ncRNA gene on the same allele, and effects in *trans*, which might target loci over long distances and on separate chromosomes.

### Non-coding RNAs interact with proteins

A popular theme of how ncRNAs potentially could exert their regulatory function is through binding of proteins and recruitment of epigenetic modifications to target regions, even though in most of these cases, specifically for the proposed *trans*-regulation, it is not clear at all how specificity could be achieved (Ptashne, 2013). In addition to the large scale RNA immunoprecipitation strategies mentioned above (Khalil et al., 2009; Zhao et al., 2010), several dedicated studies of well-described ncRNAs have shown their association with protein complexes. Prominent examples from imprinted clusters are Airn, which appears to be necessary for G9A recruitment to its target genes (Nagano et al., 2008), and Kcnq1ot1 which in a similar way interacts with G9A and PRC2 (Pandey et al., 2008). Moreover, during initiation of XCI, PRC2 recruitment to



the Xi is closely linked to the ncRNA Xist (Kaneko et al., 2010; Maenner et al., 2010; Plath et al., 2003; Silva et al., 2003; Zhao et al., 2008b). Amongst others, also ANRIL (Yap et al., 2010), HOTAIR (Rinn et al., 2007) and COLDAIR in plants (Heo and Sung, 2011) interact with PRC2 and are thought to recruit silencing activities to their target loci. Mistral and HOTTIP on the other hand associate with MLL and direct chromatin modifiers associated with activation to their target genes (Bertani et al., 2011; Wang et al., 2011). Corroborating the importance of protein binding to ncRNAs, studies on Xist have shown that different domains of the RNA confer different functions and that these functions are separable (Jeon and Lee, 2011; Wutz et al., 2002). Similarly, yeast telomerase binds different proteins with different domains (Zappulla and Cech, 2004) and HOTAIR tethers PRC2 and the histone H3K4 demethylase LSD1 by different interaction domains (Tsai et al., 2010). These findings have led to the attractive hypothesis that ncRNAs can act as modular scaffolds to bridge different protein complexes, RNA and DNA (reviewed in Guttman and Rinn, 2012). How exactly ncRNAs interact with proteins, other RNAs and DNA is not entirely clear. Base pairing between RNA-RNA, DNA-RNA and triple-helices are possibilities (Buske et al., 2011; Schmitz et al., 2010) and findings indicating secondary structure rather than sequence conservation in ncRNAs (Smith et al., 2013) suggest that the 2- or 3-D conformation of an RNA might be crucial for protein binding. Interestingly, *Drosophila* roX1 and roX2 ncRNAs undergo conformational changes upon stem-loop-mediated binding to the helicase MLE, which subsequently facilitates the formation of the full DCC complex, possibly by exposing additional protein binding modules on the RNA (Ilik et al., 2013; Maenner et al., 2013). However, a caveat of models implying protein and specifically chromatin factor binding

to ncRNAs is that many studies employ RIP-based assays that are very sensitive for unspecific binding of RNAs to RNA-binding proteins (Friedersdorf and Keene, 2014). It has also been shown that purified PRC2 components bind RNAs *in vitro* non-specifically (Davidovich et al., 2013).

In addition to directly recruiting chromatin modifying activities, ncRNAs are also involved in the organization of the nucleus. Again, this is best studied for the case of Xist which forms a very stable nuclear domain (Clemson et al., 1996) that is transcriptionally silent (Chaumeil et al., 2006) and interacts with the matrix protein hnRNP U/SAF-A (Hasegawa et al., 2010; Pullirsch et al., 2010) and a chromatin organizer, SATB1 (Agrelo et al., 2009). Similarly, Kcnq1ot1 appears to be restricted to a defined nuclear domain (Redrup et al., 2009). Even though functionally not well-defined, Neat1 and Malat1 bind to several proteins in paraspeckles and nuclear speckles, respectively (Clemson et al., 2009; Hutchinson et al., 2007; Murthy and Rangarajan, 2010), and Neat1 appears to be essential for paraspeckle formation (Mao et al., 2011). In general, different RNA species are able to induce the formation of and are highly enriched in sub-nuclear structures (Shevtsov and Dundr, 2011). Moreover, several eRNAs have been proposed to induce looping events between enhancers and promoters by binding to Mediator complex or cohesin (Lai et al., 2013; Li et al., 2013). However, distinguishing between inducing conformational changes and acting on existing conformations is challenging and has not been thoroughly addressed. In support of a role for ncRNAs in the induction of instructive topological changes of DNA also CTCF, a protein important for chromosomal organization and transcriptional insulation, has been shown to interact with ncRNAs. One study demonstrated that the ncRNA SRA binds CTCF together with p68 there-

by facilitating cohesin interaction and insulator function (Yao et al., 2010). Another study proposes that Jpx functions as a decoy and titrates CTCF away, with functional implications for gene expression at that locus (Sun et al., 2013b). A similar mode of titration of proteins could potentially be another mechanism for ncRNA function as exemplified by the Gas5 ncRNA and the glucocorticoid receptor (Kino et al., 2010), but the low expression levels of most ncRNAs render stoichiometric roles unlikely in many cases (Derrien et al., 2012). Along the same line as CTCF-SRA-p68-mediated cohesin binding, transcription factors can be bound and allosterically regulated by ncRNAs. The transcriptional repressor TLS binds to and represses its target gene *CCND1* only if associated with a ncRNA (Wang et al., 2008), and the interaction of the ncRNA NRON with several proteins including nuclear transport factors has been implicated in nuclear trafficking of the transcription factor NFAT (Willingham et al., 2005). As mentioned above, another notable example where ncRNAs act as an allosteric-type regulator are roX1 and roX2 in *Drosophila* which are essential for building the proper platform for DCC assembly, possibly by modifying protein structure and thus allowing a stepwise assembly process (Ilik et al., 2013; Maenner et al., 2013).

### ***Non-coding RNAs function in cis and in trans***

An important question regarding transcriptional regulation by ncRNAs is whether they exert their function in *cis* or in *trans*. Classical examples of ncRNAs in genomic imprinting and XCI suggest a prominent role in *cis*. In addition to Xist's silencing activity in *cis*, genetic studies deleting ncRNAs located in the Xic (Anguera et al., 2011; Chureau et al., 2011; Ogawa

and Lee, 2003; Tian et al., 2010) show that they have an activating effect on neighboring genes within the same TAD (Nora et al., 2012). This hints at a co-regulatory mechanism in *cis* delineated by the topological conformation of DNA reminiscent of the association of co-regulated genes in erythroid cells (Schoenfelder et al., 2010). A similar mechanism could be at work at enhancers as knockdown of eRNAs resulted in concomitant repression of neighboring genes in *cis* (Orom et al., 2010) or loss of chromosomal interactions between enhancers and promoters (Lai et al., 2013; Li et al., 2013). Another ncRNA activates the human growth hormone enhancer HS1 from which it is also transcribed. This activation is solely dependent on the strength of transcription and can be emulated by an unrelated RNA (Yoo et al., 2012). Moreover, many ncRNA loci show synteny without sequence conservation indicating that genomic location and transcription might be more important than the product of transcription (Ulitsky et al., 2011). This is also exemplified by a process called germline transcription at the antigen receptor genes and intergenic transcription at the globin locus. Noncoding transcription across the antigen receptor genes is essential for opening up chromatin which in turn allows targeting of V(D)J recombination machinery (Abarrategui and Krangel, 2006; reviewed in Schatz and Ji, 2011). Similarly, intergenic noncoding transcription at the globin locus demarcates active regions, and remodeling of chromatin structure by transcription is necessary for globin gene activation (Gribnau et al., 2000). Taken together, these studies hint at a mechanism of *cis*-activation involving mass action, activator recruitment or remodeling of chromatin structure. In the case of imprinted ncRNAs like Airn, Kcnq1ot1, Nespas and others repressive effects on gene activity are strictly in *cis*, as well (Sleutels et al., 2002; Thakur et al., 2004; Williamson et al., 2011), and appear

to be based on both the act of transcription alone to silence the direct antisense partners, e.g. by transcriptional interference (Latos et al., 2012), and recruitment of chromatin modifiers to more distant genes (Nagano et al., 2008; Pandey et al., 2008). The same mechanism of gene regulation in *cis* was observed for other ncRNAs like ANRIL, Six3Os, Mistral and HOTTIP (Bertani et al., 2011; Rapicavoli et al., 2011; Wang et al., 2011; Yap et al., 2010) and might also explain low expression levels of ncRNAs because tethering and recruitment of chromatin factors in *cis* by ncRNAs does not necessarily need high numbers of mature transcripts.

In contrast to *cis*-acting ncRNAs, HOTAIR was the first long ncRNA shown to be involved in chromatin modulation in *trans* (Rinn et al., 2007). Depletion of HOTAIR located on chromosome 12 resulted in loss of PRC2 and H3K27me<sub>3</sub>, and de-repression of *HOXD* genes on chromosome 2. Surprisingly, the mouse homologue Hotair does not seem to have such a function in *trans* (Schorderet and Duboule, 2011). In support of the discovery of a *trans*-acting ncRNA, novel techniques developed to map genomic binding sites of RNAs (Chu et al., 2011; Engreitz et al., 2013; Mariner et al., 2008; Simon et al., 2011) suggest that long ncRNAs bind to many sites in the genome. Particularly the “proximity transfer” model proposed by Engreitz et al. is appealing, because its proposed direct transfer of a ncRNA to a distal site by spatial proximity due to chromosomal conformation constitutes a mixture of *cis*-like and *trans*-like mechanisms and could reconcile *trans*-active ncRNA function with their observed low expression levels (Engreitz et al., 2013). However, though highly interesting, these new mapping methods cannot distinguish between direct and indirect interactions and suffer from several technical limitations like cross-linking of genomic regions that

are close by in the 3-D space of the nucleus or unspecific binding of oligonucleotides used for purification of DNA-bound RNA. In addition, the binding of a molecule to DNA per se does not necessarily need to be functional. For example, an eRNA from a *FOXC1* enhancer binds to several genes all across the genome with no functional consequences (Li et al., 2013). Even transcription factor binding only affects expression levels in a small subset of bound genes (Cusanovich et al., 2014), indicating that binding and functional consequence can be uncoupled. The care one must take when correlating expression data with binding profiles is yet again highlighted by another set of studies. The de-repression of PRC2 bound genes in PRC2 knockout ES cells has been interpreted as a *de novo* silencing activity for PRC2 (Boyer et al., 2006). However, a recent study investigating the role of PRC2 binding in initiation of silencing on a genome-wide level at high temporal resolution showed that PRC2 is only recruited to target genes after these genes are repressed (Riising et al., 2014). Therefore, as shown for *Drosophila* homeotic genes (Struhl and Akam, 1985), PcG proteins appear to be important for maintenance and memory of gene repression rather than actually initiating repression. In a more systematic loss-of-function study in mouse ES cells depletion of ca. 150 ES cell specific ncRNAs resulted in major changes of expression patterns across the entire genome (Guttman et al., 2011). Since only very modest effects on genes directly neighboring the depleted ncRNAs were detected, the authors proposed that most ncRNAs have a function in *trans*. Other genome-wide studies have also failed to find extensive correlation of expression levels between ncRNAs and their closest neighbors (Cabili et al., 2011; Derrien et al., 2012) supporting a predominant role in *trans*. There are several additional examples of ncRNAs, amongst others Paupar, Braveheart and Firre, which



by using knockdown and pulldown strategies have been implicated in regulating genes in *trans* (Hacisuleyman et al., 2014; Klattenhoff et al., 2013; Vance et al., 2014). However, on top of the issues concerning RNA pull-down experiments mentioned above, there are several other drawbacks in RNAi-mediated loss-of-function studies. This is mostly due to off-target effects and insufficient knowledge about miRNA binding specificities (reviewed in Hausser and Zavolan, 2014) and is exemplified by two examples in which *cis* and *trans* functions are contested. *Jpx* was initially shown to activate *Xist* in *trans* (Tian et al., 2010), but this effect could not be reproduced in another laboratory (Barakat et al., 2014). In a second study *lincRNA-p21* was identified as a p53-induced repressor of several genes in *trans* (Huarte et al., 2010), while a conditional knockout mouse model shows that *lincRNA-p21* acts mainly as a co-activator in *cis* (Dimitrova et al., 2014). Taken together, it appears that ncRNAs are capable of functioning in *cis* and in *trans*, but while several models of how they might work have been proposed, additional functional studies are necessary to define their modes of action in detail.

### The role of antisense transcription in gene regulation

A common theme in many of the discussed examples is antisense-transcriptional overlap with a sense partner, which prompted more comprehensive studies on a genome-wide scale. Analysis of cDNA libraries showed widespread antisense transcription in the mammalian genome in 5% to 30% of known protein coding genes (Chen et al., 2004; Katayama et al., 2005; Kiyosawa et al., 2003), and subsequent studies using novel technologies confirmed these findings in basically all metazoans studied (He, 2008;

Sigova, 2013; Pelechano, 2013; reviewed in Numata and Kiyosawa, 2012). However, no link between antisense transcription and organismal complexity was detected (Sun et al., 2006b). Interestingly, many of the described antisense transcripts are non-overlapping and arise from divergent transcription at promoters (Core et al., 2008; Seila et al., 2008). Even though abundant divergent co-expression at promoters of protein-coding genes (Sigova et al., 2013) does not support a direct regulatory role for antisense transcription in many cases, in-depth analysis of datasets revealed features of antisense transcription that do not fit with a simple transcriptional noise model (Struhl, 2007). Variation among cell lines and a non-random distribution of antisense transcripts mainly from protein-coding genes argue against entirely promiscuous transcription (He et al., 2008). This can be partially attributed to epigenetic differences between cell lines and a general increase of transcription from cryptic promoters in active genomic regions, but the finding that expression of sense and antisense transcripts can be either positively or inversely correlated suggests the presence of some degree of active regulation (Chen et al., 2005a; He et al., 2008). Positive correlation, moreover, could be a corollary of population-based studies because a subset of cells might express the sense and another subset the antisense transcript (Pelechano and Steinmetz, 2013). In addition, evolutionary conservation of antisense genes could indicate functional importance (Chen et al., 2005b; Dahary et al., 2005). Support for a function of antisense genes comes from studies in prokaryotes and budding yeast. In these systems antisense transcription is important at two levels. First, for a bimodal on-off gene activation switch, because sense transcription would first need to overcome a threshold set by antisense transcription, and second for the dampening of basal leakiness in initiation of sense

transcription (Duhring et al., 2006; Lege-  
wie et al., 2008; Xu et al., 2011b). A major  
mechanistic question is whether the prod-  
uct or the process of antisense transcrip-  
tion per se is exerting a function. Exam-  
ples for long ncRNAs which bind proteins  
(reviewed in Pelechano and Steinmetz,  
2013) are outlined above. An attractive hy-  
pothesis is that double-stranded RNA de-  
rived from antisense transcription could  
feed into RNAi-like pathways which target  
chromatin factors to homologous stretch-  
es of DNA. Despite initial indications that  
this might be the case for XCI (Ogawa  
et al., 2008), silencing in XCI (Nestero-  
va et al., 2008), and silencing at the *AN-  
RIL-INK4a-INK4b* locus (Yu et al., 2008)  
and in the *Kcnq1ot1* cluster (Redrup et  
al., 2009) occurs independent of Dicer.

Transcriptional interference is an-  
other possible mechanism for repression  
of overlapping genes in *cis* (reviewed in  
Shearwin et al., 2005). In this case the  
RNA molecule is not needed for function.  
Many studies in prokaryotes and lower  
eukaryotes give insight into the regulation  
and mechanisms of transcriptional inter-  
ference while experimental data is scarce  
for higher eukaryotes. Exploiting report-  
er constructs in various arrangements  
it has been demonstrated in *S.cerevisiae*  
that convergent transcription units allow  
initiation of transcription but overlap-  
ping transcription results in a reduction  
of mRNA levels (Prescott and Proudfoot,  
2002). Similarly, transcriptional repres-  
sion of two transcription units in mam-  
malian cells is most severe if they converge  
(Eszterhas et al., 2002). In a prokaryotic  
system, passage of RNAP through the pro-  
moter of a distal gene reduced expression  
from this promoter approximately 5- to  
6-fold (Callen et al., 2004). It was proposed  
that one transcribing RNA polymerase  
complex could displace another one by  
collision as shown by atomic force micros-  
copy for bacterial polymerases (Crampton

et al., 2006) and by *in vitro* transcription  
assays in yeast (Hobson et al., 2012). This  
model implies high levels of transcribing  
polymerase complexes, but RNA polymer-  
ase pausing increases the effect also in low  
expression systems (Palmer et al., 2009).

In addition to direct transcription-  
al interference, the transcribing polymer-  
ase complex has been shown to be able to  
modulate gene expression by virtue of its  
association with chromatin factors. Also  
in this scenario, the RNA molecule per se  
is dispensable for function. Most studies  
have been conducted in *S. cerevisiae*, but  
some data is also available for mammalian  
systems. In budding yeast, antisense tran-  
scription induces H3K36me3 in the body  
of sense genes because the histone meth-  
yltransferase Set2 travels with the elongat-  
ing Pol2 complex (Li, 2002). Subsequent  
recruitment of the histone deacetylase  
complex Rpd3 to H3K36me3 results in  
transcriptional repression of cryptic tran-  
scription from within the gene body and  
possibly silencing of overlapping sense  
promoters (Carrozza et al., 2005; House-  
ley et al., 2008). Independent of the Set2/  
Rpd3 pathway, but in a similar fashion,  
the histone deacetylase Set3 binds to  
H3K4me2 (Kim and Buratowski, 2009).  
This histone modification is put down by  
polymerase associated Set1 and is sup-  
posed to repress transcription both from  
cryptic initiation sites in gene bodies and  
from *bona fide* promoters of overlapping  
genes, e.g. by histone deacetylation (Kim  
et al., 2012; Margaritis et al., 2012; Pin-  
skaya et al., 2009). Interestingly, a similar  
mechanism might be very common in the  
yeast genome as hundreds of antisense  
transcripts, called XUTs, are involved in  
Set1-dependent silencing of their sense  
counterparts (van Dijk et al., 2011). Stud-  
ies in mouse and human ES cells indicate  
that KDM5B, a H3K4me3 demethylase,  
is recruited by H3K36me3 and suppress-  
es intragenic cryptic transcription (Xie

et al., 2011), highlighting parallels to the pathways in budding yeast. Moreover, a heritable form of anemia was shown to be caused by a deletion juxtaposing *LUC7L* to the  $\alpha$ -globin gene *HBA2* resulting in antisense transcription through and silencing of *HBA2*. Using a transgenic mouse model and differentiating ES cells it was demonstrated that, as a consequence of antisense transcription, the CpG island of *HBA2* becomes fully methylated during early development (Barbour et al., 2000; Tufarelli et al., 2003). Further support for a role for transcription per se in silencing an overlapping antisense gene comes from a series of elegant genetic experiments focusing on the imprinted *Igf2r* cluster. Truncations and promoter shifting of endogenous *Airn* indicated that only the transcriptional overlap with the *Igf2r* promoter was necessary for silencing and DNA methylation of *Igf2r* (Latos et al., 2012). Truncation of *Airn* before the *Igf2r* promoter did not result in reduction of *Igf2r* levels, suggesting that interference mostly occurred at the level of transcription initiation. The authors favored a model of strict transcriptional interference, because they did not detect any hallmarks of repressive chromatin during initial silencing at the *Igf2r* promoter, while H3K4me3 was still present (Latos et al., 2012). However, it is possible that a specific chromatin signature was initially present at the *Igf2r* promoter, since only a small subset of histone modifications was tested. Subsequent studies using an inducible system during differentiation of ES cells showed that antisense transcription was necessary to maintain silencing until the *Igf2r* promoter acquired DNA methylation (Santoro et al., 2013). In contrast to Xist's ability to silence the X chromosome, which is confined to a narrow developmental time window (Wutz and Jaenisch, 2000), *Airn* transcription was capable of silencing *Igf2r* during the entire time frame tested (Santoro et al., 2013). Given the abundance

of ncRNAs and antisense or overlapping transcription, these studies provide an appealing two-step mechanism for gene silencing, particularly in a developmental context. First, a sense gene is suppressed by antisense mediated transcriptional interference or chromatin remodeling, and subsequently this repressed state is locked in by DNA methylation. Taken together, the examples from prokaryotes, yeast and mammalian systems presented above, regardless of whether transcriptional interference is direct or mediated by polymerase-associated factors, emphasize that antisense transcription has a prominent impact on gene regulation.

### Long noncoding RNAs and antisense transcription in disease

The implications of noncoding RNAs and antisense transcription in gene regulation inevitably lead to their involvement in human disease. The list of ncRNAs that have been reported to be dysregulated in diseased states is ever-growing. Most notably, aberrant expression of ncRNAs was found in many types of cancer and ncRNA expression patterns are increasingly used as prognostic markers for disease outcome or therapy (reviewed in Gibb et al., 2011; Farazi et al., 2013). A prominent example is *MALAT1*, whose over-expression is found in a variety of tumors (reviewed in Gutschner et al., 2013). In most cases the causal relationship between ncRNA expression and pathogenesis is not clear, and it seems likely that most of these findings are of a merely correlative nature. However, ncRNA function in gene regulation was shown to be, amongst others, responsible for silencing of tumor suppressors, as in the case of *ANRIL* and *PTENP1*, by loss of targeting of PRC2 (Yap et al., 2010) and loss of a miRNA decoy (Poliseno et al., 2010), respectively. Sim-

ilar models are attractive for any kind of protein-binding RNA like HOTAIR, that theoretically could target repressive or activating complexes to certain genomic regions. A very strong example is the targeted deletion of *Xist* in hematopoietic stem cells in a mouse model resulting in partial reactivation of the *Xi* and the development of hematologic cancers with 100% penetrance (Yildirim et al., 2013). A role for the RNA molecule itself in the pathogenesis has been shown for neurologic disorders caused by the expansion of trinucleotide repeats resulting in the production of toxic RNAs. Interestingly, several of the genes responsible for these disorders also display antisense transcription (reviewed in Batra et al., 2010). A highly informative case which offers an attractive mechanistic explanation of trinucleotide repeat mediated toxicity has been recently reported for fragile X syndrome (FXS). FXS is caused by the expansion of a CGG repeat in the 5' UTR of the *FMR1* gene whose loss of function results in the disease phenotype (reviewed in Tabolacci et al., 2013). An antisense transcript has been proposed to be involved in the pathogenesis of FXS (Ladd et al., 2007). Using a small molecule that specifically stabilizes CGG hairpin structures and RNA mapping technology it was shown in differentiating human ES cells that the CGG repeat of the *FMR1* mRNA directly interacts with the CGG repeat of the *FMR1* locus (Colak et al., 2014). Uncharacterized downstream events then resulted in epigenetic silencing of *FMR1* during a defined window of time in neuronal differentiation. Another interesting example in which the process of transcription through a promoter or the RNA itself is responsible for the pathomechanism comes from an  $\alpha$ -thalassemia patient. The juxtaposition of *LUCL7* to *HBA2*, resulting in antisense transcription through *HBA2*, leads to silencing of *HBA2* and DNA methylation of its associated CpG island by an unknown mechanism (Tu-

farelli et al., 2003). Also here, DNA methylation and silencing are supposed to occur during a specific developmental time window. These examples highlight the close relationship between transcription, ncRNAs and gene regulation with human disease in a developmental context.

## Conclusions

The list of ncRNAs involved in gene regulation is ever growing and a substantial amount of evidence suggests that some of their regulatory functions are mediated via the chromatin template. However, the exact mechanisms of specifically targeting transcriptional regulation by ncRNAs, their interaction with proteins, their general prevalence and their role in development and disease remain highly active fields of research that will certainly yield exciting insights in the coming years.

## X Chromosome Inactivation in Stem Cells and Development

### Sex-chromosomes and gene dosage differences

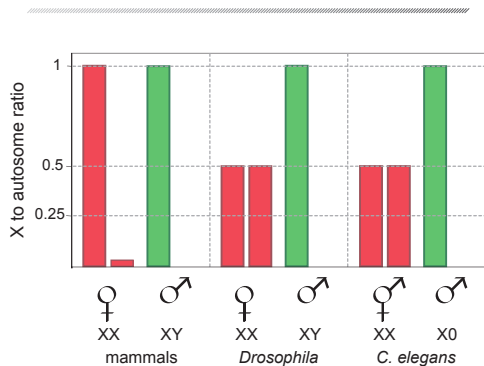
The origin of sexual reproduction with all its advantageous effects on the fitness of offspring was accompanied by the evolution of complex sex determination systems. Sex determination can be regulated by environmental cues such as temperature, however, these systems are highly sensitive to changes in 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. The evolution of heteromorphic sex chromosomes, as is found in a plethora of organisms, 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 haplo-insufficiency of sex chromosome-linked genes (Figure 3; reviewed in Disteche, 2012; Livernois et al., 2012). 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*, prompted the evolution of X chromosome inactivation (XCI). The fact that in humans only very few trisomies and no monosomies are viable, and aberrant gene regulation resulting in altered gene dosage and cancer clearly illustrate the potentially devastating effects of improper gene dosage (reviewed in Disteche, 2012).

### 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 X0 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



**Figure 3. Dosage compensation strategies in different organisms**

In mammals, the loss of Y chromosomal 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 over-expressed 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 males.



up-regulation, in XX hermaphrodites this up-regulation would be detrimental, and is counter-acted 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 (reviewed in 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, expression of genes located on 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 (reviewed in Conrad and Akhtar, 2011; Gelbart and Kuroda, 2009; Straub and Becker, 2011). 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 question. SXL is the key repressor of DCC, and DCC accumulates only in male cells where SXL is absent. 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 2) 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 2 activity at X-linked promoters and concluded that transcription initiation is the most important determinant of up-regulation of X-linked genes in male cells (Conrad et al., 2012). In *Drosophila*, the DCC is initially recruited to the X by so-called high affinity sites (HAS), which are approximately 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., 2006), 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 initiating testes development, while the W-linked genes *FET1* and *ASW* are necessary for female development (Nanda et al., 2002; Nanda et al., 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., 2013).

### **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 X0 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 Disteché, 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; Pessia et al., 2012).

XCI has mainly been studied in mice, which show two forms of XCI (Figure 4). 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 (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), 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 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 random XCI (rXCI) process is clearly visible in the calico cat, where the brown and black fur color are determined by different alleles of 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 mouse, 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 of XCI 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 different 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 blastocyst, and capture the moment just after reactivation of the paternal X chromosome, containing two active X chromosomes, one of which is in-

activated 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 inactivation hypothesis had been established two key concepts of XCI were formulated. The first 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 XXXX 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. 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

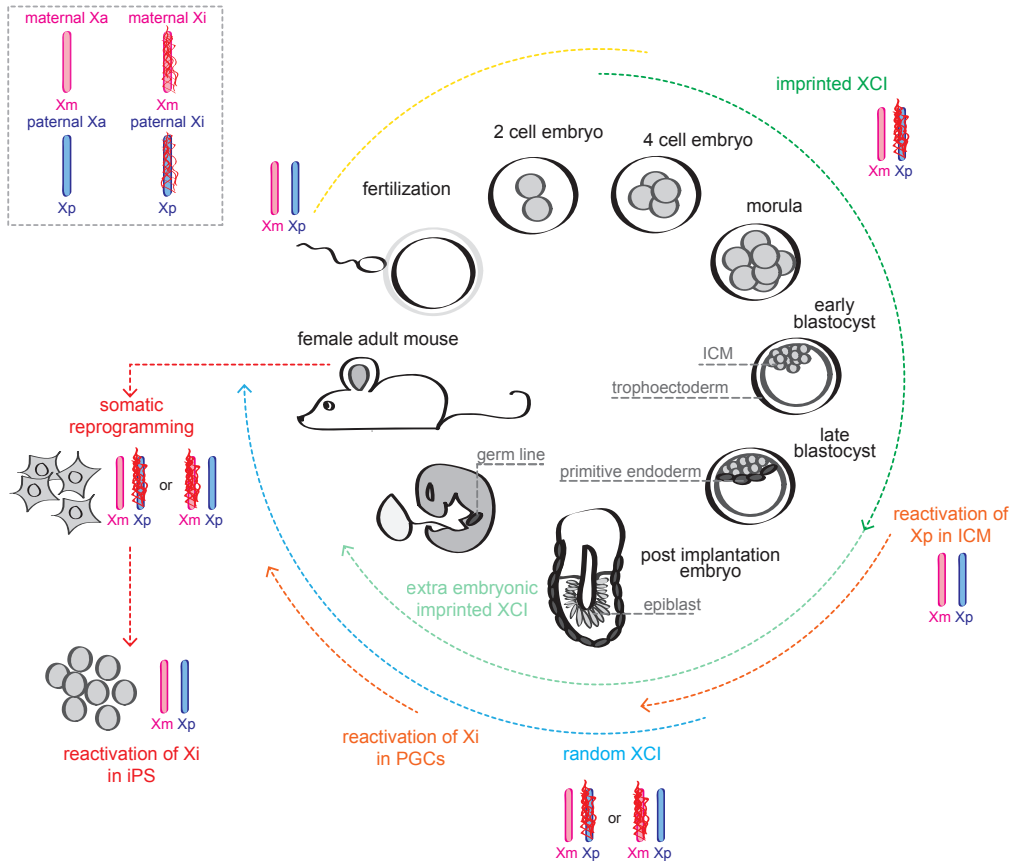


variegation in an X-to-autosome translocation model (Cattanach and Isaacson, 1967). Crossings of inbred mouse strains showed that X chromosomes from certain strains are more resistant to XCI than X chromosomes from other strains (Cattanach et al., 1972). This locus, which was mapped to a 1.85 Mb region on the X chromosome (Chadwick et al., 2006), is called X chromosome controlling element (*Xce*). 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), however, 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 mice with X:auto-some translocations and murine embryonic stem cells with truncated X chromosomes, a 8cM (10-20Mb) region on the X chromosome, delineated by the T16H and the HD3 breakpoints, was shown to harbor the *Xic* (Rastan and Robertson, 1985). 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).

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 15-17kb 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., 2001), 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., 1996a). 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.

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



**Figure 4. The cycle of life and XCI**

In female mouse embryos, both X chromosomes are active after zygotic genome activation. 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 cell reprogramming.

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., 1999a). Heterozygous 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 and Lu, 1999). In addition, *Tsix* has also been implicated in iXCI (Lee, 2000; Sado et al., 2001). *TSIX*, the human homologue, appears to have a

similar expression pattern as mouse *Tsix* as it is only expressed in cells of embryonic origin, 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 in mouse.

How *Tsix* represses *Xist* remains an open question, and 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., 2008; Sado et al., 2005; Sun et al., 2006a). 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 5). *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., 2013b; Tian et al., 2010). Interestingly, chromatin conformation capture studies examining the higher order chromatin

structure of the *Xic* indicate 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* (Dixon et al., 2012; Nora et al., 2012; Spencer et al., 2011). A recent polymer model for the *Xist* and *Tsix* TADs, based on deconvolution of population based 5C data into a range of chromatin conformations present in single cells, extended these findings (Giorgetti et al., 2014). Modeling together with FISH analysis of wild type and mutant cells showed the presence of several conformations and provided a direct correlation of transcription and TAD conformation in single cells. Moreover, it was demonstrated that certain areas coinciding with CTCF binding sites act as master regulators of TAD structure. Taken together, these findings suggest a co-activation 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.

### Models for XCI Initiation

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. 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 heterozygous *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 (Gribnau et al., 2005; Lee, 2005; Lee et al., 1999b; Marahrens et al., 1998; Monkhorst et al., 2008). 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., 2013b). So far there is no evidence for the presence of a counting element, through which the CF and BF would exert their 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 are not yet clear. Also a 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 this X-pairing process 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.

### ***Trans-acting Factors in XCI***

Female specific initiation of XCI involves the tight orchestration of expression of XCI-activators and -inhibitors during embryonic development. 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 providing a beautiful link between

loss of pluripotency and initiation of XCI (Figure 5) (Donohoe et al., 2009; Ma et al., 2011; Navarro et al., 2008; Navarro et al., 2010). In addition, using ES cells with different chromosomal compositions and *Xist* deletion or induction systems, it has been demonstrated that a twofold dosage of X-linked genes, as present in undifferentiated female ES cells, blocks exit from the pluripotent state via MAPK and Gsk3 $\beta$  inhibition, which yet again emphasizes the close relationship between pluripotency and dosage compensation (Schulz et al., 2014). 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 also have been reported to 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).

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 *Rnf12* over-expression 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). RNF12 indirectly activates *Xist* by targeting the XCI-inhibitor REX1 for proteasomal degradation (Gontan et al., 2012). RNF12 is X-encoded, and dose-dependent degradation of autosomally encoded REX1 ensures female specific initiation of XCI (Gontan et al., 2012). Female mice heterozygous for a null *Rnf12* mutation are not viable when the mutated allele is maternally transmitted showing impaired iXCI (Shin et al., 2010). These findings have been extended by a recent study using a Sox2-Cre mediated conditional deletion of *Rnf12* in the ICM (Shin et al., 2014). Female *Rnf12*<sup>-/-</sup> mice are born and show hallmarks of XCI, suggesting that the main function of RNF12 is the regulation of iXCI. However, a different system revealed rXCI-related phenotypes of *Rnf12* deletions *in vivo* (Barakat, Mira, Dupont, unpublished). This finding calls the specificity of the conditional knockout and the non-allelic nature of analysis into question, but might be explained by differences in the genetic background of the mice studied, as pluripotency factor expression levels vary between mouse strains (Battlle-Morera et al., 2008). In addition other XCI-activators might be in play. Indeed, *Rnf12*<sup>+/-</sup> ES cells show reduced initiation of XCI upon differentiation, but XCI is not completely abolished, suggesting the existence of one or more additional X-encoded XCI activators (Barakat et al., 2011). Interestingly, female *Rnf12*<sup>-/-</sup> ES cells and *Xist*<sup>+/-</sup>*Rnf12*<sup>+/-</sup>-*cis* heterozygous ES cells show a severe XCI phenotype, indicating



that RNF12 acts at two levels (Barakat et al., 2011; Barakat et al., 2014). A twofold dosage of RNF12 ensures female specific initiation of XCI, whereas requirement for one active copy of *Rnf12* provides a robust feedback mechanism preventing XCI of one X too many. *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 RNAs, 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., 2013b). 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 Silencing of the X Chromosome**

The discoveries that *Xist* is essential for XCI (Marahrens et al., 1997; Penny et al., 1996b) and that *Xist* RNA “paints” the entire X chromosome from which it is transcribed (Clemson et al., 1996) spawned two inter-twined key questions which still remain partially 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?

Efforts to characterize the chromatin landscape of the Xi yielded a long list of chromatin features 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, stay 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 on the Xi following *Xist* spreading is the exclusion of most hallmarks of active transcription such as RNA Polymerase II and general transcription factors (Chaumeil et al., 2006). A subset of “active” histone modifications (Chaumeil et al., 2002; Heard et al., 2001; Jeppesen and Turner, 1993) is excluded from the Xi, as well. Later on, the Xi acquires many chromatin marks that are associated with silent chromatin (Chadwick and Willard, 2003; Heard et al., 2001), most notably H3K27me3 and H2AK119u1 whose deposition is catalyzed by the polycomb repressive complexes PRC2 and PRC1, respectively (de Napoles et al., 2004; Mak et al., 2002; Plath et al., 2003; Silva et al., 2003; Wang et al., 2001). 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. Finally, CpG methylation in promoters of genes and other regulatory sequences, such as CpG islands, on the Xi is acquired at a late stage during development (Lock et al., 1987) and necessary to maintain the inactive state (Sado et al., 2000), suggesting that DNA methylation terminally locks silencing in place. 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, which are involved in *C. elegans* dosage compensation. In a similar fashion, knockdown of

ATF7IP results in reactivation of X-linked silenced transgenes, which suggests that this protein is important for bridging DNA methylation and H3K9me3 via its binding partners MBD1 and SETDB1 (Minkovsky et al., 2014). It has to be emphasized, however, that despite the extensive knowledge of chromatin signatures present on the Xi, the causal relationship between chromatin marks and silencing remains elusive, possibly due to the highly locus- and tissue-specific nature of the effects of these epigenetic features.

A first link between Xist localization and recruitment of PRC1 and PRC2 was found when their co-localization was observed on Xi metaphase chromosomes (Mak et al., 2002; Silva et al., 2003). In addition, studies using truncated Xist transgenes showed that different domains of Xist confer different functions. 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). The matrix attachment protein hnRNP U/SAF-A is enriched on the Xi (Pullirsch et al., 2010), binds to Xist via repeat C and is necessary for correct Xist localization (Hasegawa et al., 2010). Repeat F on Xist DNA has been proposed to function as a “nucleation center” for Xist RNA binding (Jeon and Lee, 2011). Moreover, biochemical analyses suggested that repeat A directly recruits PRC2 (Kaneko et al., 2010; Maenner et al., 2010; Zhao et al., 2008a). It should be noted, however, that Xist RNA deficient for this conserved domain is still able to induce PRC2 and H3K27me3 enrichment on the Xi and that Xist expression at later stages of development does not lead to H3K27me3 enrichment (Kohlmaier et al., 2004). Most likely Xist RNA is thus able to recruit PRC2 by redundant mechanisms. In support of such a redun-

dant process to generate and maintain the Xi, a recent study demonstrated that JARID2 associates with the Xi via Xist repeats B and F and subsequently recruits PRC2 (da Rocha et al., 2014). In addition, these redundant mechanisms might include developmentally regulated accessory factors, because the developmental context in general 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 murine ES (mES) cells and during early differentiation of ES cells (Wutz and Jaenisch, 2000), and that Xist-dependent induction of H3K27me3 follows the same pattern (Kohlmaier et al., 2004). So far, only one factor, which enables Xist-mediated silencing in a developmental context, has been identified. Using an elegant screen, SATB1 has been shown to be necessary and sufficient for Xist-mediated XCI in lymphoma cells (Agrelo et al., 2009). SATB1 is expressed early during mES cell differentiation and might help in relaying chromatin changes to actual silencing. Nevertheless, *Satb1*<sup>-/-</sup> female mice are born (Alvarez et al., 2000), again highlighting redundancies in the system. Similarly, ectopic Xist expression leads to reversible silencing in undifferentiated mES 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 cause reactivation of the Xi. Thus, maintenance of XCI is developmentally regulated, as well, and appears to be independent of Xist RNA. Interestingly, an exception from Xist-independent maintenance of iXCI is found in the extra-embryonic lineages (Ohhata et al., 2011), yet again emphasizing the plasticity and heterogeneity of XCI.

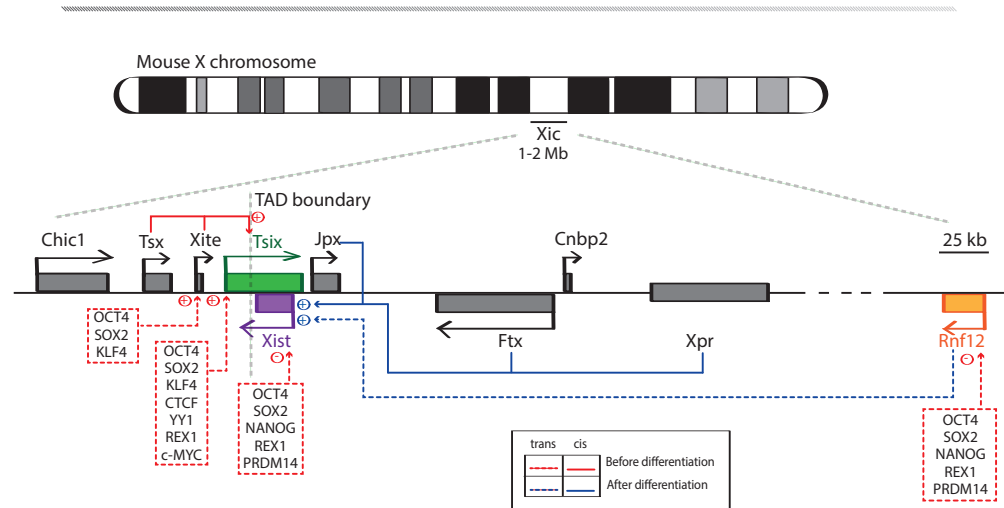
The observation that the Xi is frequently found in the nuclear periphery (Rego et al., 2008) and in close proximity



1  
ty 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 maintenance of silencing. The Xi seems to form a specialized compartment, evidenced by its interactions with the matrix attachment protein hnRNP U/SAF-A (Hasegawa et al., 2010; Pullirsch et al., 2010) and the chromosome organizing SATB1 protein (Agrelo et al., 2009), and, additionally, from the fact that removal of DNA does not affect the Xi associated nuclear matrix and the Xist RNA domain (Clemson et al., 1996). Xist has thus been proposed to have a structural role in stabilizing the Xi domain. Indeed, Xist is capable of creating a transcriptionally silent domain in the nucleus into which genes are displaced upon silencing (Chaumeil et al., 2006), even though inactive genes can be found outside and active genes inside the Xist domain (Calabrese et al., 2012), raising the question of cause and consequence and showing that the posi-

tion relative to the Xist territory is not an absolute determinant of activity or repression. The core of the Xist coated domain of X chromosomal sequences is comprised of repetitive elements (Clemson et al., 2006), of which LINEs have been implicated in facilitating the formation of a silent compartment and the silencing process per se (Chow et al., 2010). These observations suggest that the formation of a transcriptionally silent compartment and the silencing of genes by displacement into that domain are uncoupled. They might also explain why X-linked genes are silenced only after appearance of an Xist- and H3K27me3-domain on the Xi.

The observed two-fold enrichment of LINEs on the human X chromosome (Bailey et al., 2000; Boyle et al., 1990) supported earlier ideas of LINEs acting as way stations in the spreading process of Xist (Lyon, 1998). This hypothesis was put forward to explain the incomplete spreading of Xist into nearby autosomal sequences in X to autosome translocations (Duthie



**Figure 5. 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 continuous lines. In undifferentiated ES cells XCI is repressed by the action of the XCI inhibitors repressing Xist and activating Tsix expression. Upon differentiation, loss of XCI inhibitors concomitant with the combined action of XCI activators triggers Xist upregulation.

et al., 1999; Popova et al., 2006; White et al., 1998) and the silencing gradient with the most effective silencing taking place closest to the break point (Russell and Montgomery, 1970). However, several transgene studies showed that Xist RNA per se is able to spread into and silence autosomal sequences (Heard et al., 1999; Lee and Jaenisch, 1997; Lee et al., 1996) and a recent study succeeded in inactivating an entire chromosome 21 in iPSCs derived from a Down-Syndrome patient thereby rescuing the accompanying phenotype in neurons (Jiang et al., 2013). These data rule out an absolute requirement of X chromosome-specific DNA elements for the spreading of Xist. In addition, XCI occurs in some rodent species that are devoid of LINEs (Cantrell et al., 2009) and the lack of 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. 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., 2012), 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., 2012) are thought to promote initial Xist binding.

Another long-standing question has been how Xist stays exclusively on the chromosome it is transcribed from. Inducible Xist transgenes carrying a series of truncations showed that several Xist domains contribute to Xist RNA binding to the Xi (Wutz et al., 2002). Using different Xist truncation mutants, *trans* diffusion of Xist RNA to X chromosomes other than the one it originated from was observed (Jeon and Lee, 2011). Massive over-expression from multi-copy transgenes introduced into terminally differentiated cells might explain this finding which has never been observed before (Jonkers et al., 2008; Lee et al., 1996) and is reminiscent of the situation in *Drosophila*, where roX RNA over-expression leads to roX binding to autosomal sequences (Ilik et al., 2013). YY1, a transcription factor with activator and repressor function, has been implicated in tethering Xist RNA to the *Xist* locus by its ability to bind both RNA and DNA via different domains. YY1 has thus been proposed to form a “nucleation center” for the spreading of Xist RNA (Jeon and

Lee, 2011), but may also be involved in “locking in” Xist at other X-linked sequences facilitating Xist spreading *in cis*.

### Induced Pluripotent Stem Cells as a Model for X Chromosome Reactivation

In 2006, Yamanaka and colleagues published their seminal findings on the *in vitro* generation of ES cell-like induced pluripotent stem cells (iPSCs). In this straightforward experiment, a pool of pluripotency-associated candidate transcription factors were tested for their ability to reprogram mouse fibroblasts into iPSCs (Takahashi and Yamanaka, 2006). In years since, the iPSCs technology has been strongly improved providing iPSCs that closely resemble ES cells both morphologically and functionally. One of the typical features associated with the naïve pluripotent state of mouse ES cells is the presence of two active X chromosomes. Also mouse iPSCs retain two active X chromosomes indicating that reprogramming of female mouse fibroblasts into iPSCs results in reactivation of the Xi. Notably, reactivation of the Xi is not an iPSC specific phenomenon. 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). X chromosome reactivation (XCR) therefore represents a powerful tool for studying epigenetic changes and chromatin dynamics and for defining fully reprogrammed cells, thus improving *in vitro* reprogramming methods.

### X Chromosome Reactivation in Embryonic Development

In the mouse, XCI is initiated between the 4 to 8 cells stage of embryonic

development. This initial phase of XCI is imprinted and is characterized by selective silencing of the paternal allele (Xp). Initially it was thought that iXCI occurred in a lineage-specific manner, with Xp being inactivated exclusively in extra embryonic tissues and random XCI occurring in cells emanating from the inner cell mass (ICM). However, the observation of Xp reactivation in the ICM revealed the plasticity of XCI in pre-implantation embryos (Figure 4; Mak et al., 2004; Okamoto et al., 2004). Xist RNA accumulation on Xp starts at the 4 cell stage of embryonic development. Subsequently, the PRC2 complex, responsible for the establishment of histone modification H3K27me3, appears on Xp at the morula stage. At day E3.5 of mouse embryonic development, every cell of the early stage female blastocyst shows a single Xist cloud that co-localizes with accumulated H3K27me3. Moreover, allele specific transcriptional analysis of X-linked genes shows silencing of Xp, thus demonstrating that iXCI occurred in all cells of the pre-implantation embryo (Mak et al., 2004). One day later Xist clouds and PRC2 foci start to disappear in cells of the ICM concomitant with reactivation of the Xp, whereas Xist coating and silencing of the Xp remains in cells that will contribute to 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 dependent on Xist repression. iXCI thus 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). Interestingly, loss of the Xist coated Xi and Xi associated PRC2 during ICM development occurs specifically in cells that are positive for pluripotency factor NANOG, suggesting that the establishment of a pluripotent ground state is necessary for XCR to take place. This

tight linkage between pluripotency, *Xist* regulation and reversibility of Xi silencing has been extensively studied in ES cells (Navarro et al., 2008; Navarro et al., 2005; Wutz and Jaenisch, 2000), and the role of some pluripotency-specific candidate genes has also been investigated in XCR.

Xp specific induction of *Tsix* in the developing blastocyst has been reported to repress *Xist* expression 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). Notably, OCT4, SOX2, NANOG and PRDM14 transcription factors all bind intron 1 of *Xist* and have been shown to inhibit random XCI by repressing *Xist* expression in ES cells (Ma et al., 2011; Navarro et al., 2008). XCR in the ICM appears to be restricted to NANOG and PRDM14 positive cells. NANOG is required for establishment of the naïve pluripotent state of ES cells, whereas PRDM14 is involved in the 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). 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). ChIP-seq studies also reveal binding of PRDM14 and NANOG

in *Rnf12* regulatory regions, and *Rnf12* expression is up-regulated in *Prdm14* knockout ESCs (Payer et al., 2013). These findings indicate that PRDM14 and possibly NANOG act on XCR through repression of *Rnf12*. This 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 including initiation of genome-wide DNA de-methylation (Saitou and Yamaji, 2012). By day E7.5, around 40 PGCs expressing *Blimp1*, *Stella* and *Prdm14* start migrating through the embryonic gut to 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. *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 methyltransferase genes *Dnmt3a/b* (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 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 reactivation of the DNA-methylation dependent inactive state of Xi in PGCs.

### ***In Vitro X Chromosome Reactivation: Somatic Cell Reprogramming***

*In vitro* XCR can be achieved by several methods. Mouse embryos generated by somatic cell nuclear transfer (SCNT)

show reactivation of an X-linked GFP transgene 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 present in the somatic cells is reactivated (Tada et al., 2001; Takagi et al., 1983; Ying et al., 2002). For these methods, XCR was assessed as a measure of erasure of epigenetic memory but neither of the two provide an efficient and appropriate model for studying dynamics of XCR *in vitro*. The generation of induced pluripotent stem cells (iPSCs) from somatic cells provided a more convenient model system to study XCR. Forced expression of four genes encoding the transcription factors, KLF4, SOX2, c-MYC and OCT4, appears to be sufficient to reprogram somatic cells to pluripotent stem cells (Takahashi and Yamanaka, 2006). 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). XCR occurs relative late during reprogramming, when iPSCs become independent of the exogenous reprogramming factors (Stadtfeld et al., 2008). As discussed above, the *Xist* intron 1 pluripotency factors binding site has been initially proposed to be the linkage between *Xist* regulation and pluripotency (Navarro et al., 2008). How-



ever, lack of the *Xist* intron 1 binding site for OCT4, SOX2 and NANOG does not result in up-regulation of *Xist* in ES or differentiated cells and does not affect XCR upon reprogramming (Minkovsky et al., 2013). Moreover, stem cells derived from the post-implantation epiblast (EpiSCs) show *Oct4* and *Sox2* expression and retain the inactive state of Xi, excluding a critical role of these factors in triggering Xi reprogramming (Guo et al., 2009). One ES cell specific pluripotency factor, not expressed in EpiSCs is REX1. As a target of RNF12, REX1 plays a key role in XCI initiation, and coinciding up-regulation of *Rex1* and down-regulation of *Rnf12* may be the trigger to execute XCR during reprogramming. Interestingly, KLF4 alone is able to reprogram EpiSCs into iPSCs, suggesting that KLF4 might also be implicated in XCR (Guo et al., 2009). In mouse embryonic development, XCR occurs in NANOG positive cells (Mak et al., 2004). Moreover, NANOG is initially dispensable during reprogramming but becomes essential at later stages of reprogramming promoting the transition from a pre-pluripotent state to a naïve ground state pluripotency typical of ES cells (Silva et al., 2009). In addition, *Nanog* null cells in the ICM fail to reactivate the Xi whereas *Nanog* over-expression triggers the reprogramming of EpiSCs to iPSCs, together with reactivation of Xi (Silva et al., 2009). Taken together, these observations point at NANOG as a candidate for the regulation of XCR. Another good candidate identified in embryonic development studies is PRDM14 (Yamaji et al., 2008). *Prdm14* greatly accelerates reprogramming of EpiSCs into iPSCs when expressed in combination with *Klf2*, and is required for self-renewal of iPSCs (Gillich et al., 2012). Moreover, *Prdm14* null iPSCs partially fail to down-regulate *Xist* during reprogramming (Payer et al., 2013). Interestingly, PRDM14 appears to act through repression of *Rnf12* and not through *Xist* intron 1, which is dispen-

sable for reprogramming mediated XCR (Minkovsky et al., 2013). These findings suggest that acquirement of the naïve pluripotent state results in expression of pluripotency factors acting as repressors of *Rnf12*. Whether *Xist* repression is a necessary step of epigenetic reprogramming as well as the molecular mechanism by which this is achieved is not yet clear. Importantly, XCR in the ICM and PGCs will likely not mimic XCR during iPSC reprogramming and further studies are necessary to better understand these processes.

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 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 examining a wide range of

hiPSCs indicate that similar to hESCs, hiPSCs show heterogeneous XCI patterns and even tend to lose *XIST* expression after prolonged passaging in 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 additional 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, 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. Although further studies will have

to confirm whether rXCI is properly initiated, and two recent studies converting conventional hESCs to a naïve state show different outcomes regarding the state of the Xi (Takashima et al., 2014; Theunissen et al., 2014), these two studies are of great importance and seem to provide us with an *in vitro* model to explore initiation and maintenance of XCI in humans. Moreover, since stringent pluripotency assays such as chimerism or tetraploid complementation are ethically not applicable to human cells, being able to characterize the epigenetic status of X chromosomes in female cells represents a powerful indirect method to generally assess the degree of pluripotency.

## Conclusions

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. Also, several of the known regulators have been reported to play a role in rXCI, but whether they also act in iXCI or XCR remains elusive. Furthermore, the mechanism of *Xist* mediated gene silencing still needs to be resolved, and the cues that ensure the developmental regulation of XCI and its irreversibility upon differentiation await identification. XCI represents an extraordinary epigenetic paradigm, and understanding how two genetically identical X chromosomes become different epigenetic entities within the same nucleus is fascinating and challenging. New insights in XCI research will shed more light on gene regulation and chromatin remodeling mechanisms, with an impact that reaches far beyond the field of dosage compensation.



## Aim and scope of this thesis

In heterogametic species, dosage compensation evolved to account for the difference in expression of sex chromosome-linked genes. In mammals dosage compensation is achieved by inactivation of one X chromosome during early female embryogenesis in a process called X chromosome inactivation (XCI). Central players in this process are two overlapping antisense transcribed noncoding genes, *Xist* and *Tsix*. *Xist* expression is required to silence the X chromosome in *cis*, by recruiting chromatin remodeling complexes, while *Tsix* expression is involved in the repression of *Xist*. The nature of this XCI process places this field of research at the interface between stem cell biology, epigenetics and gene regulation.

The aim of this thesis is to shed further light onto the different levels of regulation that ensure faithful initiation and maintenance of XCI in a developmental context. These different levels include transcription factors, noncoding RNAs, antisense transcription, and epigenetic processes including DNA methylation, histone modifications and chromosomal conformation.

In **chapter 2** we analyze *trans*- and *cis*-acting networks that regulate the initiation of XCI. Classical transcription factors and their regulators, most prominently the ubiquitin ligase RNF12 and its primary target in ES cells, REX1, have an essential function during initiation of XCI and are also needed to maintain *Xist* expression at early stages of differentiation. Moreover, our deletion and transgene studies argue that many of the noncoding RNAs located in the *Xic* predominantly function *in cis*, and that X chromosome pairing events are not necessary for XCI to occur.

In **chapter 3** we utilize *Xist* and *Tsix* reporter lines to study their regulation on an uncoupled allele which allows us to distinguish between direct and indirect effects, revealing antagonistic roles for both genes. We also use these cells to investigate the dynamics of *Xist* and *Tsix* regulation during ES cell differentiation and find it to be more stable than previously anticipated. Reduction of the X:autosome ratio reveals semi-stable states which might correspond to distinct higher order chromatin conformations of the *Xic*. These studies also indicate that X-encoded factors are involved in *Tsix* activation.

In **chapter 4** we present evidence for a special regulatory phase during early embryonic development in which dedicated chromatin modulators act to confer stable gene repression. A reporter construct containing an inducible antisense promoter showed that antisense transcription mediated repression in ES cells was reversible and dependent on properly assembled chromatin. In contrast, transcriptional interference did not seem to have a major repressive effect in our system. Repression was locked in by DNA methylation as soon as antisense transcription through the promoter of the reporter was induced during differentiation. Since the promoter of the reporter acquired H3K36me3 upon induction of antisense transcription, and interfering with removal of this histone modification resulted in enhanced silencing, we propose that H3K36me3 might be a crucial component in the cascade of stably silencing antisense transcribed genes. The results presented in **appendix A** support the notion of a developmental stage, in which certain characteristics of chromatin are recognized. Here, reprogramming of fibroblasts derived from a carrier of a full but unmethylated *FMR1* mutation resulted in targeting of DNA methylation to the mutated allele during the reprogramming process.

Finally, in the General Discussion in **chapter 5**, we discuss our findings and those of other's in the light of the different layers of regulation of XCI and in regard to the special developmental environment this fascinating phenomenon takes place in.

## References

- Abarregui, I., and Krangel, M.S. (2006). Regulation of T cell receptor-alpha gene recombination by transcription. *Nat Immunol* 7, 1109-1115.
- 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.
- Alvarez, J.D., Yasui, D.H., Niida, H., Joh, T., Loh, D.Y., and Kohwi-Shigematsu, T. (2000). The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. *Genes Dev* 14, 521-535.
- Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmid, C., Suzuki, T., *et al.* (2014). An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455-461.
- Anguera, M.C., Ma, W., Clift, D., Namekawa, S., Kelleher, R.J., 3rd, 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.
- Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* 316, 744-747.
- Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat Rev Mol Cell Biol* 10, 91-103.
- 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., 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.
- Badeaux, A.I., and Shi, Y. (2013). Emerging roles for chromatin as a signal integration and storage platform. *Nat Rev Mol Cell Biol* 14, 211-224.
- 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 USA* 97, 6634-6639.
- Ball, M.P., Li, J.B., Gao, Y., Lee, J.H., LeProust, E.M., Park, I.H., Xie, B., Daley, G.Q., and Church, G.M. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol* 27, 361-368.
- Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299-308.
- Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res* 21, 381-395.
- Barakat, T.S., Gunhanlar, N., Gontan Pardo, C., Achame, E.M., Ghazvini, M., Bowers, R., Kenter, A., Rentmeester, E., Grootegoed, J.A., and Gribnau, J. (2011). RNF12 Activates Xist and Is Essential for X Chromosome Inactivation. *PLoS Genet* 7, e1002001.
- Barakat, T.S., Loos, F., van Staveren, S., Myronova, E., Ghazvini, M., Grootegoed, J.A., and Gribnau, J. (2014). The Trans-Activator RNF12 and Cis-Acting Elements Effectuate X Chromosome Inactivation Independent of X-Pairing. *Mol Cell* 53, 965-978.
- Barbour, V.M., Tufarelli, C., Sharpe, J.A., Smith, Z.E., Ayyub, H., Heinlein, C.A., Sloane-Stanley, J., Indrak, K., Wood, W.G., and Higgs, D.R. (2000). alpha-thalassemia resulting from a negative chromosomal position effect. *Blood* 96, 800-807.
- Barlow, D.P. (2011). Genomic imprinting: a mammalian epigenetic discovery model. *Annu Rev Genet* 45, 379-403.
- Barr, M.L., and Bertram, E.G. (1949). A morphological distinction between neurones of the male and female, and the behav-

jour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature* 163, 676.

Bartke, T., Vermeulen, M., Xhemalce, B., Robson, S.C., Mann, M., and Kouzarides, T. (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 143, 470-484.

Batlle-Morera, L., Smith, A., and Nichols, J. (2008). Parameters influencing derivation of embryonic stem cells from murine embryos. *Genesis* 46, 758-767.

Batra, R., Charizanis, K., and Swanson, M.S. (2010). Partners in crime: bidirectional transcription in unstable microsatellite disease. *Hum Mol Genet* 19, R77-82.

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.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 405, 482-485.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.

Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Mol Cell* 43, 1040-1046.

Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., *et al.* (2004). Global identification of human transcribed sequences with genome tiling arrays. *Science* 306, 2242-2246.

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev* 23, 2484-2489.

Bird, A., Taggart, M., Frommer, M., Miller, O.J., and Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* 40, 91-99.

Blackledge, N.P., Zhou, J.C., Tolstorukov, M.Y., Farcas, A.M., Park, P.J., and Klose, R.J. (2010). CpG islands recruit a histone H3 lysine 36 demethylase. *Mol Cell* 38, 179-190.

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). SmcHDD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet* 40, 663-669.

Bock, C., Beerman, I., Lien, W.H., Smith, Z.D., Gu, H., Boyle, P., Gnirke, A., Fuchs, E., Rossi, D.J., and Meissner, A. (2012). DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Mol Cell* 47, 633-647.

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.

Bologna, N.G., and Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. *Annu Rev Plant Biol* 65, 473-503.

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

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.

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science* 294, 2536-2539.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., *et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441, 349-353.

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 karyo-

typing by fluorescence in situ hybridization. *Proc Natl Acad Sci U S A* 87, 7757-7761.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., and Cedar, H. (1994). Sp1 elements protect a CpG island from de novo methylation. *Nature* 371, 435-438.

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.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128, 1089-1103.

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.

Brower-Toland, B., Findley, S.D., Jiang, L., Liu, L., Yin, H., Dus, M., Zhou, P., Elgin, S.C., and Lin, H. (2007). *Drosophila* PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev* 21, 2300-2311.

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., 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.

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.

Buhler, M., Haas, W., Gygi, S.P., and Moazed, D. (2007). RNAi-dependent and -independent RNA turnover mechanisms contribute to heterochromatic gene silencing. *Cell* 129, 707-721.

Buhler, M., Verdel, A., and Moazed, D. (2006). Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* 125, 873-886.

Buschhausen, G., Wittig, B., Graessmann, M., and Graessmann, A. (1987). Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. *Proc Natl Acad Sci USA* 84, 1177-1181.

Buske, F.A., Mattick, J.S., and Bailey, T.L. (2011). Potential in vivo roles of nucleic acid triple-helices. *RNA Biol* 8, 427-439.

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

Calabrese, J.M., Sun, W., Song, L., Mugford, J.W., Williams, L., Yee, D., Starmier, 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.

Callen, B.P., Shearwin, K.E., and Egan, J.B. (2004). Transcriptional interference between convergent promoters caused by elongation over the promoter. *Mol Cell* 14, 647-656.

Cam, H.P., Sugiyama, T., Chen, E.S., Chen, X., FitzGerald, P.C., and Grewal, S.I. (2005). Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat Genet* 37, 809-819.

Campanero, M.R., Armstrong, M.I., and Flemington, E.K. (2000). CpG methylation as a mechanism for the regulation of E2F activity. *Proc Natl Acad Sci U S A* 97, 6481-6486.

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.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones,



R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298, 1039-1043.

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.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., *et al.* (2005). The transcriptional landscape of the mammalian genome. *Science* 309, 1559-1563.

Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., *et al.* (2003). Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia. *Genome Res* 13, 1273-1289.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., *et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123, 581-592.

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

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.

Cedar, H., and Bergman, Y. (2012). Programming of DNA methylation patterns. *Annu Rev Biochem* 81, 97-117.

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, 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.

Chang, S.S., Zhang, Z., and Liu, Y. (2012). RNA interference pathways in fungi: mechanisms and functions. *Annu Rev Microbiol* 66, 305-323.

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.

Chen, J., Sun, M., Hurst, L.D., Carmichael, G.G., and Rowley, J.D. (2005a). Genome-wide analysis of coordinate expression and evolution of human cis-encoded sense-antisense transcripts. *Trends Genet* 21, 326-329.

Chen, J., Sun, M., Hurst, L.D., Carmichael, G.G., and Rowley, J.D. (2005b). Human antisense genes have unusually short introns: evidence for selection for rapid transcription. *Trends Genet* 21, 203-207.

Chen, J., Sun, M., Kent, W.J., Huang, X., Xie, H., Wang, W., Zhou, G., Shi, R.Z., and Rowley, J.D. (2004). Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res* 32, 4812-4820.

Chen, T., and Dent, S.Y. (2014). Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat Rev Genet* 15, 93-106.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol* 23, 5594-5605.

Chodavarapu, R.K., Feng, S., Bernatavichute, Y.V., Chen, P.Y., Stroud, H., Yu, Y., Hetzel, J.A., Kuo, F., Kim, J., Cokus, S.J., *et al.* (2010). Relationship between nucleosome positioning and DNA methylation. *Nature* 466, 388-392.

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

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., and Chang, H.Y. (2011). Genomic maps of long noncoding RNA

occupancy reveal principles of RNA-chromatin interactions. *Mol Cell* 44, 667-678.

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.

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.

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.

Clark, M.B., Amaral, P.P., Schlesinger, F.J., Dinger, M.E., Taft, R.J., Rinn, J.L., Ponting, C.P., Stadler, P.F., Morris, K.V., Morillon, A., *et al.* (2011). The reality of pervasive transcription. *PLoS Biol* 9, e1000625; discussion e1001102.

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

Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H., Chess, A., and Lawrence, J.B. (2009). An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* 33, 717-726.

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

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

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.

Cohen, N.M., Kenigsberg, E., and Tanay, A. (2011). Primate CpG islands are maintained by heterogeneous evolutionary regimes involving minimal selection. *Cell* 145, 773-786.

Colak, D., Zaninovic, N., Cohen, M.S., Rosenwaks, Z., Yang, W.Y., 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.

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

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.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57-74.

Consortium, E.P., Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., *et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799-816.

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.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322, 1845-1848.

Costanzi, C., and Pehrson, J.R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* 393, 599-601.

Crampton, N., Bonass, W.A., Kirkham, J., Rivetti, C., and Thomson, N.H. (2006). Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. *Nucleic Acids Res* 34, 5416-5425.

Csankovszki, G., Panning, B., Bates, B.,



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

Cusanovich, D.A., Pavlovic, B., Pritchard, J.K., and Gilad, Y. (2014). The functional consequences of variation in transcription factor binding. *PLoS Genet* 10, e1004226.

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

Dahary, D., Elroy-Stein, O., and Sorek, R. (2005). Naturally occurring antisense: transcriptional leakage or real overlap? *Genome Res* 15, 364-368.

Darrow, E.M., and Chadwick, B.P. (2013). Boosting transcription by transcription: enhancer-associated transcripts. *Chromosome Res* 21, 713-724.

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

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.

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 Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragousis, J., Wei, C.L., and Natoli, G. (2010). A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol* 8, e1000384.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev* 25, 1010-1022.

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

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*. *Nat Genet* 43, 1179-1185.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., *et al.* (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* 22, 1775-1789.

Deshpande, G., Calhoun, G., and Schedl, P. (2005). *Drosophila argonaute-2* is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. *Genes Dev* 19, 1680-1685.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399, 491-496.

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

Dimitrova, N., Zamudio, J.R., Jong, R.M., Soukup, D., Resnick, R., Sarma, K., Ward, A.J., Raj, A., Lee, J.T., Sharp, P.A., *et al.* (2014). LincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint. *Mol Cell* 54, 777-790.

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 Ctf and also controls X-chromosome pair-

ing and counting. *Nature* 460, 128-132.

Duhring, U., Axmann, I.M., Hess, W.R., and Wilde, A. (2006). An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proc Natl Acad Sci U S A* 103, 7054-7058.

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.

Ebert, M.S., Neilson, J.R., and Sharp, P.A. (2007). MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat Methods* 4, 721-726.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoepfner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D., Buetow, K.H., *et al.* (2008). Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell* 2, 437-447.

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

Ellegren, H., Hultin-Rosenberg, L., Brunstrom, B., Dencker, L., Kulima, 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., Sirotkin, 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*.

Epsztejn-Litman, S., Feldman, N., Abu-Remaileh, M., Shufaro, Y., Gerson, A., Ueda, J., Deplus, R., Fuks, F., Shinkai, Y., Cedar, H., *et al.* (2008). De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* 15, 1176-1183.

Eszterhas, S.K., Bouhassira, E.E., Martin, D.I., and 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.

Evans, M.J., and Kaufman, M.H. (1981).

Establishment in culture of pluripotent cells from mouse embryos. *Nature* 292, 154-156.

Farazi, T.A., Hoell, J.I., Morozov, P., and Tuschl, T. (2013). MicroRNAs in human cancer. *Adv Exp Med Biol* 774, 1-20.

Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* 8, 188-194.

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

Ficz, G., Hore, T.A., Santos, F., Lee, H.J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y.L., Walter, J., *et al.* (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* 13, 351-359.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-811.

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

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

Friedersdorf, M.B., and Keene, J.D. (2014). Advancing the functional utility of PAR-CLIP by quantifying background binding to mRNAs and lncRNAs. *Genome Biol* 15, R2.

Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T., and Oshimura, M. (2004). Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat Cell Biol* 6, 784-791.

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.

Galindo, M.I., Pueyo, J.I., Fouix, S., Bishop, S.A., and Couso, J.P. (2007). Peptides encoded by short ORFs control development and define a new eukaryotic gene family. *PLoS Biol* 5, e106.

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

Gendrel, A.V., Apedaile, A., Coker, H., Termanis, A., Zvetkova, I., Godwin, J., Tang, Y.A., Huntley, D., Montana, G., Taylor, S., *et al.* (2012). Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Dev Cell* 23, 265-279.

Gibb, E.A., Brown, C.J., and Lam, W.L. (2011). The functional role of long non-coding RNA in human carcinomas. *Mol Cancer* 10, 38.

Gifford, C.A., Ziller, M.J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A.K., Kelley, D.R., Shishkin, A.A., Issner, R., *et al.* (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* 153, 1149-1163.

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.

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

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., Vandenberg, J.L., Renfree, M.B., Taylor, W., *et al.* (2012). Rxs is

a metatherian RNA with Xist-like properties in X-chromosome inactivation. *Nature*.

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

Green, M.R. (2005). Eukaryotic transcription activation: right on target. *Mol Cell* 18, 399-402.

Gribnau, J., Diderich, K., Pruzina, S., Calzolari, R., and Fraser, P. (2000). Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. *Mol Cell* 5, 377-386.

Gribnau, J., Luikenhuis, S., Hochedlinger, K., Monkhorst, K., and Jaenisch, R. (2005). X chromosome choice occurs independently of asynchronous replication timing. *J Cell Biol* 168, 365-373.

Griffiths, D.S., Li, J., Dawson, M.A., Trotter, M.W., Cheng, Y.H., Smith, A.M., Mansfield, W., Liu, P., Kouzarides, T., Nichols, J., *et al.* (2011). LIF-independent JAK signalling to chromatin in embryonic stem cells uncovered from an adult stem cell disease. *Nat Cell Biol* 13, 13-21.

Grishok, A. (2013). Biology and Mechanisms of Short RNAs in *Caenorhabditis elegans*. *Adv Genet* 83, 1-69.

Gruenbaum, Y., Cedar, H., and Razin, A. (1982). Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* 295, 620-622.

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.

Grunberg, S., and Hahn, S. (2013). Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem Sci* 38, 603-611.

Gu, T.P., Guo, F., Yang, H., Wu, H.P., Xu, G.F., Liu, W., Xie, Z.G., Shi, L., He, X., Jin, S.G., *et al.* (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* 477, 606-610.

Guang, S., Bochner, A.F., Burkhart, K.B., Burton, N., Pavelec, D.M., and Kennedy, S. (2010). Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. *Nature* 465, 1097-1101.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130, 77-88.

Guibert, S., and Weber, M. (2013). Functions of DNA methylation and hydroxymethylation in Mammalian development. *Curr Top Dev Biol* 104, 47-83.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* 136, 1063-1069.

Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., *et al.* (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464, 1071-1076.

Gutschner, T., Hammerle, M., and Diederichs, S. (2013). MALAT1 -- a paradigm for long noncoding RNA function in cancer. *J Mol Med (Berl)* 91, 791-801.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., *et al.* (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458, 223-227.

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., *et al.* (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 477, 295-300.

Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., *et al.* (2010). Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol* 28, 503-510.

Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. *Nature* 482, 339-346.

Guttman, M., Russell, P., Ingolia, N.T., Weissman, J.S., and Lander, E.S. (2013). Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell* 154, 240-251.

Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C., *et al.* (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* 13, 360-369.

Hacisuleyman, E., Goff, L.A., Trapnell, C., Williams, A., Henao-Mejia, J., Sun, L., McClanahan, P., Hendrickson, D.G., Sauvageau, M., Kelley, D.R., *et al.* (2014). Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat Struct Mol Biol* 21, 198-206.

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

Hall, I.M., Shankaranarayana, G.D., Noma, K., Ayoub, N., Cohen, A., and Grewal, S.I. (2002). Establishment and maintenance of a heterochromatin domain. *Science* 297, 2232-2237.

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. *J Cell Physiol* 216, 445-452.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293-296.

Hangauer, M.J., Vaughn, I.W., and McManus, M.T. (2013). Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet* 9, e1003569.

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.

Hannon, G.J. (2002). RNA interference. *Nature* 418, 244-251.

Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N., and Grosveld, F. (1991). Importance of globin gene order for correct developmental expression. *Genes Dev* 5, 1387-1394.

Hansen, T.B., Jensen, T.I., Claus-



en, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* 495, 384-388.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405, 486-489.

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

Hashimshony, T., Zhang, J., Keshet, I., Bustin, M., and Cedar, H. (2003). The role of DNA methylation in setting up chromatin structure during development. *Nat Genet* 34, 187-192.

Hausser, J., and Zavolan, M. (2014). Identification and consequences of miRNA-target interactions-beyond repression of gene expression. *Nat Rev Genet* 15, 599-612.

He, Y., Vogelstein, B., Velculescu, V.E., Papadopoulos, N., and Kinzler, K.W. (2008). The antisense transcriptomes of human cells. *Science* 322, 1855-1857.

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., 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.

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

Henikoff, S., and Shilatifard, A. (2011). Histone modification: cause or cog? *Trends Genet* 27, 389-396.

Heo, J.B., and Sung, S. (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* 331, 76-79.

Hilfiker, A., Hilfiker-Kleiner, D., Pan-nuti, 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.

Hobson, D.J., Wei, W., Steinmetz, L.M., and Svejstrup, J.Q. (2012). RNA polymerase II collision interrupts convergent transcription. *Mol Cell* 48, 365-374.

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.

Holler, M., Westin, G., Jiricny, J., and Schaffner, W. (1988). Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev* 2, 1127-1135.

Hong, L., Schroth, G.P., Matthews, H.R., Yau, P., and Bradbury, E.M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *J Biol Chem* 268, 305-314.

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

Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., et al. (2010). A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* 142, 409-419.

Hutchinson, J.N., Ensminger, A.W., Clemson, C.M., Lynch, C.R., Lawrence, J.B., and Chess, A. (2007). A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* 8, 39.

Ietswaart, R., Wu, Z., and Dean, C. (2012). Flowering time control: another window to the connection between antisense RNA and chromatin. *Trends Genet* 28, 445-453.

Ilik, I.A., Quinn, J.J., Georgiev, P., Tavares-Cadete, F., Maticzka, D., Toscano, S., Wan, Y., Spitale, R.C., Luscombe, N., Backofen, R., et al. (2013). Tandem stem-loops in roX RNAs act together to mediate X chromosome dosage compensation in *Drosophila*. *Mol Cell* 51, 156-173.

- Illingworth, R.S., Gruenewald-Schneider, U., Webb, S., Kerr, A.R., James, K.D., Turner, D.J., Smith, C., Harrison, D.J., Andrews, R., and Bird, A.P. (2010). Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet* 6, e1001134.
- Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147, 789-802.
- 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.
- Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol Cell Biol* 24, 8862-8871.
- Janowski, B.A., Huffman, K.E., Schwartz, J.C., Ram, R., Nordsell, R., Shames, D.S., Minna, J.D., and Corey, D.R. (2006). Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. *Nat Struct Mol Biol* 13, 787-792.
- 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.
- Jeon, Y., and Lee, J.T. (2011). YY1 Tethers Xist RNA to the Inactive X Nucleation Center. *Cell* 146, 119-133.
- Jeong, S., Liang, G., Sharma, S., Lin, J.C., Choi, S.H., Han, H., Yoo, C.B., Egger, G., Yang, A.S., and Jones, P.A. (2009). Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol Cell Biol* 29, 5366-5376.
- 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.
- Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503, 290-294.
- 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.
- Juven-Gershon, T., Hsu, J.Y., Theisen, J.W., and Kadonaga, J.T. (2008). The RNA polymerase II core promoter - the gateway to transcription. *Curr Opin Cell Biol* 20, 253-259.
- Juven-Gershon, T., and Kadonaga, J.T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev Biol* 339, 225-229.
- Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H., and Razin, A. (1992). Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev* 6, 705-714.
- Kaji, K., Caballero, I.M., MacLeod, R., Nichols, J., Wilson, V.A., and Hendrich, B. (2006). The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nat Cell Biol* 8, 285-292.
- Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* 429, 900-903.
- 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.

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., *et al.* (2005). Antisense transcription in the mammalian transcriptome. *Science* 309, 1564-1566.

Kato, H., Goto, D.B., Martienssen, R.A., Urano, T., Furukawa, K., and Murakami, Y. (2005). RNA polymerase II is required for RNAi-dependent heterochromatin assembly. *Science* 309, 467-469.

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.

Keshet, I., Lieman-Hurwitz, J., and Cedar, H. (1986). DNA methylation affects the formation of active chromatin. *Cell* 44, 535-543.

Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., *et al.* (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A* 106, 11667-11672.

Kim, D.H., Villeneuve, L.M., Morris, K.V., and Rossi, J.J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* 13, 793-797.

Kim, J., Kollhoff, A., Bergmann, A., and Stubbs, L. (2003). Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. *Hum Mol Genet* 12, 233-245.

Kim, T., and Buratowski, S. (2009). Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. *Cell* 137, 259-272.

Kim, T., Xu, Z., Clauder-Munster, S., Steinmetz, L.M., and Buratowski, S. (2012). Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. *Cell* 150, 1158-1169.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptevich, M., Barbara-Haley, K., Kuersten, S., *et al.* (2010). Widespread transcription at neuronal ac-

tivity-regulated enhancers. *Nature* 465, 182-187.

Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Non-coding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal* 3, ra8.

Kiyosawa, H., Yamanaka, I., Osato, N., Kondo, S., Hayashizaki, Y., Group, R.G., and Members, G.S.L. (2003). Antisense transcripts with FANTOM2 clone set and their implications for gene regulation. *Genome Res* 13, 1324-1334.

Klattenhoff, C.A., Scheuermann, J.C., Surface, L.E., Bradley, R.K., Fields, P.A., Steinhäuser, M.L., Ding, H., Butty, V.L., Torrey, L., Haas, S., *et al.* (2013). Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell* 152, 570-583.

Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 31, 89-97.

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.

Kondo, T., Plaza, S., Zanet, J., Benrabah, E., Valenti, P., Hashimoto, Y., Kobayashi, S., Payre, F., and Kageyama, Y. (2010). Small peptides switch the transcriptional activity of Shavenbaby during *Drosophila* embryogenesis. *Science* 329, 336-339.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868-871.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. *Science* 184, 865-868.

Krumlauf, R. (1994). Hox genes in vertebrate development. *Cell* 78, 191-201.

Ladd, P.D., Smith, L.E., Rabaia, N.A., Moore, J.M., Georges, S.A., Hansen, R.S., Hagerman, R.J., Tassone, F., Tapscott, S.J., and Filippova, G.N. (2007). An antisense transcript spanning the CGG repeat region of *FMRI* is upregulated in premutation carriers but silenced in full mutation individuals. *Hum Mol Genet* 24, 3174-3187.

Lagarkova, M.A., Volchikov, P.Y., Lyaki-



- sheva, 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.
- Lai, F., Orom, U.A., Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A., and Shiekhattar, R. (2013). Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* 494, 497-501.
- Landolin, J.M., Johnson, D.S., Trinklein, N.D., Aldred, S.F., Medina, C., Shulha, H., Weng, Z., and Myers, R.M. (2010). Sequence features that drive human promoter function and tissue specificity. *Genome Res* 20, 890-898.
- Larschan, E., Alekseyenko, A.A., Lai, W.R., Park, P.J., and Kuroda, M.I. (2006). MSL complex associates with clusters of actively transcribed genes along the *Drosophila* male X chromosome. *Cold Spring Harb Symp Quant Biol* 71, 385-394.
- 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.
- Latos, P.A., Pauler, F.M., Koerner, M.V., Senergin, H.B., Hudson, Q.J., Stocsits, R.R., Allhoff, W., Stricker, S.H., Klement, R.M., Warczok, K.E., *et al.* (2012). Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* 338, 1469-1472.
- Lauberth, S.M., Nakayama, T., Wu, X., Ferris, A.L., Tang, Z., Hughes, S.H., and Roeder, R.G. (2013). H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* 152, 1021-1036.
- Lee, J., and Lu, N. (1999). Targeted mutagenesis of Tsix leads to non-random X inactivation. *Cell* 99, 47-57.
- 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., Davidow, L.S., and Warshawsky, D. (1999a). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* 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.
- Lee, J.T., Lu, N., and Han, Y. (1999b). 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., 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, R.C., and Ambros, V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862-864.
- Legewie, S., Dienst, D., Wilde, A., Hertz, H., and Axmann, I.M. (2008). Small RNAs establish delays and temporal thresholds in gene expression. *Biophys J* 95, 3232-3238.
- 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.
- Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71, 865-873.
- Lettice, L.A., Heaney, S.J., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet* 12, 1725-1735.
- Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. *Nature* 424, 147-151.
- Levo, M., and Segal, E. (2014). In pursuit of design principles of regulatory sequences. *Nat Rev Genet* 15, 453-468.
- Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.
- Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature* 366, 362-365.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in em-

bryonic lethality. *Cell* 69, 915-926.

Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., *et al.* (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498, 516-520.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289-293.

Lin, H., Wang, Y., Wang, Y., Tian, F., Pu, P., Yu, Y., Mao, H., Yang, Y., Wang, P., Hu, L., *et al.* (2010). Coordinated regulation of active and repressive histone methylations by a dual-specificity histone demethylase *ceKDM7A* from *Caenorhabditis elegans*. *Cell Res* 20, 899-907.

Lin, W., Srajer, G., Evrard, Y.A., Phan, H.M., Furuta, Y., and Dent, S.Y. (2007). Developmental potential of *Gcn5*(-/-) embryonic stem cells in vivo and in vitro. *Dev Dyn* 236, 1547-1557.

Livernois, A.M., Graves, J.A., and Waters, P.D. (2012). The origin and evolution of vertebrate sex chromosomes and dosage compensation. *Heredity (Edinb)* 108, 50-58.

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

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.

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.

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.

Maenner, S., Muller, M., Frohlich, J., Langer, D., and Becker, P.B. (2013). ATP-dependent roX RNA remodeling by the helicase maleless enables specific association of MSL proteins. *Mol Cell* 51, 174-184.

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

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137, 522-535.

Mancini-Dinardo, D., Steele, S.J., Levorse, J.M., Ingram, R.S., and Tilghman, S.M. (2006). Elongation of the *Kcnqlot1* transcript is required for genomic imprinting of

neighboring genes. *Genes Dev* 20, 1268-1282.

Mao, Y.S., Sunwoo, H., Zhang, B., and Spector, D.L. (2011). Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol* 13, 95-101.

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.

Margaritis, T., Oreal, V., Brabers, N., Maestroni, L., Vitaliano-Prunier, A., Benschop, J.J., van Hooff, S., van Leenen, D., Dargemont, C., Geli, V., *et al.* (2012). Two distinct repressive mechanisms for histone 3 lysine 4 methylation through promoting 3'-end antisense transcription. *PLoS Genet* 8, e1002952.

Mariner, P.D., Walters, R.D., Espinoza, C.A., Drullinger, L.F., Wagner, S.D., Kugel, J.F., and Goodrich, J.A. (2008). Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. *Mol Cell* 29, 499-509.

Marks, H., Kalkan, T., Menafrá, R., Denisov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A., *et al.* (2012). The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149, 590-604.

Marques, A.C., and Ponting, C.P. (2009). Catalogues of mammalian long non-coding RNAs: modest conservation and incompleteness. *Genome Biol* 10, R124.

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

Martin GR, E.C., Travis B, Tucker G, Yatiziv S, Martin DW Jr, Clift S, Cohen S. (1978). X-chromosome inactivation during differentiation of female teratocarcinoma stem cells in vitro. *Nature* (5643), 329-333.

Matzke, M.A., and Mosher, R.A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* 15, 394-408.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., *et al.* (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454, 766-770.

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.

Meng, L., Person, R.E., and Beaudet, A.L. (2012). Ube3a-ATS is an atypical RNA polymerase II transcript that represses the paternal expression of Ube3a. *Hum Mol Genet* 21, 3001-3012.

Merkenschlager, M., and Odom, D.T. (2013). CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* 152, 1285-1297.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* 10, 105-116.

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

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.

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

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 com-

pensation in *C. elegans*. *Cell* 55, 167-183.

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.

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.

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

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

Monk, M., Boubelik, M., and Lehnert, S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* 99, 371-382.

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.

Morris, K.V., Chan, S.W., Jacobsen, S.E., and Looney, D.J. (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* 305, 1289-1292.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to non-coding centromeric RNAs. *Cell* 119, 789-802.

Murthy, U.M., and Rangarajan, P.N. (2010). Identification of protein interaction regions of VINC/NEAT1/Men epsilon RNA. *FEBS Lett* 584, 1531-1535.

Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA

epigenetically silences transcription by targeting G9a to chromatin. *Science* 322, 1717-1720.

Nan, X., Campoy, F.J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* 88, 471-481.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393, 386-389.

Nanda, I., Haaf, T., Scharl, 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.

Nanda, I., Shan, Z., Scharl, 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. *Nat Genet* 21, 258-259.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* 2, 279-289.

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 regulated by pluripotency factors in embryonic stem cells. *Hum Genet*.

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

Navarro, P., Pichard, S., Ciaudo, C., Avner, P., and Rougeulle, C. (2005). *Tsix* transcription across the *Xist* gene alters chro-



matin conformation without affecting Xist transcription: implications for X-chromosome inactivation. *Genes Dev* 19, 1474-1484.

Naveh-Mani, T., and Cedar, H. (1981). Active gene sequences are undermethylated. *Proc Natl Acad Sci U S A* 78, 4246-4250.

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

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

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.

Ng, H.H., Robert, F., Young, R.A., and 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.

Ng, S.Y., Johnson, R., and Stanton, L.W. (2012). Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. *EMBO J* 31, 522-533.

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

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

Nichols, J., Zevnik, B., Anastasiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998).

Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95, 379-391.

Noma, K., Sugiyama, T., Cam, H., Verdel, A., Zofall, M., Jia, S., Moazed, D., and Grewal, S.I. (2004). RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nat Genet* 36, 1174-1180.

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.

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*.

Nord, A.S., Blow, M.J., Attanasio, C., Akiyama, J.A., Holt, A., Hosseini, R., Phouanavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., *et al.* (2013). Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* 155, 1521-1531.

Numata, K., Kanai, A., Saito, R., Kondo, S., Adachi, J., Wilming, L.G., Hume, D.A., Hayashizaki, Y., Tomita, M., Group, R.G., *et al.* (2003). Identification of putative noncoding RNAs among the RIKEN mouse full-length cDNA collection. *Genome Res* 13, 1301-1306.

Numata, K., and Kiyosawa, H. (2012). Genome-wide impact of endogenous antisense transcripts in eukaryotes. *Front Biosci (Landmark Ed)* 17, 300-315.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene *Ezh2* is required for early mouse development. *Mol Cell Biol* 21, 4330-4336.

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

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

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

Ohinata, Y., Payer, B., O'Carroll, D., Ancellin, 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.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247-257.

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

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., *et al.* (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* 448, 714-717.

Orom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., *et al.* (2010). Long noncoding RNAs with enhancer-like function in human cells. *Cell* 143, 46-58.

Orom, U.A., and Shiekhattar, R. (2013). Long noncoding RNAs usher in a new era in the biology of enhancers. *Cell* 154, 1190-1193.

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DN-

MT3L domain. *EMBO Rep* 10, 1235-1241.

Palmer, A.C., Ahlgren-Berg, A., Egan, J.B., Dodd, I.B., and Shearwin, K.E. (2009). Potent transcriptional interference by pausing of RNA polymerases over a downstream promoter. *Mol Cell* 34, 545-555.

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

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

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. *Mol Cell* 52, 805-818.

Pelechano, V., and Steinmetz, L.M. (2013). Gene regulation by antisense transcription. *Nat Rev Genet* 14, 880-893.

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

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

Pera, M.F., and Tam, P.P. (2010). Extrinsic regulation of pluripotent stem cells. *Nature* 465, 713-720.

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.

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*.

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.

Peters, J. (2014). The role of genomic imprinting in biology and disease: an expanding view. *Nat Rev Genet* 15, 517-530.

Pinskaya, M., Gourvennec, S., and Morillon, A. (2009). H3 lysine 4 di- and tri-methylation deposited by cryptic transcription attenuates promoter activation. *EMBO J* 28, 1697-1707.

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., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, 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.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W.J., and Pandolfi, P.P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 465, 1033-1038.

Ponjavic, J., Ponting, C.P., and Lunter, G. (2007). Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. *Genome Res* 17, 556-565.

Popova, B.C., Tada, T., Takagi, N., Brockdorff, N., and Nesterova, T.B. (2006). Attenuated spread of X-inactivation in an X;autosome translocation. *Proc Natl Acad Sci USA* 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. *Dev Cell* 9, 339-349.

Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* 433, 434-438.

Prescott, E., and Proudfoot, N. (2002). Transcriptional collision between convergent genes in budding yeast. *Proc Natl Acad Sci U S A* 99, 8796-8801.

Ptashne, M. (2005). Regulation of transcription: from lambda to eukaryotes. *Trends Biochem Sci* 30, 275-279.

Ptashne, M. (2013). Epigenetics: core misconcept. *Proc Natl Acad Sci USA* 110, 7101-7103.

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.

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P.V., Grimaldi, G., Riccio, A., *et al.* (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell* 44, 361-372.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* 138, 114-128.

Rapicavoli, N.A., Poth, E.M., Zhu, H., and Blackshaw, S. (2011). The long noncoding RNA Six3OS acts in trans to regulate retinal development by modulating Six3 activity. *Neural Dev* 6, 32.

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.

Ravasi, T., Suzuki, H., Pang, K.C., Katayama, S., Furuno, M., Okunishi, R., Fukuda, S., Ru, K., Frith, M.C., Gongora, M.M., *et al.* (2006). Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res* 16, 11-19.

Redrup, L., Branco, M.R., Perdeaux, E.R., Krueger, C., Lewis, A., Santos, F., Nagano, T., Cobb, B.S., Fraser, P., and Reik, W. (2009). The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing. *Development* 136, 525-530.

Rego, A., Sinclair, P.B., Tao, W., Kireev, I., and Belmont, A.S. (2008). The facultative heterochromatin protein 1 (FHP1) is a

tative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure. *J Cell Sci* 121, 1119-1127.

Reynolds, N., Latos, P., Hynes-Al-len, A., Loos, R., Leaford, D., O'Shaughnessy, A., Mosaku, O., Signolet, J., Brennecke, P., Kalkan, T., *et al.* (2012). NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. *Cell Stem Cell* 10, 583-594.

Riising, E.M., Comet, I., Leblanc, B., Wu, X., Johansen, J.V., and Helin, K. (2014). Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol Cell* 55, 347-360.

Rinn, J., Kertesz, M., Wang, J., Squazzo, S., Xu, X., Bruggmann, S., Goodnough, L., Helms, J., Farnham, P., Segal, E., *et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129, 1311-1323.

Roeder, R.G. (1998). Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. *Cold Spring Harb Symp Quant Biol* 63, 201-218.

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.

Ruthenburg, A.J., Li, H., Milne, T.A., Dewell, S., McGinty, R.K., Yuen, M., Ueberheide, B., Dou, Y., Muir, T.W., Patel, D.J., *et al.* (2011). Recognition of a mononucleoso-

mal histone modification pattern by BPTF via multivalent interactions. *Cell* 145, 692-706.

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

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

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

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.

Sanford, J.P., Clark, H.J., Chapman, V.M., and Rossant, J. (1987). Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev* 1, 1039-1046.

Santoro, F., Mayer, D., Klement, R.M., Warczok, K.E., Stukalov, A., Barlow, D.P., and Pauler, F.M. (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., and Brutlag, D.L. (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.

Schatz, D.G., and Ji, Y. (2011). Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol* 11, 251-263.

Schmidt, J.V., Levorse, J.M., and Tilghman, S.M. (1999). Enhancer competition between H19 and Igf2 does not mediate their imprinting. *Proc Natl Acad Sci U S A* 96, 9733-9738.

Schmitz, K.-M., Mayer, C., Postepska, A., and Grummt, I. (2010). Interaction of non-

coding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev* 24, 2264-2269.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S., *et al.* (2010). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet* 42, 53-61.

Schorderet, P., and Duboule, D. (2011). Structural and functional differences in the long non-coding RNA *hotair* in mouse and human. *PLoS Genet* 7, e1002071.

Schultz, D.C., Ayyanathan, K., Negor-ev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev* 16, 919-932.

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

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. *Science* 322, 1849-1851.

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

Shearwin, K.E., Callen, B.P., and Egan, J.B. (2005). Transcriptional interference--a crash course. *Trends Genet* 21, 339-345.

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.

Shevtsov, S.P., and Dundr, M. (2011). Nucleation of nuclear bodies by RNA. *Nat Cell Biol* 13, 167-173.

Shibata, S., and Lee, J.T. (2004). Tsix transcription- versus RNA-based mechanisms in Xist repression and epige-

netic 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.

Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311, 844-847.

Siegfried, Z., Eden, S., Mendelsohn, M., Feng, X., Tsuberi, B.Z., and Cedar, H. (1999). DNA methylation represses transcription in vivo. *Nat Genet* 22, 203-206.

Sigova, A.A., Mullen, A.C., Molinie, B., Gupta, S., Orlando, D.A., Guenther, M.G., Almada, A.E., Lin, C., Sharp, P.A., Giallourakis, C.C., *et al.* (2013). Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. *Proc Natl Acad Sci U S A* 110, 2876-2881.

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.

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-chro-

mosome inactivation. *Nature* 504, 465-469.

Simon, M.D., Wang, C.I., Kharchenko, P.V., West, J.A., Chapman, B.A., Alekseyenko, A.A., Borowsky, M.L., Kuroda, M.I., and Kingston, R.E. (2011). The genomic binding sites of a noncoding RNA. *Proc Natl Acad Sci U S A* 108, 20497-20502.

Sims, R.J., 3rd, Millhouse, S., Chen, C.F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L., and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol Cell* 28, 665-676.

Sleutels, F., Zwart, R., and Barlow, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 415, 810-813.

Smith, M.A., Gesell, T., Stadler, P.F., and Mattick, J.S. (2013). Widespread purifying selection on RNA structure in mammals. *Nucleic Acids Res* 41, 8220-8236.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 484, 339-344.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nat Rev Genet* 14, 204-220.

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 Ctfc and contributes to initiation of X-chromosome inactivation. *Genetics* 189, 441-454.

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.

Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A., and Cedar, H. (1982a). Clonal inheritance of the pattern of DNA methylation in mouse cells. *Proc Natl Acad Sci U S A* 79, 61-65.

Stein, R., Razin, A., and Cedar, H. (1982b). In vitro methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells. *Proc Natl Acad Sci U S A* 79, 3418-3422.

Stephenson, R.O., Rossant, J., and Tam, P.P. (2012). Intercellular interactions, position, and polarity in establishing blastocyst cell lineages and embryonic axes. *Cold Spring Harb Perspect Biol* 4.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.

Straub, T., and Becker, P.B. (2011). Transcription modulation chromosome-wide: universal features and principles of dosage compensation in worms and flies. *Curr Opin Genet Dev* 21, 147-153.

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.

Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., Simon, I., Yakhini, Z., and Cedar, H. (2009). Developmental programming of CpG island methylation profiles in the human genome. *Nat Struct Mol Biol* 16, 564-571.

Struhl, G., and Akam, M. (1985). Altered distributions of Ultrabithorax transcripts in extra sex combs mutant embryos of *Drosophila*. *EMBO J* 4, 3259-3264.

Struhl, K. (2007). Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat Struct Mol Biol* 14, 103-105.

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. (2006a). A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol Cell* 21, 617-628.

Sun, L., Goff, L.A., Trapnell, C., Alex-



ander, R., Lo, K.A., Hacısuleyman, E., Sauvageau, M., Tazon-Vega, B., Kelley, D.R., Hendrickson, D.G., *et al.* (2013a). Long noncoding RNAs regulate adipogenesis. *Proc Natl Acad Sci U S A* 110, 3387-3392.

Sun, M., Hurst, L.D., Carmichael, G.G., and Chen, J. (2006b). Evidence for variation in abundance of antisense transcripts between multicellular animals but no relationship between antisense transcription and organismic complexity. *Genome Res* 16, 922-933.

Sun, S., DelRosario, B.C., Szanto, A., Ogawa, Y., Jeon, Y., and Lee, J.T. (2013b). Jpx RNA activates Xist by evicting CTCF. *Cell* 153, 1537-1551.

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

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.

Tabolacci, E., and Chiurazzi, P. (2013). Epigenetics, Fragile X Syndrome and Transcriptional Therapy. *Am J Med Genet Part A* 161A:2797-2808.

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.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., *et al.* (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* 16, 1779-1791.

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.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., *et al.* (2009). Conversion of 5-methylcytosine to

5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930-935.

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., and Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* 256, 640-642.

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

Takahashi, K., 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.

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.

Takahashi, N., Okamoto, A., Kobayashi, R., Shirai, M., Obata, Y., Ogawa, H., Sotomaru, Y., and Kono, T. (2009). Deletion of *Gtl2*, imprinted non-coding RNA, with its differentially methylated region induces lethal parent-origin-dependent defects in mice. *Hum Mol Genet* 18, 1879-1888.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficiz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., *et al.* (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* 158, 1254-1269.

Taverna, S.D., Ilin, S., Rogers, R.S., Tanny, J.C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L.P., Chait, B.T., *et al.* (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol Cell* 24, 785-796.

Tay, Y., Rinn, J., and Pandolfi, P.P. (2014). The multilayered complexity of ceRNA cross-talk and competition. *Nature* 505, 344-352.

Thakur, N., Tiwari, V.K., Thomassin, H., Pandey, R.R., Kanduri, M., Gondor, A., Grange,

T., Ohlsson, R., and Kanduri, C. (2004). An antisense RNA regulates the bidirectional silencing property of the *Kcnq1* imprinting control region. *Mol Cell Biol* 24, 7855-7862.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., *et al.* (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* 15, 471-487.

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

Thomson, J.P., Skene, P.J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A.R., Deaton, A., Andrews, R., James, K.D., *et al.* (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* 464, 1082-1086.

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.

Torres-Padilla, M.E., and Chambers, I. (2014). Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. *Development* 141, 2173-2181.

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689-693.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R., *et al.* (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases *Dnmt1*, *Dnmt3a* and *Dnmt3b*. *Genes Cells* 11, 805-814.

Tufarelli, C., Stanley, J.A., Garrick, D., Sharpe, J.A., Ayyub, H., Wood, W.G., and Higgs, D.R. (2003). Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet* 34, 157-165.

Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* 147, 1537-1550.

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 Bakel, H., Nislow, C., Blencowe, B.J., and Hughes, T.R. (2010). Most "dark matter" transcripts are associated with known genes. *PLoS Biol* 8, e1000371.

van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N., and Stuitje, A.R. (1990). Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* 2, 291-299.

van Dijk, E.L., Chen, C.L., d'Aubenton-Carafa, Y., Gourvennec, S., Kwapisz, M., Roche, V., Bertrand, C., Silvain, M., Legoux, P., Loeillet, S., *et al.* (2011). XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* 475, 114-117.

Vance, K.W., Sansom, S.N., Lee, S., Chalei, V., Kong, L., Cooper, S.E., Oliver, P.L., and Ponting, C.P. (2014). The long non-coding RNA *Paupar* regulates the expression of both local and distal genes. *EMBO J* 33, 296-311.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* 303, 672-676.

Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G., *et al.* (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142, 967-980.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., *et al.* (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854-858.

Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell* 136, 669-687.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 297, 1833-1837.

Voncken, J.W., Roelen, B.A., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., and van Lohuizen, M.



- (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc Natl Acad Sci U S A* 100, 2468-2473.
- Wake, N., Takagi, N., and Sasaki, M. (1976). Non-random inactivation of X chromosome in the rat yolk sac. *Nature* 262, 580-581.
- Walsh, C.P., Chaillet, J.R., and Bestor, T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet* 20, 116-117.
- Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A., and Magnuson, T. (2001). Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat Genet* 28, 371-375.
- Wang, J., Zhang, J., Zheng, H., Li, J., Liu, D., Li, H., Samudrala, R., Yu, J., and Wong, G.K. (2004). Mouse transcriptome: neutral evolution of 'non-coding' complementary DNAs. *Nature* 431, 1 p following 757; discussion following 757.
- Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., *et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472, 120-124.
- Wang, S.H., and Elgin, S.C. (2011). *Drosophila* Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. *Proc Natl Acad Sci U S A* 108, 21164-21169.
- Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* 454, 126-130.
- 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.
- Watanabe, T., Tomizawa, S., Mitsuya, K., Totoki, Y., Yamamoto, Y., Kuramochi-Miyagawa, S., Iida, N., Hoki, Y., Murphy, P.J., Toyoda, A., *et al.* (2011). Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus. *Science* 332, 848-852.
- 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 postimplantation homozygous tetraploid mouse embryos. *Genet Res* 59, 205-214.
- Weber, C.M., and Henikoff, S. (2014). Histone variants: dynamic punctuation in transcription. *Genes Dev* 28, 672-682.
- Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet* 41, 246-250.
- Wendt, K.S., and Grosveld, F.G. (2014). Transcription in the context of the 3D nucleus. *Curr Opin Genet Dev* 25, 62-67.
- 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.
- 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.
- Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855-862.
- Williams, K., Christensen, J., and Helin, K. (2012). DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep* 13, 28-35.
- Williams, L.H., Kalantry, S., Starmeyer, 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.
- Williamson, C.M., Ball, S.T., Dawson, C., Mehta, S., Beechey, C.V., Fray, M., Teboul, L., Dear, T.N., Kelsey, G., and Peters, J. (2011). Uncoupling antisense-mediated silencing and DNA methylation in the im-

printed gnas cluster. *PLoS Genet* 7, e1001347.

Willingham, A.T., Orth, A.P., Batalov, S., Peters, E.C., Wen, B.G., Aza-Blanc, P., Hogenesch, J.B., and Schultz, P.G. (2005). A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* 309, 1570-1573.

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

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

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., *et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442, 86-90.

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

Xie, W., Schultz, M.D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., Whitaker, J.W., Tian, S., Hawkins, R.D., Leung, D., *et al.* (2013). Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* 153, 1134-1148.

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. *Nat Genet* 42, 1043-1047.

Xu, C.R., Cole, P.A., Meyers, D.J., Kormish, J., Dent, S., and Zaret, K.S. (2011a). Chromatin "prepattern" and histone modifiers in a fate choice for liver and pancreas. *Science* 332, 963-966.

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

Xu, W., Edmondson, D.G., Evrard, Y.A., Wakamiya, M., Behringer, R.R., and Roth, S.Y. (2000). Loss of Gcn5l2 leads to increased

apoptosis and mesodermal defects during mouse development. *Nat Genet* 26, 229-232.

Xu, Z., Wei, W., Gagneur, J., Claudermunster, S., Smolik, M., Huber, W., and Steinmetz, L.M. (2011b). Antisense expression increases gene expression variability and locus interdependency. *Mol Syst Biol* 7, 468.

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

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. *Nat Genet* 40, 1016-1022.

Yamaji, M., Ueda, J., Hayashi, K., Ohta, H., Yabuta, Y., Kurimoto, K., Nakato, R., Yamada, Y., Shirahige, K., and Saitou, M. (2013). PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* 12, 368-382.

Yao, H., Brick, K., Evrard, Y., Xiao, T., Camerini-Otero, R.D., and Felsenfeld, G. (2010). Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev* 24, 2543-2555.

Yap, K.L., Li, S., Munoz-Cabello, A.M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M.J., and Zhou, M.M. (2010). Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* 38, 662-674.

Yildirim, E., Kirby, J.E., Brown, D.E., Mercier, F.E., Sadreyev, R.I., Scadden, D.T., and Lee, J.T. (2013). Xist RNA is a potent suppressor of hematologic cancer in mice. *Cell* 152, 727-742.

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

Yoo, E.J., Cooke, N.E., and Liebhaber, S.A. (2012). An RNA-independent linkage of noncoding transcription to long-range enhancer function. *Mol Cell Biol* 32, 2020-2029.

Yu, J., Vodyanik, M.A., Smuga-Ot-

to, 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.

Yu, W., Gius, D., Onyango, P., Muldoon-Jacobs, K., Karp, J., Feinberg, A.P., and Cui, H. (2008). Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* 451, 202-206.

Yuan, P., Han, J., Guo, G., Orlov, Y.L., Huss, M., Loh, Y.H., Yaw, L.P., Robson, P., Lim, B., and Ng, H.H. (2009). Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev* 23, 2507-2520.

Zappulla, D.C., and Cech, T.R. (2004). Yeast telomerase RNA: a flexible scaffold for protein subunits. *Proc Natl Acad Sci U S A* 101, 10024-10029.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* 25, 2227-2241.

Zhang, H.M., Liu, T., Liu, C.J., Song, S., Zhang, X., Liu, W., Jia, H., Xue, Y., and Guo, A.Y. (2014). AnimalTFDB 2.0: a resource for expression, prediction and functional study of animal transcription factors. *Nucleic Acids Res.*

Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I. (2008). Roles of the Ctr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* 15, 381-388.

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., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., and Lee, J.T. (2010). Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol Cell* 40, 939-953.

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

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., and Lee, J.T. (2008b). 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.

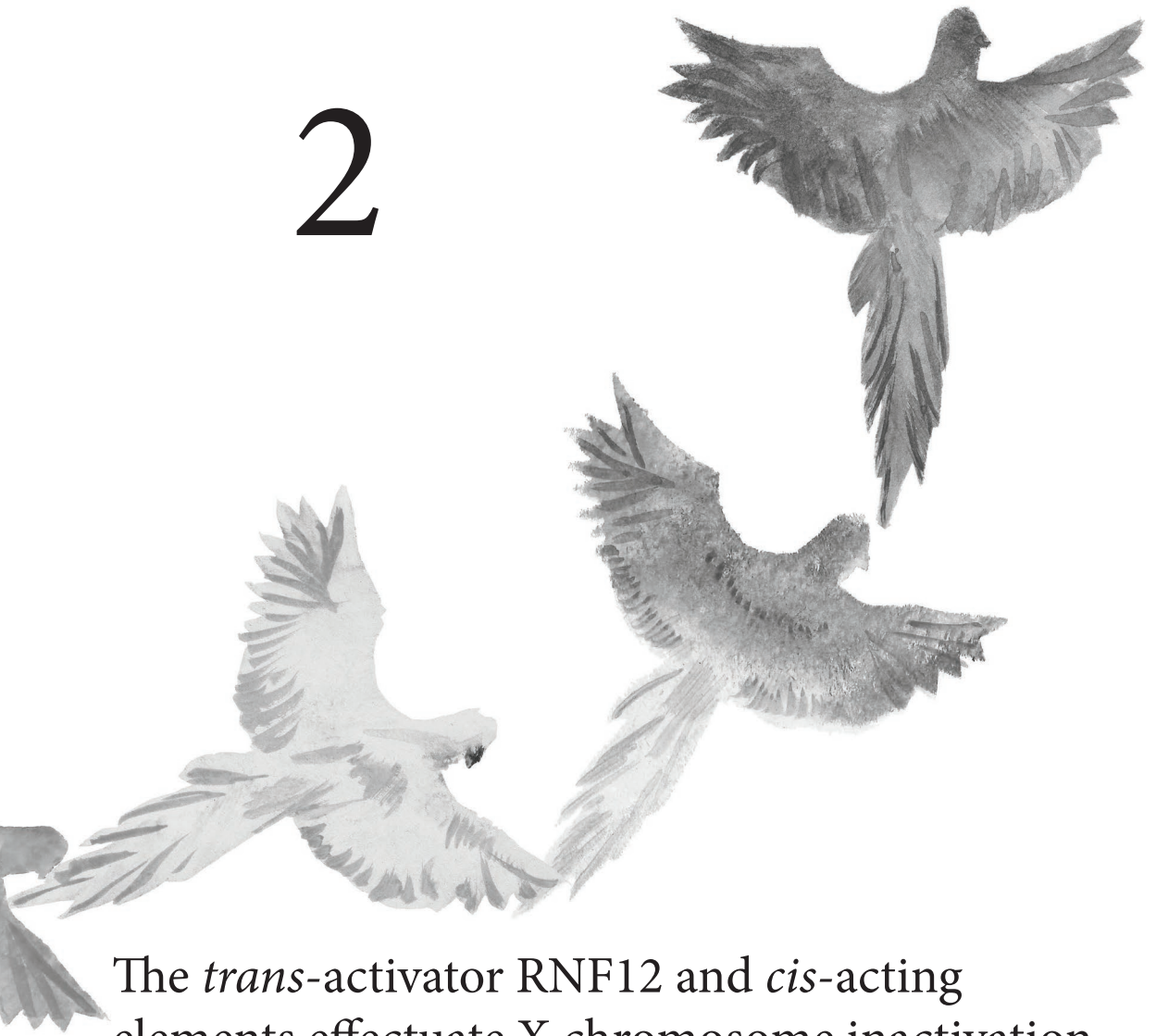
Zhou, Y., Cheunsuchon, P., Nakayama, Y., Lawlor, M.W., Zhong, Y., Rice, K.A., Zhang, L., Zhang, X., Gordon, F.E., Lidov, H.G., *et al.* (2010). Activation of paternally expressed genes and perinatal death caused by deletion of the Gtl2 gene. *Development* 137, 2643-2652.

Zhu, J., Adli, M., Zou, J.Y., Verstaapen, G., Coyne, M., Zhang, X., Durham, T., Miri, M., Deshpande, V., De Jager, P.L., *et al.* (2013). Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* 152, 642-654.

# CHAPTER



# 2



The *trans*-activator RNF12 and *cis*-acting elements effectuate X chromosome inactivation independent of X-pairing

*published in:*

Barakat, T.S., Loos, F., van Staveren, S., Myronova, E., Ghazvini, M., Grootegoed, J.A., and Gribnau, J. (2014). Mol Cell 53, 965-978.

# The *trans*-activator RNF12 and *cis*-acting elements effectuate X chromosome inactivation independent of X-pairing

Barakat, T.S.<sup>1</sup>, Loos, F.<sup>1,3</sup>, van Staveren, S.<sup>1,3</sup>, Myronova, E.<sup>1</sup>, Ghazvini, M.<sup>1,2</sup>, Grootegoed, J.A.<sup>1</sup>, and Gribnau, J.<sup>1,\*</sup>

<sup>1</sup> Department of Reproduction and Development

<sup>2</sup> Erasmus Stem Cell Institute

Erasmus MC, University Medical Center, Rotterdam, The Netherlands

<sup>3</sup>These authors contributed equally to this work

\*correspondence: [j.gribnau@erasmusmc.nl](mailto:j.gribnau@erasmusmc.nl)

Mol Cell 53, 965-978, March 20, 2014



## Abstract

X chromosome inactivation (XCI) in female placental mammals is a vital mechanism for dosage compensation between X-linked and autosomal genes. XCI starts with activation of *Xist* and silencing of the negative regulator *Tsix*, followed by *cis*-spreading of *Xist* RNA over the future inactive X chromosome (Xi). Here, we show that XCI does not require physical contact between the two X chromosomes (X-pairing), but is regulated by *trans*-acting diffusible factors. We found that the X-encoded *trans*-acting and dose-dependent XCI-activator RNF12 acts in concert with the *cis*-regulatory region containing *Jpx*, *Ftx*, and *Xpr*, to activate *Xist* and to overcome repression by *Tsix*. RNF12 acts at two subsequent steps; two active copies of *Rnf12* drive initiation of XCI, and one copy needs to remain active to maintain XCI towards establishment of the Xi. This two-step mechanism ensures that XCI is very robust and fine-tuned, preventing XCI of both X chromosomes.

## Introduction

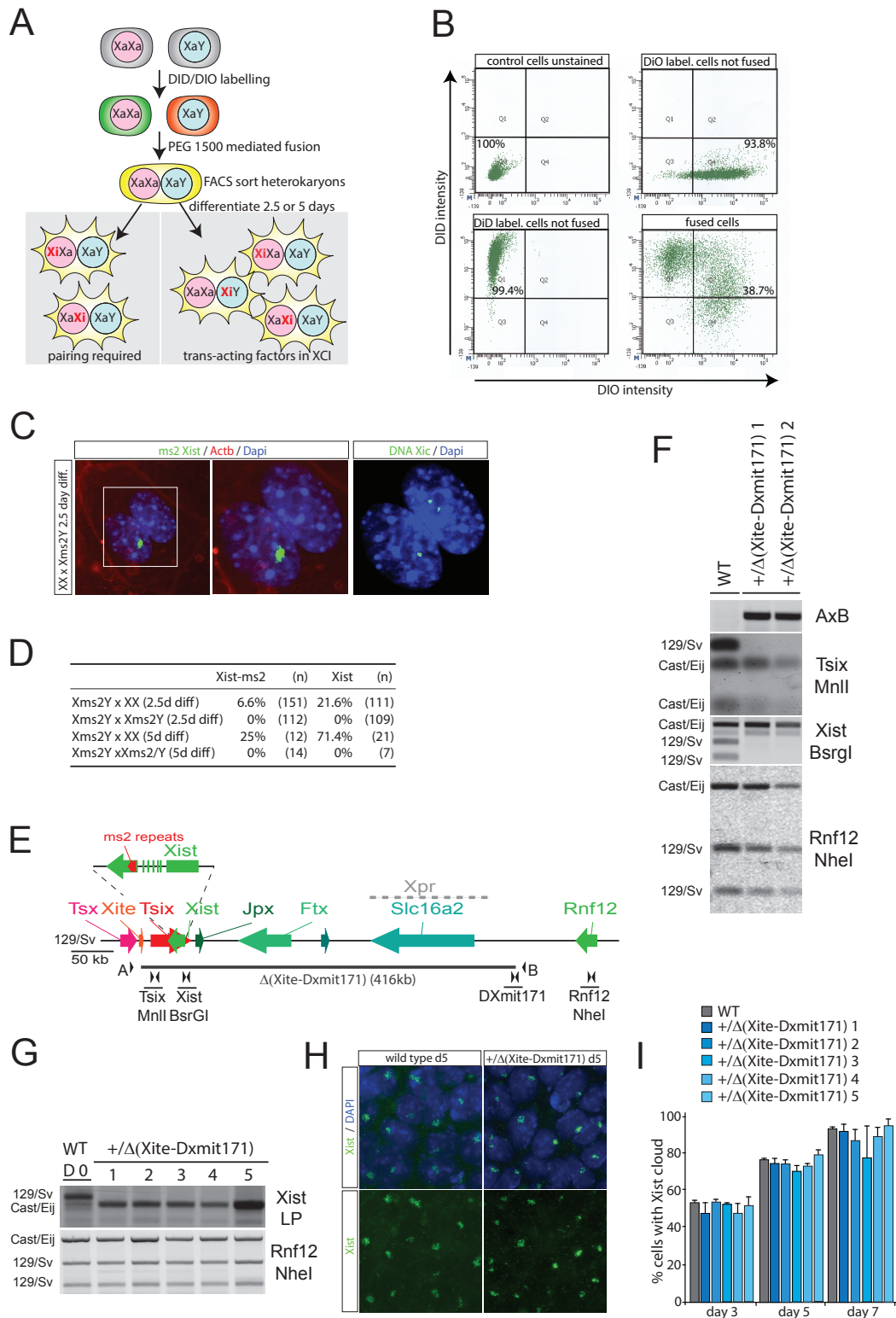
Evolution of the mammalian sex chromosomes over a period of 160 million years is associated with increased transcription of the X chromosome and with XCI, together resulting in gene dosage compensation between the X chromosome and the autosomes both in XX female and in XY male cells (Gribnau and Grootegoed, 2012; Wutz, 2011). XCI is controlled by the X inactivation centre (Xic), which is required for XCI (Augui et al., 2011). Two non-coding and *cis*-acting genes, *Xist* and *Tsix*, map to the Xic and play key roles in initiation of XCI. *Xist*

RNA coats the Xi from which it is transcribed, thereby attracting chromatin modifying complexes involved in silencing (Borsani et al., 1991; Brockdorff et al., 1991). *Tsix* is a negative regulator of *Xist*, generating transcripts in antisense direction and in conflict with *Xist* transcription, which prevents XCI of the future Xa (Lee et al., 1999). *Xist* and *Tsix* represent the master switch locus proposed to be regulated by *cis*- and *trans*-acting regulatory elements and genes involved in the repression or activation of XCI.

One of the most intriguing questions related to the XCI process is how a cell determines the number of X chromosomes and initiates XCI if more than one X chromosome is present per diploid genome. Our previous studies have indicated that initiation of XCI is mediated by at least one X-linked activator and autosomally encoded inhibitors of the XCI process (Monkhorst et al., 2008). We and others have identified X-encoded RNF12 as an important *trans*-acting XCI-activator (Barakat et al., 2011a; Jonkers et al., 2009; Shin et al., 2010). *Rnf12* is located just 500 kb upstream of *Xist* and encodes an E3 ubiquitin ligase, which catalyzes dose-dependent breakdown of REX1 by targeting REX1 for proteasomal degradation (Gontan et al., 2012). When present at an effective concentration, REX1 inhibits *Xist* transcription and stimulates *Tsix* transcription (Gontan et al., 2012; Navarro et al., 2010), thereby blocking initiation of XCI. Breakdown of REX1 is most prominent in differentiating female cells, which still have two active copies of *Rnf12*, resulting in female specific initiation of XCI.

*Rnf12* is a crucial player in initiation of XCI illustrated by the loss of XCI in female *Rnf12*<sup>-/-</sup> embryonic stem cells (ESCs) (Barakat et al., 2011a). The observation that XCI is less robust but still occurring in *Rnf12*<sup>+/-</sup> female ESCs, in contrast to the

2



## ◀ Figure 1. X pairing is not required for initiation of random XCI

(A) Male and female ESCs were DiD- and DiO-labelled, and fused to form XY-XX heterokaryons, which were FACS-sorted prior to differentiation. Initiation of XCI in the male nucleus of XY-XX heterokaryons indicates that pairing is not required for XCI initiation. (B) FACS analysis after fusion of DiD and DiO labelled XX and XY cells. (C) Left panel: Immuno-RNA FISH analysis detecting ms2 tagged *Xist* (FITC) and *Actb* (Rhodamine Red), in the XY nucleus of a XY-XX heterokaryon, differentiated for 2.5 days. Right panel shows sequential DNA-FISH of the same cell shown on the left, detecting a probe of the *Xic* (FITC). (D) Quantification of ms2 *Xist* or total *Xist* in the indicated heterokaryon experiments. N=number of heterokaryons analysed. (E) Schematic representation of part of the *Xic*. Line underneath the scheme indicates the region deleted in the +/Δ(*Xite-Dxmit171*) ESC lines. Primers and enzymes for RFLP digestion are indicated. The ms2 repeat integrated in *Xist* is shown above. (F) PCR analysis of genomic DNA from wild type parental ESCs and two +/Δ(*Xite-Dxmit171*) ESC lines with primers shown in (E). (G) Allele specific RT-PCR analysis of cDNA from undifferentiated wild type and +/Δ(*Xite-Dxmit171*) ESC lines. (H) *Xist* RNA-FISH analysis of day 5-differentiated female wild type or +/Δ(*Xite-Dxmit171*) cells. *Xist*, FITC; DNA is stained with DAPI (blue). (I) *Xist* RNA-FISH quantified at different time points after start of differentiation of female wild type and five individual +/Δ(*Xite-Dxmit171*) ESC lines. Average of two independent differentiation experiments is shown. Error-bars represent standard deviation, n>200 cells counted per time point.

virtual absence of XCI in male cells, suggested the presence of additional activators of XCI acting in parallel with RNFI2 to direct initiation of the dosage compensation process (Barakat et al., 2011b; Jonkers et al., 2009). Several non-coding genes and elements, located in close proximity to *Xist*, have been implicated to exert XCI-activator activity. These include the non-coding genes *Jpx* and *Ftx*, which are co-activated with *Xist* and up-regulated upon ES cell differentiation (Chureau et al., 2011; Tian et al., 2011). Deletion of *Jpx* has been reported to result in a severe loss of XCI, leading to massive cell death (Tian et al., 2011), and it has been suggested that *Jpx* mediates female specific *Xist* activation in *trans*, by dose-dependent eviction of CTCF at the *Xist* promoter (Sun et al., 2013). A deletion of *Ftx* negatively affects expression of both *Xist* and *Jpx* (Chureau et al., 2011), but this has been analysed only for male cells, which impedes making a distinction between *cis* and *trans* effects. Other elements located within the *Xic* have also been implicated in *trans*-activation of XCI, acting by mediating spatial movements and direct interactions of the *Xic*'s of the two X chromosomes in female cells at the onset of XCI. This X-pairing process is regulated by multiple elements including *Xite*, *Tsix* and the *Xpr* region, and requires active transcription

and the factors CTCF and OCT4 (Augui et al., 2007; Bacher et al., 2006; Donohoe et al., 2009; Xu et al., 2007; Xu et al., 2006). X-pairing might play a role in counting the number of X chromosomes present in a nucleus, and in XCI choice and initiation (Augui et al., 2007; Bacher et al., 2006; Donohoe et al., 2009; Xu et al., 2007; Xu et al., 2006), but may also be a consequence of initiation of XCI (Barakat et al., 2010).

In the present study, we have addressed the question if direct physical interaction between the two X chromosomes indeed is required for proper execution of XCI. We find that counting and initiation of XCI do not require direct physical interaction of wild type X chromosomes, indicating that XCI is regulated by diffusible *trans*-acting factors that are transported through the cytoplasm. We provide evidence that these factors may not include gene products from *Jpx* and *Ftx*, which appear to activate *Xist* mainly, if not exclusively, in *cis*, rather than acting in *trans*. This reinforces the model for a mechanism in which *Rnf12* is required and sufficient for XCI, acting in two steps to initiate XCI and establish the Xi.

## Results

### Counting and initiation of XCI in heterokaryons independent of X-pairing

Initiation of XCI may involve *trans*-acting diffusible factors in interplay with *cis*-regulatory sequences. In addition, initiation of XCI may also be directed by physical pairing of the X chromosomes (Bacher et al., 2006; Xu et al., 2006). Direct physical contact between the two X chromosomes either may be a mechanistic factor at the start of XCI initiation, or could be a consequence of the XCI process. To discriminate between these options, we studied XCI in a situation which precludes any direct interaction between a future Xi and Xa within one and the same diploid nucleus, by making use of XX-XY heterokaryons. In previous studies we found that XXXY tetraploid ESCs never initiate XCI upon differentiation, whereas XXXY tetraploid ESCs initiate XCI on one of the X chromosomes, in agreement with XCI leading to one active X per diploid genome (Monkhorst et al., 2008). Here, we generated female-male (XX-XY) ESC heterokaryons (Figure 1A), to investigate if XCI in these XX-XY heterokaryons with two diploid nuclei would proceed as observed for the XXXY tetraploid synkaryons, with random XCI of any one of the three X chromosomes. To generate XX-XY and XY-XY heterokaryons, DiD and DiO labelled cells were fused with polyethylene glycol, followed by FACS to sort double positive heterokaryons (Figure 1B), which were then differentiated. The male ESC line used for this fusion harbours an ms2 tagged *Xist* gene (Figure 1E). Immuno-RNA-FISH was performed detecting *Actb* (beta-actin) to discriminate between synkaryons and heterokaryons, and *Xist*-ms2 to identify XY nuclei that initiated

XCI. The number of heterokaryons obtained was 7.9% and 1.5%, at day 2.5 and day 5 of differentiation, respectively. In the XX-XY heterokaryons, we detected male nuclei with *Xist*-ms2 clouds, which were all found in XY-XY heterokaryons (Figure 1C and data not shown). Using sequential RNA-DNA FISH it was confirmed that in the male nucleus of the XX-XY heterokaryons the *Xist*-ms2 clouds were restricted to the X chromosome, excluding the possibility that complex heterokaryons with sex chromosomal aneuploidies had formed. We counted 21.6% and 71.4% of the XX-XY heterokaryons having one *Xist* cloud (either *Xist* or *Xist*-ms2), at day 2.5 and day 5, respectively (Figure 1D). Quantification of the *Xist*-ms2 positive nuclei showed that in approximately one out of three of these XX-XY heterokaryons, 6.6% and 25.0% at day 2.5 and day 5, respectively, it was the male nucleus which had initiated XCI (Figure 1D). This result is in full agreement with a mechanism for control of random XCI based on an independent probability for each one of the three X chromosomes to initiate XCI (Monkhorst et al., 2008). Hence, in the XX-XY heterokaryons, XCI is not restricted to XX nuclei, but is randomly distributed over all three X chromosomes present, including the one X chromosome in the male nucleus which is unable to engage in XX pairing. XCI initiation and counting in these heterokaryons appears to be under control of *trans*-acting factors which can pass across nuclear membranes.

### Deletion of X-pairing elements on one of the two X chromosomes does not obstruct XCI

The above observation led us to re-investigate whether the known X-pairing elements within the *Xic* are required for XCI. To this end, we generated ES cell lines

which lack all the known X-pairing elements, namely *Xite*, *Tsix*, and the *Xpr* region (Augui et al., 2007; Bacher et al., 2006; Donohoe et al., 2009; Xu et al., 2007; Xu et al., 2006) from one of the two X chromosomes. To delete these elements, we introduced an *Xite* targeting cassette flanked by lox sites in the 129/Sv X chromosome by homologous recombination in wild type F1 2-1 (129/Sv:Cast/Eij) ESCs (Figure 1E and Supplementary Figure 1A-B). We also introduced a neo cassette flanked by lox sites into the Dxmit171 marker located between *Xpr* and *Rnf12*, of the 129/Sv X chromosome, by homologous recombination with a BAC targeting vector (Supplementary Figure 1C-E). Next, Cre mediated deletion of the region from *Xite* up to and including the *Xpr* region of the 129/Sv X chromosome yielded +/Δ(*Xite*-Dxmit171) ESC lines, with a wild type Cast/Eij X chromosome (Figure 1F-G). Investigating these cells using *Xist* RNA-FISH, and in comparison to a wild type female F1 129/Sv:Cast/Eij ESC line, we found that the frequency and overall appearance of *Xist* cloud formation in five different +/Δ(*Xite*-Dxmit171) ESC lines was not at all impaired (Figure 1H-I). Quantitative RT-PCR showed upregulation of *Xist* RNA, and proper down-regulation of the pluripotency factor genes *Nanog* and *Klf4* upon differentiation of the +/Δ(*Xite*-Dxmit171) ESC lines compared to wild type ESC lines (Supplementary Figure 1F-H). In addition, no defects in cell proliferation, survival, morphology, and karyotype were observed for these mutant ESC lines (Supplementary Figure 2I, Supplementary Figure 2A, and data not shown). Inescapably, therefore, our findings lead to the conclusion that the known X-pairing elements can be deleted from one of the two X chromosomes in a female nucleus, without loss of mechanisms controlling X chromosome counting and initiation of XCI.

### ***Jpx*, *Ftx*, and *Xpr* cooperate with *Rnf12*, in activation of *Xist***

The observed robust initiation of XCI in the +/Δ(*Xite*-Dxmit171) ESCs indicates that *Jpx* and *Ftx*, present in only one copy in these female cells, do not act as predominant *trans*-activators of XCI, at least not as powerful as RNF12, as in *Rnf12*<sup>+/-</sup> cells the rate of XCI is strongly reduced (Barakat et al., 2011a). A relatively subtle effect of *Jpx* and *Ftx* in *trans* might be masked by the presence of two *Rnf12* alleles in the +/Δ(*Xite*-Dxmit171) ES cells. Hence, we generated female ESC lines lacking the region covering *Jpx*, *Ftx*, *Xpr*, and also *Rnf12*, on one of the two X chromosomes. To this end, *Xist*<sup>2lox/+</sup> ESCs were targeted with an *Rnf12* BAC vector introducing a lox sequence in *Rnf12* (Figure 2A-B and Supplementary Figure 2B), followed by transient expression of Cre recombinase to remove all sequences from in between *Xist* and *Jpx* up to and including *Rnf12*, yielding +/Δ(*Jpx*-*Rnf12*) ESCs in which *Xist* and a 5 kb upstream region was still present on the targeted 129/Sv X chromosome (Figure 2C and Supplementary Figure 3A).

For the +/Δ(*Jpx*-*Rnf12*) ESCs we observed a relatively strong reduction of XCI, compared to the *Rnf12*<sup>+/-</sup> female ESCs, at all time-points analysed using *Xist* RNA-FISH (Figure 2E-F and Supplementary Figure 2F). As observed for the +/Δ(*Xite*-Dxmit171) ESC lines, the undifferentiated and embryonic body (EB)-differentiated +/Δ(*Jpx*-*Rnf12*) ESCs did not display differences in morphology, cell viability or proliferation, compared to the wild type parental ESCs (Figure 2D and Supplementary Figure 2C-E), contrasting a previous study indicating massive cell death upon differentiation of *Jpx*<sup>+/-</sup> ES cells, attributed to loss of XCI (Tian et al., 2011). Although we observed some heterogeneity in the outgrowth of differenti-



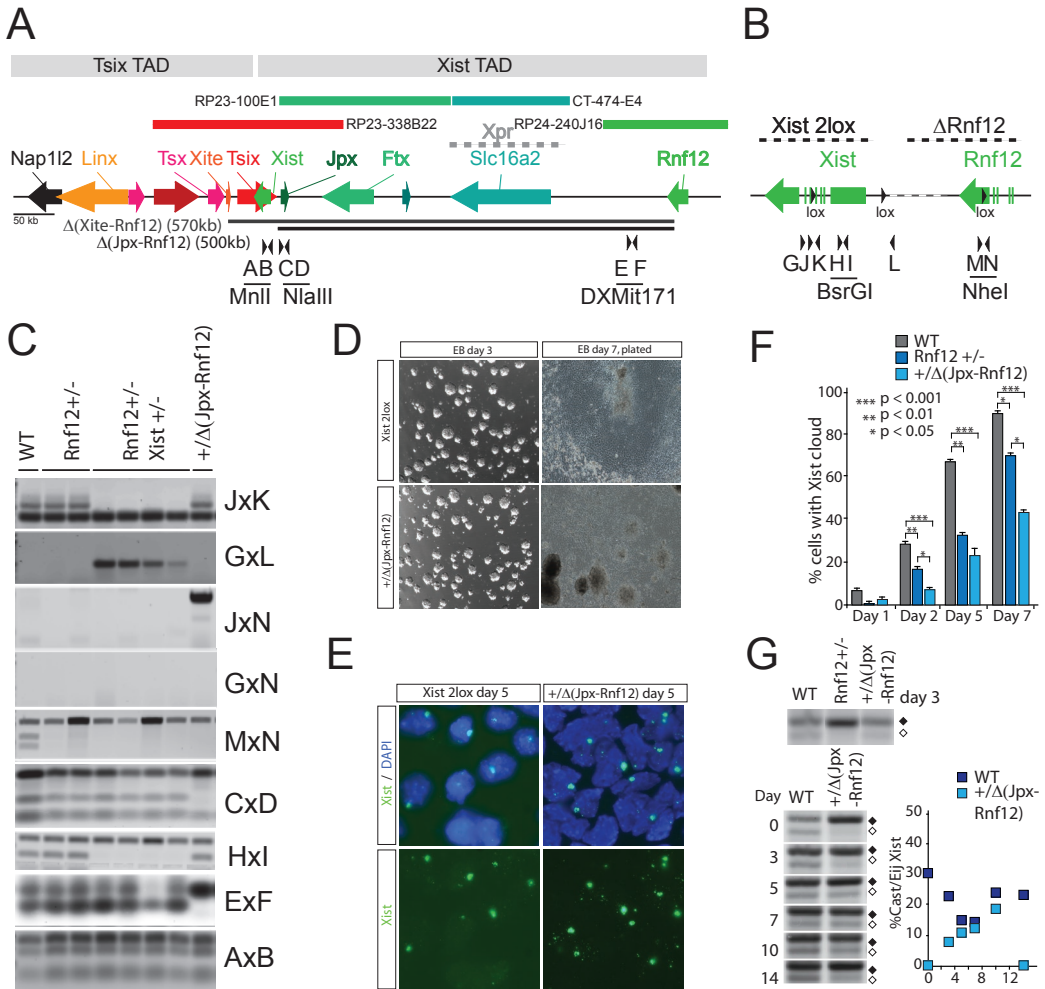


Figure 2. Generation and analysis of +/Δ(Jpx-Rnf12) ESCs

(A, B) Map showing parts of the Xic, and location of the primers used to genotype the ESC clones. Line underneath the scheme in (A) indicates the region deleted in the +/Δ(Jpx-Rnf12) and +/Δ(Xite-Rnf12) ESC lines. (C) PCR analysis on genomic DNA of Xist<sup>2lox/+</sup> Rnf12<sup>+/-</sup> ESC clones after transient Cre expression using primer combinations indicated in (A) and (B). (D) Pictures of day 3- and day 7-differentiated +/Δ(Jpx-Rnf12) and Xist<sup>2lox/+</sup> control ESCs. (E) Xist RNA FISH (FITC) on day 5-differentiated +/Δ(Jpx-Rnf12) and Xist<sup>2lox/+</sup> control ESCs (DAPI is blue). (F) Quantification of relative number of Xist clouds in +/Δ(Jpx-Rnf12), Rnf12<sup>+/-</sup>, and wild type ESCs at different time points of differentiation. Average of three independent differentiation experiments is shown, and the error bars represent standard deviation (t test, \* p<0.05, \*\* p<0.01, \*\*\* p<0.001). (G) Allele specific expression analysis of Xist by amplification of an Xist length polymorphism comparing day 3-differentiated wild type, Rnf12<sup>+/-</sup>, and +/Δ(Jpx-Rnf12) ESCs (in the top panel, the black rectangle represents 129/Sv and the open rectangle is Cast/Eij). The bottom panel and graph show the relative expression and quantification of Xist emanating from the targeted Δ(Jpx-Rnf12) 129/Sv allele and the wild type Cast/Eij allele, both analysed at different time points after differentiation.

ating +/Δ(Jpx-Rnf12) ESC subclones, we obtained no differences in Xist induction in between cell lines and clones (Supple-

mentary Figure 2C-F). In addition, allele specific PCR amplification of Xist on genomic DNA isolated from individual



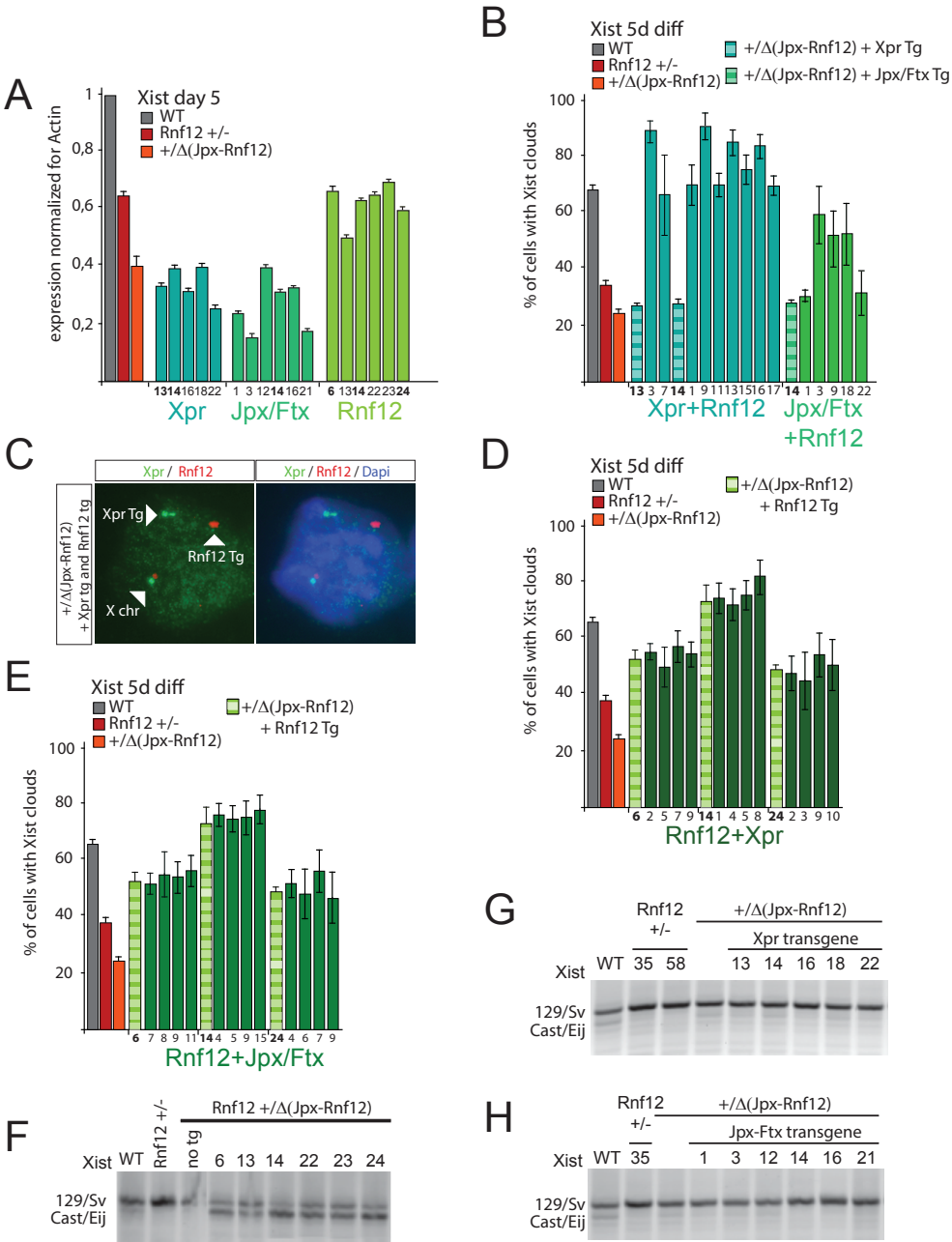
+/ $\Delta$ (Jpx-Rnf12) ESC clones indicated a stable X chromosome karyotype, without loss of an X chromosome upon differentiation. Also the original +/ $\Delta$ (Jpx-Rnf12) ESC line used in all further experiments displayed a stable karyotype, as assessed by regular karyotyping and combined RNA-DNA-FISH (Supplementary Figure 3C and data not shown). Importantly, competition experiments with differentiating Rnf12<sup>-/-</sup> and wild type ESCs did not indicate significant selection against Rnf12<sup>-/-</sup> cells, which only sporadically initiate XCI, indicating that loss of XCI does not affect cell survival *in vitro* (data not shown). Hence, it seems unlikely that a single copy of *Jpx* in female cells is incompatible with cell viability. In agreement with this, +/ $\Delta$ (Jpx-Rnf12) ES cells injected in blastocysts could contribute to adult chimaeras, presenting with high coat-color contribution (Supplementary Figure 3B). However, as XCI in the +/ $\Delta$ (Jpx-Rnf12) ES cells is more strongly reduced, compared to Rnf12<sup>+/-</sup> ESCs (Figure 2F), these findings suggest that *Jpx*, and *Ftx*, and possibly *Xpr*, act in conjunction with Rnf12, in the *cis*- or *trans*-activation of XCI. On the other hand, the results indicate that XCI is still initiated in the presence of only one copy of Rnf12, *Jpx*, *Ftx*, and the *Xpr*, albeit at a lower frequency.

### ***Jpx*, *Ftx*, and *Xpr* act in *cis* in the activation of *Xist***

Activation of XCI by RNF12 is mediated in *trans* through proteasomal degradation of REX1, which acts as a repressor of *Xist* (Gontan et al., 2012). To test whether the reduction of XCI in +/ $\Delta$ (Jpx-Rnf12) cells compared to Rnf12<sup>+/-</sup> cells represents loss of control either in *cis* or in *trans*, we analysed allele specific *Xist* expression at different time points of ESC differentiation. To this end, we used

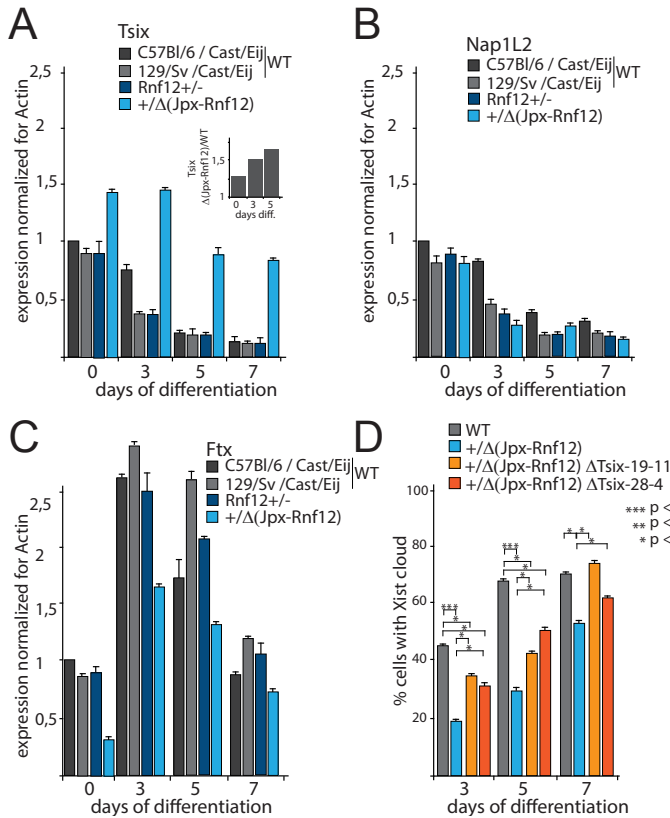
ESCs carrying X chromosomes from two genetic backgrounds, Cast/Eij and 129/Sv. In wild type female F1 Cast/Eij / 129/Sv ESCs, XCI is skewed with a 65%:35% preferential inactivation of the 129/Sv X chromosome (Figure 2G). When the 129/Sv X chromosome was modified to carry a  $\Delta$ Rnf12 deletion, the Rnf12<sup>+/-</sup> female ESCs, upon differentiation, showed completely skewed XCI of the  $\Delta$ Rnf12 129/Sv X chromosome (Barakat et al., 2011b; Jonkers et al., 2009) (Figure 2G). This might be explained, because inactivation of the wild type Cast/Eij X chromosome in these cells would lead to an Rnf12 null cell, which may be selected against, or may be unable to maintain *Xist* expression due to the absence of RNF12. When we investigated XCI skewing for cells carrying the large  $\Delta$ (Jpx-Rnf12) deletion on the 129/Sv X chromosome, we detected a relative increase in *Xist* from the wild type Cast/Eij X chromosome compared to Rnf12<sup>+/-</sup> cells, at day 3 of differentiation (Figure 2G). RNA-DNA FISH analysis detecting *Xist* RNA and the *Xpr* region, not present on the  $\Delta$ (Jpx-Rnf12) allele revealed that 18% of the cells showed an *Xist* cloud on the wild type Cast/Eij X chromosome after 7 days of differentiation (Supplementary Figure 3C). This expression of *Xist* RNA emanating from the wild type allele was lost at day 14 of differentiation (Figure 2G). The transient Cast/Eij derived *Xist* expression which we observed in the +/ $\Delta$ (Jpx-Rnf12) cells may point to a role for the *Jpx-Rnf12* region in the activation of *Xist* in *cis*, leading to a further reduction in total *Xist* expression in the +/ $\Delta$ (Jpx-Rnf12) cells and therefore a relative increase in *Xist* RNA transcribed from the wild type Cast/Eij X chromosome.

To test this, we introduced BAC transgenes covering either Rnf12 or both *Jpx* and *Ftx* (*Jpx/Ftx*), into the +/ $\Delta$ (Jpx-Rnf12) ESCs (Figure 2A). We also introduced a BAC covering the *Xpr* into these



**Figure 3. Rnf12, but not Xpr and Jpx/Ftx activate XCI in trans**

(A) RT-qPCR examining *Xist* expression normalized to *Actin*, for *Xpr*, *Jpx/Ftx*, and *Rnf12* transgenic +/-Δ(*Jpx-Rnf12*) ESC clones differentiated for 5 days, compared to wild type female parental cells, *Rnf12*<sup>-/-</sup> cells and +/-Δ(*Jpx-Rnf12*) cells. Error bars represent standard deviation (clones used in double transgenic studies are indicated in bold). (B) Quantification of the percentage of *Xpr+Rnf12* and *Jpx/Ftx+Rnf12* double transgenic +/-Δ(*Jpx-Rnf12*) ESCs with an *Xist* cloud at day 5 of differentiation. Error bars represent the 95% confidence interval, n> 200 cells per ESC line. Also shown are the percentage of *Xist* clouds in the founder cell lines with the *Xpr* and *Jpx/Ftx* transgenes alone (dashed bars), and wild type, *Rnf12*<sup>-/-</sup>,



**Figure 4 The Jpx-Rnf12 region acts in cis to repress Tsix**

(A) RT-qPCR expression analysis of *Tsix* normalized to *Actin*, at different time points after initiation of differentiation of C57Bl/6 / Cast/Eij and 129/Sv / Cast/Eij wild type, Rnf12<sup>+/-</sup> and +/Δ(Jpx-Rnf12) ESCs. Error bars represent standard deviation. Inset: allele specific expression ratio of *Tsix* transcribed from the 129/Sv (Δ(Jpx-Rnf12) or wild type) and Cast/Eij (wild type) alleles in +/Δ(Jpx-Rnf12) versus wild type ESCs at different time points of differentiation. (B,C) As (A), but now *Nap1L2* (B) and *Ftx* (C) are assessed. (D) Quantification of relative number of Xist RNA clouds in wild type, +/Δ(Jpx-Rnf12), and two +/Δ(Jpx-Rnf12) Δ*Tsix* ES cell lines, at different time points after start of differentiation. Average of three independent differentiation experiments is shown. Error bars represent standard deviation, n>200 per time point (t test, \* p<0.05, \*\* p<0.01, \*\*\* p<0.001).

cells. Positive clones were identified by qPCR on genomic DNA (data not shown), and by RFLP RT-qPCR analysis to verify transgenic expression of *Ftx*, *Jpx*, and *Rnf12* (Supplementary Figure 4A-C). For the *Jpx/Ftx* transgenic ESC lines we found that three of the six *Jpx/Ftx* expressed both *Ftx* and *Jpx*, and that three clones expressed *Jpx* alone. *Slc16a2*, which overlaps with the *Xpr* region, is not expressed in our ES cells, and verification of *Xpr* positive clones was therefore performed by DNA FISH (Supplementary Figure 3D).

For day 5 differentiated cells, using qRT-PCR, we observed upregulation of *Xist* expression only in the +/Δ(Jpx-Rnf12) ESC clones carrying an *Rnf12* transgene, but no consistent upregulation was found for the *Jpx/Ftx* or *Xpr* transgenic cell lines (Figure 3A and Supplementary Figure 3E-H). Introduction of the *Rnf12* transgene as a second transgene, next to either *Jpx/Ftx* or *Xpr*, resulted in upregulation of *Xist* in most of these double transgenic +/Δ(Jpx-Rnf12) ESC lines (Figure 3B-C), with a positive correlation between *Xist*

and +/Δ(Jpx-Rnf12) ESC lines. Averages of three independent differentiation experiments for these controls are shown; error bars represent standard deviation, n>200 per cell line. (C) DNA-FISH analysis on Δ(Jpx-Rnf12), *Xpr+Rnf12* double transgenic undifferentiated ESCs, with *Xpr* (FITC) and *Rnf12* (Rhodamine Red) probes (FISH signals are indicated with triangles). DNA is stained with DAPI (blue). (D) As in (B), but now *Rnf12+Xpr* double transgenic +/Δ(Jpx-Rnf12) cell lines are shown. (E) As (D), but now *Rnf12+Jpx/Ftx* double transgenic +/Δ(Jpx-Rnf12) cell lines are assessed. (F) Allele specific expression of *Xist* in day 5-differentiated *Rnf12* transgenic +/Δ(Jpx-Rnf12) clones. Upper band represents *Xist* from the 129/Sv allele; lower band represents *Xist* from the Cast/Eij X chromosome. (G) As (F), but now *Xpr* transgenic +/Δ(Jpx-Rnf12) clones are assessed. (H) as (F), but here for *Jpx/Ftx* transgenic +/Δ(Jpx-Rnf12) clones.

upregulation and the expression of transgenic *Rnf12* (Supplementary Figure 5A-B and Supplementary Table 1). A predominant effect of RNF12 is also indicated by a reciprocal experiment, where we did not find an additional effect of either *Jpx/Ftx* or *Xpr* transgenes in  $+\Delta(\text{Jpx-Rnf12})$  ESC cell lines harbouring an *Rnf12* transgene (Figure 3D-E and Supplementary Figure 5E-H). Hence, our results indicate that the region encompassing *Jpx*, *Ftx* and *Xpr* does not have a pronounced *trans*-regulating function in XCI, but is mainly involved in activation of *Xist* in *cis*, in contrast to the prominent *trans*-acting activity of RNF12.

### ***Jpx*, *Ftx*, and *Xpr* lower the threshold for RNF12-mediated activation of *Xist***

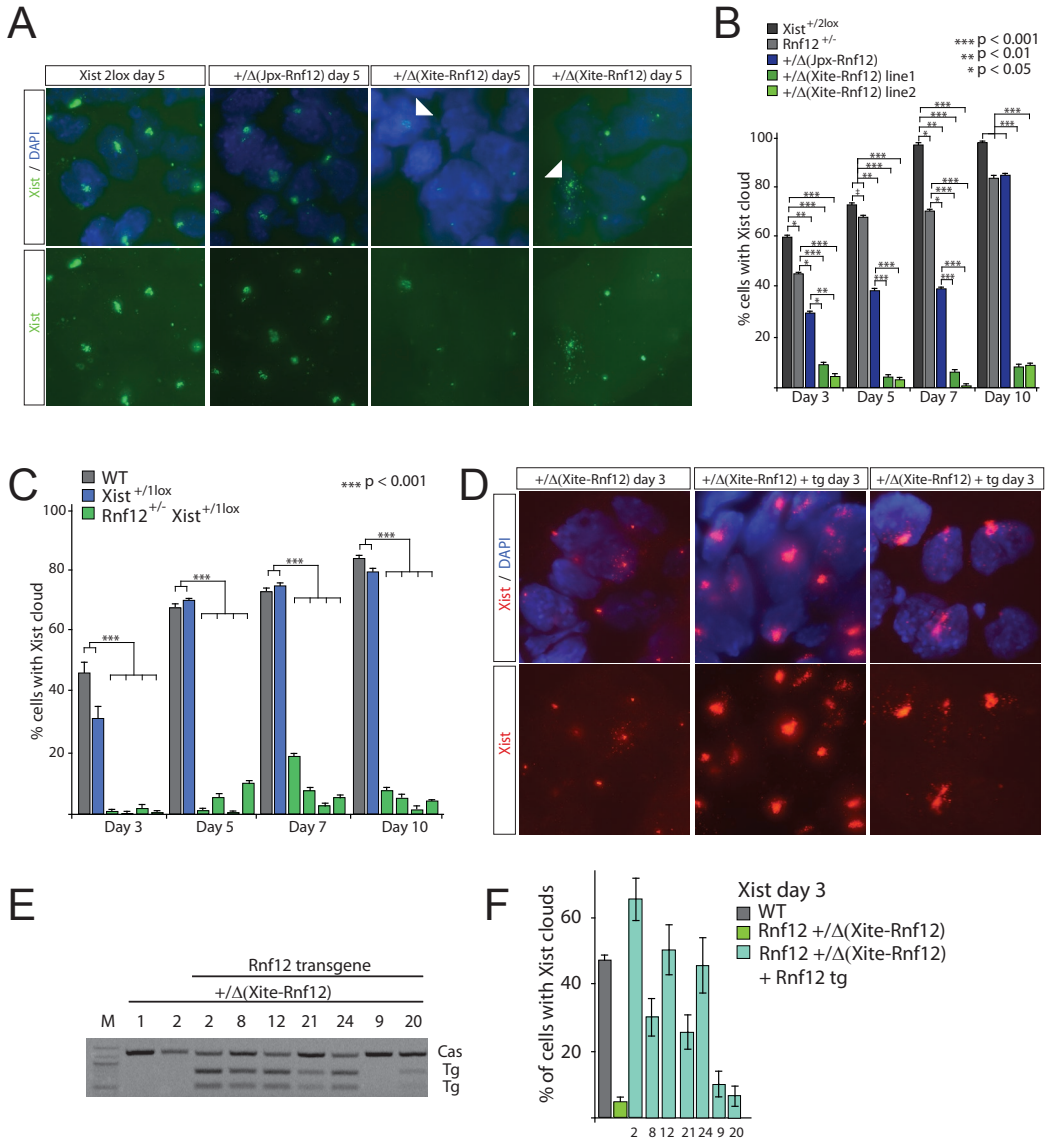
Our findings predict that in a cell carrying a mutant X chromosome lacking a significant part of the *cis*-regulatory region of the *Xic*, next to a wild type X chromosome, *trans*-activation by RNF12 of the *Xist* gene located on the mutant X will be severely compromised. Indeed, when we determined the allelic origin of *Xist* RNA in *Rnf12* transgenic  $+\Delta(\text{Jpx-Rnf12})$  day 5 differentiated ESCs, we found preferential upregulation of *Xist* expression from the Cast/Eij wild type X chromosome (Figure 3F and Supplementary Figure 5B). Combination of the quantitative RT-PCR data with allele specific expression data indicates that the upregulation of *Xist* can be solely attributed to increased activation of the Cast/Eij *Xist* allele (Supplementary Figure 5C-D). Such an effect was not observed upon expression of *Jpx/Ftx* and *Xpr* transgenes in  $+\Delta(\text{Jpx-Rnf12})$  ESC lines (Figure 3G-H), in concordance with a lack of *trans*-activity of these regions.

What could be an explanation for the lack of RNF12 mediated activation of XCI on the mutant X chromosome? To

study this in more detail, a relevant observation was made for the  $+\Delta(\text{Jpx-Rnf12})$  female ESCs at different time points of differentiation, where RT-qPCR analysis indicated upregulation of *Tsix* (Figure 4A), no change in expression of *Nap1L2*, a gene located 300 kb centromeric to *Tsix* (Figure 2A and Figure 4B), and reduced *Ftx* expression (Figure 4C). Such an upregulation of *Tsix* was absent in *Rnf12*<sup>+/-</sup> cells. Allele specific expression analysis indicated an increase of expression of *Tsix* from the  $\Delta(\text{Jpx-Rnf12})$  allele (Figure 4A), suggesting a role for the *Jpx-Rnf12* region in overcoming *Tsix*-mediated repression of *Xist*, possibly through a *cis*-acting mechanism that involves co-activation of *Xist*, or by direct repression of *Tsix*. Abolishing the function of *Tsix* on the mutant  $\Delta(\text{Jpx-Rnf12})$  allele might rescue the XCI phenotype in the  $+\Delta(\text{Jpx-Rnf12})$  female ESCs. To test this, we replaced *Tsix* major exon 1 linked to the  $\Delta(\text{Jpx-Rnf12})$  allele by a mCherry open reading frame (Supplementary Figure 6A-D). Analysis of XCI during differentiation of these  $+\Delta\text{Tsix}\Delta(\text{Jpx-Rnf12})$  ESCs indicated that premature abortion of *Tsix* transcription indeed restored initiation of XCI in *cis* on the  $\Delta(\text{Jpx-Rnf12})$  allele (Figure 4D and 6B). Hence, when *Jpx*, *Ftx*, and *Xpr* are setting up a *cis*-environment, the results indicate that this *cis*-action overcomes *Tsix*-mediated repression of *Xist* activation. If such a *cis*-environment is lacking, *Tsix* expression will be enhanced, and this will result in a strong inhibition of *Xist*.

### **RNF12- mediated feedback is sufficient for proper initiation of XCI**

Our findings underscore the important role of *cis*-interactions of elements and genes in the activation of *Xist* or *Tsix*. However, the *cis*-interacting landscape cannot distinguish between male



**Figure 5. RNF12 is required for initiation of XCI and establishment of Xi**

**(A)** *Xist* RNA FISH (FITC) on day 5-differentiated wild type *Xist*<sup>+2lox</sup>, *+/Δ(Jpx-Rnf12)*, and two *+/Δ(Xite-Rnf12)* ESC lines (DAPI is blue). **(B)** Quantification of relative number of *Xist* clouds during a differentiation time course (days 3, 5, 7 and 10) in cells shown in (A) and in *Rnf12*<sup>+/-</sup> cells. Average of two independent differentiation experiments is shown. Error bars represent standard deviation, *n*>200 cells per time point (t test, \* *p*<0.05, \*\* *p*<0.01, \*\*\* *p*<0.001). **(C)** As in (B), but now female wild type *Xist*<sup>+2lox</sup>, *Xist*<sup>+1lox</sup>, and four *Xist*<sup>+1lox</sup> *Rnf12*<sup>+/-</sup> cell lines are shown. **(D)** *Xist* RNA FISH (Rhodamine Red) on day 3-differentiated *+/Δ(Xite-Rnf12)* ESCs and two *Rnf12* transgenic *+/Δ(Xite-Rnf12)* ES cell lines (DAPI is blue). **(E)** Allele-specific RT-PCR detecting *Rnf12* in two *+/Δ(Xite-Rnf12)* ES cell lines, and *Rnf12* transgenic *+/Δ(Xite-Rnf12)* ES cell lines. Upper band (Cas) represents Cast/Eij-derived *Rnf12* RNA from the wild type X chromosome. Lower two bands (Tg) represent *Rnf12* RNA derived from the autosomally integrated 129/Sv *Rnf12* BAC. **(F)** Quantification of RNA-FISH results shown in (D), for female wild type *Xist*<sup>+2lox</sup>, *+/Δ(Xite-Rnf12)* ESCs, and *Rnf12* transgenic *+/Δ(Xite-Rnf12)* ES cell lines, differentiated for three days. Error bars represent 95% confidence interval. >200 cells analysed per cell line.



► **Figure 6. *Rnf12* is expressed from the Xa throughout XCI**

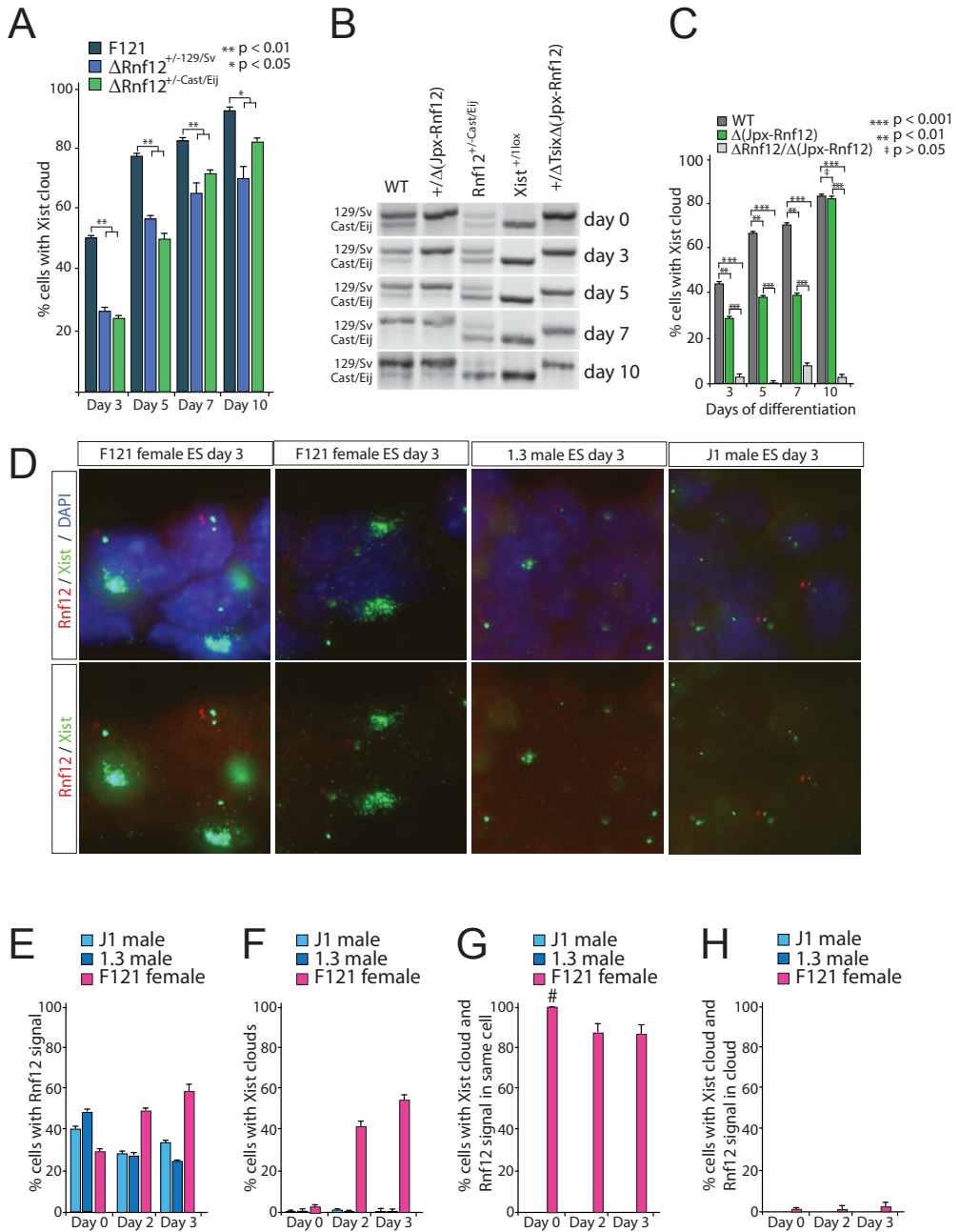
(A) Percentage of cells with *Xist* clouds at different time points of differentiation, in wild type and *Rnf12*<sup>-/-</sup> cells with a targeted 129/Sv or Cast/Eij allele. Average of three independent differentiation experiments is shown. Error bars represent standard deviation, n>200 cells per time point (t test, \* p<0.05, \*\* p<0.01). (B) Allele specific expression analysis of *Xist* detecting a length polymorphism, comparing wild type, +/Δ(Jpx-*Rnf12*), *Rnf12*<sup>+/-</sup> (Cast/Eij), *Xist*<sup>+/-</sup>/lox, +/ΔTsixΔ(Jpx-*Rnf12*) ES cell lines at different time points of differentiation. Upper band represents 129/Sv derived *Xist*, whereas lower band represents Cast/Eij derived *Xist*. (C) Quantification of relative number of wild type, Δ(Jpx-*Rnf12*), and Δ*Rnf12*/Δ(Jpx-*Rnf12*) ESCs with *Xist* clouds during differentiation (day 3, 5, 7 and 10). Average of three independent differentiation experiments is shown. Error bars represent standard deviation, n>200 cells per time point (t test, \*p>0.05, \*\* p<0.01 \*\*\*< 0.001). (D) RNA-FISH detecting *Xist* (FITC) and *Rnf12* transcription foci (Rhodamine Red) in day 3-differentiated female (F121) and male (1.3 and J1) wild type ESCs. (E) Quantification of *Rnf12* transcription foci in female (F121) and male (1.3 and J1) wild type ESCs during a short differentiation time course (day 0, 2 and 3). Shown are averages of two independent differentiation experiments, and plotted is the percentage of cells with *Rnf12* transcription foci. Error bar represents standard deviation, n>200 per time point and ESC line. (F) As in (E), but depicting the percentage of *Xist* clouds. (G) As in (E), showing the percentage of cells with an *Xist* coated Xi and a *Rnf12* transcription focus associated with the Xa. (# only few cells showed an *Xist* cloud at day 0 of differentiation). (H) As in (E), but showing the percentage of cells with a *Xist* cloud co-localizing with a *Rnf12* transcription focus.

XY and female XX cells. In contrast, the *trans*-acting factor RNF12, and putative other XCI-activators, are instrumental in the determination of the number of X chromosomes present in a nucleus, and subsequent initiation of XCI if more than one X is present per diploid genome (Barakat et al., 2010; Monkhorst et al., 2008). To test whether *Rnf12* is essential for the initiation of XCI in our Δ(Jpx-*Rnf12*) cells, we removed the second *Rnf12* allele by BAC-mediated gene targeting (Supplementary Figure 6G-H). Differentiation of these ESC lines showed severe down-regulation of XCI at all time points compared to wild type controls (Figure 6C). This reinforces that RNF12 is a key *trans*-acting activator of XCI, and also indicates that its presence might be continuously required for XCI to be established.

To test this continuous requirement, we generated ESC lines in which *Xist* and *Rnf12* are removed in *cis*. In these cell lines, *Rnf12* is expressed from the wild type X chromosome prior to and during the early stages of XCI. This expression will be lost when XCI is initiated on that wild type chromosome, and if expression of RNF12 is continuously required for maintaining the XCI initiation, such *Rnf12* silencing should result in a failure of

maintaining XCI. To remove *Xist*, we first introduced a *Xite* targeting cassette into the +/Δ(Jpx-*Rnf12*) ESC line to allow Cre mediated deletion of the region from *Xite* through to and including part of *Rnf12* (Supplementary Figure 7A). Cre-mediated loopout of the intervening sequences in properly targeted clones yielded +/Δ(*Xite*-*Rnf12*) ESC lines (Supplementary Figure 7B-C). *Xist* RNA-FISH indicated that *Xist* cloud formation in differentiating ESC lines was severely compromised by this heterozygous deletion (Figure 5A-B). In addition, the few *Xist* clouds present looked dispersed, as if *Xist* targeting to the Xi was disturbed. This cannot be explained by an absence of all elements involved in X-pairing, as +/Δ(*Xite*-Dxmit171) ESCs undergo normal XCI (Figure 1). Furthermore this cannot be explained by loss of one *Rnf12* allele, because the XCI phenotype is more pronounced than what we observed in the +/Δ(Jpx-*Rnf12*) cells. Hence, this might suggest that RNF12 is continuously required to establish the Xi.

To preclude a possible effect of the combined deletion of many elements on the XCI phenotype, and to test whether RNF12-mediated feedback is indeed sufficient to prevent XCI on the Xa in XaXi cells, we generated *Xist*<sup>+/-</sup>/lox *Rnf12*<sup>+/-</sup> *cis*



compound knockout ESC lines with single gene mutations located on the same X chromosome. Targeting of *Xist*<sup>2lox/+</sup> ESCs with an *Rnf12*-lox cassette targeting the same 129/Sv allele was verified, followed by Cre mediated loopout of *Xist* (Figure

2B-C). XCI analysis of several *Xist*<sup>+1lox</sup> *Rnf12*<sup>+/-</sup> ESC lines indicated a severely reduced percentage of *Xist* clouds, compared to wild type and *Xist*<sup>+1lox</sup> *Rnf12*<sup>+/-</sup> ESC lines (Figure 5C). In addition, as in the +/Δ(*Xite*-*Rnf12*) cells, *Xist* clouds in

**2** *Xist*<sup>+/-1lox</sup> *Rnf12*<sup>+/-</sup> cells were dispersed, as if *Xist* could not properly accumulate, or this might represent cells which have accumulated *Xist* followed by a down-regulation of *Xist* expression. Control cells, in which *Xist* was mutated on the 129/Sv X chromosome and *Rnf12* was mutated on the Cast/Eij X chromosome, showed normal XCI of the X chromosome harbouring the *Rnf12* mutation (data not shown). Therefore, the results of both the  $+/Δ(Xite-Rnf12)$  and the *Xist*<sup>+/-1lox</sup> *Rnf12*<sup>+/-</sup> ESCs suggest that RNF12 is continuously required to establish the Xi. Indeed, when we introduced an autosomal copy of *Rnf12* in  $Δ(Xite-Rnf12)$  ESCs, these EB differentiated *Rnf12* rescued  $Δ(Xite-Rnf12)$  ESCs showed robust upregulation of *Xist* with normal appearing *Xist* clouds localized at the Xi (Figure 5D-F and Supplementary Table 1C). In contrast, introduction of a BAC covering either *Ftx* and *Jpx* or the *Xpr* did not result in restoration of *Xist* expression in these cells (Supplementary Figure 7D-E).

A dual role for RNF12 in initiation of XCI and establishment of the inactive state of the Xi, explains the reported skewing towards inactivation of the X chromosome carrying an *Rnf12* mutation in Cast/Eij/129/Sv heterozygous *Rnf12*<sup>+/-129/Sv</sup> cells (Barakat et al., 2011a; Jonkers et al., 2009). In this situation, the wild type copy of *Rnf12* remains active, which predicts that, in ESC lines with a mutated Cast/Eij *Rnf12* allele, the Cast/Eij X chromosome will be preferentially inactivated. Indeed, when we obtained such *Rnf12*<sup>-Cast/Eij/+</sup> ESCs through BAC-mediated targeting, we observed reciprocal skewing in these *Rnf12*<sup>-Cast/Eij/+</sup> ESCs compared to *Rnf12*<sup>+/-129/Sv</sup> cells upon differentiation (Figure 6A-B and Supplementary Figure 6E-F). From this, one would expect that, in differentiating wild type ES cells which have started XCI, *Rnf12* expression is always maintained from the future Xa. This was tested using RNA-FISH detecting *Xist* and *Rnf12*.

We found that almost 90% of the female cells with an *Xist* cloud also contained an *Rnf12* transcription signal located on the Xa (Figure 6D and 6E, G). Also, we found that *Rnf12* transcription was almost never observed within an *Xist* cloud, even during early stages of *Xist* cloud formation (Figure 6D and 6F-H). This confirms that there is a highly efficient feedback mechanism, to ensure that *Rnf12* does not escape from XCI and that RNF12 is down-regulated as soon as XCI is initiated, preventing inactivation of one X too many.

## Discussion

The genomic region around the *Xist* gene is composed of a complex series of elements and genes, which act together to achieve well-controlled transcription of *Xist* RNA, leading to inactivation of either one of the two X chromosomes in female cells of placental mammals within a defined developmental time window. Herein, we have studied the initiation of random XCI in mouse ES cells, and the first steps towards establishment of the Xi. Our findings reveal a mechanism for XCI which involves *cis*-action of the genomic region nearby *Xist* in conjunction with *trans*-action of X-linked gene products. This mechanism provides dosage-mediated initiation of XCI and feedback, protecting against silencing of too many X chromosomes.

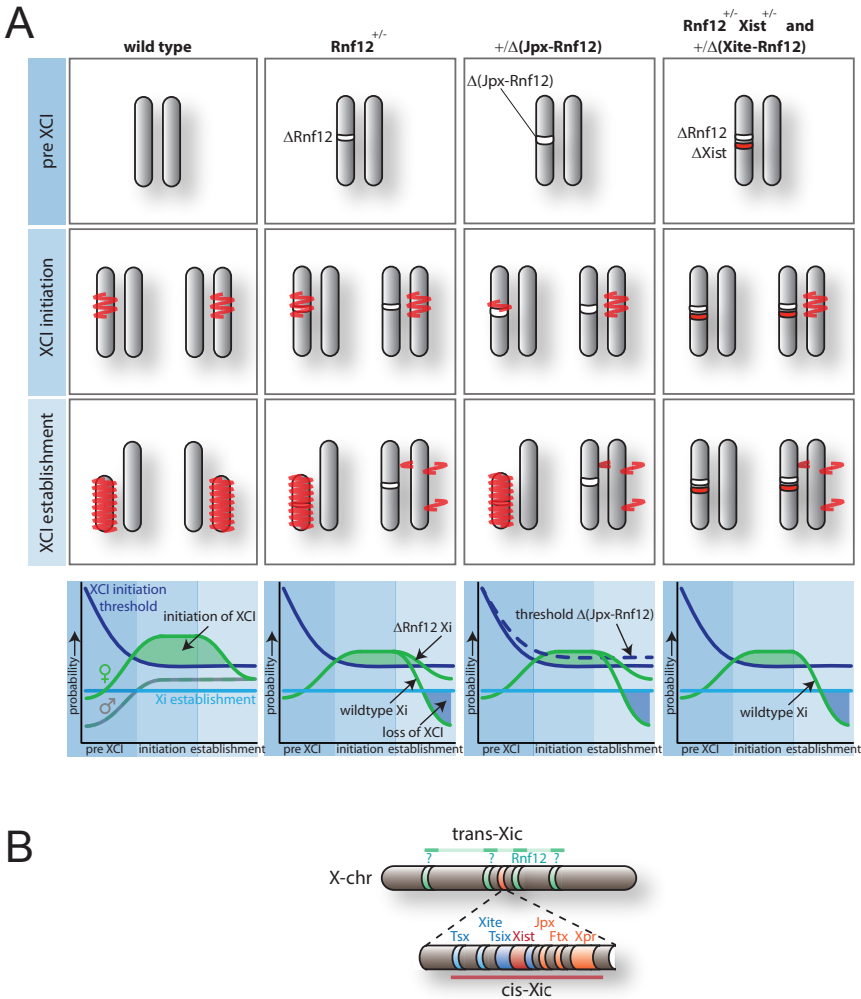
## Co-activation of *Xist*, *Jpx* and *Ftx*

*Cis*-acting elements and genes are important regulators of *Xist* activity by regulating *Xist* directly or indirectly through *Tsix*. However, *cis*-acting information will not allow the cell to discriminate between male and female, which is determined by

the dose-dependent action of *trans*-acting XCI-activators and inhibitors. In recent years, several *cis*-acting sequences have been identified to regulate *Xist* or *Tsix*, and 3C and 5C studies have indicated that *Xist* and *Tsix* reside in different neighbouring chromatin interaction domains, or topologically associated domains (TAD) (Nora et al., 2012; Tsai et al., 2008). *Jpx*, *Ftx* and *Xpr* reside in the *Xist* TAD, and previous experimental data confirmed a positive regulatory role for these elements in activation of *Xist* (Augui et al., 2007; Chureau et al., 2011; Sun et al., 2013; Tian et al., 2011). Our study involving  $+\Delta(\text{Xite-Dxmit171})$  and  $+\Delta(\text{Jpx-Rnf12})$  ESCs indicates that this part of the *Xist* TAD activates *Xist* in *cis*. The *Rnf12* gene is located near the telomeric end of the *Xist* TAD, but it is unlikely that this gene plays a role in *cis*-activation of XCI, as *Rnf12*<sup>+/-</sup> cells preferentially inactivate the  $\Delta\text{Rnf12}$  X chromosome. In addition, the XCI phenotype in the  $+\Delta(\text{Jpx-Rnf12})$  cells, which also affected the wild type allele, can be rescued by introduction of a randomly integrated *Rnf12* transgene. Evidence for a strong *trans*-activating role for RNF12 in *Xist* activation is provided by many different experimental approaches. If other genes within the *Xist* TAD have *trans*-activating activity, such an activity is weak, compared to the effect exerted by RNF12. No detectable *trans*-acting activity of *Jpx* and *Ftx* was found in our previous studies (Jonkers et al., 2009). The present transgene studies involving *Jpx*, *Ftx* and *Xpr* introduced in  $+\Delta(\text{Jpx-Rnf12})$  and  $+\Delta(\text{Xite-Rnf12})$  cells did not reveal any *trans*-acting activity of the region deleted in  $\Delta(\text{Jpx-Rnf12})$  ES cells. This, and our finding that XCI is unaffected in  $+\Delta(\text{Xite-Dxmit171})$  ESCs, contrasts two recent studies suggesting a role for *Jpx* in the regulation of XCI in *trans* (Sun et al., 2013; Tian et al., 2011). In comparing the outcome of different studies, one needs to take into account that the genetic background of ESCs might

play some role. On the other hand, we arrive at our conclusion based on a comparison between the *trans* effects of *Jpx/Ftx* and *Jpx* alone in comparison to that of *Rnf12*, for ESCs on the same genetic background. Such a comparison was not made in the other studies (Sun et al., 2013; Tian et al., 2011). Differentiation of our *Rnf12*<sup>-/-</sup> ESCs did not result in massive cell death, as reported for the *Jpx*<sup>+/-</sup> ESCs (Tian et al., 2011), despite the fact that the *Rnf12*<sup>-/-</sup> cells only sporadically initiate XCI (Barakat et al., 2011a). This indicates that loss of XCI, when induced by homozygous loss of *Rnf12*, does not result in massive cell death *in vitro* upon ESC differentiation. Hence, loss of *Jpx*<sup>+/-</sup> ESCs by cell death (Tian et al., 2011) might be caused by some mechanism unrelated to XCI.

One key question is how *Jpx*, *Ftx*, and *Xpr* manage to regulate *Xist* in *cis*. In the present study, we find that *Xist* expression is severely compromised on the X chromosome with a  $\Delta(\text{Jpx-Rnf12})$  deletion. Introduction of a randomly integrated *Rnf12* transgene shows preferential upregulation of *Xist* from the wild type X chromosome, indicating that the RNF12 concentration threshold to initiate XCI on the  $\Delta(\text{Jpx-Rnf12})$  X chromosome is very high (Figure 7A). In agreement with this, we found enhanced *Tsix* expression in *cis*. By removing *Tsix* in *cis* we were able to reduce the RNF12 threshold, allowing an increased rate of initiation of XCI. These results indicate that *Jpx*, *Ftx*, and *Xpr* counteract the repressive action of *Tsix* by activating *Xist* in *cis*. Recent studies have indicated that genes located within a TAD are highly co-regulated (Dixon et al., 2012; Nora et al., 2012). This TAD restricted *cis*-activating environment controlling *Xist* and *Tsix* may exert its action by means of promoter and enhancer co-activation through mass action (Schoenfelder et al., 2010). Such a model is supported by findings that *Jpx* and *Ftx* form an inter-



**Figure 7. Cis- and trans-acting factors in XCI**

(A) Overview of progression of XCI in wild type (left panels), *Rnf12*<sup>+/-</sup> (middle left panels), +/ $\Delta$ (*Jpx-Rnf12*) (middle right panels), and *Rnf12*<sup>+/-</sup> *Xist*<sup>+/-</sup> and +/ $\Delta$ (*Xite-Rnf12*) (right panels) cells. XCI is initiated when the concentration of RNF12 and other putative XCI-activators (green line, bottom panels) crosses the threshold (dark blue line), which does not happen in male cells. For establishment of the Xi one active copy of *Rnf12* is required. In *Rnf12* mutant cells (middle and right panels) loss of one copy of *Rnf12* results in reduced initiation of XCI (lower green line). Loss of the *Jpx-Rnf12* region results in a higher threshold (dashed dark blue line) to initiate XCI on the mutant X chromosome, resulting in a further reduction of XCI. In *Rnf12*<sup>+/-</sup>, +/ $\Delta$ (*Jpx-Rnf12*), *Rnf12*<sup>+/-</sup> *Xist*<sup>+/-</sup>, and +/ $\Delta$ (*Xite-Rnf12*) cells inactivation of the wild type X chromosome leads to inactivation of the single active copy of *Rnf12*, followed by a drop in the RNF12 concentration below the threshold (light blue line) required for maintained *Xist* expression and establishment of the Xi. (B) Location of the *cis*- and *trans*-Xic. The *cis*-Xic covers all *cis*-acting regions regulating *Xist*, whereas the *trans*-Xic includes the regions on the X encoding all trans-acting activators of XCI.

action hub together with *Xist* (Tsai et al., 2008), and that disruption of *Ftx* results in down-regulation of both *Jpx* and *Xist* (Chureau et al., 2011). This would mean

that the additive activity of promoters and regulatory elements within the *Xist* TAD is crucial in the regulation of *Xist*. Co-activation through promoter contacts may



indicate that there is no major function for the *Jpx* and *Ftx* RNAs in activation of *Xist*, a suggestion which explain the poor conservation of the exon-intron structure of these genes among placental mammals.

### XCI in the absence of X-pairing

In female cells both X chromosomes come in close proximity prior to the initiation of XCI. Initial reports showed a significant percentage of cells with both Xic's located within a 2  $\mu$ m distance at the onset of XCI (Bacher et al., 2006; Xu et al., 2006). More recently, time lapse imaging studies revealed a short complete overlap of both Xic's prior to *Xist* spreading (Masui et al., 2011). We addressed the question whether pairing is required for XCI to be initiated. The present observations on initiation of XCI in male nuclei in XX-XY ESC heterokaryons, and initiation of XCI in ESCs that lack all elements identified to be crucial for the pairing process, indicate that pairing is not required for XCI initiation and counting. X-pairing likely reflects changes in transcriptional activity of genes located within the *Xist* and *Tsix* TADs, and might be a consequence of the XCI process. This is supported by the finding that transcription is required for pairing to occur (Xu et al., 2007), and the absence of X-pairing after CTCF knockdown (Xu et al., 2007) might reflect dys-regulation of XCI activation and initiation. Active genes are preferentially located in the nuclear interior whereas silent gene loci reside in the nuclear periphery, and transient activation of *Xist*, and other genes within the *Xist* TAD, on both X chromosomes may lead to relocation of the loci. In addition, association of co-regulated genes has been reported, for example, for the erythroid and stem cell specific genes in developing red cells and ESCs, respectively, showing preferential recruit-

ment of these genes to a limited number of transcription factories enriched for shared transcription factors (de Wit et al., 2013; Schoenfelder et al., 2010). Similarly, *Xist* and possibly *Jpx*, *Ftx*, and *Xpr* could transiently share the same, in number restricted, transcriptional interactomes. These inter-X chromosomal associations will happen by chance, and may explain that XCI is associated with X-pairing.

### RNF12 expression crucial for maintained *Xist* expression

The near complete loss of XCI in *Rnf12*<sup>-/-</sup> ESCs indicates a crucial role for RNF12 in the initiation of XCI. Other authors have suggested a less prominent role for RNF12 in initiation of random XCI as compared to imprinted XCI in mice (Shin et al., 2010). However, our previous and present findings involving the *Rnf12*<sup>-/-</sup> and  $\Delta$ *Rnf12*/ $\Delta$ (*Jpx*-*Rnf12*) ESC lines, all support a critical role for RNF12 in initiation of random XCI. XCI was also severely affected in the +/ $\Delta$ (*Xite*-*Rnf12*), and *Xist*<sup>+/<sup>lox</sup></sup> *Rnf12*<sup>+/-</sup> (cis) ESC lines, to a much higher extend than in *Rnf12*<sup>+/-</sup> and *Xist*<sup>lox/+</sup> *Rnf12*<sup>+/-</sup> (trans) ESCs. This indicates that the effect of the respective mutations in the +/ $\Delta$ (*Xite*-*Rnf12*), and *Xist*<sup>+/<sup>lox</sup></sup> *Rnf12*<sup>+/-</sup> (cis) ESC lines on XCI cannot be attributed to loss of XCI initiation, a conclusion that is supported by our observation of dispersed *Xist* clouds upon differentiation of these ESC lines. Therefore, the results obtained with the +/ $\Delta$ (*Xite*-*Rnf12*), and *Xist*<sup>+/<sup>lox</sup></sup> *Rnf12*<sup>+/-</sup> ESC lines indicate that RNF12 is required at a next step, after *Xist* upregulation, to maintain *Xist* expression towards establishment of the Xi.

Our data indicate that *Rnf12* is acting at two subsequent steps to regulate XCI. In differentiating wild type female ESCs the X-encoded XCI-activator activ-

ity (including Rnf12) will be twice as high as in male ESCs driving initiation of XCI, which is stochastic and can happen on either one of the two X chromosomes. This leads to silencing of all activators of XCI in *cis*, which will prevent XCI on the second X chromosome. As a consequence of the stochastic nature of the process initiation of XCI may happen on the two X chromosomes at the same time, or within the time window where the Xi has not been established yet. In the RNF12-based XCI model, where RNF12 is involved also in maintenance of XCI, initiation of XCI on both X chromosomes results in loss of XCI, which is a reversible process early during ESC differentiation, and may therefore allow cells to restart the XCI process. This continuous requirement for RNF12 throughout XCI is probably an underlying cause for complete skewing of XCI in *Rnf12*<sup>+/-</sup> cells, and may explain loss of imprinted XCI in female *Rnf12*<sup>+/-</sup> embryos after inheritance of the null allele through the maternal germ line (Shin et al., 2010). In +/Δ(Xite-Rnf12) and *Xist*<sup>+/<sup>lox</sup></sup>*Rnf12*<sup>+/-</sup> cells, accumulation of *Xist* on the wild type X chromosome will shut down the single functional *Rnf12* gene in *cis* (Figure 7A) followed by loss of *Xist* coating, visible in the +/Δ(Xite-Rnf12) and *Xist*<sup>+/<sup>lox</sup></sup>*Rnf12*<sup>+/-</sup> cells as dispersed signals after *Xist* RNA-FISH. This may result in reactivation of the Xi, allowing RNF12 to increase and start XCI again, but in most of the +/Δ(Xite-Rnf12) and *Xist*<sup>+/<sup>lox</sup></sup>*Rnf12*<sup>+/-</sup> cells this vicious circle never allows spreading of *Xist*. The low amount of cells that do show dispersed spreading of *Xist* underscores the robust feedback mechanism involving close proximity of *Rnf12* to *Xist*, but also the high turnover of both RNF12 and its target REX1.

## A *cis*- and *trans*-X inactivation center

The Xic was determined by genetic studies delineating a 10 Mb region in mouse, or a 700 kb region in human, to be required for XCI (Augui et al., 2011). Most of the gene interactions are restricted within the TADs and are much more abundant than inter TAD interactions, suggesting that the *Xist* and *Tsix* TADs most likely represent the *cis*-regulatory region required for XCI, representing the *cis*-Xic (Figure 7B). Our findings support this, and indicate that all important *Xist* regulatory elements reside within the *Xist* TAD. However, to be able to regulate *Xist* and *Tsix* in a sex-specific manner, *trans*-acting cues are crucial. The *trans*-Xic may encompass multiple regions on the X chromosome involved in activation of the XCI process. Genes generating activators of XCI can be located anywhere on the X chromosome, although close proximity to *Xist* will facilitate a rapid feedback mechanism preventing XCI on both X chromosomes. Parallel pathways in the activation of XCI may be present, as indicated by the presence of a percentage of *Xist* clouds in *Rnf12*<sup>+/-</sup> cells, which is higher than the percentage found for male cells. In addition, XCI-activator activity may also come from X-linked genes acting in the same pathway as RNF12, by affecting *Rnf12* gene or protein activity.

Although putative XCI-activators acting on *Rnf12* expression or function may still be identified, our present studies support a critical role for the RNF12-REX1 pathway in XCI. The double dose of RNF12 expression in female cells drives female-specific initiation of XCI, whereas the continued requirement for one active copy of *Rnf12* to establish the Xi provides a robust and fast feedback mechanism which takes care that once XCI has started, this does not result in oscillation between off-on states.

## Experimental Procedures

### ES cell culture, generation of knockout cell lines, and transgenesis

Wild type ESC and culture media for ESC culture and differentiation have been described (Barakat et al., 2011a). To generate the  $+\Delta(\text{Xite-Dxmit171})$  cell lines, a pXite DTA Hygro TK vector was used to insert a lox site upstream of *Xite* (Monkhorst et al., 2008) in wild type female F121 ESCs, and correct targeting was verified by Southern analysis. Then, the Dxmit171 length polymorphism was targeted by a BAC targeting vector replacing the Dxmit171 length polymorphism on the 129/Sv allele by a floxed neomycin cassette. Transient Cre expression resulted in the loopout of the region between the lox sites, resulting in the  $\Delta(\text{Xite-Dxmit171})$  allele. To generate the  $+\Delta(\text{Jpx-Rnf12})$  deletion cell lines, a female *Xist* 2lox ESC line, containing a wild type Cast/Eij X chromosome and a 129/Sv X chromosome with a floxed *Xist* allele (Csankovszki et al., 1999) was used. *Rnf12* was targeted using the previously described BAC targeting vectors and methods (Barakat et al., 2011a; Jonkers et al., 2009), thereby introducing an additional lox site on the 129/Sv X chromosome. Correct targeting was verified using an RFLP based screening method and Southern blotting (Barakat et al., 2011b). Transient expression of Cre was used to delete either only *Xist* (*Xist*<sup>+/-lox</sup> *Rnf12*<sup>+/-</sup> cell lines) or the *Jpx/Ftx/Xpr* region ( $+\Delta(\text{Jpx-Rnf12})$  cell lines). Correct deletion was verified by PCR, using the primers described in Supplementary Table 2, and confirmed by DNA FISH. To target *Tsix*, a BAC targeting vector was created which replaced the transcriptional start site of *Tsix* by mCherry, thereby abolishing *Tsix* transcription. Correct targeting was verified by PCR

(resulting in  $+\Delta(\text{Jpx-Rnf12}) \Delta\text{Tsix}$  cell lines). To remove *Xite* and *Tsix* sequences, a pXite DTA hygroTK vector was used to insert a lox site upstream of *Xite* (Monkhorst et al., 2008) in the cell line already deleted for *Jpx*, *Ftx*, and *Xpr*, and mutated for *Rnf12*. After verification of correct targeting by Southern blotting (Monkhorst et al., 2008), transient Cre expression was used to loopout *Xite*, *Tsix*, and *Xist* ( $\Delta(\text{Xite-Rnf12})$  allele), which was detected by PCR analysis. For rescue experiments, the previously modified BAC transgenes covering *Rnf12*, *Jpx/Ftx* or the *Xpr* (Jonkers et al., 2009), with either neomycin or puromycin selection, were used. Copy number of transgenes was estimated using qPCR on genomic DNA, as previously described (Barakat et al., 2011a; Jonkers et al., 2009).

### Expression analysis

Expression analysis was performed by qRT-PCR, as previously described (Barakat et al., 2011a), using the primers listed in Supplementary Table 2.

### Generation of experimental heterokaryons

Male (1.3) (Jonkers et al., 2008) and female (F121) mouse ESCs were labelled with Vibrant 1,1'-dioctadecyl-3,3,3',3'-tetramethylindodicarbocyanine (DiD) and 3,3'-dioctadecyloxacarbocyanine perchlorate (DiO) cell labelling solutions (Molecular Probes, Invitrogen), respectively. Cells were resuspended at  $1 \cdot 10^6$  cells/ml in DMEM and labelled with 5  $\mu\text{l/ml}$  of dye at 37°C for 15 min. After washing, cells were allowed to recover for two days in ES cell medium on MEFs. ESCs were then preplated and mixed in a 1:1 ratio, washed, and fused with 50%

polyethylene glycol (pH 7.4) (PEG 1500; Roche Diagnostics) at 37°C over 1 min before dilution. Cells were washed and cultured in ESC-media overnight in gelatinised culture dishes. Differentiation was started after 12 hours, by washing and addition of EB-medium. After 12 hours of differentiation, DiD+DiO+ cells were FACS-sorted using a FACSARIA cell sorter (BD Bioscience). The DiO+DiD+ sorted cells were plated out on chamber slides in EB-medium and differentiated for additional periods of 48 hours (day 2.5) or 108 hours (day 5) prior to fixation.

### Fluorescent in situ hybridization and immunofluorescence

Procedures, probe labelling and probes for RNA-FISH have been described (Barakat et al., 2011a; Jonkers et al., 2009). For DNA-FISH, BAC CT7-474E4 and RP23-100E1 were used to detect the *Xpr* and *Ftx/Jpx* regions, respectively. For immuno-FISH, cells were fixed for 10 minutes using 4% v/v PFA/PBS, permeabilized using 0.1% v/v Triton-X100 for 5 minutes, and post-fixed for 5 minutes with 4% v/v PFA/PBS. Detection of MS2 *Xist* in male nuclei occurred with a digoxigenin-labelled MS2 probe (Jonkers et al., 2008). Actb was detected using a mouse  $\beta$ -ACTIN antibody (Sigma).

### Cell proliferation assays

For cell proliferation analysis, equal amounts of cells were allowed to differentiate on gelatinised culture dishes. Cells were washed, trypsinized, and viable cells were counted at the different time points indicated. For EB differentiation, cells were collected at different time points, and DNA was isolated and concentra-

tion measured. All measurements and countings were performed in triplicate, on three independent differentiations.

## Supplemental Information

Supplemental Information includes seven figures and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2014.02.006>.

## Acknowledgements

We would like to thank Agnese Loda and Bas de Hoon for helpful comments. Reinier van der Linden is acknowledged for help with FACS analysis. Also we thank all department members for stimulating discussions. This work was supported by grants from NWO (NWO-TOP and NWO-VICI) and the ERC to J.G.

## References

- 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* (New York, NY) **318**, 1632-1636.
- Augui, S., Nora, E.P., and Heard, E. (2011). Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat Rev Genet* **12**, 429-442.
- 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. *Nature cell biology* **8**, 293-299.
- 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. (2011a). RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS genetics* 7, e1002001.
- Barakat, T.S., Jonkers, I., Monkhorst, K., and Gribnau, J. (2010). X-changing information on X inactivation. *Experimental cell research* 316, 679-687.
- Barakat, T.S., Rentmeester, E., Sleutels, F., Grootegoed, J.A., and Gribnau, J. (2011b). Precise BAC targeting of genetically polymorphic mouse ES cells. *Nucleic Acids Res.*
- 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.
- 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. *Human molecular genetics* 20, 705-718.
- 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.
- de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M., *et al.* (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature* 501, 227-231.
- 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.
- Gontan, C., Achame, E.M., Demmers, J., Barakat, T.S., Rentmeester, E., van, I.W., Grootegoed, J.A., and Gribnau, J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature* 485, 386-390.
- Gribnau, J., and Grootegoed, J.A. (2012). Origin and evolution of X chromosome inactivation. *Current opinion in cell biology* 24, 397-404.
- 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. *Molecular and cellular biology* 28, 5583-5594.
- Lee, J.T., Davidow, L.S., and Warshawsky, D. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nature genetics* 21, 400-404.
- Masui, O., Bonnet, I., Le Baccon, P., Brito, I., Pollex, T., Murphy, N., Hupe, P., Barillot, E., Belmont, A.S., and Heard, E. (2011). Live-cell chromosome dynamics and outcome of X chromosome pairing events during ES cell differentiation. *Cell* 145, 447-458.
- 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.
- 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.
- 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.
- Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S., *et al.* (2010). Preferential associations between



co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet* 42, 53-61.

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.

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

Tian, D., Sun, S., and Lee, J.T. (2011). The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* 143, 390-403.

Tsai, C.L., Rowntree, R.K., Cohen, D.E., and Lee, J.T. (2008). Higher order chromatin structure at the X-inactivation center via looping DNA. *Developmental biology* 319, 416-425.

Wutz, A. (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nature reviews* 12, 542-553.

Xu, N., Donohoe, M.E., Silva, S.S., and Lee, J.T. (2007). Evidence that homologous X-chromosome pairing requires transcription and Ctf protein. *Nature genetics* 39, 1390-1396.

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



# CHAPTER



# 3



*Xist* and *Tsix* transcription dynamics in mouse embryonic stem cells is regulated by the X-to-autosome ratio and predicted by semi-stable transcriptional states

Loos, F., Lehmann, J., ten Berge, D., Grootegoed, J.A., and Gribnau, J.

*(manuscript in preparation)*

# **Xist and Tsix transcription dynamics in mouse embryonic stem cells is regulated by the X-to-autosome ratio and predicted by semi-stable transcriptional states**

Loos, F.<sup>1</sup>, Lehmann, J.<sup>2</sup>, ten Berge, D.<sup>2</sup>, Grootegoed, J.A.<sup>1</sup>, and Gribnau, J.<sup>1,\*</sup>

<sup>1</sup> Department of Reproduction and Development

<sup>2</sup> Erasmus MC Stem Cell Institute

Erasmus MC, University Medical Center, Rotterdam, The Netherlands

\*correspondence: j.gribnau@erasmusmc.nl

*(manuscript in preparation)*



## Summary

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 master locus 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*. We generated a series of *Xist* and *Tsix* reporter mouse embryonic stem (ES) cell lines, to study the genetic and dynamic regulation of these genes upon ES cell differentiation, and uncoupled from each other. Our results reveal mutually antagonistic roles for *Tsix* on *Xist* and vice versa, and, by analysis of different X-to-autosome (X:A) ratios, provide evidence for additional X-linked genes involved in their regulation. In addition, we observed independent stochastic but concerted regulation of the two genes, which is rather stable during ES cell differentiation. Interestingly, we also found two semi-stable transcriptional states of the *Xic*, which are delineated by the *Xist* and *Tsix* topologically associating domains (TADs), and which might represent higher order conformations predicting the outcome of XCI.

## 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), and *cis*-spreading of *Xist* leads to recruitment of chromatin remodeling complexes that render the X inactive. *Tsix* is transcribed anti-sense to *Xist* and fully overlaps with *Xist* (Lee et al., 1999). *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 and Lu, 1999; Navarro et al., 2006; Ohhata et al., 2008; Sado et al., 2005).

*Xist* and *Tsix* are key components of the master switch locus that is regulated by X-encoded XCI-activators and autosomally encoded 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 loss of pluripotency 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 fulfill redundant roles in the regulation of XCI (Barakat et al., 2011a; Nesterova et al., 2011).

The X-linked gene *Rnf12* encodes the only XCI-activator that has been described so far (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). Overexpression of *Rnf12* results in ectopic XCI in differentiating

transgenic embryonic stem cells (ESCs). 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).

**3** *Rnf12* knockout studies revealed a reduction of XCI in differentiating female *Rnf12*<sup>+/-</sup> ES cells, and a near loss in XCI initiation in *Rnf12*<sup>-/-</sup> ES cells (Barakat et al., 2011b). However, remained initiation of XCI in a subpopulation of *Rnf12*<sup>+/-</sup> 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 (Chureau et al., 2011; Tian et al., 2010). Both genes are located in a region 10-100kb distal to *Xist*, and knockout studies indicated that both genes are involved in *Xist* activation. Although transgene studies implicated *Jpx* as a *trans*-activator of *Xist*, recent studies involving a knockout of a region from *Xite* up to the *Xpr* region did not reveal any effects *in trans* suggesting that *Ftx* and *Jpx* are *cis*-activators 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) (Nora et al., 2012). Positive regulators of *Xist*, including *Jpx* and *Ftx* are located in the same TAD, and similarly *Tsix* positive regulators are located in the *Tsix* TAD, suggesting that these two TADs represent the minimal X inactivation center covering all *cis*-regulatory elements, which are regulated by *trans*-acting XCI-activa-

tors and -inhibitors. 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).

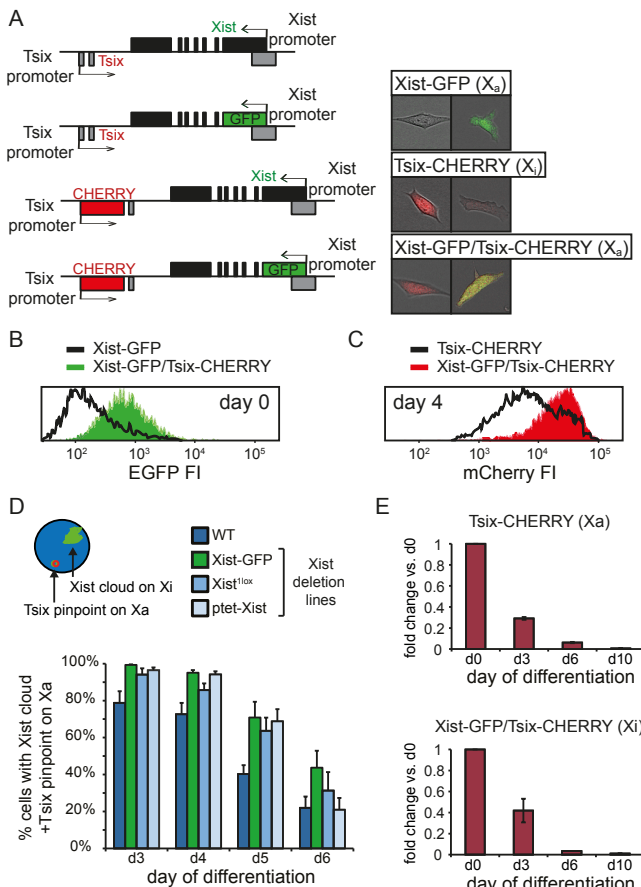
The overlapping gene bodies of *Xist* and *Tsix* and the mutually antagonistic roles of these two genes hamper clear insights in the regulatory mechanisms that govern *Xist* and *Tsix* transcription. To be able to study the independent pathways directing *Xist* and *Tsix* transcription we have generated *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 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, and display 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 RNAs 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 and Lu, 1999; Stavropoulos 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 anti-

sense 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., 2011c), exons 1 of *Xist* and *Tsix* were replaced with EGFP and mCherry coding sequences, respectively (Fig. S1). Expression of the reporters was thus controlled by the endogenous promoters of these two non-coding genes. 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



**Figure 1. *Xist* and *Tsix* Reporter Lines Reveal Antagonistic Roles for *Xist* and *Tsix***

(A) Map of the *Xist*/*Tsix* locus showing design of the reporter cell lines and exemplary pictures of undifferentiated and differentiated cells. (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. Cells scored show *Xist* cloud identifying the Xi and a *Tsix* pinpoint from the Xa. Dark blue bars represent wild type female ES cells, green bar *Xist*-GFP line and light blue bars two independent *Xist* deletion lines. Error bars indicate 95% confidence interval,  $n > 150$  for all time points and cell lines. (E) Analysis of *Tsix* transcript levels at different time points of differentiation by quantitative RT-PCR. Quantification of *Tsix* emanating from the Xa in *Tsix*-CHERRY and the Xi in *Xist*-GFP/*Tsix*-CHERRY is depicted as fold change as compared to undifferentiated cells. Error bars represent SD of two independent experiments. See also Figure S1 and Figure S2.

were removed and polyA signals downstream of the reporters terminated transcription (Fig. S1). By successive rounds of targeting and 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 allele was unperturbed (Fig. S2A). Moreover, FACS analysis of EGFP and mCherry expression for all three ES cell lines during differentiation shows faithful recapitulation of the behaviour of wild-type *Xist* and *Tsix* during the first days of differentiation (Fig. S2B). EGFP is initially up-regulated until the *Xist* promoter on the mutant allele becomes repressed around day 3 of differentiation because this *Xist* deletion allele is destined to become the active X chromosome (Xa). Conversely, mCherry is highly expressed in undifferentiated ES cells and is gradually down-regulated upon differentiation. As expected, comparison of *Xist*-GFP/*Tsix*-CHERRY ES cells, which allows independent track-

ing of *Xist*/*Tsix*, with *Xist*-GFP ES cells shows EGFP de-repression in undifferentiated cells if *Tsix* is deleted in *cis* (Fig.1B). Comparison of *Xist*-GFP/*Tsix*-CHERRY with *Tsix*-CHERRY revealed delayed down-regulation of the mCherry reporter in the double knockin, which might indicate a role for *Xist* in silencing *Tsix*. Interestingly, quantification of *Tsix* levels emanating from the wild type X chromosome in *Xist*-GFP/*Tsix*-CHERRY (*Tsix* only on Xi) and *Tsix*-CHERRY (*Tsix* only on Xa) by qPCR showed similar kinetics of *Tsix* down-regulation on Xi and Xa (Fig. 1E). Therefore, 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* (Fig.1C). 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*<sup>1lox</sup>

### ► Figure 2. Time-Lapse Imaging of Live Cells

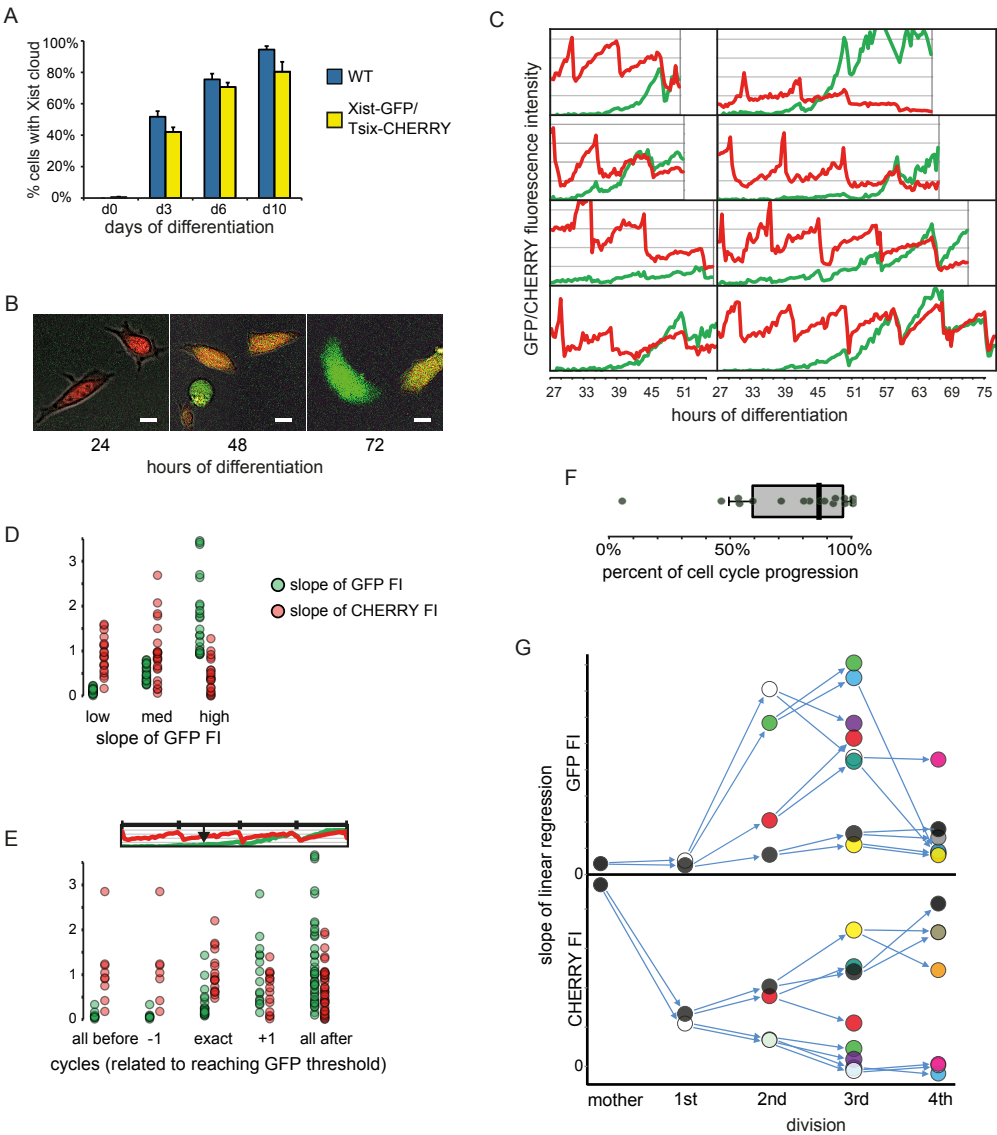
(A) Quantification of *Xist* RNA FISH in differentiating wild type and *Xist*-GFP/*Tsix*-CHERRY cells. Error bars indicate 95% confidence interval,  $n > 100$  for day 0,  $n > 350$  for day 3 and 6,  $n > 150$  for day 10. (B) Exemplary pictures of *Xist*-GFP/*Tsix*-CHERRY cells taken at different time points of differentiation during time-lapse imaging. Scale bar is 5  $\mu\text{m}$ . (C) Whole cell integrated FI values of EGFP (green) and mCherry (red) plotted over time for several exemplary cells during time-lapse imaging. (D) 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. (E) Same as in (D), but 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). (F) EGFP FI reaching a threshold as defined in (E) in relation to cell cycle progression. Dots show time point in cell cycle when EGFP FI threshold is passed. Box plot summarizes data, with thick bar representing median, edges of box the 25<sup>th</sup> and 75<sup>th</sup> percentiles and whiskers the 9<sup>th</sup> and 91<sup>st</sup> percentiles. (G) Pedigree of an exemplary cell followed through four cell divisions. In top panel, slope of linear regression as described in (D) 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.

(Csankovszki et al., 1999) and ptet-Xist (A. Loda, unpublished), show persisting *Tsix* transcription from Xa compared to wild-type ES cells (Fig.1D & S2C). 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 un-

coupled allele in Xist-GFP/*Tsix*-CHERRY.

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

The Xist-GFP/*Tsix*-CHERRY line allowed us to monitor the regulation of *Xist* and *Tsix* independently and disconnected from their antagonistic silenc-





ing capabilities at single cell level. Initial analysis of FACS plots showed that *Xist* up-regulation precedes *Tsix* down-regulation (Fig. S2B). To verify proper regulation of the wild-type X chromosome *Xist* RNA FISH was performed on *Xist*-GFP/*Tsix*-CHERRY ES cells at different time points of differentiation. Consistent with qPCR data (Fig. S2A), *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 (Fig. 2A). These data, together with RNA quantification and FACS analysis, demonstrate that both wild-type and mutant allele are properly regulated also on single cell level, and that cloud formation occurs early during differentiation.

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. 2B). The integrated EGFP and mCherry fluorescence intensities (FI) of entire single cells were measured, resulting in oscillating patterns due to accumulation of fluorescent reporters followed by dilution upon cell division (Fig. 2C). 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. 2D). Next, we set a threshold for mean EGFP FI to estimate at which point EGFP FI rises above background noise. This serves as a proxy for activation of the *Xist* promoter and allowed us to bin cell cycles in relation to the time point of

*Xist* activation. 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 (Fig. 2E). This independence is also apparent from the original EGFP/mCherry FI plots, in which mCherry down-regulation in different cells occurs before, exactly at or after EGFP activation (Fig. 2C). Analysis of the timing of EGFP up-regulation demonstrates that passing the threshold predominantly occurs in the second half of the cell cycle (Fig. 2F). This clustering of threshold-passing events suggests that *Xist* is activated preferentially during a defined window of time in the cell cycle, even though an arbitrarily set threshold and uncertainties regarding the kinetics of translation and maturation of EGFP preclude inference of the exact time of *Xist* activation.

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. 2G). 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. It also hints at major changes in reporter activity predominantly taking place during or just after mitosis, because shifts in the slope of FI over time rarely occur between cell divisions. Taken together, live cell imaging and fate mapping suggest that on an uncoupled allele, *Xist* and *Tsix* are antagonistically regulated in a develop-

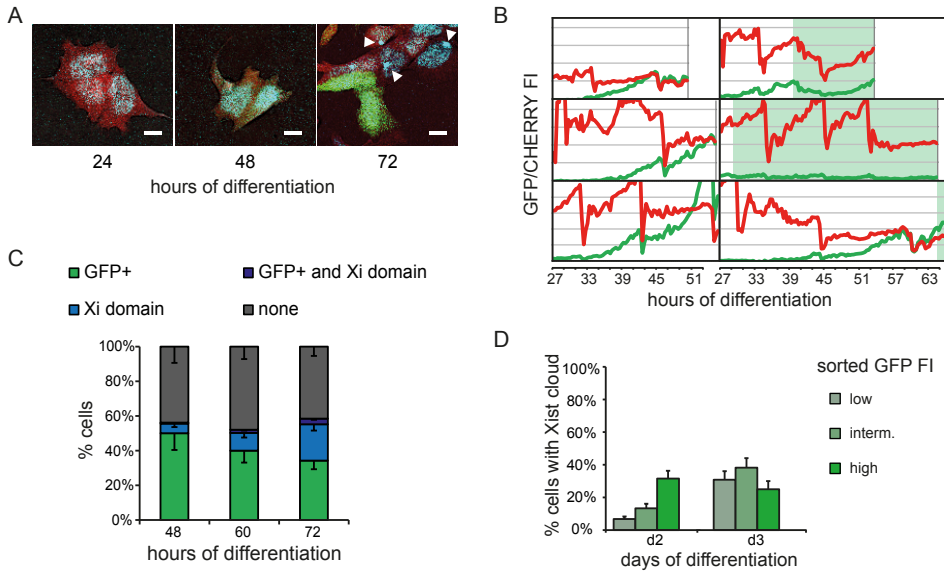


mentally concerted manner, even though up- and down-regulation of both genes per se are independent and probably stochastic. We also find evidence for a specific time point in the cell cycle at which *Xist* activation occurs and observe that fate decisions, once established, are quite stable.

### Interdependent regulation of the *Xic*'s

Even though the EGFP and mCherry reporters provide an estimate of *Xist* and *Tsix* promoter activity, we cannot directly relate them to silencing of an X chromosome, as only the wild-type allele harbors an intact copy of *Xist* and can therefore initiate XCI. To address this problem, an *Ezh2* transgene fused to mTagBFP2-FLAG (Subach et al., 2011) was introduced into *Xist*-GFP/*Tsix*-CHERRY ES cells. EZH2 is part of the PRC2 complex that transiently co-localizes with the Xi during initiation of XCI (Plath et al., 2003) and serves as a marker for actual silencing of one X chromosome. Immuno-fluorescence staining confirmed that transgenic *Ezh2*-FLAG co-localizes with H3K27me3 (Fig. S3A). In addition, the kinetics of association of the fluorescent *Ezh2* transgene with the Xi as determined by fixation and direct imaging of differentiating ES cells (Fig. S3B) were highly similar to earlier studies (de Napoles et al., 2004; Plath et al., 2003). To unravel the relation between *Xist* and *Tsix* activation and silencing of the Xi, we performed live cell imaging of differentiating *Xist*-GFP/*Tsix*-CHERRY cells containing mTagBFP2-*Ezh2* by confocal microscopy for extended periods of time. As expected, a distinct mTagBFP2-EZH2 focus was discernible in a fraction of cells starting at around 48 hours of differentiation (Fig. 3A). Analysis was hampered by the fact that Xi domains appeared late during imaging when high cell densities partially precluded proper tracking and FI

measurements of cells. Moreover, cell viability was reduced, and cell cycle length increased as compared to the previous experiments, probably due to the use of high energy wavelengths necessary to excite mTagBFP2. Nonetheless, exploiting the same strategy as outlined before we were able to follow single cells over time and to correlate the appearance of an Xi domain with EGFP and mCherry levels (Fig. 3B). As noted for the correlation between up-regulation of EGFP and down-regulation of mCherry, Xi domains appear to emerge loosely around the time of EGFP activation and mCherry repression with no tight order of events being discernible (Fig. 3B, right panel). Close examination of later time points of differentiation, however, revealed that cells containing an EZH2 focus basically never displayed high levels of EGFP (Fig. 3A, right panel). 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 for EGFP level and presence of an EZH2/Xi domain (Fig. 3C). The results show the expected percentages of EZH2 focus formation and moreover confirm that high EGFP levels almost never concur with an Xi domain. This suggests that a sub-population of cells cannot up-regulate *Xist* on the wild type allele, although our experimental setup does not preclude the possibility that EGFP<sup>high</sup> cells gain an Xi beyond the imaging time window we were able to study. In addition, despite the relatively high EGFP half-life of approximately 10 hours, EGFP down-regulation may precede *Xist* cloud and EZH2 domain formation (Figure S3C). Comparison of *Xist* RNA-FISH data and EZH2 domain formation indeed indicates a lag of about a day between *Xist* accumulation and the time required to acquire an Xi that is visible by live imaging (Fig. 2A and Fig. S3B). To study the correlation between EGFP levels and formation of *Xist* clouds in more detail,



**Figure 3. Correlation between Xist-GFP activation and Xist cloud formation**

(A) Exemplary pictures of Xist-GFP/Tsix-CHERRY+ mTagBFP2-Ezh2 cells taken at different time points of differentiation during time-lapse imaging. White arrowheads denote mTagBFP2-Ezh2 focus on Xi. Scale bar is 5  $\mu$ m. (B) Whole cell integrated FI values of EGFP (green) and mCherry (red) plotted over time for several exemplary cells during time-lapse imaging. Green background shading indicates presence of repressing a silenced Xi. (C) 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. (D) Day two and three differentiating Xist-GFP/Tsix-CHERRY cells were FACS-sorted into EGFP low, intermediate and high fractions. Graphs show quantification of Xist RNA FISH in these fractions. Error bars indicate 95% confidence interval,  $n > 250$  for all time points and fractions. See also Figure S3.

and to obtain independent proof for the observed mutual exclusive occurrence of high EGFP levels and factual initiation of XCI on the wild-type X chromosome, we sorted Xist-GFP/Tsix-CHERRY cells based on their EGFP FI and scored the percentage of cells with Xist clouds by RNA FISH (Fig. 3D). At day 2 of differentiation we observed a clear correlation between EGFP fluorescent intensity and percentage of Xist clouds, demonstrating that both *Xist* promoters become activated and that EGFP up-regulation and XCI initiation scale at least to some degree. Interestingly, at day 3 the EGFP<sup>high</sup> fraction of cells contained less Xist clouds than the EGFP<sup>intermediate</sup> fraction. This suggests that EGFP<sup>high</sup> cells indeed down-regulate

EGFP before EZH2-BFP becomes visible, but also suggests 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. These data indicate that, while EGFP is a suitable proxy not only for activation of the *Xist* promoter but also for Xist cloud formation, some cells might get locked in an active state of the Xist-GFP allele and do not switch to initiating XCI on the wild-type X chromosome.

## Activators and Inhibitors of XCI

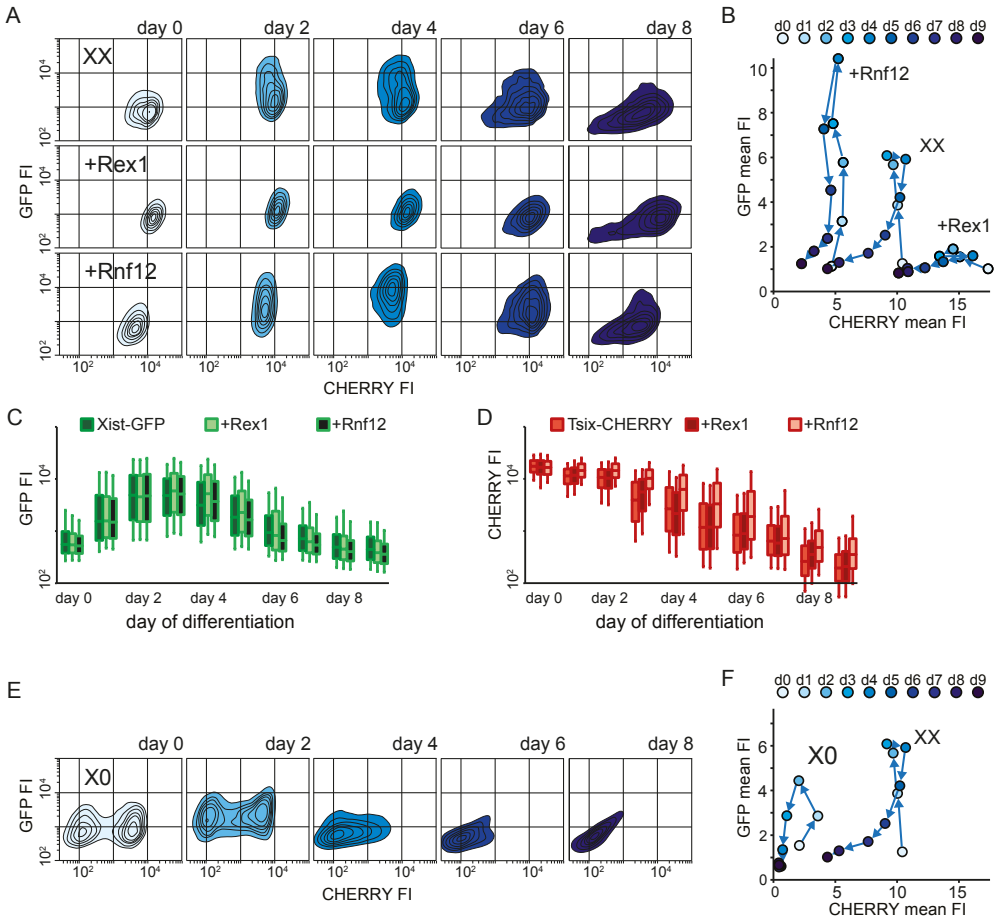
RNF12 functions as a *trans*-activator of XCI (Jonkers et al., 2009) by targeting REX1, a repressor of XCI, for proteasomal degradation (Gontan et al., 2012). Previous work has indicated that REX1 might have a dual role in the activation of XCI by activating *Tsix* and repressing *Xist* (Barakat et al., 2011a; Gontan et al., 2012; Navarro et al., 2010). 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 (Fig. S4A). 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* over-expression 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 (Fig. S4B). 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 (Fig. 4C, D). This finding indicates that antisense transcription or the antisense transcript represses transcription of the *Xist* and *Tsix* promot-

er 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 (Fig. S5A). Comparison of these X0 lines (XGTC-X0), 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. 4E,F). Increased *Rex1* expression, as a consequence of loss of one copy of *Rnf12* in XGTC-X0 cells, might explain reduced Xist-GFP up-regulation upon differentiation, as was found for the *Rex1* transgenic XX Xist-GFP/Tsix-CHERRY lines. An increase in REX1 levels, however, does not explain the reduced mCherry expression level throughout ES cell differentiation, which was not found in the XX Xist-GFP/Tsix-CHERRY *Rex1* transgenic cell lines. 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 (Fig. S5B). This finding indicates 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

Interestingly, the XGTC-X0 cells display a strikingly bimodal mCherry distribution, indicating that in similar proportions of cells the *Tsix* promoter is either



**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). Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events. (B) Same as in (A), but mean FI for EGFP and mCherry is plotted. (C) Box plots showing EGFP FI as determined by FACS analysis at different time points of differentiation for Xist-GFP, Xist-GFP+Rex1 and Xist-GFP+Rnf12. Center bar in box plots represents median, edges of box the 25<sup>th</sup> and 75<sup>th</sup> percentiles and whiskers the 9<sup>th</sup> and 91<sup>st</sup> percentiles. (D) Same as in (C), but mCherry FI for Tsix-CHERRY, Tsix-CHERRY+Rex1 and Tsix-CHERRY+Rnf12 is depicted. (E) Contour plots of FACS analysis showing EGFP and mCherry FI at different time points of differentiation for XGTC-X0 line. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events. (F) Same as in (E), but mean FI for EGFP and mCherry is plotted. See also Figure S4 and Figure S5.

on or off. 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 (Fig. S5C). We obtained similar results with a male Tsix-CHERRY

only knock-in cell line showing a similar bimodal mCherry distribution (Fig. S5D). 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 pos-

itive 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. S5E). As we noticed that also female *Xist*-GFP/*Tsix*-CHERRY cells grown in 2i+LIF conditions displayed two separable mCherry populations, we wanted to test if the additional X in these cells increases the rate of switching between the two populations. Therefore, we sorted out mCherry low and high cells and monitored them for two weeks under 2i+LIF conditions (Fig. 5B). Again, recovery of the mixed population of mCherry positive and negative cells did not occur within this time frame. We next attempted to reset the state of the *Xic* by developmental cues and started to differentiate the sorted XGTC-X0 and *Xist*-GFP/*Tsix*-CHERRY cells. Since *Xist*-GFP/*Tsix*-CHERRY cells were grown in 2i+LIF we also transferred them to serum+LIF, which has been reported to trigger a “primed” state in ES cells (Marks et al., 2012). Before complete shutdown of the *Tsix* promoter upon differentiation, XGTC-X0 cells from both low and high mCherry populations show recovery of the bimodal distribution to some degree (Fig. 5A). In differentiating *Xist*-GFP/*Tsix*-CHERRY cells, mCherry levels stay mostly stable and similar results are obtained from cells transferred to serum+LIF (Fig. 5B & S5F). Intriguingly however, the mCherry low populations activates the *Xist* promoter-driven EGFP reporter much more strongly than the mCherry high population (Fig. 5B). 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 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. 5C). To find the basis of the difference between the two populations, RNA sequencing was performed on FACS-sorted mCherry positive and negative XGTC-X0 cells. This analysis indicated that both populations have highly similar expression profiles (Pearson correlation coefficient  $r=0.9832$ ; Fig. S5G), and confirmed that expression of the pluripotency factors was indifferent between the two cell populations. 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 (Fig. S5H). Most of the genes showing expression changes were restricted to the two flanking topological associated domains that *Xist* and *Tsix* are embedded in (Fig. 5D). 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. 5B,C). These differential epigenetic states might be 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 fulfills 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.

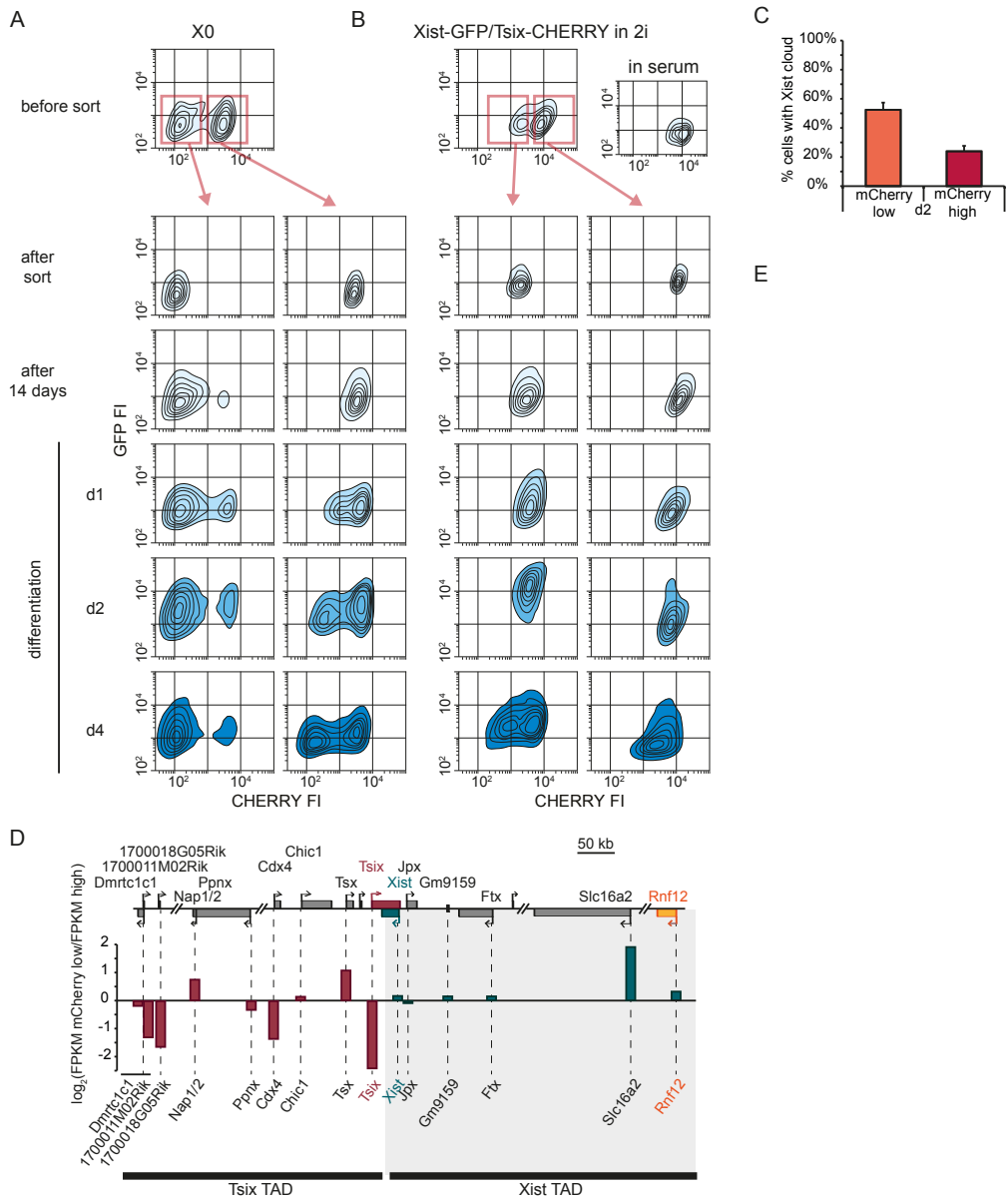


Interestingly, our studies also indicate that *Xist* acts locally facilitating the shutdown of *Tsix* on the future Xa, as we observed sustained *Tsix* expression comparing three different *Xist* knockout ES cell lines with wild type cells during ES cell differentiation. As down-regulation of the mCherry reporter shows equal kinetics on the Xa and Xi, this also suggests a role for *Xist* in silencing *Tsix* on the future Xi. Although this effect is likely mediated through *Xist* RNA instructed local recruitment of chromatin remodeling complexes, we cannot exclude a transcriptional interference mechanism to be involved.

Live cell imaging of XX cells harboring *Xist*/*Tsix* fluorescent reporters indicated that also in the absence of sense-antisense overlapping transcription expression of *Xist* and *Tsix* is anti-correlated. Nevertheless, this anti-correlation is not strict, and we find *Xist* up-regulation prior to *Tsix* down-regulation and vice versa. This suggests a mechanism of stochastic expression of both genes, where initiation of *Xist* expression is increased during differentiation until a level is reached which is sufficient to spread *in cis*, leading to *Tsix* silencing thereby providing a feed forward loop facilitating further *Xist* transcription initiation, accumulation and spreading. Interestingly, our results also point at a specific time window in the cell cycle where *Xist* up-regulation is initiated. Although our experimental setup precludes precise pinpointing of this time point of *Xist* up-regulation, the slopes of increasing *Xist*-GFP only seem to change after cell division suggesting that the M-G1 transition may be instructive in facilitating changes in the transcriptional state of *Xist*. In support of this, exclusion of transcription factors from chromatin during mitosis has been proposed to be important for resetting transcriptional states (Martinez-Balbas et al., 1995).

The present live cell imaging studies also 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-X0 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 different semi-stable states of higher order chromatin structure that can be propagated through many cell divisions. A recently developed polymer model predicted such different states of higher order chromatin structure (Giorgetti et al., 2014). Sorted cell populations provide the exciting opportunity to examine these different chromatin structural states in detail by DNA FISH and chromatin capture technologies. We predict these transcriptional states to be maintained independent of DNA methylation which is nearly absent in 2i conditions (Habibi et al., 2013), and does not seem to be involved in regulation of *Tsix* until after implantation (Prissette et al., 2001). 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 cells are observed which might indicate 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. This does not necessarily mean that different transcriptional states





**Figure 5. Two Stable States of the *Xic* Predict XCI Potential**

(A) Contour plots of FACS analysis showing EGFP and mCherry FI for XGTC-X0 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) Exactly the same as in (A), but shown for female Xist-GFP/Tsix-CHERRY line grown in 2i+LIF. Small inset on right depicts same line grown on feeders in serum. (C) Quantification of Xist RNA FISH in female Xist-GFP/Tsix-CHERRY cells at day two of differentiation after sorting into mCherry low and high populations. Error bars indicate 95% confidence interval,  $n = 313$  for mCherry low and  $n = 305$  for mCherry high populations. (D) Expression levels of genes located in the *Xic* as determined by RNA sequencing of XGTC-X0 mCherry low and high populations. Top indicates location of genes along the X chromosome, bars show  $\log_2(\text{FPKM mCherry low}/\text{FPKM mCherry high})$ . See also Figure S5.

as represented by the Tsix-mCherry medium and high subpopulations are intrinsically stable. Rather, we favor a scenario in which chromatin conformation is fluctuating but exists preferentially in one or the other conformation. 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, 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 medium than high cells. This might be related to differences in the expression level of activators and inhibitors of XCI, coordinated with the transcriptional state of the *Xic*. RNF12 is the prime candidate factor in the communication of the *Xic*'s *in trans*, as this activator of XCI is located within the *cis-Xic* that shows distinct transcriptional states. 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* (Navarro et al., 2011; Payer et al., 2013), 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. Our results also indicate that the 2i culture conditions are suboptimal for studying the XCI process, and that serum/LIF culture conditions can be further optimized to allow a more robust but well titrated XCI process.

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 (Gontan et al., 2012). 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). 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 X0 reporter cell lines dur-

ing 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.

## Experimental procedures

### 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 (gift from G.-J. Kremers) 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

Culture media and conditions for ES cell culture and differentiation have been described (Barakat et al., 2011a). 2i+LIF conditions were: DMEM supplemented with 100 U/ml penicillin/streptomycin, 20% KnockOut Serum Replacement (Gibco), 0.1mM NEAA, 0.1mM 2-mercaptoethanol, 5000 U/ml LIF, 1μM MEK inhibitor PD0325901 (Stemgent) and 3

μM GSK3 inhibitor CH99021 (Stemgent). Transgenic ES cell lines were generated using wild-type female line F1 2-1 (129/Sv-Cast/Ei) and wild-type male line J1 (129/Sv). BAC targeting strategy was used as has been described (Barakat et al., 2011c). In short, *Xist* knockin was created as follows: EGFP/neomycin-resistance-cassette flanked by lox sites was targeted by homologous recombination in bacteria to a BAC (Barakat et al., 2011c). 5' and 3' targeting arms were amplified from BAC using primers 1 + 2 and 5 + 6, respectively. With the modified BAC wild-type ES cells were targeted, and resistance cassette was removed by transient Cre transfection, resulting in cell line *Xist*-GFP. A *Xist* ScrFI RFLP with primers 4 + 20 was used to screen drug-resistant clones for correct recombination events. *Tsix* knockin was created as follows: 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. XGTC-X0 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 PflMI RFLP located in the X-linked gene *Atrx* using primers 68 + 69.

## FACS Analysis and Cell Sorting

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

## Expression Analysis

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

## Live Cell Imaging and Image Analysis

Cells were preplated to remove feeders and differentiation was initiated 12 hours prior to start of imaging. Cells were seeded at low density ( $10^4$  cells/well) in a 6-well glass bottom dish (MatTek P06G-1.5-20-F) coated with human plasma fibronectin (Millipore). Imaging was performed on a Leica SP5 AOBs at 37°C and 5% CO<sub>2</sub> using adaptive focus control to keep cells in focus during the entire experiment. Pictures were taken every 20 minutes for a total of 68 hours. Tiled images were acquired and automatically stitched

to record a large field of view at sufficient resolution to resolve subcellular structures and follow cells over time. Average projection of Z-stack was generated in Fiji (version 1.45b) and background corrected integrated fluorescence intensities for EGFP and mCherry were measured for single cells over the entire time frame that a given cell was clearly discriminable. Based on recorded values, linear regression by least squares method was performed to calculate the straight line that best fits the data. The slope of this function with fluorescence intensity being dependent on time was used as a proxy for *Xist* or *Tsix* promoter activity. Threshold for *Xist* activation was calculated by using 3.29 standard deviations (corresponding to 99.9% within confidence interval) of mean EGFP FI values measured in cells within the first six hours of time-lapse experiment.

## Fluorescent In Situ Hybridization and Immunofluorescence

For *Xist*/*Tsix* RNA-FISH and immunofluorescence stainings cells were grown on or absorbed to poly-lysinated coverslips. For RNA-FISH, cells were fixed for 10 minutes with 4% paraformaldehyde (PFA)-PBS at room temperature, washed with 70% EtOH, permeabilized 4 minutes with 0.2% pepsin at 37°C and post-fixed with 4% PFA-PBS for 5 minutes at room temperature. Coverslips were washed twice with PBS and dehydrated in a gradient of 70%, 90%, and 100% EtOH. Nick-labeled DNA probes (digoxigenin for *Xist*/*Tsix* probe, biotin for *Tsix* probe) were dissolved in hybridization mixture (50% formamide, 2XSSC (1XSSC: 0.15 M NaCl, 0.015 M sodium citrate), 50 mM phosphate buffer (pH 7.0), 10% dextran sulfate) and 100 ng/µl mouse Cot DNA to a final concentration of 1 ng/µl. Probe was denatured for 5 min, pre-hybridized

for 45 min at 37°C, and coverslips were incubated overnight in a humid chamber at 37°C. After hybridization, coverslips were washed once in 2XSSC, three times in 50% formamide-2X SSC, both at 37°C and twice in TST (0.1 M Tris, 0.15 M NaCl, 0.05% Tween 20) at room temperature. Blocking was done in BSA-TST for 30 minutes at room temperature. Detection was done by subsequent steps of incubation with anti-digoxigenin (Boehringer) and two FITC-labeled antibodies for Xist/Tsix RNA detection or anti-biotin (Roche) and two rhodamine-labeled antibodies for Tsix RNA detection in blocking buffer for 30 min at room temperature. Coverslips were washed twice with TST between detection steps and once finally with TS (0.1 M Tris, 0.15 M NaCl). Dehydrated coverslips were mounted with ProLong Gold Antifade with DAPI (Molecular Probes). For immunofluorescence, cells were fixed for 10 minutes at room temperature in 4% PFA-PBS followed by three washes in PBS and permeabilization in 0.25% Triton-X100-PBS. Blocking was done in blocking solution (0.5% BSA, 1% Tween20 in PBS) for 1 hour at room temperature. All antibody incubation steps were done for 1 hour at room temperature in blocking solution, followed by three washes in blocking solution. Primary antibodies were used at the following concentrations: anti-Flag-M2(1:1000), anti-H3K27me3 (1:500). Secondary antibodies used were conjugated to Alexa Fluor 488 or Alexa Fluor 546 (Molecular Probes; 1:500).

## RNA sequencing

RNA samples were 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 ref-

erence 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.

## References

- Barakat, T.S., Gunhanlar, N., Gontan Pardo, C., Achame, E.M., Ghazvini, M., Boers, R., Kenter, A., Rentmeester, E., Grootegoed, J.A., and Gribnau, J. (2011a). RNF12 Activates Xist and Is Essential for X Chromosome Inactivation. *PLoS Genet* 7, e1002001.
- 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. (2011b). RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS Genet* 7, e1002001.



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

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

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.

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

de Napolés, 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.

Donohoe, M.E., Zhang, L.F., Xu, N., Shi, Y., and Lee, J.T. (2007). Identification of a Ctfc cofactor, Yy1, for the X chromosome binary switch. *Mol Cell* 25, 43-56.

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.

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

Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C., *et al.* (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, T.S., Achame, E.M., Monkhorst, K., Kenter, A., Rent-

meester, 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.

Lee, J., and Lu, N. (1999). Targeted mutagenesis of Tsix leads to non-random X inactivation. *Cell* 99, 47-57.

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

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.

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

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., Kalkan, T., Menafra, R., Denisov, S., Jones, K., Hofmeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A., *et al.* (2012). The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149, 590-604.

Martinez-Balbas, M.A., Dey, A., Rabin-drans, S.K., Ozato, K., and Wu, C. (1995). Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* 83, 29-38.

Masui, O., Bonnet, I., Le Baccon, P., Brito, I., Pollex, T., Murphy, N., Hupe, P., Barillot, E., Belmont, A.S., and Heard, E. (2011). Live-Cell Chromosome Dynamics and Outcome of X Chromosome Pairing Events during ES Cell Differentiation. *Cell* 145, 447-458.

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 regulated by pluripotency factors in embryonic stem cells. *Hum Genet*.

Navarro, P., Oldfield, A., Legoupi, J., Fes-

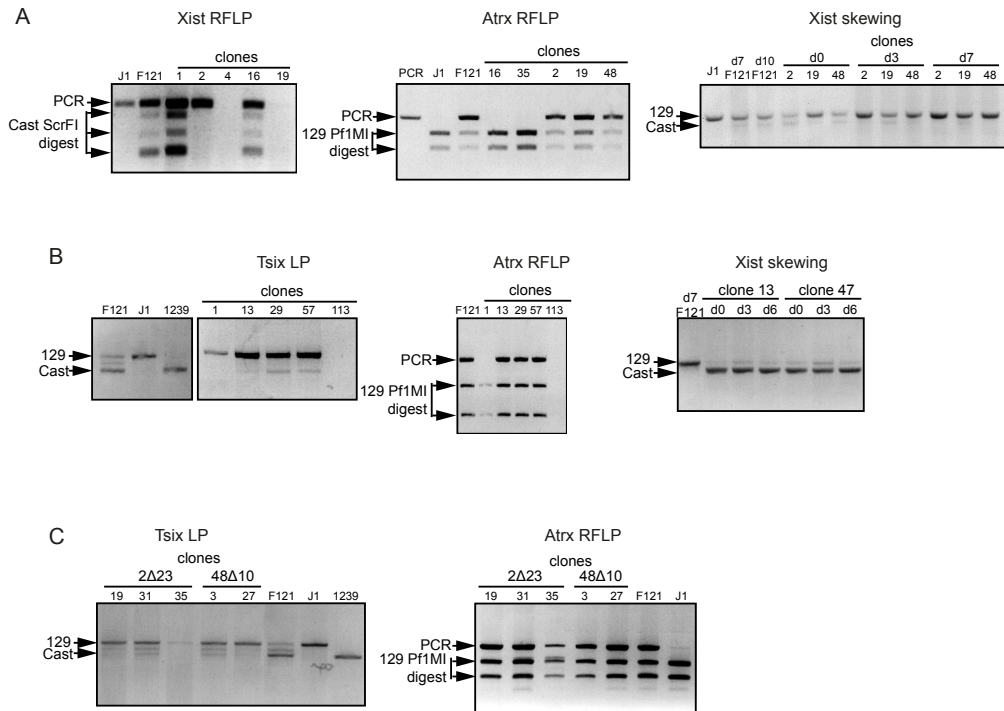


- tuccia, N., Dubois, A., Attia, M., Schoorlemmer, J., Rougeulle, C., Chambers, I., and Avner, P. (2010). Molecular coupling of Tsix regulation and pluripotency. *Nature* 468, 457-460.
- Navarro, P., Page, D.R., Avner, P., and Rougeulle, C. (2006). Tsix-mediated epigenetic switch of a CTCF-flanked region of the Xist promoter determines the Xist transcription program. *Genes Dev* 20, 2787-2792.
- Nesterova, T.B., 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.
- 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*.
- Ohhata, T., Hoki, Y., Sasaki, H., and 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., and Lee, J.T. (2013). Tsix RNA and the Germline Factor, PRDM14, Link X Reactivation and Stem Cell Reprogramming. *Mol 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.
- Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, 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.
- Prisette, M., El-Maarri, O., Arnaud, D., Walter, J., and Avner, P. (2001). Methylation profiles of DXPas34 during the onset of X-inactivation. *Hum Mol Genet* 10, 31-38.
- 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., Bluthgen, N., and Heard, E. (2014). The Two Active X Chromosomes in Female ESCs Block Exit from the Pluripotent State by Modulating the ESC Signaling Network. *Cell Stem Cell* 14, 203-216.
- 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.
- Stavropoulos, N., Lu, N., and Lee, J.T. (2001). A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice. *Proc Natl Acad Sci U S A* 98, 10232-10237.
- Subach, O.M., Cranfill, P.J., Davidson, M.W., and Verkhusha, V.V. (2011). An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. *PLoS One* 6, e28674.
- Sun, S., Del Rosario, B.C., Szanto, A., Oga, Y., Jeon, Y., and Lee, J.T. (2013). Jpx RNA activates Xist by evicting CTCF. *Cell* 153, 1537-1551.
- 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.

## Supplemental Information

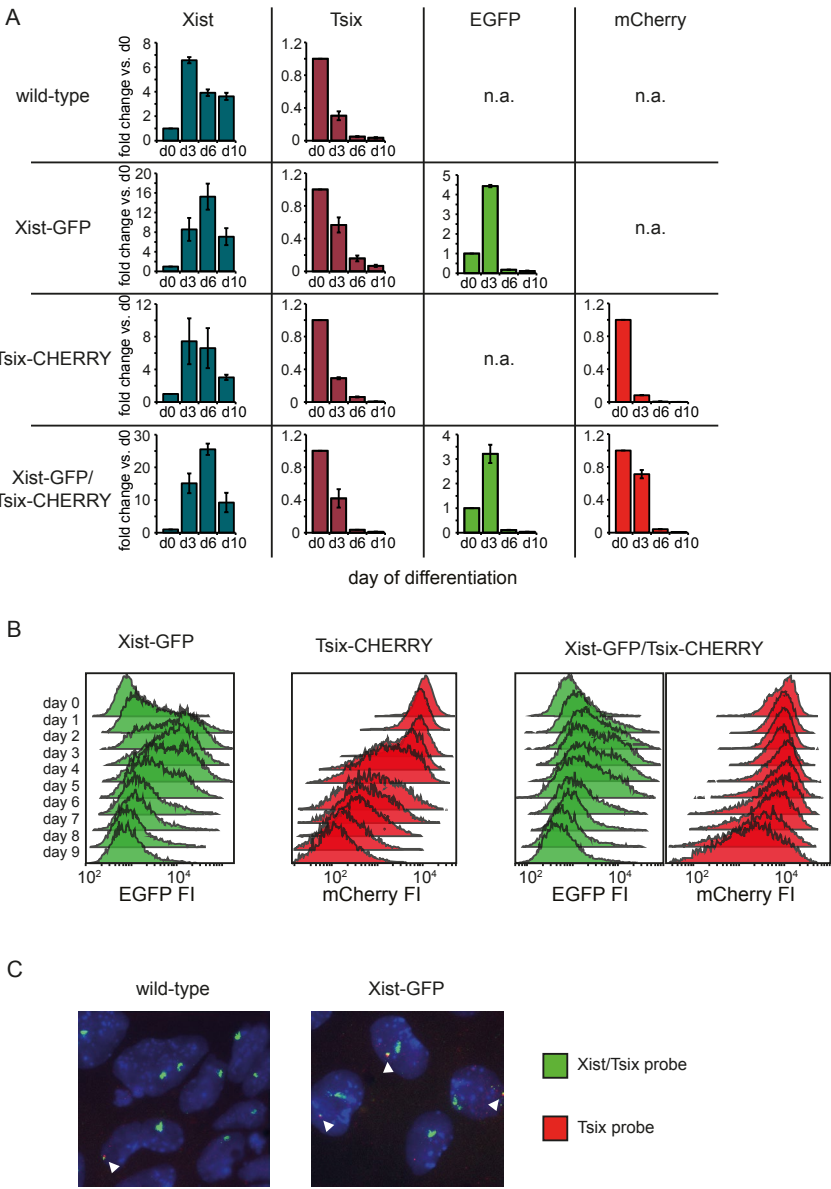
Xist and Tsix transcription dynamics are regulated by the X to autosome ratio and predicted by semi-stable transcriptional states

3



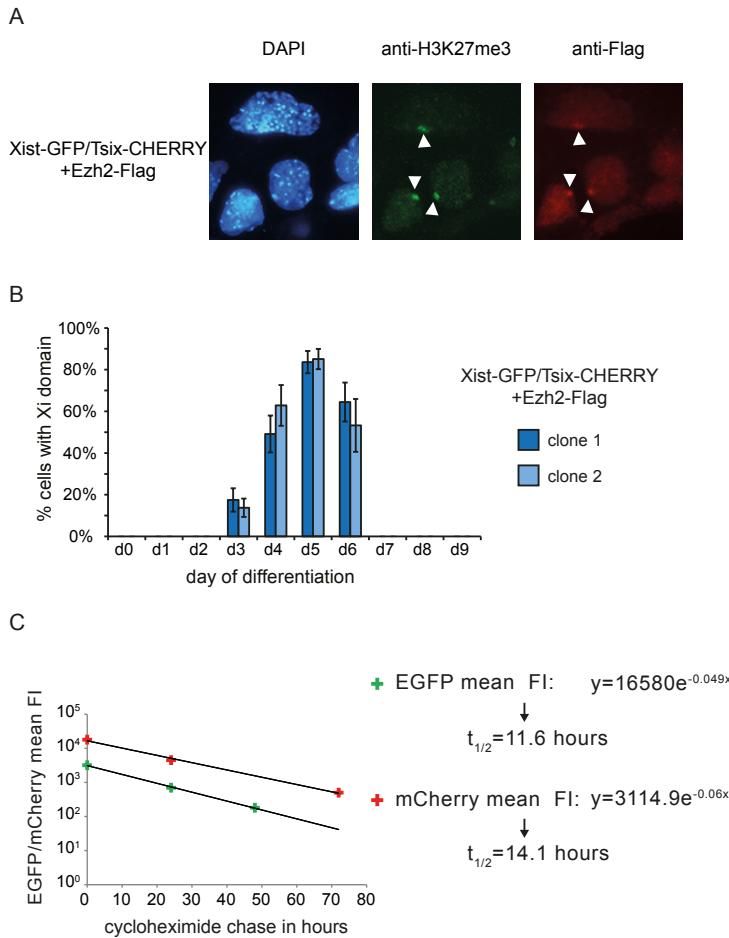
**Figure S1, related to Figure 1. Targeting of Cell Lines**

**(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. 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 polymorphism 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 223.



**Figure S2, related to Figure 1. Behavior of Wild Type and Mutant Alleles of *Xist* and *Tsix***

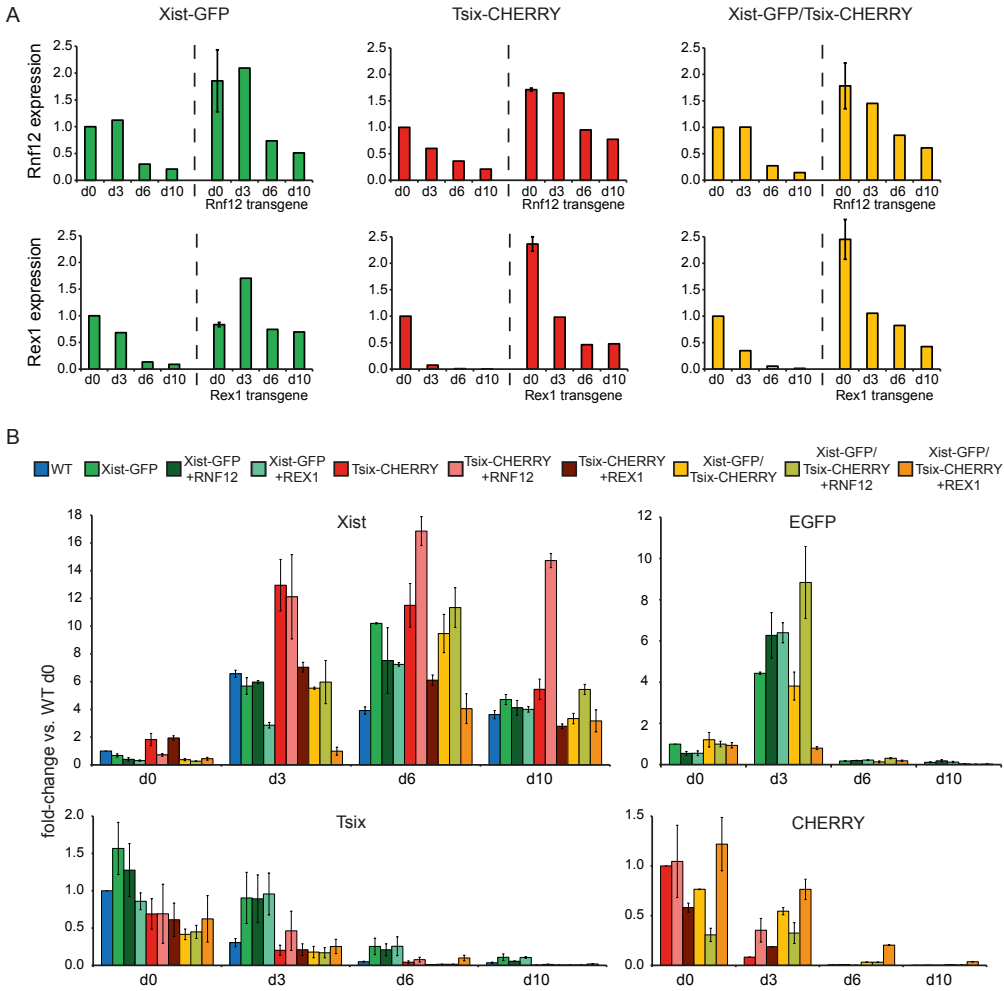
**(A)** Expression analysis of *Xist*, *Tsix*, EGFP, mCherry expression levels at different time points of differentiation by quantitative RT-PCR. Quantification is depicted as fold change as compared to undifferentiated cells. Of note, in wild type cells *Xist* and *Tsix* levels arise from both the 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)** 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. **(C)** *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.



**Figure S3, related to Figure 3. Analysis of Ezh2 Fusion Proteins and Reporter Half-life**

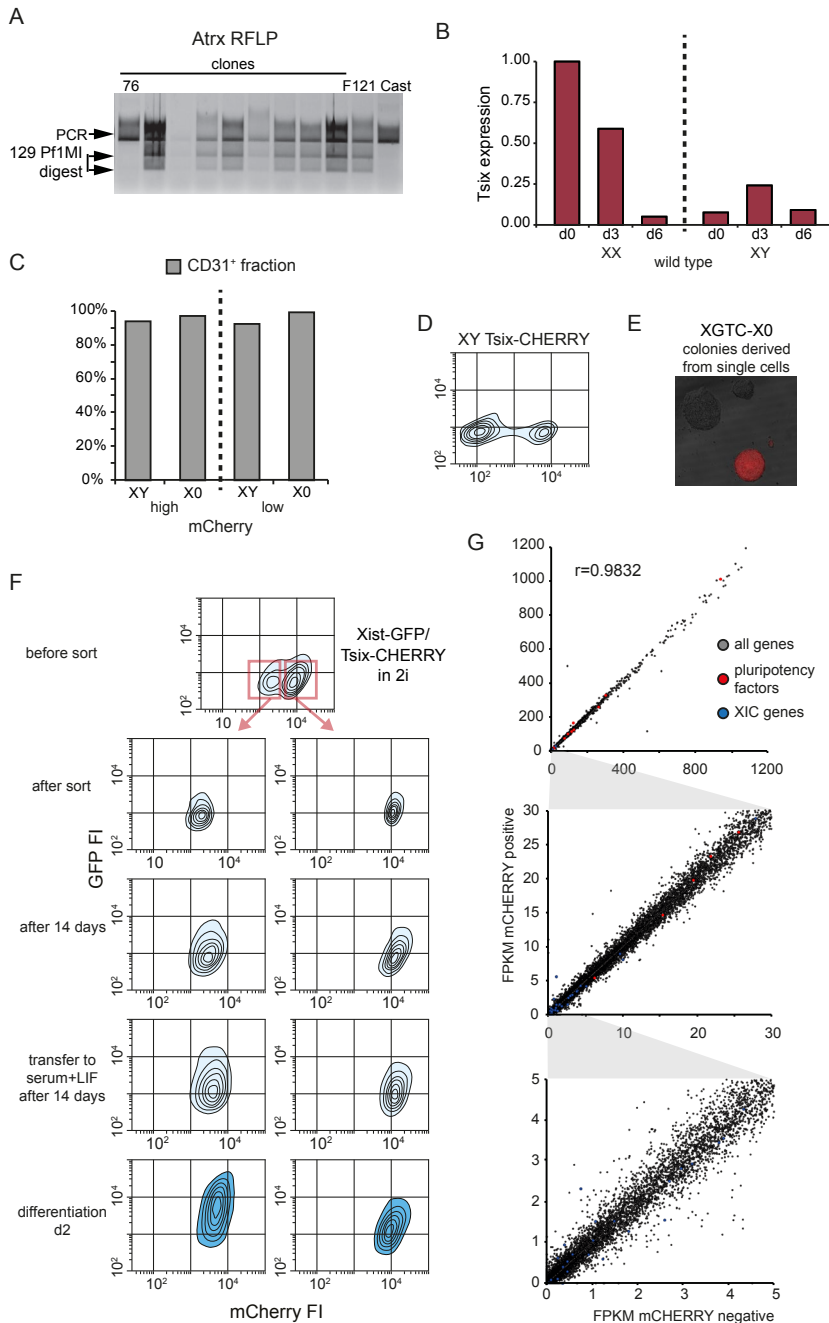
(A) Immunofluorescence staining for H3K27me3 and Flag in Xist-GFP/Tsix-CHERRY line at day 3 of differentiation. White arrowheads indicate Xi domain as identified by H3K27me3 and Ezh2-Flag staining. (B) Quantification of Xist-GFP/Tsix-CHERRY cells showing transient enrichment of Ezh2-Flag on the Xi during 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. (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  $100\mu\text{g/ml}$  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:  $t_{1/2} = \ln(2)/k$ .





**Figure S4, related to Figure 4. Expression Analysis of Rnf12 and Rex1 Transgenic 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.



**Figure S5, related to Figure 4 and Figure 5. Analysis of Cell Lines Showing Bimodal *Tsix* activity**

(A) 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. For example, clone 76 had lost the wild type 129/Sv X chromosome. (B) Expression analysis of *Tsix* at different time points of differentiation by

Table S1. Primers used in this study as listed in the Experimental procedures section

#	SEQUENCE	DESCRIPTION
1	AGGTACCTCCCAAGGTATGGAGTCACC	Forward primer 5' targeting arm for <i>Xist</i>
2	TACCGGTAGGAGAGAAACCGGAAGAA	Reverse primer 5' targeting arm for <i>Xist</i>
5	TAGTACTCAATGGCTTGACCCAGACTT	Forward primer 3' targeting arm for <i>Xist</i>
6	TAGTACTGTGCCAGAAGAGGGAGTCAG	Reverse primer 3' targeting arm for <i>Xist</i> ; Reverse primer ScrFI RFLP in <i>Xist</i>
20	GCTGGTTCGTCTATCTTGTTGG	Forward primer ScrFI RFLP in <i>Xist</i>
25	CTTTGGTCTCTGGGTTTCCA	Forward primer 5' targeting arm for <i>Tsix</i>
27	TACCGGTAGCTGGCTATCACGCTCTTC	Reverse primer 5' targeting arm for <i>Tsix</i>
29	GAGGGCAGATGCCTAAAGTG	Forward primer 3' targeting arm for <i>Tsix</i>
30	CGCAGGCATTTACCTTCAT	Reverse primer 3' targeting arm for <i>Tsix</i>
36	AGTGCAGCGCTTGTGTCA	Forward primer <i>Tsix</i> length polymorphism, for DNA
41	TATTACCCACGCCAGGCTTA	Reverse primer <i>Tsix</i> length polymorphism, for DNA
68	TCCCCAATTA AAGGTGTTGA	Forward primer Pf1MI RFLP in <i>Atrx</i>
69	AATTACGTTCTCCTCTTTCAT	Reverse primer Pf1MI RFLP in <i>Atrx</i>
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 <i>Tsix</i> expression
138	GCTGCTTGGCAATCACTTTA	Reverse primer <i>Tsix</i> expression
157	AACCCTAAGGCCAACCGTAAAAAG	Forward primer <i>Actb</i> expression
158	CATGGCTGGGGTGTGAAGGTCTC	Reverse primer <i>Actb</i> expression
159	GGATCTGCTTGAACACTGTC	Forward primer <i>Xist</i> expression
160	CAGGCAATCCTT CTTCCTGAG	Reverse primer <i>Xist</i> expression
1445	ACTGGGTCTTCAGCGTGA	Forward primer <i>Xist</i> length polymorphism exon 6-7, for RNA
1446	GCAACAACGAATTAGACAACAC	Reverse primer <i>Xist</i> length polymorphism exon 6-7, for RNA

Figure S5 continued

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. **(C)** FACS analysis of mCherry levels and pluripotency marker CD31 in XY *Tsix*-CHERRY (XY) and XGTC-X0 (X0). Percentage of CD31<sup>+</sup> 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. **(D)** Contour plots of EGFP and mCherry levels in undifferentiated XY *Tsix*-CHERRY line as determined by FACS analysis. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events. **(E)** Image of colonies of XGTC-X0 cell line grown out from single cells after single cell dilution plating. Colonies were either mCherry low or high. **(F)** Contour plots of EGFP and mCherry levels of *Xist*-GFP/*Tsix*-CHERRY line grown in 2i+LIF. From top to bottom: Undifferentiated, unsorted cells; cells just after sorting (as indicated by red bounding box and arrows); cells 14 days after sorting; sorted cells transferred to serum+LIF for 14 days; cells transferred to serum+LIF for 14 days and differentiated for 2 days. Starting from outermost contour, lines represent 7.5%, 22.5%, 37.5%, 52.5%, 67.5%, 82.5% of total events. **(G)** RNA sequencing of XGTC-X0 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  $r=0.9832$ .



# CHAPTER



# 4



Chromatin mediated reversible silencing  
of sense-antisense gene pairs in ESCs is  
consolidated upon differentiation

Loos, F., Loda, A., van Wijk, L., Grootegoed, J.A., and  
Gribnau, J.

*(submitted)*



# Chromatin mediated reversible silencing of sense-antisense gene pairs in ESCs is consolidated upon differentiation

Loos, F., Loda, A., van Wijk, L., Grootegoed, J.A., and Gribnau, J.\*

Department of Reproduction and Development

Erasmus MC, University Medical Center, Rotterdam, The Netherlands

\*correspondence: j.gribnau@erasmusmc.nl

*(submitted)*

## 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 (1). 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* (2, 3). 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 (4, 5). 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* (6, 7).

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 RNAs (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 (8), but is dependent on *Tsix* antisense transcription through the *Xist* promoter, which leads to *Xist* promoter associated changes in histone modifications and CpG methylation (9, 10). 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 (11), 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 (12). One form of alpha-thalassemia has been associated with juxtaposition of *LU-C7L* 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 (13). 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. 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 (14). 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 (15). 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 (16) 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

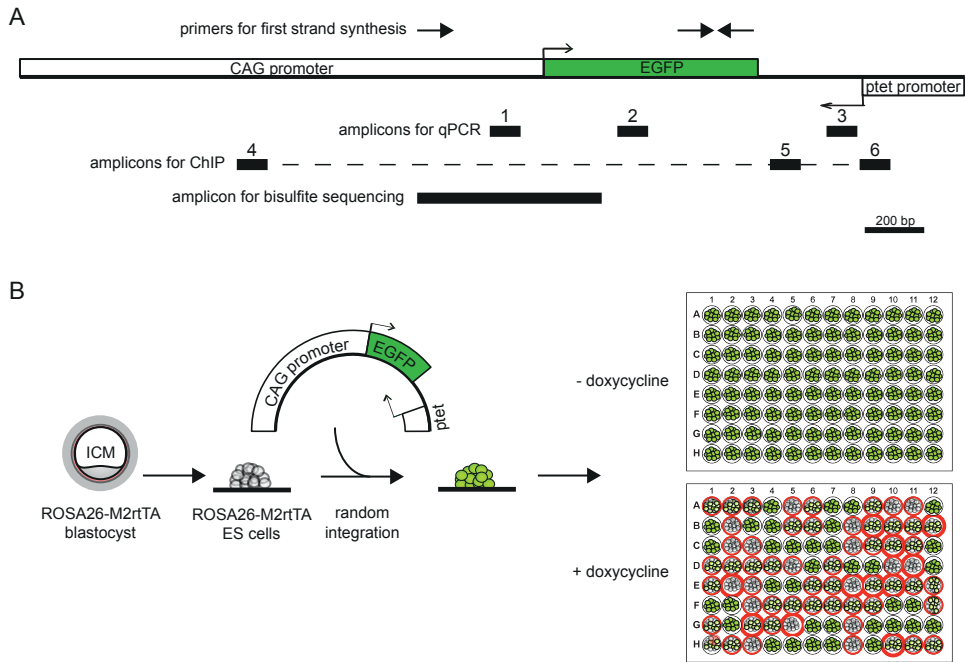
Single cell suspensions were prepared by TE treatment for 7 minutes at 37°C and 30 minutes pre-plating to remove feeder cells if necessary. Duplets were excluded by appropriate gating and dead/dying cells by Hoechst 33258 staining (1 µg/ml, Molecular Probes). Relative fluorescence intensities (FI) were determined for EGFP and mCherry. Cell analysis was performed on LSRFortessa and cell sorting on FACSaria III (BD Biosciences) with FACS Diva 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 S1. Results were normalized to Actin, using the  $\Delta C_T$  method (17) and mostly represented as fold-change versus undifferentiated no doxycycline control.

### Chromatin Immunoprecipitation

In short, approximately  $5 \times 10^6$  cells were cross-linked in dish for 10 minutes at room temperature by adding 1/10 volume 11% buffered formaldehyde solution (50



**FIG 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)** 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 (red circles denote clones with a decreased level of EGFP expression).

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 re-suspended 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  $2^{-CT(\text{pull-down})} / 2^{-CT(\text{input})}$ .

### 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 (18).

### Statistical Methods

Standard deviation was calculated as:

$$\sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

with  $\bar{x}$  for the sample mean and  $n$  for sample size.

## RESULTS

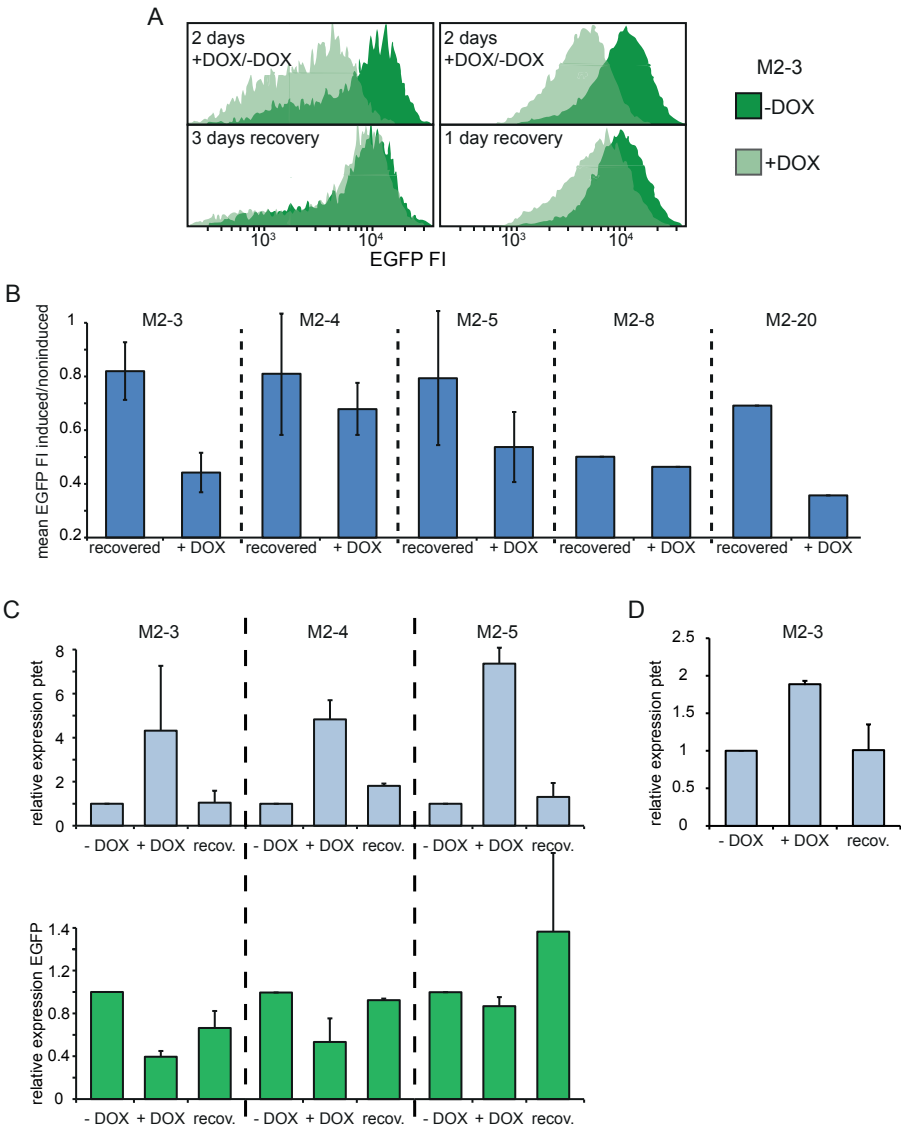
### Inducible antisense transcription reversibly silences an EGFP reporter

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 (19) 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 (16), allowing us to induce doxycycline (DOX)-dependent antisense transcription through the EGFP coding sequence and its promoter. EGFP positive 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 (Fig.1B).

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). 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). Comparable results were found with several other M2 clones. 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





**FIG 2 Antisense transcription and EGFP reporter expression in undifferentiated ESCs**

**(A)** Histograms of EGFP fluorescence intensity (FI) distribution 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). Error bars, where present, represent SD of two or three independent experiments. **(C)** Strand-specific expression analysis in non-induced (-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. Error bars represent SD of two or three independent experiments. **(D)** Same as in top panel of (C) but results for the ptet distal amplicon are shown (amplicon 1).

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 (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.

### **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 (3), but for other such gene pairs a requirement for antisense ncRNA to recruit chromatin modifying complexes has been reported (6). 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 RNA polymerase II (POL2) 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 POL2 associated proteins, SET1 (20) and SET2 (21), 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). Upon DOX addition, however, H3K36me3 accumulated at the CAG promoter with a concomitant decrease in H3K4me3, 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. Analogous to fluorescence and mRNA abundance measurements, this effect was 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 (22). 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 (23, 24). To test if DOX-induced repression of the CAG promoter that drives EGFP expression and contains a CGI in-

4 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, without containing any specific RNA sequences and with no particular locus requirements, antisense transcription generates a special chromatin state at an unrelated promoter, located nearby and transcribed in a coding sense direction, and reversibly silences its experimental sense partner in *cis*.

### 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 (25, 26). 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 transcrip-

### ► FIG 3 Chromatin structure in the CAG and ptet promoter regions in undifferentiated ESCs

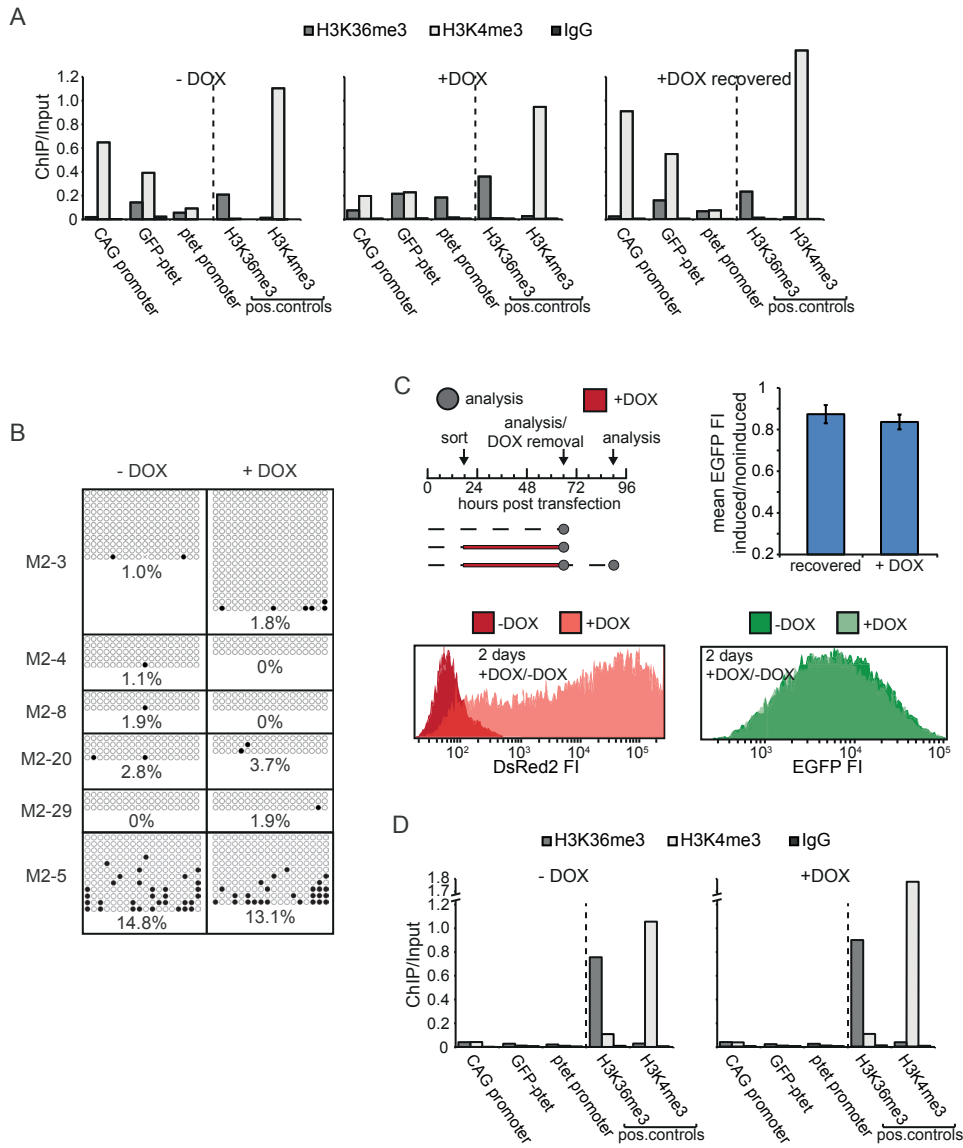
(A) ChIP-qPCR analysis of the region encompassing CAG and ptet promoters in the M2-3 clone for non-induced (-DOX), induced (+DOX), and recovered (+DOX recovered) conditions as outlined in main text. Antibodies against H3K36me and H3K4me3, and whole IgG as mock control, were used as indicated. Loci analyzed, 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. (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.

tion, 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 p<sub>tet</sub> induction are not hampered on a transiently transfected template devoid of the histone modifications H3K4me3 and H3K36me3, repression of EGFP mediated by antisense tran-

scription does not occur in that situation.

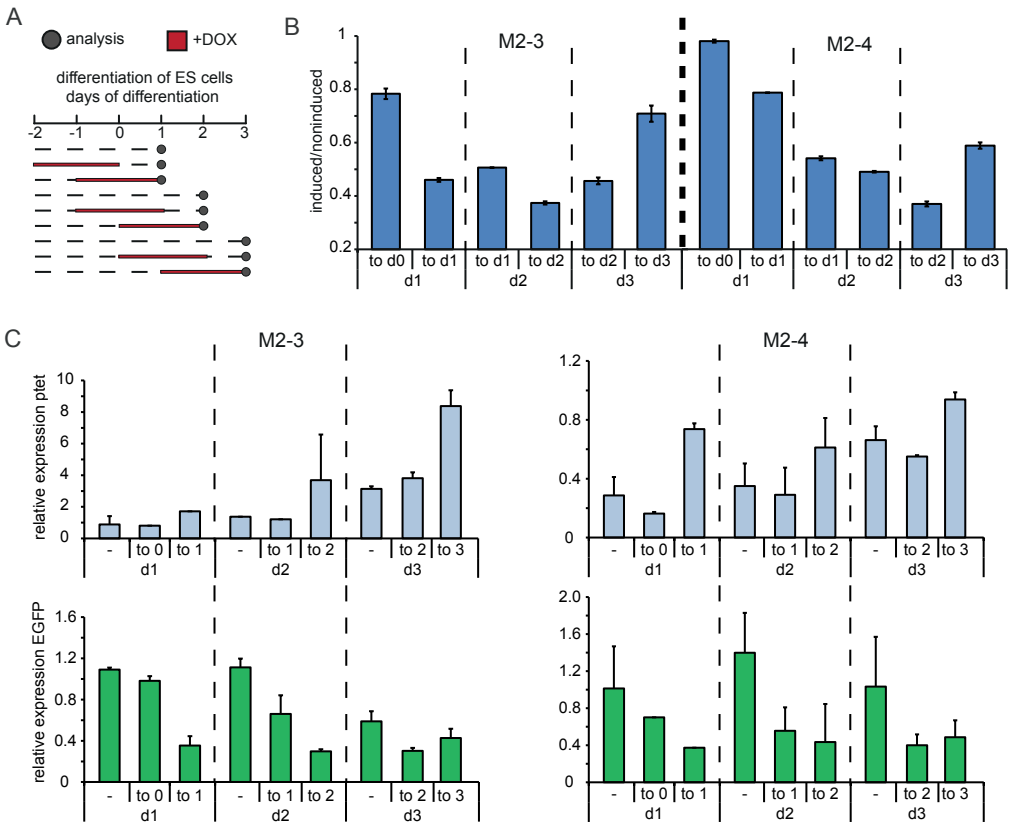
### Antisense transcription-mediated repression during ESC differentiation

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



**FIG 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. Error bars represent SD of three independent experiments. (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 plet 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. Error bars represent SD of two or three independent experiments.

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). 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). 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.

### **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). 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 non-induced state was less complete than in undifferentiated cells (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 (27) and several other genes (24), is highly dynamic during and essential for embryonic development

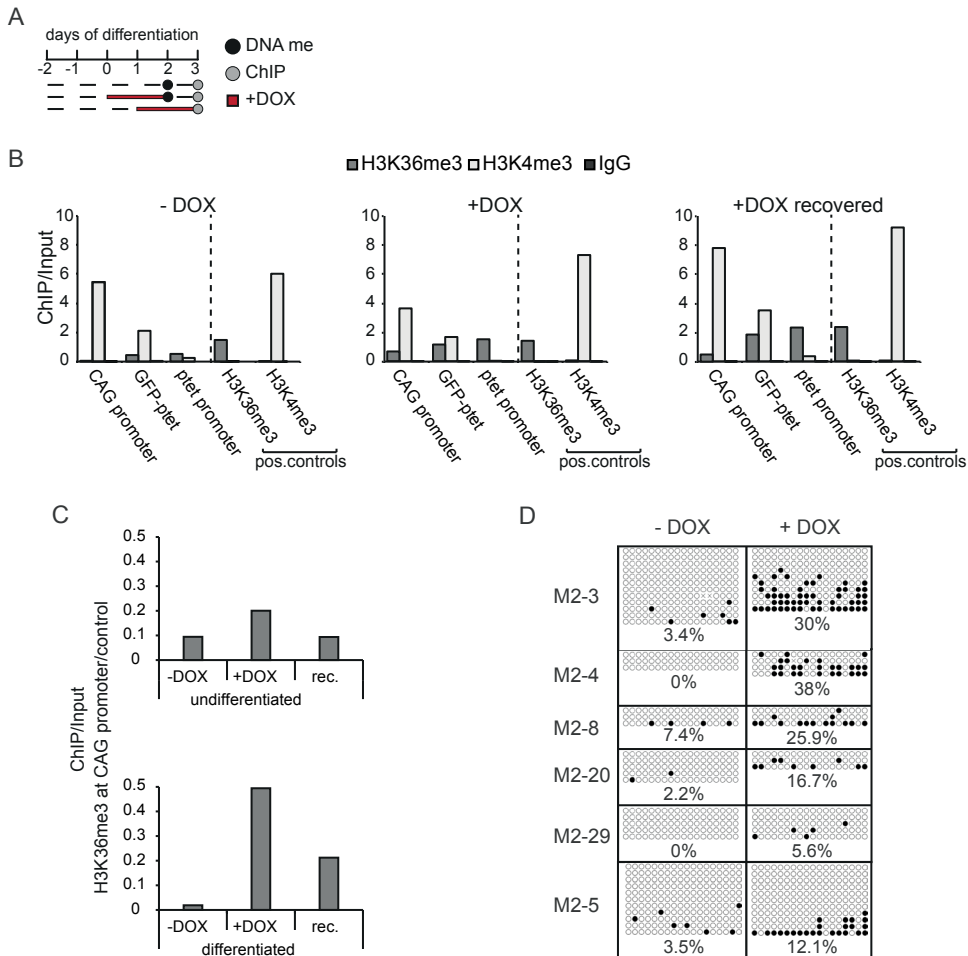


(28, 29). We therefore tested if DNA methylation is involved in the repression of the EGFP reporter, by bisulfite sequencing of cells differentiated for 2 days. Strikingly, all clones analyzed displayed a marked increase in DNA methylation in the CAG promoter which was strictly dependent on antisense 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.

### **Perturbation of chromatin modifications and silencing of the EGFP reporter**

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 small molecule inhibitors to interfere with enzymes that catalyze DNA methylation or the addition or removal of histone modifications. M2-3 cells were differentiated for two days and samples treated for two days with or without DOX, where parallel samples were allowed to recover for 24 hours after two days of DOX regimen. In addition, the following small molecule inhibitors were added to the medium (Table 1): i) 5-aza-2'-deoxycytidine (5-aza-dC), an inhibitor of DNA methylation; ii) suberanilohydroxamic acid (SAHA), an HDAC inhibitor; iii) 2,4-pyridinedicarboxylic acid (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 (30); iv) tranylcypromine (2-PCPA), which preferentially inhibits the monoamine oxidase LSD1 H3K4me1/2 demethylation activity (31); v) curcumin (32), a histone acetyltransferase inhibitor; vi) pargyline (PAR), an LSD1 inhibitor with a preference for inhibition of H3K9me1/2 demethylation (31); and vii) JQ1 which inhibits BRD4 and thereby transcriptional elongation (33). Following these treatments, the cells were analyzed by FACS. Most inhibitors had only minimal effects on repression and recovery in undifferentiated cells (Fig. 6A). SAHA and 5-aza-dC treatment, in line with their reported activating properties, resulted in slightly less efficient repression and subsequent enhanced recovery from repression of the EGFP reporter. In the presence of 2,4-PDCA, antisense transcription-mediated repression was slightly more efficient, which suggests that H3K4me3 and H3K36me3 might be involved in down-regulation of the reporter. Surprisingly, PAR addition increased overall EGFP reporter activity, possibly because inhibition was not preferentially on LSD1 H3K9me1/2 methylation activity, but rather on its H3K4me1/2 demethylation activity. JQ1 treatment resulted in complete loss of recovery from antisense transcription-mediated repression of the reporter. This is consistent with its repressive effect on transcriptional elongation and probably caused by complete ablation of transcription. Upon differentiation, most epigenetic drugs cause partial loss of DOX-inducible EGFP repression and increased recovery after DOX removal (Fig. 6B), showing that these inhibitors interfere to some degree with antisense transcription-mediated repression and stable silencing occurring during differentiation. De-repression was most pronounced in the presence of 5-aza-dC, indicating that DNA methylation might be



**FIG 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. (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 control for normalization. (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.

involved in this stable silencing. Also JQ1 treatment led to marked de-repression of EGFP during differentiation, indicating

that transcriptional elongation plays an important role in establishing the stably repressed state. However, in contrast to

all other small molecule inhibitors used in this study, 2,4-PDCA strongly enhanced both the direct repressive effect and stable silencing by antisense transcription during differentiation. 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

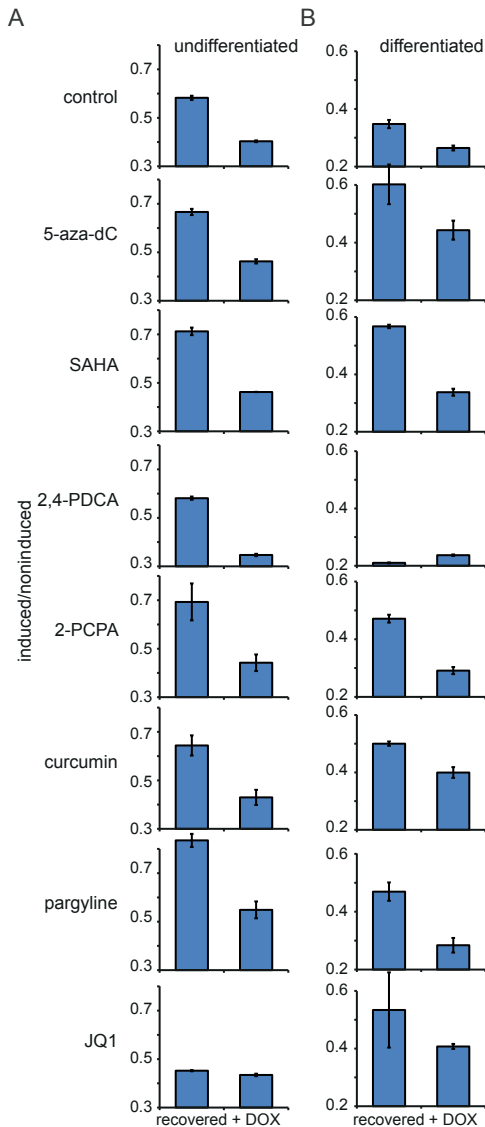
Transcriptional interference mechanisms have been thoroughly studied in prokaryotes and yeast [reviewed in (34)]. These studies indicate inhibition of transcription of the sense gene of a sense-antisense gene pair, which might involve transcription invoked torsional effects or transcriptional collision. Torsional or topological effects have also been implicated in transcriptional interference in higher eukaryotes (35). However, our findings indi-

cate that antisense-mediated repression of a sense gene is absent on transiently transfected templates, precluding an important role for 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 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 differenti-

**TABLE 1 Small molecule inhibitors and their effects on histone modifying enzymes**  
First column gives name, second abbreviation, third enzymes targeted and fourth biological process inhibited.

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



**FIG 6 Interference of chromatin modifiers with small molecule inhibitors**

(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.

ating 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 where the act of transcription and RNA-polymerase 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 RNA-polII, and the resulting accumulation of H3K36me3 in the gene body is important for recruitment of the histone deacetylase Rpd3 (21, 36, 37). In mammals, H3K36me3 is involved in recruitment of the H3K4me3 demethylase KDM5B (38), and the DNA de novo methyltransferases DNMT3A (39), 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 which 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 GFP 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 (9, 10, 40). In addition, forced *Tsix* expression during development results in *Xist* promoter methylation (41). *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 (42). Also for the imprinted *Igf2r/Airn* locus, *Airn* antisense transcription through the *Igf2r* promoter is required for silencing of *Igf2r* (5). 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 (43). 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

[42]. This antisense transcription has been implicated in providing thresholds that need to be overcome for sense genes to be expressed (44). 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 relies on transcription mediated accumulation of histone modifications in promoters leading to gene silencing. 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.

## 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.

## REFERENCES

1. 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 I, 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.
2. 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 Kcnql imprinting control region. *Mol Cell Biol* **24**:7855-7862.
3. Latos PA, Pauler FM, Koerner MV, Sen-gerin 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.
4. 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.
5. Sleutels F, Zwart R, Barlow DP. 2002. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415**:810-813.
6. Pandey R, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C. 2008. Kcnqlot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* **32**:232-246.
7. 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.
8. Nesterova TB, Popova BC, Cobb BS, Norton S, Senner CE, Tang YA, Spruce T, Rodriguez TA, Sado T, Merkschlager M, Brockdorff N. 2008. Dicer regulates Xist promoter methylation in ES cells indirectly through transcriptional control of Dnmt3a. *Epigenetics Chromatin* **1**:2.
9. Sado T, Hoki Y, Sasaki H. 2005. Tsix silences Xist through modification of chromatin structure. *Dev Cell* **9**:159-165.
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. 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.
18. Kumaki Y, Oda M, Okano M. 2008. QUMA: quantification tool for methylation analysis. *Nucleic Acids Res* **36**:W170-W175.
19. Alexopoulou AN, Couchman JR, Whitford 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.

20. **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.
21. **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.
22. **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.
23. **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.
24. **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.
25. **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.
26. **Hebbbar PB, Archer TK.** 2008. Altered histone H1 stoichiometry and an absence of nucleosome positioning on transfected DNA. *J Biol Chem* **283**:4595-4601.
27. **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.
28. **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.
29. **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.
30. **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.
31. **Lee MG, Wynder C, Schmidt DM, McCafferty DG, Shiekhhattar R.** 2006. Histone H3 lysine 4 demethylation is a target of nonselective antidepressive medications. *Chem Biol* **13**:563-567.
32. **Jiao B, Ma H, Shokhirev MN, Drung A, Yang Q, Shin J, Lu S, Byron M, Kalantry S, Mercurio AM, Lawrence JB, Hoffmann A, Bach I.** 2012. Paternal RLIM/Rnf12 Is a Survival Factor for Milk-Producing Alveolar Cells. *Cell* **149**:630-641.
33. **Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, Morse EM, Keates T, Hickman TT, Felletar I, Philpott M, Munro S, McKeown MR, Wang Y, Christie AL, West N, Cameron MJ, Schwartz B, Heightman TD, La Thangue N, French CA, Wiest O, Kung AL, Knapp S, Bradner JE.** 2010. Selective inhibition of BET bromodomains. *Nature* **468**:1067-1073.
34. **Shearwin KE, Callen BP, Egan JB.** 2005. Transcriptional interference--a crash course. *Trends Genet* **21**:339-345.
35. **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.
36. **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.

37. 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.
38. 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.
39. 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.
40. 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.
41. 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.
42. Beard C, Li E, Jaenisch R. 1995. Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev* **9**:2325-2334.
43. 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.
44. 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.

## Supplemental Information

Chromatin mediated reversible silencing of sense-antisense gene pairs in ESCs is consolidated upon differentiation

**TABLE S1 Primers used in this study as listed in the Materials and Methods section**

#	SEQUENCE	DESCRIPTION
228	ACACGCAGCTCATTGTAG	First strand primer for strand-specific <i>Actb</i> expression
226	GATATCGCTGCGCTGGTCGT	FOR primer <i>Actb</i> expression
227	AGATCTTCTCCATGTCGTCC	REV primer <i>Actb</i> expression
217	CTTCTCGTTGGGGTCTTTGC	First strand primer for strand-specific EGFP expression
106	AGGGCATCGACTCAAGGAG	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	CTGCAGAAATTCTAGAGCCGC	REV primer ptet distal expression
195	CTCTGACTGACCGCGTTACT	FOR primer ChIP for CAG promoter
196	TTTCACGCAGCCACAGAAAA	REV primer ChIP for CAG promoter
323	CACATTGTAGAGGTTTACTTGCT	FOR primer ChIP for GFP-ptet
324	AGCTGCAATAAACAAAGTTAACAACA	REV primer ChIP for GFP-ptet
317	GAGCTCGAATTCTCCAGGCG	FOR primer ChIP for ptet promoter
318	GTATGTCGAGGTAGGCGTG	REV primer ChIP for ptet promoter
343	TCTCCAGCATCCTCTACACA	FOR primer ChIP for H3K36me3 control in <i>Mcm2</i>
344	CTATGGTATGTGTGGTGGGCA	REV primer ChIP for H3K36me3 control in <i>Mcm2</i>
331	GCTAGGTTAGGAGAGCCCAGA	FOR primer ChIP for H3K4me3 control in <i>Mrps14</i>
332	AGGTCTCAATCATCCGACTCTC	REV primer ChIP for H3K4me3 control in <i>Mrps14</i>
204	GGATTTTTTTTGTTTAAATTTGTG	FOR primer for bisulfite sequencing amplicon
207	AAATAAACTTCAAAATCAACTTACC	REV primer for bisulfite sequencing amplicon
209	GTA AACGACGGCCAG	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)

# CHAPTER



# 5



## General Discussion

Friedemann Loos



5

## General Discussion

Dosage compensation has evolved independently in several taxa to account for the gene dosage differences in heterogametic species (reviewed in Disteche, 2012). This together with the highly deleterious nature of aneuploidies and polysomies for example in humans demonstrates how essential proper gene dosage is for normal development and cell homeostasis. To solve the potential gene dosage problem related to the evolution of sex chromosomes, in mammals a mechanism called X chromosome inactivation (XCI) has co-evolved. This process leads to transcriptional inactivation of one X chromosome in every female cell that is initiated around pre-implantation stage of embryogenesis. The inactive X (Xi), which can be of maternal or paternal origin, is subsequently clonally propagated through cell division, implying that female mammals essentially represent a mosaic organism of cells in which either one or the other X chromosome is silenced (reviewed in Loda et al., 2015/chapter 1). In summary, more than 50 years after XCI has first been postulated, this process is still a fascinating paradigm for epigenetics and gene regulation in general, and specifically for the developmental context it occurs in.

The following paragraphs discuss the different levels of regulation that underlie XCI and relate these processes to a special window of time during early embryonic development in which major epigenetic changes take place.

## X chromosome inactivation is governed by multiple levels of gene regulation

XCI is a multifaceted process which has to ensure not only the timely and robust initiation of silencing during early embryonic development, but must also guarantee its propagation and stability in diverse tissues over the entire life-time of an animal. Therefore the regulation of XCI relies on several interwoven and sometimes redundant principles which only together can provide error free execution of the entire program.

At the heart of the regulation of initiation of XCI lies the X chromosome inactivation center (*Xic*). This concept has originally been introduced to explain the phenotypes of X-chromosomal translocations, deletions and truncations and delineates the region necessary for XCI to occur *in cis* (Rastan and Robertson, 1985). Several genes in the *Xic* have been implicated in the regulation of XCI, most importantly *Xist*, whose noncoding RNA product is necessary to silence the X chromosome (Borsani et al., 1991; Brockdorff et al., 1991; Penny et al., 1996) and *Tsix* which is a negative regulator of *Xist* (Lee and Lu, 1999; Lee et al., 1999). The presence of several DNase I-hypersensitive sites in the *Xist* promoter suggests that *Xist* regulation, just as for other genes, involves the action of specific transcription factors (Sheardown et al., 1997). However, the first X-linked activator of XCI, whose existence was predicted from studies in ES cells demonstrating that the X to autosome ratio determines the probability to undergo XCI (Monkhorst et al., 2009), was found to be the E3 ubiquitin ligase RNF12 (Jonkers et al., 2009). RNF12 activates *Xist* by targeting REX1, a transcriptional repressor of *Xist*, for degradation (Gontan et al., 2012). The pluripotent state of ES cells is moreover linked to repression of XCI by several pluripotency

factors which repress *Xist* and/or activate *Tsix* (Donohoe et al., 2009; Ma et al., 2011; Navarro et al., 2008; Navarro et al., 2010). Interestingly, female *Rnf12*<sup>+/-</sup> ES cells do not completely resemble male ES cells in that a substantial percentage of cells still initiates XCI, which suggests the existence of additional X-linked *trans*-activators. In addition, comparison of ES cells with different sex chromosome compositions indicates differential activity of signaling pathways between cells harboring one and two X chromosomes, which results in blockage of exit from the pluripotent state in XX cells (Schulz et al., 2014). This differential signaling environment might also transmit cues influencing the XCI process.

Our deletion and transgenic rescue experiments in chapter 2 did not reveal any additional *trans*-acting factors, but rather provide evidence that several additional genes located in the *Xic* most likely exert their influence strictly *in cis*. *Jpx*, *Ftx* and *Xpr* have been reported to activate *Xist* expression (Augui et al., 2007; Chureau et al., 2011; Tian et al., 2010), however, in our hands loss of *Jpx*/*Ftx* or *Xpr* does not influence XCI *in trans*, in addition BAC transgenes expressing these genes or introducing additional *Xpr* elements do not affect XCI, which contrasts with results obtained using *Rnf12* transgenes. Therefore, we propose that *Jpx*, *Ftx* and *Xpr* are part of a *cis*-acting *Xic*, while *Rnf12* is the only so far identified gene of the *trans*-acting *Xic*. Since *Xite*, *Tsx* and *Ppnx/Linx* activate or are co-activated with *Tsix* (Anguera et al., 2011; Nora et al., 2012; Ogawa and Lee, 2003) analogously to *Jpx/Ftx/Xpr*-dependent *cis*-activation of *Xist*, these genes belong to the *cis*-*Xic*, as well. Interestingly, recent studies identifying the topological conformation of the *Xic* have shown that genes belonging to the *Xist* or *Tsix* *cis*-regulatory region are embedded in two different, adjacent TADs, and deletion of a boundary element between them re-

sults in dysregulation of the *Xic* (Nora et al., 2012; Spencer et al., 2011). Along the same lines, the configuration of the *Tsix* TAD appears to correspond to a defined transcriptional output of genes embedded in it (Giorgetti et al., 2014). It is thus tempting to speculate that genes located in the same TADs share common regulatory elements responsible for a certain degree of co-regulation. This also implies that the *Xist* and *Tsix* TADs might represent the entire *cis*-regulatory region involved in timely and robust execution of XCI.

Another interesting finding presented in chapter 2 regards the temporally extended requirement for at least one active copy of *Rnf12*. This was shown by preferential inactivation of the mutant allele in *Rnf12*<sup>+/-</sup> ES cells and by a compound deletion of *Xist* and *Rnf12* *in cis*. This indicates that inactivation of the single intact *Rnf12* copy would lead to recurrent repression of *Xist* via the REX1 axis and failure to properly establish an Xi. It also explains the imprinted XCI (iXCI) defect observed in the case of maternal transmission of a *Rnf12* deletion (Shin et al., 2010), because the single intact *Rnf12* copy is of paternal origin and would inevitably be silenced upon iXCI. However, as *Rex1* is quickly down-regulated upon differentiation and *Rex1*<sup>-/-</sup> female mice are born (Masui et al., 2008), it remains to be determined in how far REX1 is indeed the only effector of *Rnf12* deletions. Moreover, even though an almost absolute requirement exists for RNF12 in random XCI in ES cells (Barakat et al., 2011), a recent study using a Sox2-driven conditional *Rnf12* knockout *in vivo* demonstrates that this knockout is viable and that *Rnf12*<sup>-/-</sup> females show hallmarks of XCI (Shin et al., 2014). This effect might be attributed to the fact that XCI *in vivo* might be much more robust than in ES cells, as is also evident from the presence of a significant proportion of differentiated ES cells with no *Xist* clouds

(Monkhorst et al., 2008). For example, on top of specificity, timing and penetrance issues regarding the Sox2-driver, additional activators of XCI might compensate in such a *Rnf12* knockout situation.

As shown by various genetic studies in mice and ES cells, the mutual interplay of *Xist* and *Tsix* is essential for the regulation of XCI (reviewed in Loda et al., 2015/ chapter 1). However, as these two genes completely overlap and have opposing effects on the outcome of XCI, the analysis of their mutual regulation is severely complicated. This is exemplified by our results in chapter 2, demonstrating that a *Tsix* deletion *in cis* to a *Jpx/Ftx/Xpr/Rnf12* deletion rescues the corresponding XCI phenotype. Similar results have been obtained in a different study which exclusively examined the role of *Jpx* in XCI (Tian, 2010). Loss of *Xist* activation as a consequence of a *Jpx* knockout can be rescued by a deletion of *Tsix in cis*. This suggests that *Tsix* expression might potentially mask any regulatory effects exerted on the *Xist*-regulatory network. In chapter 3 we therefore employ a series of fluorescent reporter constructs for *Xist* and *Tsix* to investigate their early regulation on an uncoupled allele. That is, we replaced the first exons of *Xist* and *Tsix* with an EGFP and mCherry cassette, respectively. Both reporters are thus under control of the endogenous promoters and a polyA-signal terminates transcription downstream of the reporter. Comparing the single EGFP and mCherry knock-in's to the double knockin on the same allele, we observed increased *Xist*-promoter driven GFP expression in double knockin ES cells confirming the described *Tsix*-mediated *Xist* repression. In addition we found a repressive effect of *Xist* on *Tsix* not only on the Xi but also on the future active X. We also observed transient bi-allelic *Xist* up-regulation which is consistent with a role of *Xist* in the repression of *Tsix* in that this transient up-regulation might

be sufficient to silence *Tsix in cis*. These findings suggest that mono-allelic *Xist* expression and accumulation which finally lead to inactivation of one X chromosome might get established later than currently believed. At least the visual accumulation of *Xist* in a cloud on the Xi occurs after an initial bi-allelic up-regulation and appears to differ from *Xist*'s very local action on the Xa. How such a mechanism might work or how differential *Xist* expression could be achieved at a later time point in development remains to be elucidated. Yet again, these findings also emphasize the need to investigate their regulation independent of interfering aspects of the corresponding antisense partner. In this context, we also show that known activators and inhibitors of XCI can only properly function on balanced alleles, that is, or on the wild type allele or the *Xist*-GFP/*Tsix*-CHERRY double knockin allele. The effects of *Rex1* and *Rnf12* transgenes on the single knockin's were severely attenuated, suggesting that the process of antisense transcription or the antisense transcript itself overrules any regulatory impact these proteins might have. For future experiments the *Xist*-GFP/*Tsix*-CHERRY line might be very useful to discover novel activators and inhibitors of XCI, because i) it provides a robust and easy readout of activation or inhibition and ii) it enables distinction between direct and indirect effects on the promoters of *Xist* and *Tsix*. Of note, the deletion of a few REX1 and pluripotency factor binding sites due to the process of generating the knockin cell lines does not appear to abolish proper developmental regulation of *Xist* and *Tsix*. Recent studies have indicated competitive binding of REX1 and YY1 to this region, REX1 being involved in repression of *Xist*, whereas YY1 acts as an *cis*-activator of *Xist* (Makhlouf, 2014). However, our results, indicate that REX1 and RNF12 likely act in a redundant fashion on multiple sites. We also derived an X0 *Xist*-GFP/

Tsix-CHERRY ES cell line from the original XX *Xist*-GFP/*Tsix*-CHERRY line. Analysis of this line not only confirmed reduced dynamics of the *Xist* promoter, most likely due to lower levels of RNF12 and therefore increased levels of REX1, but also showed a striking reduction in *Tsix* promoter activity. These data indicate that *Tsix* activity is also dependent on the X to autosome ratio, which, in addition to RNF12-REX1-mediated activation of *Xist* upon differentiation, might provide a fail-proof mechanism to keep *Xist* activity in check in undifferentiated female ES cells. Indeed, mutation of *Tsix* in undifferentiated female ES cells results in ectopic up-regulation of *Xist*, albeit that it does not cause *Xist* accumulation and XCI on the mutated allele (Lee, 1999; Luikenhuis, 2001).

5 Easy readout of reporter activity in *Xist*-GFP/*Tsix*-CHERRY cells also allowed detailed analysis of the dynamic patterns of *Xist* and *Tsix* regulation during differentiation. It is important to note, however, that due to reporter half-lives of ca. 10–14 hours, we do not extract information about small scale events like transcriptional bursting but rather obtain estimates of general promoter activity by integration over time. FACS and live imaging of single cells showed that *Xist* up- and down-regulation precede *Tsix* down-regulation on the Xa. Without interference from the antisense partner, this pattern is developmentally concerted, but not strictly dependent on each other. *Tsix* repression is loosely correlated to *Xist* activation but can occur before, after or right at the time point of *Xist* activation. This implies that *Xist* up-regulation and *Tsix* down-regulation depend on a stochastic mechanism most likely influenced by concentration gradients of *trans*-acting factors like RNF12, CTCF, YY1, REX1 or other pluripotency factors rather than depending on defined sequential events. It also shows that *Tsix* will be repressed eventually, even in the ab-

sence of a *Xist*-mediated silencing mechanism. In addition, following single cells over several cell divisions during ES cell differentiation and scoring rates of EGFP and mCherry production, we were able to construct pedigrees of cell lineages and correlate them to *Xist* and *Tsix* promoter activities. Interestingly, the two daughter cells of any given cell division behave very similarly with regard to activation levels of the *Xist* and *Tsix* promoters and mirror the state of the mother cell as compared to other cells of that generation. Only occasionally daughter cells deviate from each other, which can be interpreted as switches in promoter activity. Therefore, states of promoter activity are surprisingly stable and switch only slowly or rarely. This observation is supported by the presence of cells that strongly activate the *Xist*-GFP reporter, but never manage to switch to activation of the wild-type *Xist* allele and therefore never accumulate an *Xist* domain on the wild-type X chromosome in the time span that we were able to monitor the cells. In general, switches might occur in a certain phase of the cell cycle, because events of *Xist* activation appear to cluster. Most likely they occur around mitosis, because we rarely observe dents in the slopes of fluorescent intensity raw data tracks from live imaging experiments from G1 through G2. This points to a role for mitosis in the establishment of transcriptional states at the *Xic*, as has been proposed based on findings that transcription factors are displaced during mitosis thereby providing a window of opportunity for cell state transitions (Martinez-Balbas et al., 1995; reviewed in Egli et al., 2008).

The surprising stability of transcriptional states is also evident from the presence of two mCherry populations in the XX *Xist*-GFP/*Tsix*-CHERRY line grown under 2i plus LIF conditions and the *Xist*-GFP/*Tsix*-CHERRY X0 ES cell line. Careful analysis of these cell lines indicates

that indeed the two cell populations are not explained by heterogeneity of the cell line or its pluripotent state. Therefore, they represent two truly distinct states which can only be based on epigenetic phenomena. We rule out DNA methylation to be the cause of these distinct states because we detected only very low levels of CpG methylation in the *Tsix*-associated CpG island, consistent with general hypomethylation in ES cells (Zvetkova et al., 2005), particularly under 2i conditions (Habibi, 2013). Moreover, DNA methylation at the *Tsix* regulatory *DxPas34* region and its CpG island is generally believed to be acquired only at post-implantation stages (Prissette et al., 2001). Instead, we favor a model in which the different transcriptional states that we observe represent distinct 3D conformational states of the *Xic*, which would correspond to the two major conformational classes predicted by a recently introduced polymer model (Giorgetti et al., 2014). These two conformational classes proposed on the basis of *in silico* analysis of 5C data are associated with low or high *Tsix* expression consistent with our observation of mCherry low and high populations. High-resolution 3D DNA-FISH and 3C-type experiments will help to determine if these conformational classes correlate with the observed mCherry low and high states. Close examination of the *Xic* expression profile in these populations resembles predictions from this model and moreover mirrors the reported co-regulation within and the anti-correlation between the *Tsix* and *Xist* TADs (Nora et al., 2012). We therefore envision a model in which a given *Xic* can adopt distinct conformations which favor either *Tsix* or *Xist* expression. These conformations most likely switch. However, the presence of stable populations indicates that cells showing low mCherry levels have a higher probability to assume a *Xic* conformation favorable for *Xist* expression whereas cells showing high mCherry levels spend more

time in the *Tsix* expression conformation. Based on the observation that in sorted populations only developmental cues like transfer to differentiation medium result in at least partial recovery of the initial two populations, we propose that only a very short time window at the onset of differentiation might allow cells to switch states. This phase of highly dynamic chromatin movements has also been described using a different system of live imaging (Masui et al., 2011). The fact that we do not see these two distinct populations in female ES cells grown in serum conditions, might indicate that in these cells the *Xic* alleles rapidly fluctuate between distinct states, or that in these cells all alleles have adopted the same configuration.

The differentiation of sorted mCherry low and high populations presented in chapter 3 revealed another interesting finding. The potential to up-regulate *Xist* and undergo XCI under 2i conditions appears to be restricted already before differentiation is initiated. That is demonstrated by almost exclusive activation of the *Xist*-GFP and a markedly increased frequency of XCI initiation in the mCherry low population only. While activation of the *Xist*-GFP might be a *cis*-effect resulting from the conformational state of the mutant allele, that is, a conformation that favors low *Tsix* and higher *Xist* expression, preferential activation of the wild-type *Xist* and subsequent coating of the X chromosome suggest a *trans*-effect. This *trans*-effect might be mediated by increased RNF12 levels, because *Rnf12* belongs to the *Xist* TAD which in mCherry low cells presumably has adopted a conformation that is favorable for gene expression. Because of feedforward and feedback loops involving REX1 mediated repression of *Rnf12*, a more active state of *Rnf12* on one allele might impact *Rnf12* expression on the other allele. An implication of these data is that “choice”,



if any, has to have occurred by and large already before developmental cues initiate XCI. This reminds of a model proposing that the X chromosome exist in different epigenetic states prior to initiation of XCI (Mlynarczyk-Evans et al., 2006), albeit that we find these states to be very stable. Our data also suggest that the very short window of time right at the onset of differentiation when major topological changes occur might be too short in ES cells grown in 2i to still modulate the outcome of XCI. Our findings may also explain the population of serum grown and differentiated ES cells that do not show signs of XCI because some cells might harbor two *Xic* alleles in a *Tsix*-favorable conformation. However, how different states that are predictive for the result of XCI are established in ES cells, and if cell-cell signaling plays a role in this process, remain open questions.

Taken together our results described in chapter 2 and chapter 3 highlight that XCI is governed by different principles of transcriptional regulation: Transcription factors repress and activate genes in the *Xic*, antisense transcription at the *Xist*/*Tsix* locus results in mutual antagonistic interference and epigenetic mechanisms including the 3D conformation of the *Xic* determine the activation potential of genes embedded within separate TADs. Moreover, we provide evidence that the regulation of initiation of XCI might be less dynamic and flexible than previously believed.

### **A unique developmental window of time for chromatin-mediated gene regulation**

One of the most striking features of XCI is its close dependency on developmental states. On top of *Xist* and *Tsix* regulation by pluripotency factors, this is impressively highlighted by elegant ex-

periments forcing *Xist* expression at three different stages in development. In undifferentiated cells, silencing, if at all, is mild and reversible (Loda, unpublished; Wutz and Jaenisch, 2000). In contrast, during early differentiation, *Xist*-mediated silencing is fixed and becomes irreversible. Finally, forced *Xist* expression at later stages of differentiation is incapable of inducing gene repression (Wutz and Jaenisch, 2000). These findings clearly show that some developmental accessory factors or a specific environment are necessary to “lock in” silencing. Most evidence suggests that this effect is mediated via epigenetic modifications, since the Xi displays a very particular chromatin state (reviewed in Nora and Heard, 2010) with DNA methylation being acquired at a late stage (Lock et al., 1987; Sado et al., 2004) and in a peculiar pattern, namely mostly at CpG islands and not at inter- and intragenic CpG’s (Weber et al., 2007). Moreover, DNA methylation appears to be required for stable silencing, because interfering with DNA methylation results in reactivation of the Xi and of *Xist* on the Xa (Beard et al., 1995; Mohandas et al., 1981; Sado et al., 2000).

In appendix A we present a curious case of developmentally induced DNA methylation, which emphasizes the presence and importance of chromatin-mediated gene regulation in a specific developmental context. Fragile X syndrome (FXS) is a genetic disease associated with mental retardation and caused by a trinucleotide repeat expansion in the *FMR1* gene (reviewed in Brouwer et al., 2009). A human ES cell line carrying the mutation showed that *FMR1* was normally expressed in undifferentiated cells and became silenced and methylated only upon differentiation (Eiges et al., 2007). This indicates that while the presence of the CGG expansion is not sufficient to cause silencing in the undifferentiated state, a specific environment at later stages of development induc-

es gene repression. In another study, generation of iPS cells from FXS fibroblasts did not revert DNA methylation or silencing of *FMR1*, demonstrating that the epigenetic changes induced during early embryonic development are not reversible, at least not by the protein complexes present during the reprogramming process (Urbach et al., 2010). Alternatively, these cells might have never passed through that special regulatory phase due to difficulties in reprogramming human cells to a naïve pluripotent state. In appendix A we report a third case in which we generated iPS cells from carriers of the full *FMR1* mutation that nonetheless was unmethylated and therefore still expressed in neurons. Upon reprogramming, the mutated *FMR1* acquired DNA methylation and was silenced. These results suggest that the reprogrammed cells passed through a developmental stage in which the fully mutated *FMR1* allele was “properly” recognized and targeted for silencing. However, it raises the question how the very same allele was originally protected during normal embryonic development of the carriers. Since two brothers show the same phenotype of escaping from methylation of the mutated *FMR1*, it is likely that a genetic component is responsible for it. For example, the brothers might both lack a functional protein complex that recognizes a specific, repression-inducing signature at the mutated *FMR1* allele which results in loss of silencing at that locus. It is not clear, though, why this protective function is lost later on when fibroblast are reprogrammed to iPS cells. One possible explanation could be that a similar complex is ectopically activated during the reprogramming process. Whatever the reason for this remarkable genetic protection from silencing of a fully mutated *FMR1* allele and loss thereof upon reprogramming, thorough analysis of the described fibroblast and iPS cell lines might yield important insights into

the nature of chromatin factors that govern gene regulation during development.

We further pursued the idea of a special regulatory phase during early development by a series of experiments described in chapter 4. We were prompted by an interesting characteristic of the *Xic*, i.e. the overlap between *Xist* and *Tsix* and their antagonistic roles in the process of XCI. Several models of how *Tsix* expression could repress *Xist* *in cis* have been proposed, including transcriptional interference, RNAi-related pathways and transcription-mediated modulation of chromatin structure in the *Xist* promoter (reviewed in Loda et al., 2015/chapter 1). Strikingly, the transcriptional overlap of *Tsix* with the *Xist* promoter is responsible for the generation of a repressive chromatin environment at the *Xist* promoter (Ohhata et al., 2008; Sado et al., 2005) and generally modulates chromatin structure over the entire *Tsix-Xist* transcription unit (Navarro et al., 2005). Moreover, the silenced alleles of both *Xist* and *Tsix* acquire DNA methylation at later stages of development (Norris et al., 1994), which represents a special case as only a subset of promoters including those of germ line, pluripotency and lineage-specific factors are methylated while most remain unmethylated (Borgel et al., 2010; Smith et al., 2012). However, DNA methylation appears to be dispensable for the early establishment of *Xist* expression patterns and initiation of XCI (Sado et al., 2004). A similar situation was reported for the imprinted *Igf2r/Airn* locus, in which the transcriptional overlap alone is sufficient for *Igf2r* repression and silencing is stabilized at later stages by DNA methylation (Latos et al., 2012; Santoro et al., 2013). Based on these findings and the aforementioned developmental “window of opportunity” for *Xist*-mediated silencing, we hypothesized that antisense transcription through an active promoter might be a more wide-

spread mechanism to repress genes during development. This idea is especially tempting in the light of the large amount of noncoding RNA transcripts present in mammals. Our inducible reporter system described in chapter 4 was devised to test this hypothesis in a general context independent of specific loci or sequences. Of note, we assumed that induced antisense transcription through the EGFP cassette would not give rise to a functional RNA species, thus precluding RNA mediated repression in this system. Interestingly, antisense transcription-mediated repression of EGFP was completely reversible in undifferentiated ES cells. This repression was dependent on an intact chromatin template, as shown by loss of repression on transiently transfected plasmids that were aberrantly chromatinized. Thus, it seems unlikely that mere transcriptional interference is causing repression of the EGFP reporter. Rather, antisense transcription induces changes in chromatin structure particularly at active promoters, as shown by ChIP. It results in a specific H3K4me<sub>3</sub>-H3K36me<sub>3</sub> signature that in turn might recruit repressive factors to the promoter. In undifferentiated cells, this process is reversible, possibly because at this stage of development an unidentified complex capable of transmitting the silencing signal is absent. This is supported by analysis of the EGFP promoter which is not gaining DNA methylation in undifferentiated cells, even if repressed. However, the situation changes completely if antisense transcription through the EGFP promoter is induced during differentiation. The EGFP reporter does not recover from repression and DNA methylation is acquired at its promoter. This clearly demonstrates the presence of differentiation-specific chromatin machinery that greatly influences the outcome of antisense transcription-mediated gene repression. Since interference with demethylation of H3K4me<sub>3</sub>/H3K36me<sub>3</sub> resulted in enhanced repression of EGFP,

particularly upon differentiation, and particularly abolished recovery of EGFP from repression, we believe that this effect is relayed by factors binding to H3K4me<sub>3</sub> and/or H3K36me<sub>3</sub>. For example, an extensive proteomics approach revealed combinatorial specificities for histone binding proteins (Vermeulen et al., 2010). Moreover, H3K36me<sub>3</sub> and H3K4me<sub>3</sub> display antagonistic effects on the recruitment of DNMT's (Dhayalan et al., 2010; Ooi et al., 2007), implying that the ratio between these two histone modifications might be instructive for recruitment of DNMT3A/B and/or additional factors. As the DNA methyltransferases DNMT3A/B interact with heterochromatin-associated proteins and transcriptional repressors (Fuks et al., 2001; Velasco et al., 2010), they could indeed be key factors in relaying a silencing cue. However, our results in ES cells showing chromatin mediated reversible repression in the absence of DNA methylation suggest that additional factors might contribute to initial gene repression. Taken together, we favor a model in which the process of transcription by virtue of the association of Pol2 with chromatin modifiers (reviewed in Smolle et al., 2013), but not its RNA product or mere transcriptional interference, is responsible for repression of an overlapping promoter. Upon differentiation, hallmarks of this repression are then recognized by complexes specific for a developmental context and relayed into stable repression including DNA methylation. Clearly, future research directed at revealing the identity of such chromatin complexes will be instrumental in understanding developmental programming of gene regulation.

In conclusion, our result described in chapter 4 and appendix A emphasize that a short window of time just at the onset of differentiation in ES cells represents a crucial phase for modulation of chromatin structure and gene regulation. In this

context, specific chromatin features which might be partially dependent on transcription are sufficient to result in sustained changes in gene expression programs.

## References

- Anguera, M.C., Ma, W., Clift, D., Namekawa, S., Kelleher, R.J., 3rd, 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.
- 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.
- Barakat, T.S., Gunhanlar, N., Gontan Pardo, C., Achame, E.M., Ghazvini, M., Bowers, 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.
- Beard, C., Li, E., and Jaenisch, R. (1995). Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev* 9, 2325-2334.
- Borgel, J., Guibert, S., Li, Y., Chiba, H., Schubeler, D., Sasaki, H., Forne, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. *Nat Genet* 42, 1093-1100.
- 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.
- Brouwer, J.R., Willemsen, R., and Oostra, B.A. (2009). The FMR1 gene and fragile X-associated tremor/ataxia syndrome. *Am J Med Genet B Neuropsychiatr Genet* 150B, 782-798.
- 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.
- Dhayan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem* 285, 26114-26120.
- Disteche, C.M. (2012). Dosage compensation of the sex chromosomes. *Annu Rev Genet* 46, 537-560.
- Donohoe, M.E., Silva, S.S., Pinter, S.F., Xu, N., and Lee, J.T. (2009). The pluripotency factor Oct4 interacts with Ctf and also controls X-chromosome pairing and counting. *Nature* 460, 128-132.
- Egli, D., Birkhoff, G., and Eggan, K. (2008). Mediators of reprogramming: transcription factors and transitions through mitosis. *Nat Rev Mol Cell Biol* 9, 505-516.
- Eiges, R., Urbach, A., Malcov, M., Frumkin, T., Schwartz, T., Amit, A., Yaron, Y., Eden, A., Yanuka, O., Benvenisty, N., *et al.* (2007). Developmental study of fragile X syndrome using human embryonic stem cells derived from preimplantation genetically diagnosed embryos. *Cell Stem Cell* 1, 568-577.
- Fuks, F., Burgers, W.A., Godin, N., Kasai, M., and Kouzarides, T. (2001). Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J* 20, 2536-2544.
- 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.
- Gontan, C., Achame, E.M., Demmers, J., Barakat, T.S., Rentmeester, E., van, I.W., Grootegoed, J.A., and Gribnau, J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature* 485, 386-390.
- 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 chro-

mosome inactivation. *Cell* 139, 999-1011.

Latos, P.A., Pauler, F.M., Koerner, M.V., Senergin, H.B., Hudson, Q.J., Stocsits, R.R., Allhoff, W., Stricker, S.H., Klement, R.M., Warczok, K.E., *et al.* (2012). Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* 338, 1469-1472.

Lee, J., and Lu, N. (1999). Targeted mutagenesis of Tsix leads to non-random X inactivation. *Cell* 99, 47-57.

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

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.

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

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.

Martinez-Balbas, M.A., Dey, A., Rabin dran, S.K., Ozato, K., and Wu, C. (1995). Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* 83, 29-38.

Masui, O., Bonnet, I., Le Baccon, P., Brito, I., Pollex, T., Murphy, N., Hupe, P., Barillot, E., Belmont, A.S., and Heard, E. (2011). Live-Cell Chromosome Dynamics and Outcome of X Chromosome Pairing Events during ES Cell Differentiation. *Cell* 145, 447-458.

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

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

Mohandas, T., Sparkes, R.S., and Shapiro, L.J. (1981). Reactivation of an inactive human

X chromosome: evidence for X inactivation by DNA methylation. *Science* 211, 393-396.

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

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.

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

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

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.

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

Norris, D.P., Patel, D., Kay, G.F., Penny, G.D., Brockdorff, N., Sheardown, S.A., and Rastan, S. (1994). Evidence that random and imprinted Xist expression is controlled by preemptive methylation. *Cell* 77, 41-51.

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. (2008). Crucial role of anti-sense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. *Development* 135, 227-235.

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., *et al.* (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* 448, 714-717.

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.

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

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.

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

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

Sado, T., Okano, M., Li, E., and Sasaki, H. (2004). De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. *Development* 131, 975-982.

Santoro, F., Mayer, D., Klement, R.M., Warczok, K.E., Stukalov, A., Barlow, D.P., and Pauler, F.M. (2013). Imprinted Igf2r silencing depends on continuous Airn lncRNA expression and is not restricted to a developmental window. *Development* 140, 1184-1195.

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

Sheardown, S.A., Newall, A.E., Norris, D.P., Rastan, S., and Brockdorff, N. (1997). Reg-

ulatory elements in the minimal promoter region of the mouse Xist gene. *Gene* 203, 159-168.

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.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 484, 339-344.

Smolle, M., Workman, J.L., and Venkatesh, S. (2013). reSETting chromatin during transcription elongation. *Epigenetics* 8, 10-15.

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

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.

Urbach, A., Bar-Nur, O., Daley, G.Q., and Benvenisty, N. (2010). Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. *Cell Stem Cell* 6, 407-411.

Velasco, G., Hubé, F., Rollin, J., Neuillet, D., Phillippe, C., Bouzinba-Segard, H., Galvani, A., Viegas-Pequignot, E., and Francastel, C. (2010). Dnmt3b recruitment through E2F6 transcriptional repressor mediates germline gene silencing in murine somatic tissues. *Proc Natl Acad Sci U S A* 107, 9281-9286.

Vermeulen, M., Eberl, H., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K., Olsen, J., Hyman, A., Stunnenberg, H., *et al.* (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142, 967-980.



Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M., and Schubeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 39, 457-466.

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.

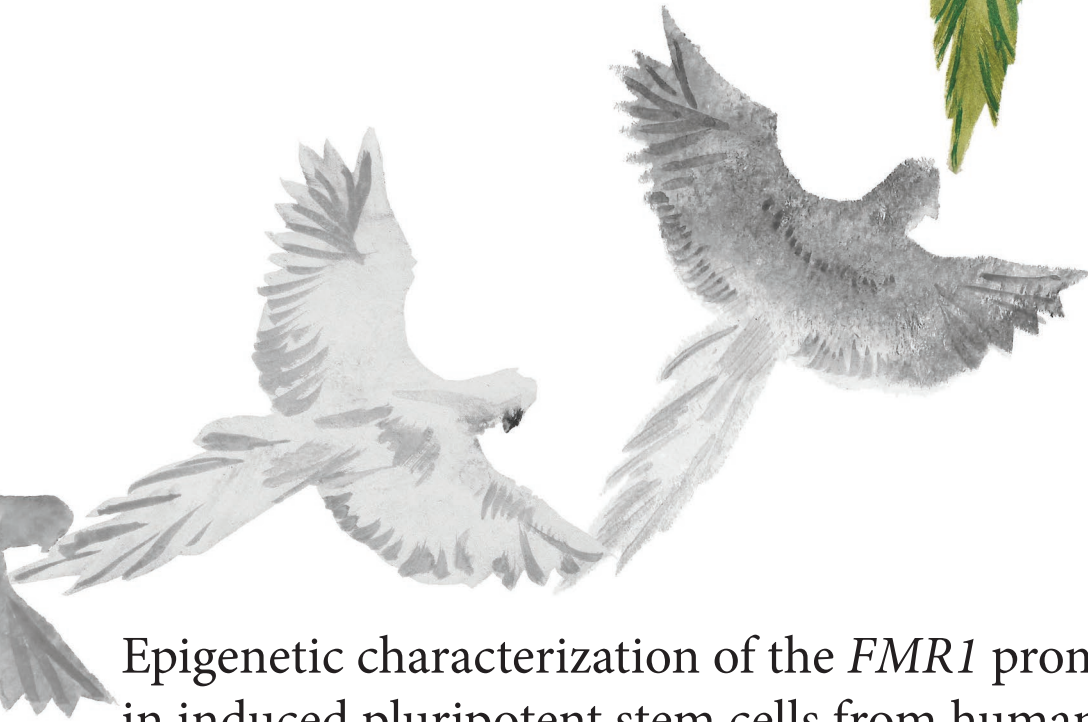
Zvetkova, I., Apedaile, A., Ramsahoye, B., Mermoud, J.E., Crompton, L.A., John, R., Feil, R., and Brockdorff, N. (2005). Global hypomethylation of the genome in XX embryonic stem cells. *Nat Genet* 37, 1274-1279.



# APPENDIX



# A



## Epigenetic characterization of the *FMR1* promoter in induced pluripotent stem cells from human fibroblasts carrying an unmethylated full mutation

*published in:*

de Esch, C.E.F., Ghazvini, M., Loos, F., Schelling-Kazaryan, N., Widagdo, W., Munshi, S.T., van der Wal, E., Douben, H., Gunhanlar, N., Kushner, S.A., Pijnappel, W.W.M.P., de Vrij, F.M.S., Geijssen, N., Gribnau, J., and Willemsen, R. (2014). *Stem Cell Reports* 3(4), 548-55.

# Epigenetic characterization of the *FMR1* promoter in induced pluripotent stem cells from human fibroblasts carrying an unmethylated full mutation

de Esch, C.E.F.<sup>1</sup>, Ghazvini, M.<sup>2,3</sup>, Loos, F.<sup>3</sup>, Schelling-Kazaryan, N.<sup>4</sup>, Widagdo, W.<sup>1</sup>, Munshi, S.T.<sup>5</sup>, van der Wal, E.<sup>6</sup>, Douben, H.<sup>1</sup>, Gunhanlar, N.<sup>5</sup>, Kushner, S.A.<sup>5</sup>, Pijnappel, W.W.M.P.<sup>6</sup>, de Vrij, F.M.S.<sup>5</sup>, Geijssen, N.<sup>4,7</sup>, Gribnau, J.<sup>3,8</sup>, and Willemsen, R.<sup>1,8</sup>

<sup>1</sup> Department of Clinical Genetics, Erasmus Medical Center, 3015 GE Rotterdam, The Netherlands

<sup>2</sup> iPS Cell Facility, Erasmus Medical Center, 3015 GE Rotterdam, The Netherlands

<sup>3</sup> Department of Reproduction and Development, Erasmus Medical Center, 3015 GE Rotterdam, The Netherlands

<sup>4</sup> KNAW Hubrecht Institute and UMC Utrecht, 3584 CT Utrecht, The Netherlands

<sup>5</sup> Department of Psychiatry, Erasmus Medical Center, 3015 GE Rotterdam, The Netherlands

<sup>6</sup> Molecular Stem Cell Biology, Department of Clinical Genetics and Department of Pediatrics, Division of Metabolic Diseases and Genetics, Center for Lysosomal and Metabolic Diseases, Erasmus Medical Center, 3015 GE Rotterdam, The Netherlands

<sup>7</sup> Department Companion Animals, Utrecht University School for Veterinary Medicine, 3508 TD Utrecht, The Netherlands

<sup>8</sup> Co-senior author

\*correspondence: [r.willemsen@erasmusmc.nl](mailto:r.willemsen@erasmusmc.nl)

Stem Cell Reports 3(4), 548-55, October 14, 2014

## Summary

Silencing of the *FMR1* gene leads to fragile X syndrome, the most common cause of inherited intellectual disability. To study the epigenetic modifications of the *FMR1* gene during silencing in time, we used fibroblasts and induced pluripotent stem cells (iPSCs) of an unmethylated full mutation (uFM) individual with normal intelligence. The uFM fibroblast line carried an unmethylated *FMR1* promoter region and expressed normal to slightly increased *FMR1* mRNA levels. The *FMR1* expression in the uFM line corresponds with the increased H3 acetylation and H3K4 methylation in combination with a reduced H3K9 methylation. After reprogramming, the *FMR1* promoter region was methylated in all uFM iPSC clones. Two clones were analyzed further and showed a lack of *FMR1* expression while the presence of specific histone modifications also indicated a repressed *FMR1* promoter. In conclusion, these findings demonstrate that the standard reprogramming procedure leads to epigenetic silencing of the fully mutated *FMR1* gene.

## Introduction

The most common inherited form of intellectual disability, fragile X syndrome (FXS), is caused by the absence of the *FMR1* gene product, the fragile X mental retardation protein (FMRP). In the majority of FXS patients, the transcriptional silencing of the *FMR1* gene is initiated by an expansion of a naturally occurring CGG repeat in the 5' untranslated region (UTR) of the *FMR1* gene, to more than ~200 units (Verkerk et al. 1991; Pear-

son et al. 2005). This so called full mutation results in hypermethylation of the cytosines in the repeat region and the *FMR1* promoter region during early human embryonic development (Sutcliffe et al. 1992; Willemsen et al. 2002). This results in a lack of *FMR1* transcription and consequently an absence of FMRP. Along with hypermethylation, the *FMR1* promoter in FXS is characterized by additional epigenetic marks specific for transcriptionally repressed chromatin including reduced histone H3 and H4 acetylation, reduced histone H3K4 methylation and increased histone H3K9 methylation (Coffee et al. 1999; Coffee et al. 2002; Pietrobono et al. 2005; Tabolacci et al. 2005). However, the timing and molecular mechanisms involved in the CGG expansion, the concomitant DNA methylation and the additional epigenetic changes that occur during embryonic development are not yet fully understood. Insights into these processes may lead to a more complete understanding of the developmental processes underlying fragile X syndrome, which in turn could lead to new therapeutic strategies.

Since murine fragile X models cannot be used to investigate epigenetic *FMR1* inactivation as methylation of the full mutations does not occur, human FXS embryonic stem cells have been studied. These studies showed that FMRP is expressed during early embryonic development, but that epigenetic silencing of *FMR1* occurs upon differentiation (Eiges et al. 2007; Gerhardt et al. 2013). A further attempt to study the epigenetic changes over time made use of induced pluripotent stem cells (iPSCs) generated from human FXS fibroblasts. In contrast to human embryonic FX stem cells these pluripotent cells were shown to already carry a fully methylated *FMR1* promoter and additional heterochromatin marks, so the epigenetic silencing mechanisms in time could not be studied (Urbach et al. 2010;





Sheridan et al. 2011; Bar-Nur et al. 2012).

In 1991, a familial case was reported in which two brothers with normal intelligence were shown to have a full *FMR1* mutation without the concomitant hypermethylation of the CGG repeat and the promoter region (Smeets et al. 1995). In order to unravel the molecular mechanisms behind the epigenetic silencing in fragile X syndrome, we derived iPSC cells from these human fibroblasts, to analyze the epigenetic characteristics of the *FMR1* promoter after reprogramming and during differentiation. Here, we report the characterization of these iPSC cells and show, unexpectedly, that the *FMR1* promoter of the unmethylated full mutation cell line becomes methylated during reprogramming and stays methylated after differentiation into neural progenitor cells.

## Results

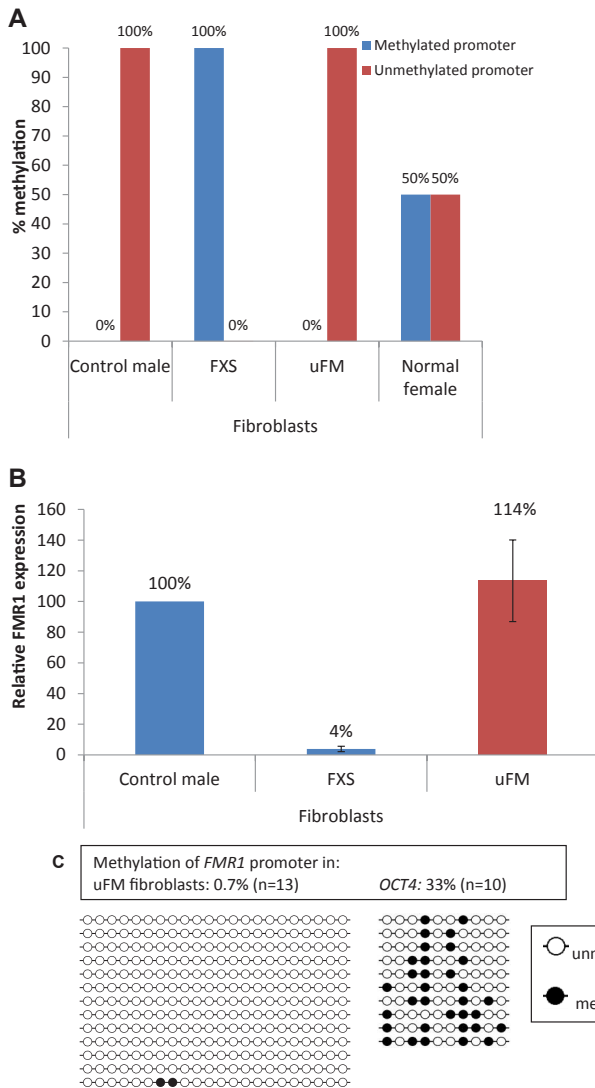
### Fibroblast characterization

Fibroblasts from a normal male carrying an unmethylated full mutation first described by Smeets *et al.* (1995) (uFM), fibroblasts from a clinically diagnosed male fragile X syndrome patient (14 years old, FXS), and an unrelated unaffected male control line (3 years old, Control) were analyzed for *FMR1* 5'UTR CGG repeat length, methylation status, *FMR1* expression and the histone marks associated with the *FMR1* promoter. As expected, the Control line showed a CGG repeat length within the normal range (< 55) while the uFM and the FXS line showed CGG repeat lengths in the full mutation range (approximately 233 and 380 repeats respectively) (Figure S1). Also as expected, the part of the *FMR1* promoter analyzed after bisulfite conversion

was not methylated in the Control and the uFM cell lines, while in the FXS cell line the *FMR1* promoter was methylated (Figure 1A, and Figure S2 for location of the primers). As the methylation status is predictive of *FMR1* expression, indeed the Control line showed normal expression levels, the uFM line showed normal to slightly increased *FMR1* expression while the FXS cell line did not express *FMR1* transcripts (Figure 1B). Additionally, bisulfite Sanger sequencing of a region of the *FMR1* promoter containing 22 CpGs was carried out which confirmed the absence of methylation of the *FMR1* promoter in the uFM fibroblast line (Figure 1C).

### Fibroblast reprogramming and iPSC characterization

The fibroblasts were reprogrammed to iPSC lines according to established protocols (Takahashi et al. 2007; Warlich et al. 2011). Firstly, four iPSC clones were generated which showed typical characteristics of pluripotent stem cells: morphology similar to that of embryonic stem cells (data not shown), expression of alkaline phosphatase (data not shown), silencing of the multi-cistronic lentiviral transgene (data not shown), reactivation of genes indicative of pluripotency (data not shown), immunoreactivity for OCT4, NANOG, Tra-1-60, Tra-1-81 and SSEA4 (Figure S3), propagation for a long time in culture (up to passage 30) and maintenance of a normal diploid karyotype (data not shown). All four cell lines generated embryoid bodies which, after differentiation *in vitro*, expressed markers of endoderm, mesoderm and ectoderm (Figure S3). These four lines were extensively characterized and the results are described below. Secondly, we generated eight additional iPSC clones from the uFM fibroblast line solely in order to confirm the methylation



**Figure 1. Methylation status and *FMR1* expression levels in the fibroblast cell lines**

(A) Methylation status of a region of the *FMR1* promoter in fibroblasts of the male control line, fragile X line (FXS) and the unmethylated full mutation line (uFM). Values were normalized to *CLK2* promoter activity first. The normalized exponential values were then presented as a percentage relative to the female fibroblast control line, for which the normalized exponential values were set to 50% for each primer set (n=2-3 separate measurements). (B) RT-qPCR data showing *FMR1* transcript levels in fibroblasts of the male control line, fragile X line (FXS) and the unmethylated full mutation line (uFM) normalized to *CLK2* expression. Values are means  $\pm$  SEM relative to appropriate male control line (n=2-3 separate measurements). (C) The percentage of methylated CpGs in the *FMR1* promoter and as a control the *OCT4* promoter, in 13 and 10 clones respectively after Sanger sequencing of bisulfite converted DNA of the uFM fibroblast line. Each line represents a clone, and each circle represents a CpG site which is methylated (closed circle) or unmethylated (open circle).

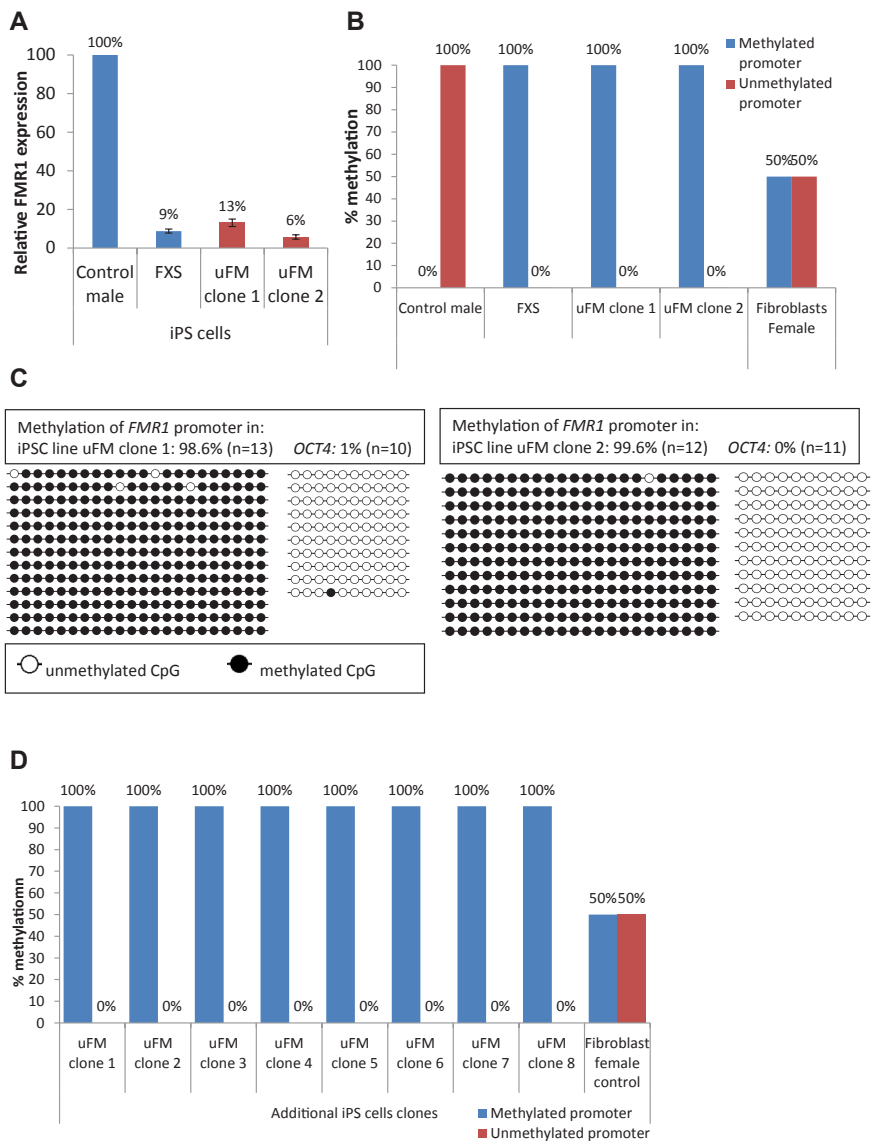
status of the *FMR1* promoter by quantitative PCR (Figure 2D). These additional iPS cell clones were generated from the uFM fibroblast line by the same methods as described, except this time we used naive human stem cell medium (WIS-NHSM) as defined by Gafni *et al*, 2013. This medium facilitates the derivation of naïve pluripotent iPS cells with properties highly similar to mouse naive ES cells.

### Reprogramming effects on CGG repeat length, *FMR1* expression and methylation

Analysis of the CGG repeat in the 5'UTR of the *FMR1* promoter indicated that the repeat length in the cell lines carrying a full mutation did not contract to levels below 200 CGGs during reprogramming (Figure S1). The iPSC clone of the Control cell line contained a CGG length under 55 repeats. Nonetheless, the

CGG repeat length contracted slightly in the FXS iPSC cell line after reprogramming, from 380 repeats to approximately 290 repeats. In contrast, the repeat was expanded in the two uFM iPSC cell clones to approximately 330 and 380 repeats (Figure S1). As expected, the iPSC clone of the Control cell line showed *FMRI* expression, in contrast to the FXS iPSC clone which did

not show *FMRI* expression. Unexpectedly, the two uFM iPSC clones did not express *FMRI* either (Figure 2A). Further analysis showed that the bisulfite converted *FMRI* promoter region was methylated in the FXS iPSC clone as well as in both uFM iPSC clones while the Control iPSC cell line did not show any methylation (Figure 2B). Bisulfite Sanger sequencing



confirmed the methylation status of the two uFM iPS clones (Figure 2C). The additional eight iPS clones generated from the uFM fibroblast line in WIS-NSHM medium also showed complete methylation of the bisulfite converted *FMR1* region (Figure 2D). Thus, the originally unmethylated extended CGG repeat found in the uFM fibroblasts became methylated at some point during the reprogramming process.

Chromatin immunoprecipitation (ChIP) experiments with the fibroblast lines showed that the *FMR1* promoter of the Control line carried active histone marks, H3 acetylation and H3K4 di-methylation with values similar to the positive control, namely the active gene *APRT*, and values much higher than the negative control *Crystalline*, which only serves as a positive control for repressed genes. The inactive mark H3K9 tri-methylation was not enriched in the Control fibroblasts (Figure 3 A, B, C). The uFM fibroblast line carried histone marks representative of an actively transcribed gene, namely H3 acetylation and H3K4 methylation at similar levels as the Control line. The inactive mark H3K9 methylation could not be detected in the uFM fibroblast line (Figure 3 A, B, C). The *FMR1* promoter of the FXS cell line only showed enrichment of the repressive mark

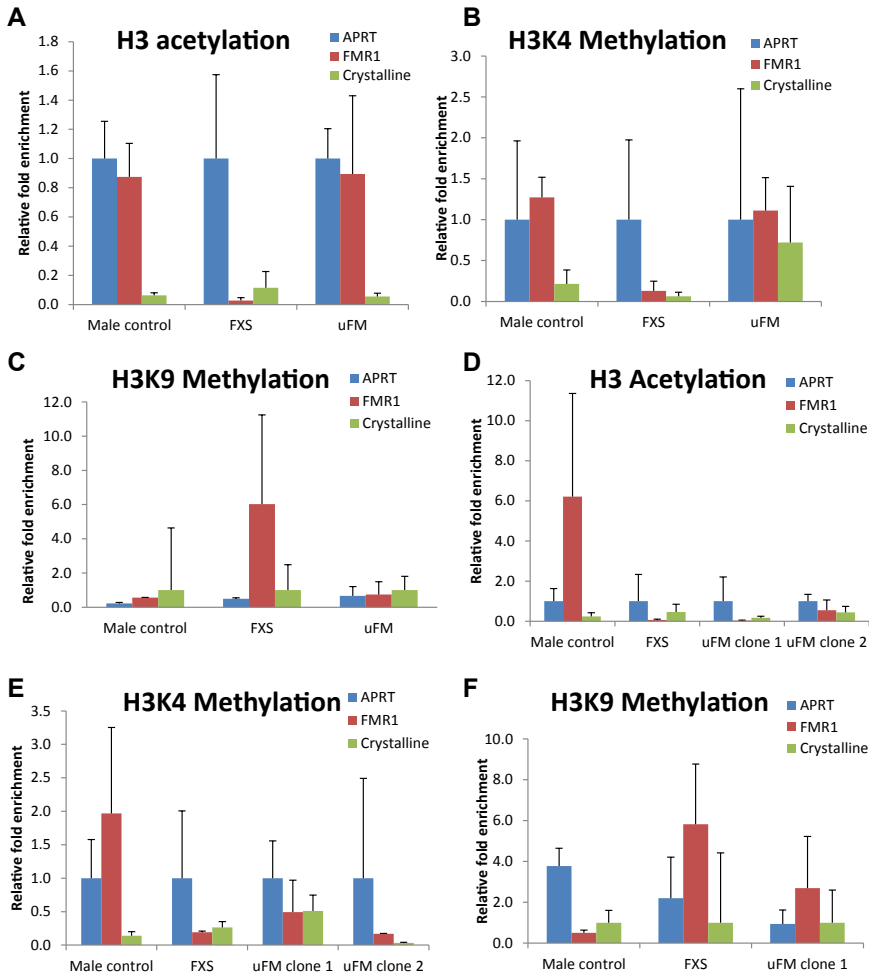
H3K9 methylation (Figure 3 A, B, C).

ChIP analysis of the *FMR1* promoter in iPS cells showed enrichment of the active marks H3 acetylation and H3K4 methylation in the Control iPS cell clone, to levels higher than the positive control *APRT*. The FXS iPS cells and clone 1 of the uFM iPS cells showed an increase of the repressive mark H3K9 methylation to values above the repressive control *Crystalline* while enrichment of the active marks could not be detected in FXS iPS cells and uFM iPS cell clones 1 and 2 (Figure 3 D, E, F). Next, we investigated the effects of differentiation into neural progenitor cells (NPCs) on *FMR1* expression and methylation (see Figure S4 for staining with marker *SOX2*). NPCs derived from the FXS and uFM iPS cells lacked *FMR1* expression and carried a methylated *FMR1* promoter. The NPCs derived from the Control iPSC clone showed clear *FMR1* expression and an unmethylated promoter region (Figure 4A and B). These findings indicate that the reprogramming process leads to methylation of the expanded *FMR1* CGG repeat sequence, which results in a stable shut down of *FMR1* gene expression.

#### ◀ Figure 2. Methylation status and *FMR1* expression levels in the induced pluripotent stem cells

(A) RT-qPCR data showing *FMR1* transcript levels in induced pluripotent stem cells (iPSCs) of the male control line, fragile X line (FXS) and the unmethylated full mutation clones (uFM clone 1 and clone 2) normalized to *CLK2* expression. Values are mean  $\pm$  SEM relative to appropriate male control line (n=2-3 separate measurements). (B) Methylation status of a region of the *FMR1* promoter in iPSCs of the male control line, fragile X line (FXS) and the unmethylated full mutation clones (uFM clone 1 and clone 2). Values were normalized to *CLK2* promoter activity first. The normalized exponential values were then presented as a percentage relative to the female fibroblast control line, for which the normalized exponential values were set to 50% for each primer set (n=2-3 separate measurements). (C) The percentage of methylated CpGs in the *FMR1* promoter and as a control the *OCT4* promoter, after Sanger sequencing of bisulfite converted DNA of the uFM iPSC clones. Each line represents a clone, and each circle represents a CpG site which is methylated (closed circle) or unmethylated (open circle). (D) Methylation status of a region of the *FMR1* promoter in additionally generated iPSC clones of the unmethylated full mutation fibroblast line in naive human stem cell medium. Values were normalized to *CLK2* promoter activity first. The normalized exponential values were then presented as a percentage relative to the female fibroblast control line, for which the normalized exponential values were set to 50% for each primer set (n=2 separate measurements).





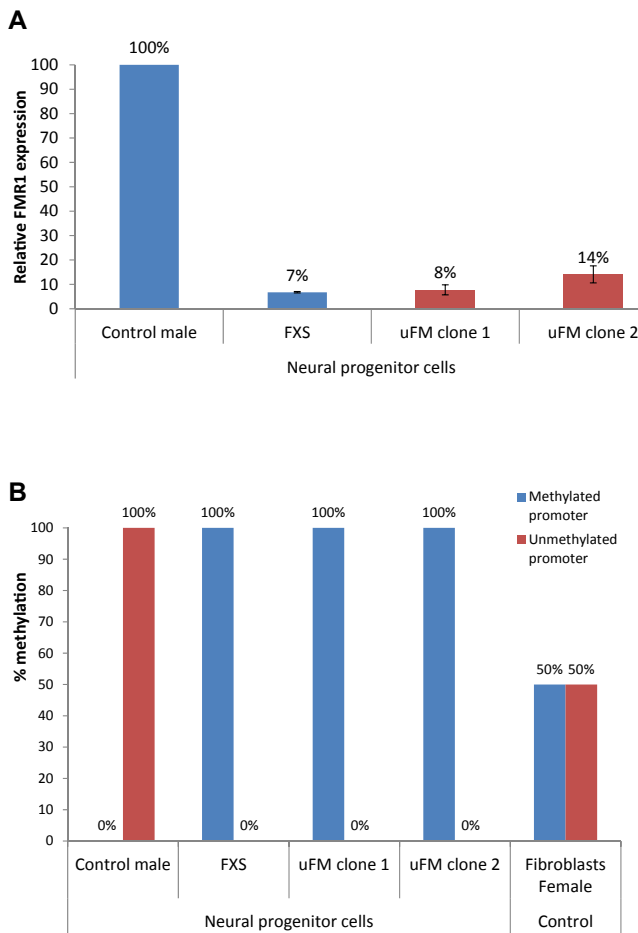
**Figure 3. Chromatin immunoprecipitation analysis of H3 acetylation, H3K4 methylation and H3K9 methylation in the *FMR1* promoter of fibroblasts and iPS cells**

Chromatin immunoprecipitation analysis of H3 acetylation, H3K4 methylation and H3K9 methylation in the *FMR1* promoter of fibroblasts (A, B and C) and iPS cells (D, E and F) respectively. Results were normalized to the appropriate positive control (APRT or Crystalline) and averaged from at least two different experiments.

## Discussion

We undertook this study in an attempt to unravel the epigenetic mechanisms involved in the silencing of the *FMR1* gene in fragile X syndrome by the use of a fibroblast line carrying an unmethylated full mutation. There have been several attempts to study epigenetic si-

lencing in fragile X syndrome. Eiges *et al.*(2007), have shown that FXS human embryonic stem cells (hESCs) still express FMRP at a level similar to that in unaffected hESCs, while the FMRP level decreases as the hESCs were differentiated. Based on these results it was expected that by reprogramming FXS fibroblasts into pluripotent stem cells, the hypermethylated state of the *FMR1* promoter region would be reversed.



**Figure 4. Methylation status and *FMR1* expression levels in neural progenitor cells**

**(A)** RT-qPCR data showing *FMR1* transcript levels in neural progenitor cells (NPCs) of the male control line, fragile X line (FXS) and the unmethylated full mutation clones (uFM clone 1 and clone 2) normalised to *CLK2* expression. Values are mean  $\pm$  SE relative to appropriate male control line (n=2 separate measurements). **(B)** Methylation status of a region of the *FMR1* promoter in NPCs of the male control line, fragile X line (FXS) and the unmethylated full mutation clones (uFM clone 1 and clone 2). Values were normalized to *CLK2* promoter activity first. The normalized exponential values were then presented as a percentage relative to the female fibroblast control line, for which the normalized exponential values were set to 50% for each primer set (n=2-3 separate measurements).

However, by now several research groups have shown that iPSCs derived from FXS patients show epigenetic marks characteristic for heterochromatin similar to the full mutation fibroblasts they originated from (Urbach et al. 2010; Sheridan et al. 2011; Bar-Nur et al. 2012). These observations could be explained by the fact that the FXS iPS cells may not have all the characteristics of early pluripotency, but that they represent a later stage of human development (Urbach et al. 2010; Sheridan et al. 2011; Bar-Nur et al. 2012; Gafni et al. 2013).

Another approach was used in studies with human fragile X lymphoblastic cells, here a fully mutated and hyper-

methylated *FMR1* gene was reactivated by treatment with 5-azadeoxycytidine, a hypomethylating agent. Although such treatment significantly reduced DNA methylation in some cells, it could not restore all remaining epigenetic marks to control levels (Chiurazzi et al. 1998; Chiurazzi et al. 1999; Coffee et al. 1999; Coffee et al. 2002). Drugs such as 4-phenylbutyrate, sodium butyrate or trichostatin A, which block the activity of histone deacetylases, did not restore *FMR1* expression to normal levels (Chiurazzi et al. 1999; Coffee et al. 1999; Coffee et al. 2002; Tabolacci et al. 2005). In addition, treatment with a compound that reduces the *in vitro* expression of the FRAXA fragile site, acetyl-L-carni-



tine, did not restore the *FMR1* expression either (Tabolacci et al. 2005). Recently, 5-azadeoxycytidine treatment was also tested on fragile X iPSC cells and it appeared to restore *FMR1* expression in both iPSC cells and differentiated neurons, which offers possibilities to use these cells as an epigenetic model (Bar-Nur et al. 2012).

The availability of a fibroblast cell line carrying an unmethylated full mutation (uFM) provided a new opportunity to study the epigenetic silencing mechanisms in time. We first characterized the uFM fibroblast cell line together with a normal male fibroblast control line and a FXS fibroblast cell line carrying a fully methylated *FMR1* promoter. Although increased *FMR1* mRNA levels (up to 5 times) were reported in lymphoblastoid cells of premutation carriers (55~200 unmethylated CGGs), our findings of normal to slightly increased *FMR1* mRNA levels in the uFM fibroblasts are similar to the findings of Pietrobono *et al.* (2005), who examined a lymphoblastic cell line from the same individual. The lack of DNA methylation ensures that the chromatin is less densely packed and more accessible for transcription, which explains the *FMR1* expression in this cell line. Our ChIP results differ from the original ChIP analysis of the uFM lymphoblastoid cell line (Tabolacci et al. 2005). We found a similar increase in H3K4 methylation; however we did not find decreased H3 acetylation levels or intermediate H3K9 levels in the uFM fibroblasts. These differences could be explained by the fact that we have analyzed a distinct cell type (fibroblasts versus lymphoblastoid cells), and by differences in the ChIP protocol (eg quantification methods and reference genes used). Since the uFM fibroblast line lacked methylation of the *FMR1* promoter site despite the high number of CGG repeats, we expected to find an unmethylated *FMR1* promoter and normal levels of

*FMR1* mRNA after reprogramming into iPSCs. Surprisingly, we found the promoter region of *FMR1* to be hypermethylated in all iPSC clones. Other epigenetic chromatin marks also indicated a repressed *FMR1* promoter similar to the marks observed in the fragile X iPSC cell line. After differentiation of these iPSC cells into neural progenitor cells, the *FMR1* promoter remained methylated and thus silenced.

There are three possible explanations for our findings. Firstly, it is possible that the reprogramming process resulted in iPSCs that were solely derived from methylated FM fibroblasts and not of the unmethylated cells. This assumes that methylated FM fibroblasts were present in our culture, which according to our bisulfite sequencing results seems highly unlikely. Second, there may be an unknown genetic factor present in this individual which was protective against DNA methylation during embryonic development but which was absent in his fibroblasts or which was altered or blocked during the reprogramming process. In our case, the brother of this individual was also carrier of an unmethylated full mutation. Being a carrier of an unmethylated full mutation is already a very rare phenomenon, but the fact that two children escaped methylation in one family clearly points towards the involvement of a maternal-paternal genetic component or environmental factors. Finally, the reprogramming process might activate genes that induce *de novo* methylation of the *FMR1* promoter. Although the *FMR1* gene in this individual escaped methylation during embryonic development, the full mutation in his fibroblasts might be recognized by epigenetic remodelers eg by histone and/or DNA methyltransferases (DNMTs) that are not recruited in embryonic development. This would also explain the unmethylated full mutation observed in human embryonic FXS stem cells because these cells never went

through this reprogramming process. A strategy to test this hypothesis would be for example to perform the reprogramming of the uFM fibroblasts as well as FXS fibroblast lines under conditions that inhibit the functioning of DNMT 3a and 3b.

In conclusion, standard reprogramming of somatic uFM fibroblasts into pluripotent stem cells by the use of 4 transcription factors, did not lead to de-methylation of the expanded CGG repeat, and even induced methylation of an unmethylated template. Very recently, Gafni *et al.* (2013) suggested that a more naïve ground state pluripotent stem cell in which epigenetic memory is completely erased could be obtained by a unique combination of cytokines and small molecule inhibitors (WIS-NHS medium). This study also demonstrated the reactivation of the *FMR1* gene in FXS iPS cells after the reprogramming of FXS fibroblast under naïve conditions. However, in contrast to these findings, the use of this WIS-NHS medium did not prevent the occurrence of the *de novo* methylation of the extended *FMR1* repeat in our uFM iPS clones. In conclusion, our results show that although this fibroblast line may offer a unique system to study the *de novo* methylation of an extended *FMR1* repeat during reprogramming, the mechanism behind the silencing of the *FMR1* gene in fragile X syndrome remains elusive.

## Experimental Procedures

### Cell culture

The rare fibroblast cell line established from a normal male carrying an unmethylated full mutation first described by Smeets *et al.* (1999) (uFM) was used. This

line has been subcloned, so that a homogeneous population of cells that carry a fully extended repeat was obtained. Fibroblasts from a clinically diagnosed male fragile X syndrome patient (14 years, FXS), and an unrelated unaffected male (3 years, Control) and female control fibroblast line (9 years) were all obtained from the cell repository of the department of Clinical Genetics, Erasmus MC, Rotterdam. The fibroblasts were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco-Invitrogen) containing 10% fetal calf serum and 1% penicillin/streptomycin.

### iPS cell generation

Reprogramming of human primary skin fibroblasts was performed as described previously (Warlich *et al.* 2011). Briefly, fibroblasts were infected with a single, multi-cistronic lentiviral vector encoding *OCT4*, *SOX2*, *KLF4*, and *MYC* and cultured on  $\gamma$ -irradiated mouse embryonic feeder (MEF) cells until iPSC colonies could be picked (Warlich *et al.* 2011). The first set of generated iPS cell lines, namely the male control line, the FXS line and the uFM iPS clone 1 and 2 were cultured in conventional ES cell culture medium containing DMEM/F12 (Gibco-Invitrogen) supplemented with 20% knock-out serum replacement (Gibco-Invitrogen), 2mM L-glutamine, 50 units of penicillin/streptomycin/glutamine, 0.1 mM MEM-non-essential aminoacids (PAA Laboratories GmbH), 0.1mM  $\beta$ -mercaptoethanol, and 10 ng/ml bFGF (Invitrogen) filtered through a 0.22 $\mu$ m filter (Corning). Human iPS lines growing on conventional medium were passaged weekly using collagenase IV (1 mg/ml, Invitrogen) on  $\gamma$ -irradiated MEFs. The second round of reprogramming of the uFM fibroblast line was done in naïve ES medium (WIS-NHSM medium) according

to Gafni *et al.*, 2013. Wis-NHSM medium containing 475 ml knockout DMEM (Invitrogen), 20% knockout serum (invitrogen), human insulin (Sigma, 12.5 µg/ml), 10 µg recombinant human Lif (Peprotech), 8 ng/ml recombinant bFGF (Peprotech) and 1 ng/ml recombinant TGF-β1 (Peprotech), 1 mM glutamine (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma), Penicillin-Streptomycin (Invitrogen) and small molecule inhibitors: PD0325901 (1 µM, ERK1/2i, Axon Medchem); CHIR99021 (3 µM, GSKβi, Axon Medchem); SP600125 (10 µM, JNKi, TORCIS) and SB203580 (10 µM, p38i, Axon Medchem) Y-27632 (5 µM, Axon Medchem) and protein kinase C inhibitor G06983 (5 µM, TOCRIS). Naïve human iPS clones were grown on γ-irradiated MEFs on gelatin coated plates and passaged by single-cell trypsinization (0.05% EDTA) every 4 days. These cells showed the main characteristics of induced pluripotent stem cells including a morphology similar to that of embryonic stem cells, silencing of retroviral transgenes and reactivation of pluripotency genes (data not shown). These cells were only used to affirm the methylation status of the *FMR1* promoter after reprogramming by methylation specific quantitative PCR.

## Differentiation of the iPS cells

### *In-vitro differentiation of embryonic bodies*

To form embryonic bodies (EBs), iPSC colonies from 2 wells per line were broken up by collagenase IV treatment and transferred to ultra-low attachment 6-wells plates (Corning). Floating EBs were then cultured in iPSC medium without bFGF for a minimum of 6 days with supplemented SB431542 for ectoderm conditions only. The EBs designated for

endoderm were then transferred to gelatin coated 12-wells plates containing the following medium: RPMI 1640 (Gibco-Invitrogen), supplemented with 20% FBS, 1: 100 dilution of penicillin/streptomycin/glutamine and alpha-thioglycerol (0.4mM). Mesoderm differentiation from the EBs was induced in gelatin coated 12-wells plates with DMEM low glucose medium supplemented with 15% FBS, 1: 100 dilution of penicillin/streptomycin/glutamine and 1:100 dilution of MEM-non-essential amino acids. The formation of ectoderm was induced in matrigel coated plates with the following medium: neurobasal medium (Gibco) and DMEM/F12 (v/v 50/50) supplemented with 1: 100 dilution of penicillin/streptomycin/glutamine and 1:100 dilution of MEM-non-essential amino acids, 0.02% BSA (Gibco), 1:200 N2 (Gibco) and 1:100 B27 (Gibco). After two weeks in culture the cells were fixed with formalin and immunostainings were performed.

### *Neural differentiation*

Human iPS cells were differentiated according to Brennand *et al.* (2011), with modifications. Briefly, iPS colonies were dissociated from MEFs with collagenase (100 U/ml) and transferred to non-adherent plates in hES cell medium on a shaker in an incubator at 37°C/5% CO<sub>2</sub>. After two days, embryonic bodies (EBs) were placed in neural induction medium (DMEM/F12, 1x N2, 2 µg/ml heparin, penicillin/streptomycin) and cultured for another four days in suspension. EBs were gently dissociated and plated onto laminin coated dishes in NPC medium (DMEM/F12, 1x N2, 1x B27-RA, 1 µg/ml laminin and 20 ng/ml FGF2, penicillin/streptomycin). After one week, NPCs were dissociated with collagenase (100 U/ml), replated, and used for staining and methylation analysis after 3-5 passages. All cell culture reagents were obtained from Invitrogen.

### Karyotype analysis

For karyotype analysis, cells in a well of a 6-wells plate were treated with colcemid (100 ng/ml) for 1 hour. Then cells were harvested with trypsin, treated with hypotonic solution and fixed. Metaphases were spread onto glass slides and stained with DAPI (Dako). Chromosomes were classified according to the International System for Human Cytogenetic Nomenclature. At least 10 metaphases were analyzed per cell line.

### Alkaline phosphatase staining and immunocytochemistry

Staining for alkaline phosphatase was carried out using the Alkaline Phosphatase kit (Sigma-Aldrich) according to the manufacturer's instructions. For immunocytochemistry, iPS cells or differentiated iPS cells were washed with PBS once, fixed with 4% formalin solution for 5 min and washed again with PBS. Cells were then incubated with 50mM glycine for 5 min, washed with PBS and permeabilized with 0.5% Triton X-100 for 5 min (only for *OCT4* and *Nanog*). After blocking for 45 min at room temperature with 0.1% PBS-Tween containing 2% fetal bovine serum (Invitrogen), primary antibody staining was performed for 1 hour in room temperature with antibodies diluted in blocking solution. Cells were then washed and incubated with the appropriate secondary Cy3 or Alexa Fluor A555 antibody (1:200, Jackson ImmunoResearch Laboratories or Invitrogen) for 45 min. Afterwards, cells were washed with twice 0.1% PBS-Tween, with a nuclear staining step in between (Hoechst or DAPI). Cells were covered with Mowiol and a glass slide. The antibodies used for pluripotency stainings or neural marker stainings were goat anti human *OCT3/4* (1:100, Santa Cruz Biotech-

nology), goat anti human *NANOG* (1:50, R&D Systems), mouse anti human *Tra-1-60*, *Tra-1-80*, and *SSEA4* (1:100 Santa Cruz Biotechnology) and rabbit anti *Sox2* (1:1000 Millipore). Antibodies used for *in vitro* differentiation stainings were anti human smooth muscle actin (1:50, DAKO), rabbit anti human alpha-fetoprotein (1:200, Dako), mouse anti human  $\beta$ -tubulin III (*Tuj1*) (1:200, Sigma-Aldrich).

### RNA isolation and *FMR1* expression analysis

RNA was isolated using the RNAeasy kit (Qiagen), and 1  $\mu$ g of RNA was reverse transcribed using iScript (BioRad). Real-time PCR was carried out in triplicate using Kappa mix and a 7300 Real-time PCR system (Applied Biosystems). A forward primer located in exon 4 was used in combination with a reverse primer located in exon 5 to measure *FMR1* expression: 5'-GGTGGTTAGCTAAAGTGAGGA-3' and 5'-GTGGCAGGTTTGTGGGATTA-3'. *CLK2* was used as reference gene with forward primer 5'-CCTACAACCTAGAGAAGAAGCGAG-3' and reverse primer 5'-CACTGCCAAAGTCTACCACC-3' (de Brouwer et al. 2006). *FMR1* expression was normalized to *CLK2* expression and data was presented as an average value from 2 to 3 independent measurements. The expression values of the male control and the female control cell lines were combined and their average relative fold enrichment was set to 100%.

### DNA isolation, CGG repeat length and methylation analysis

In order to isolate total genomic DNA, cell were treated with lysis buffer containing 100mM NaCl, 10mM Tris,

15mM EDTA, 0.5% SDS, and 5% Proteinase K. After overnight incubation at 55°C, DNA was extracted and precipitated using a standard protocol containing saturated salt solution and ethanol.

CGG repeat size was determined in a PCR reaction using the primers 5'-CGGAGGCGCCGCTGCCAGG-3' and 5'-TGCGGGCGCTCGAGGCCAG-3' with the Expand high fidelity PCR kit (Roche) supplemented with 2.5 M betaine. PCR was performed with 35 cycles of 35 seconds denaturing at 98°C, 35 seconds of annealing at 55°C, and 5 minutes elongation at 72°C. PCR products were analyzed with standard agarose gel electrophoresis.

Genomic DNA was modified by bisulfite treatment according to the EpiTect Bisulfite Kit. The diluted converted DNA was then measured using quantitative PCR with two different primer sets designed specifically for a region of *FMRI* promoter (see supplemental figure 1 for the locations). One primer set contained the methylated DNA sequence and the other contained the unmethylated DNA sequence of a region of the *FMRI* promoter after bisulfite conversion. The primers for the methylated sequence are F 5'-GGTCGAAA-GATAGACGCGC-3', R 5'-AAACAA TGCGACCTATCACCG-3'; and for the unmethylated sequence are F 5'-TGTTG-TTTTGTGTTTGTGTTAGA-3', R 5'-AACATAATTTCAATATTACACCC-3' and for the promoter of the unmethylated bisulfite converted reference gene *CLK2*: F 5'-CGGTTGATTTTGGGTGAAGT-3' and R 5'-TCCCGACTAAAATCCCA-CAA-3'. All reactions were carried out in triplicate using SYBR Green ROX mix and a 7300 Real-time PCR system. Experiments were only analyzed when the Ct values of the female control sample were under 30 for both primer sets, as an indication for an efficient bisulfite conversion and DNA recovery. For each

sample, the values for the methylated and the unmethylated sequences were normalized to *CLK2* promoter activity first to obtain delta Cts. The normalized exponential values from the measurements of both primer sets were then set to 50% for the female control cell line. These values represent the random X inactivation in female control cells. The normalized exponential data of the remaining samples was then presented as a percentage relative to the female control data. Average ratios from at least two independent measurements were used for each sample.

### Bisulfite sequencing

Genomic DNA (1000ng) was modified by bisulfite treatment according to the EpiTect Bisulfite Kit. Then a region of the *FMRI* promoter containing 22 CpGs was amplified using PlatinumTaq (Invitrogen) and the primers F1 5'-GAGTGTATTTTGTAGAAATGGG-3' and R1 5'-TCTCTCTTCAAATAACCTAAAAAC-3' (see supplemental figure 1 for location of primers), while the *OCT4* promoter containing 10 CpG sites was amplified using the forward primer 5'-GAGGGAGAGAGGGGTTGAGTAG-3' and the reverse primer 5'-CCTCCAAAAAACCTTAAAAACTTAAC-3' (based on Al-khtib et al, 2012). PCR products were cloned into pGEM-T Easy (Promega) and single clones were sequenced by Sanger sequencing.

### Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed according to the Upstate ChIP protocol with some small modifications. In short, approximately 2.5 million cells were crosslinked with



1% formaldehyde for 5 minutes at room temperature. After quenching the reaction with 125 mM glycine, cells were subsequently suspended in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris pH 8.1) containing proteinase inhibitor (Roche, Complete). Chromatin was then sonicated using the Bioruptor (Diagonide) to create 200bp-1000bp DNA fragments. All chromatin was pre-cleared by treatment with salmon sperm agarose beads (Millipore) for 0.5 hour at room temperature. Immunoprecipitation was performed overnight using 7.5 µg anti-acetylated histone H3 (Millipore), anti-dimethyl histone H3K4 (Millipore), anti-trimethyl histone H3K9 (Millipore), or anti-IgG antibody (Millipore) in dilution buffer. Next, crosslinking was reversed by incubation with 0.2M NaCl at 65°C and DNA was purified using a PCR clean-up kit (Mobio). Finally, eluted DNA fragments were used for quantitative PCR analysis using primers for the *FMR1* promoter region F 5'-AACTGGGATAACCGGATGCAT-3' and R 5'-GGCCAGAACGCCCCATTTC-3' (see supplemental figure 1 for location) as well as appropriate positive and negative controls namely *APRT* F 5'-GCCTTGACTCGCACTTTT-3', and R 5'-TAGGCGCCATCGATTTTA-3' and *Crystalline* F 5'-CCGTGGTACCAAAGCTGA-3', and R 5'-AGCCGGCTGGGGTAGAA-3'. The Ct values of both histone modifications were first normalized for the non-specific IgG antibody treatment and then for the amount of input DNA. Data was then presented in relative fold enrichment after further normalization to the *APRT* gene for H3 acetylation and H3K4 methylation and *Crystalline* for H3K9 methylation. Data from at least two separate experiments were averaged and both reference genes were previously used by Urbach *et al.* (2010) and Bar-Nur *et al.* (2012).

## Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at <http://dx.doi.org/10.1016/j.stemcr.2014.07.013>.

## Acknowledgements

The authors thank Dr. A. Schambach for kindly sharing the iPSC reprogramming vector. This work was supported by the Netherlands Organisation for Health Research and Development (RW; ZonMw; 912-07-022), E-Rare program entitled "Cure FXS" (RW; no. EU/FISPS09102673), the Prinses Beatrix Spierfonds (WWMP and EvdW, project number: W.OR13-21), by a fellowship of the Hersenstichting Nederland (FMS-dV) and the NeuroBasic-PharmaPhenomics consortium (SAK) and by NWO VICI and ERC grants (JG). All authors declare no financial conflict of interest.

## References

- Bar-Nur, O., I. Caspi, et al. (2012). "Molecular analysis of *FMR1* reactivation in fragile-X induced pluripotent stem cells and their neuronal derivatives." *J Mol Cell Biol* 4(3): 180-183.
- Chiurazzi, P., M. G. Pomponi, et al. (1999). "Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the *FMR1* gene." *Hum Mol Genet* 8(12): 2317-2323.
- Chiurazzi, P., M. G. Pomponi, et al. (1998). "In vitro reactivation of the *FMR1* gene involved in fragile X syndrome." *Hum Mol Genet* 7(1): 109-113.



Coffee, B., F. Zhang, et al. (2002). "Histone modifications depict an aberrantly heterochromatinized FMR1 gene in fragile x syndrome." *Am J Hum Genet* 71(4): 923-932.

Coffee, B., F. Zhang, et al. (1999). "Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells." *Nat Genet* 22(1): 98-101.

de Brouwer, A. P., H. van Bokhoven, et al. (2006). "Comparison of 12 reference genes for normalization of gene expression levels in Epstein-Barr virus-transformed lymphoblastoid cell lines and fibroblasts." *Mol Diagn Ther* 10(3): 197-204.

Eiges, R., A. Urbach, et al. (2007). "Developmental Study of Fragile X Syndrome Using Human Embryonic Stem Cells Derived from Preimplantation Genetically Diagnosed Embryos." *Cell Stem Cell* 1(5): 568-577.

Gafni, O., L. Weinberger, et al. (2013). "Derivation of novel human ground state naive pluripotent stem cells." *Nature* 504, 282-286.

Gerhardt, J., M. J. Tomishima, et al. (2013). "The DNA Replication Program Is Altered at the FMR1 Locus in Fragile X Embryonic Stem Cells." *Mol Cell* 53, 19-31.

Pearson, C. E., K. Nichol Edamura, et al. (2005). "Repeat instability: mechanisms of dynamic mutations." *Nat Rev Genet* 6(10): 729-742.

Petrobono, R., E. Tabolacci, et al. (2005). "Molecular dissection of the events leading to inactivation of the FMR1 gene." *Hum Mol Genet* 14(2): 267-277.

Sheridan, S. D., K. M. Theriault, et al. (2011). "Epigenetic characterization of the FMR1 gene and aberrant neurodevelopment in human induced pluripotent stem cell models of fragile x syndrome." *PLoS ONE* 6(10): e26203.

Smeets, H., A. Smits, et al. (1995). "Normal phenotype in two brothers with a full FMR1 mutation." *Hum Mol Genet* 4(11): 2103-2108.

Sutcliffe, J. S., D. L. Nelson, et al. (1992). "DNA methylation represses FMR-1 transcription in fragile X syndrome." *Hum Mol Genet* 1(6): 397-400.

Tabolacci, E., R. Petrobono, et al. (2005). "Differential epigenetic modifications in the FMR1 gene of the fragile X syn-

drome after reactivating pharmacological treatments." *Eur J Hum Genet* 13(5): 641-648.

Takahashi, K., K. Okita, et al. (2007). "Induction of pluripotent stem cells from fibroblast cultures." *Nat Protoc* 2(12): 3081-3089.

Urbach, A., O. Bar-Nur, et al. (2010). "Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells." *Cell Stem Cell* 6(5): 407-411.

Verkerk, A. J., M. Pieretti, et al. (1991). "Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome." *Cell* 65(5): 905-914.

Warlich, E., J. Kuehle, et al. (2011). "Lentiviral vector design and imaging approaches to visualize the early stages of cellular reprogramming." *Mol Ther* 19(4): 782-789.

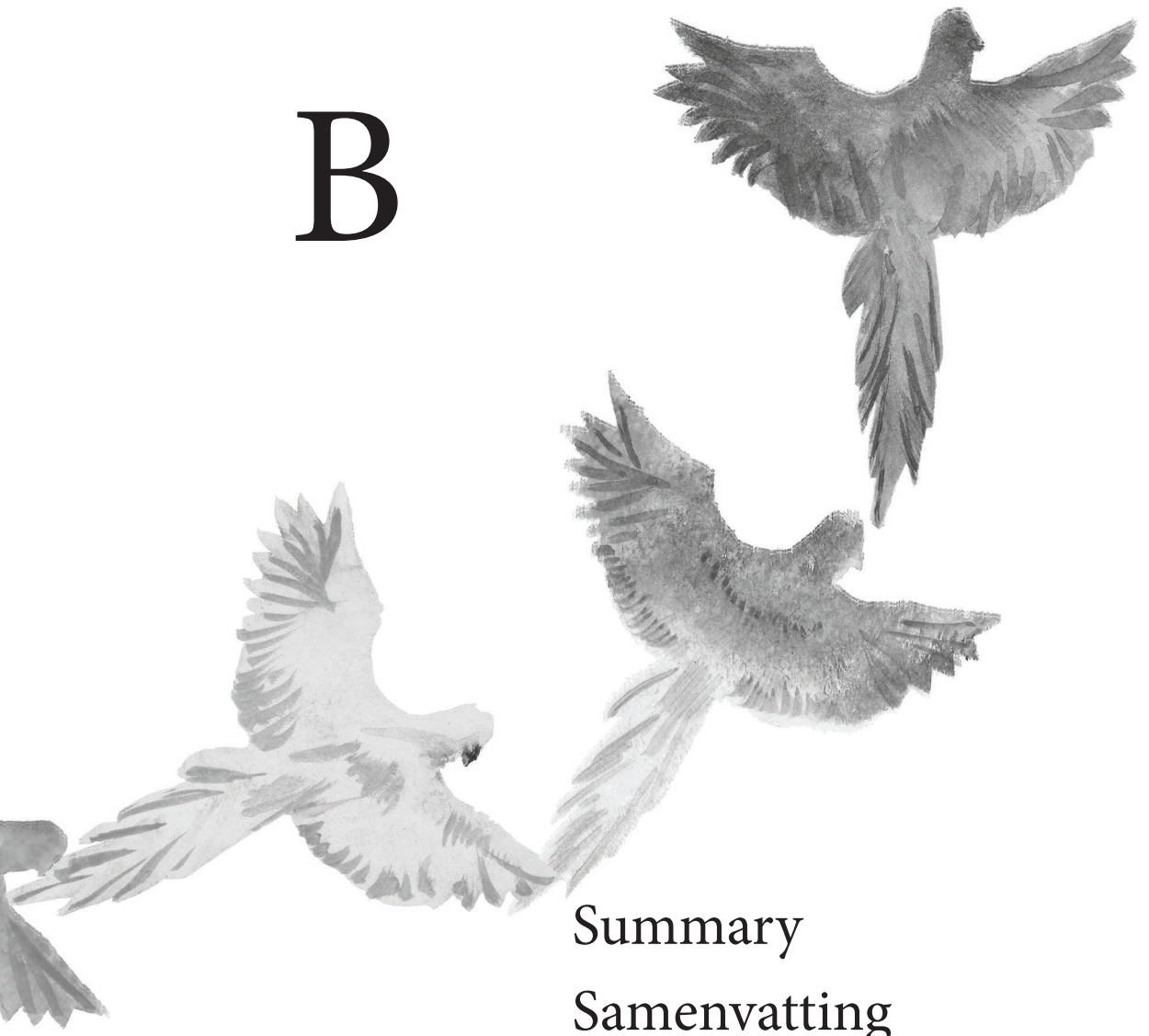
Willemsen, R., C. J. Bontekoe, et al. (2002). "Timing of the absence of FMR1 expression in full mutation chorionic villi." *Hum Genet* 110(6): 601-605.



# APPENDIX



# B



Summary

Samenvatting

Zusammenfassung

Curriculum vitae

List of Publications

PhD Portfolio

Acknowledgements



## Summary

Gene regulation is the process that defines to what extent and in which spatiotemporal context genes in a cell are turned on or off. If different sets of genes are active in two different cells, the molecular identity of these cells will differ accordingly and render the cells capable of executing diverging tasks. Thus, gene regulation is fundamental for establishing the many different identities of different cell types or tissues in multicellular organisms. This relationship is emphasized by a vast number of cases in which misregulation of gene expression causes diseases, most prominently cancer.

Gene regulation is orchestrated by several layers of transcriptional control. These include: i) proteins that directly activate or repress genes, so-called transcription factors; ii) noncoding RNAs, RNAs that influence gene expression but are not translated into proteins; iii) antisense transcription, referring to transcription that originates from both DNA strands and thus runs in opposite directions and possibly overlaps; iv) epigenetic processes, i.e. properties of the packaged DNA template that do not affect the base composition *per se* and include DNA methylation, histone modifications and the organization of DNA in the 3D space of the nucleus. In combination, these layers are responsible for the faithful execution of regulatory and developmental programs.

Another phenomenon that clearly illustrates the importance of proper gene regulation is a genetic regulatory mechanism termed dosage compensation. Dosage compensation has evolved in species in which males and females have different sex chromosome compositions. In mammals, for example, females carry two X chromo-

somes whereas males are XY, thus resulting in a theoretical two-fold excess of X-linked gene products in females. To account for this difference, a dosage compensation mechanism called X chromosome inactivation (XCI) silences almost all genes on one of the two X chromosomes in pre-implantation embryos of female mammals.

XCI is a stepwise process in which the number of X chromosomes is first “counted” to ensure that only in XX females one X chromosome is inactivated. This “counting” is effected by an X-linked activator of XCI, RNF12, which activates XCI only when present in higher concentrations, as is the case in XX females. The inactivation of one X chromosome in female cells also leads to silencing of one copy of *Rnf12*, thereby lowering the RNF12 concentration in these cells and preventing XCI of the second X chromosome. In XY male cells the RNF12 concentration is not high enough to initiate XCI and the single X chromosome remains active. RNF12 activates XCI by mediating the degradation of REX1, a transcription factor that represses the *Xist* gene. Thus, the noncoding RNA *Xist* is up-regulated. In a process termed “choice” *Xist* expression is subsequently maintained on only one X chromosome leading to *Xist* spreading along and coating of that entire X chromosome. The noncoding RNA *Xist* then recruits protein complexes that modify the chromatin structure of the X chromosome resulting in silencing of X-linked genes. At a late stage during XCI, DNA methylation adds an additional layer of stability to the repressed state of X-linked genes.

As XCI recapitulates all aspects of gene regulation in a developmental context, it has emerged as a paradigm for gene regulation as a whole. Conveniently, mouse embryonic stem cells (ES cells) and ES cell differentiation can be used as a model to study XCI. This thesis sheds





further light on the many layers of gene regulation that act on initiation and maintenance of XCI. In particular, the dynamics of regulation of genes involved in the initiation of XCI, and the role of the developmental context XCI takes place in are addressed by using ES cells as a model.

A general introduction in **chapter 1** gives an overview of the state-of-the-art of science in the fields of gene regulation, noncoding RNAs and XCI. **Chapter 2** describes a series of experiments aiming to shed further light on gene networks controlling XCI and on the role of RNF12, the only known activator of XCI to date. Our results using deletions of the *Rnf12* gene in different contexts indicate that RNF12 acts at two levels to regulate XCI. The two-fold higher concentration of RNF12 present in XX females as compared to XY males is essential for initiation of XCI. In addition, one active copy of *Rnf12* is necessary to maintain high levels of *Xist* expression and execute inactivation of the X chromosome properly. A second set of experiments deleting the region around *Xist* moreover suggests that many *Xist*-regulatory genes in close proximity to *Xist* function mostly by co-activation in *cis*. That is, these genes activate *Xist* only if they are located on the same chromosome and in close proximity to *Xist*, possibly by attracting other factors important for gene activation in a cumulative fashion. Reintroduction of these genes as transgenes into cells that lack the endogenous genes does thus not lead to *Xist* activation. In contrast, *Rnf12* transgenes activate *Xist* regardless of their position in the genome, confirming RNF12 as the only known *trans*-activator of XCI to date. We also show that “pairing” of X chromosomes, which has been implicated in the initiation of XCI, is not an absolute requirement for XCI to occur.

*Xist*, the gene responsible for effecting XCI, is negatively regulated by a sec-

ond noncoding RNA gene, *Tsix*, which is transcribed in antisense direction to and fully overlaps with *Xist*. Because of their overlapping nature it is difficult to distinguish between direct and indirect effects on their regulation. Therefore, in **chapter 3** we devise a strategy to study their early genetic regulation independently from influences of the antisense partner. To this end *Xist* was replaced by a green (EGFP) and *Tsix* by a red (mCherry) fluorescent reporter. As a result the fluorescent reporters are controlled by the endogenous promoters of *Xist* and *Tsix*, meaning that cells accumulate EGFP when the *Xist* promoter and mCherry when the *Tsix* promoter is activated. Experiments using these cell lines show that both genes display antagonistic roles, *Tsix* is repressing *Xist* and vice versa. The fluorescent reporter proteins in these cell lines allow us to easily score the activity levels of *Xist* and *Tsix* in living cells using fluorescence based techniques like microscopy or flow cytometry. Therefore we assessed the dynamics of *Xist* and *Tsix* regulation during ES cell differentiation by continuous live cell imaging. It appears that *Xist* and *Tsix* regulation are developmentally concerted but not directly dependent on each other. Expression states are also more stable than previously anticipated, meaning that once a “decision” is taken for either up- or down-regulation, this “decision” is maintained over time. For example, the mother cell and both daughters of a cell division display very similar levels of *Xist*/*Tsix* activation. The fluorescent reporters also enable careful dissection of the impact that XCI activators and inhibitors like RNF12 and REX1 have on *Xist* and *Tsix* regulation. In addition, experiments using female cells that had lost one of the two X chromosomes provide evidence that not only *Xist* but also *Tsix* is regulated by the number of X chromosomes present in a cell. Finally, under specific experimental conditions two semi-stable populations



are observed in the reporter cell lines, one population displaying lower the other higher Tsix/ mCherry levels. Careful analysis of these populations shows that they possibly correspond to two different 3D higher order conformations of the DNA surrounding *Xist* and *Tsix*. As the different populations also predict the outcome of the XCI process, that is Tsix/ mCherry low cells up-regulate *Xist* whereas Tsix/ mCherry high cells do not, they beautifully link the 3D conformation of DNA with functional consequences.

During the earliest stages of development the embryo undergoes a series of marked changes. These changes are predominantly driven by extensive modulation of gene regulatory networks and concomitant remodeling of chromatin, the packaged DNA template. It is also in this period that XCI is initiated and established. The second part of the thesis therefore focuses on the interplay of gene regulation and chromatin structure in this early phase of embryonic development. In **appendix A** a technique is used to reprogram skin fibroblast to so-called induced pluripotent stem cells (iPS cells). iPS cells resemble ES cells, for example they can differentiate into different cell types. Thus, by reprogramming fibroblasts into iPS cells these cells are forced back through differentiation in a reverse order until they become pluripotent again. The fibroblast used in **appendix A** are derived from a carrier of an *FMR1* mutation. Normally, carriers of this mutation acquire DNA methylation at the *FMR1* promoter during early development, resulting in silencing of this gene which causes the disease Fragile X Syndrome. However, in this particular case the carrier's *FMR1* promoter remained unmethylated and the gene active. Surprisingly, upon reprogramming of the fibroblasts the *FMR1* promoter became methylated. That implies that the reprogrammed cells went through a specific

developmental stage in which the mutation was “properly” recognized, leading to DNA methylation and gene silencing. Why the *FMR1* gene was not methylated and silenced in the first place, that is during early development of the mutation carrier, remains a topic for future research. Also in **chapter 4** a special regulatory phase during early embryonic development is analyzed. As many genes, e.g. *Xist* and *Tsix*, are arranged as sense-antisense gene pairs, a reporter system was created that mimicked this situation. A sequence coding for the green fluorescent protein EGFP was inserted in antisense orientation to an inducible promoter, which is a promoter controlled by the addition or removal of the drug doxycycline. This construct was introduced into ES cells such that antisense transcription through the EGFP coding sequence/promoter can be turned on and off by addition or removal of doxycycline. In undifferentiated ES cells induction of antisense transcription through the EGFP silences the EGFP reversibly. That is, as soon as doxycycline is removed from the medium and antisense transcription stops, the EGFP is re-expressed. This effect is not due to direct transcriptional interference, e.g. by collision of transcription machineries arriving from opposite directions. Rather, the effect appears to be mediated by chromatin modifications that are set down by the transcription machinery in the EGFP promoter. In contrast to the reversible silencing observed in undifferentiated ES cells, antisense transcription through the EGFP during differentiation leads to stable silencing of the EGFP. In this situation additional hallmarks of gene silencing like DNA methylation are detected. These results indicate yet again that cells go through a special phase during early development in which specific chromatin features are recognized and have a lasting effect on chromatin structure and gene regulation.



Finally, in the General Discussion in **chapter 5** the results of this thesis are discussed in the light of current developments in the fields of X chromosome inactivation and gene regulation.



## Samenvatting

Genregulatie is het proces dat bepaalt wanneer en in welke mate genen in een cel worden in- of uitgeschakeld. Als verschillende sets van genen in twee verschillende cellen actief zijn, verschilt ook de moleculaire identiteit van deze cellen en kunnen ze uiteenlopende taken uitvoeren. Genregulatie is dus fundamenteel voor het bepalen van de vele verschillende identiteiten van verschillende celtypen en weefsels in meercellige organismen. Dit wordt benadrukt door een groot aantal voorbeelden waarin misregulatie van genexpressie ziekte veroorzaakt zoals bijvoorbeeld kanker.

Genregulatie komt tot stand via diverse mechanismen van transcriptionele controle. Dit zijn: i) eiwitten die direct genen activeren of onderdrukken, zogenaamde transcription factors; ii) non-coding RNA's, RNA's die de expressie van genen beïnvloeden, maar niet in eiwitten vertaald worden; iii) antisense transcriptie, transcriptie die afkomstig is van beide DNA strengen en dus in tegengestelde richtingen loopt en eventueel ook overlapt; iv) epigenetische processen, dat wil zeggen eigenschappen van het DNA die niet op een wijziging van de DNA basen zelf slaan, zoals DNA methylering, histone modificaties en de organisatie van DNA in de 3D ruimte van de celkern. Bij elkaar zijn deze mechanismen verantwoordelijk voor de juiste uitvoering van genregulatiernetwerken en het differentiëren/specialiseren van de cellen.

Een ander fenomeen dat heel duidelijk het belang voor correcte genregulatie illustreert is een mechanisme van genetische regulatie die dosis compensatie wordt genoemd. Dosis compensatie is geëvolueerd in diersoorten met verschil-

lende geslachtschromosomen tussen mannetjes en vrouwtjes. Zoogdiervrouwtjes hebben twee X-chromosomen en mannetjes zijn XY. Theoretisch hebben vrouwtjes dus twee keer meer X-gebonden gen producten dan mannetjes. Om rekening te houden met dit verschil dempt een dosis compensatie mechanisme, X-chromosoom inactivatie (XCI) genoemd, de transcriptie van bijna alle genen op één van de twee X-chromosomen in pre-implantatie embryo's van vrouwelijke zoogdieren.

XCI is een stap voor stap proces waarbij het aantal X-chromosomen eerst wordt "geteld" zodat alleen bij XX vrouwen één X-chromosoom geïnactiveerd wordt. Dit "tellen" gebeurt door een X-gebonden activator van XCI, RNF12, die XCI activeert wanneer het aanwezig is in hoge concentraties, zoals dit het geval is in XX vrouwen. De inactivatie van een X-chromosoom in vrouwelijke cellen leidt ook tot de uitschakeling van een kopie van *Rnf12*, waardoor de RNF12 concentratie in deze cellen verlaagd wordt en zo XCI van het tweede X-chromosoom voorkomen wordt. In XY mannelijke cellen is de RNF12 concentratie niet hoog genoeg om XCI te initiëren en het enige X-chromosoom blijft actief. RNF12 activeert XCI door afbraak van REX1, een transcription factor die het *Xist* gen onderdrukt en daardoor wordt de noncoding RNA *Xist* geactiveerd. In een zogenaamd "keuzeproces" wordt *Xist* expressie vervolgens op slechts één X-chromosoom behouden en leidt tot het verspreiden van *Xist* op het gehele X-chromosoom. Het noncoding RNA *Xist* rekruteert dan eiwitcomplexen die de chromatinestructuur van het X-chromosoom veranderen. Dit resulteert uiteindelijk in de uitschakeling van X-gebonden genen. In een later stadium tijdens XCI voegt DNA methylering extra stabiliteit aan de onderdrukte toestand van X-gebonden genen.



Omdat XCI alle aspecten van genregulatie in de context van ontwikkeling omvat, wordt het als een paradigma voor genregulatie als geheel beschouwd. Een groot voordeel is dat embryonale stamcellen (ES cellen) van de muis en differentiatie van deze ES cellen als model voor XCI gebruikt kunnen worden. Deze thesis onderzoekt de vele verschillende lagen van genregulatie die werken op de initiatie en handhaving van XCI. In het bijzonder, de dynamiek van de regulatie van genen die betrokken zijn bij initiatie van XCI en de rol van de ontwikkelingscontext waar de XCI in plaatsvindt worden met ES cellen als model onderzocht.

De algemene inleiding in **hoofdstuk 1** geeft een overzicht van de state-of-the-art van de wetenschap op het gebied van genregulatie, noncoding RNAs en XCI.

**Hoofdstuk 2** beschrijft een aantal experimenten die het netwerk van genen die XCI reguleren en de rol van RNF12, de enige bekende activator van XCI, verder belichten. Onze resultaten, met behulp van de deletie van het *Rnf12* gen in verschillende contexten, tonen dat RNF12 op twee niveaus werkt om XCI te reguleren. De tweemaal hogere concentratie van RNF12 in XX vrouwen in vergelijking met XY mannen is essentieel voor initiatie van XCI. Bovendien is een actieve kopie van *Rnf12* nodig om hoge *Xist* expressie te handhaven en de inactivering van het X-chromosoom goed uit te voeren. Een tweede serie experimenten die de regio rond *Xist* verwijdert suggereert bovendien dat vele *Xist*-regulerende genen in de nabijheid van *Xist* meestal door co-activatie *in cis* werken. Dat wil zeggen dat deze genen *Xist* alleen activeren als zij zich op hetzelfde chromosoom en in de nabijheid van *Xist* bevinden, eventueel door het aantrekken van andere factoren die belangrijk zijn voor de gen activatie. Herintroductie van deze genen als trans-

genen in cellen die de endogene genen missen leidt dus niet tot *Xist* activeering. Daar tegenover staan *Rnf12* transgenen, die *Xist* ongeacht hun positie in het genoom activeren. Dit bevestigt dat RNF12 de enige bekende trans-activator van XCI is. We tonen ook aan dat “pairing” van X chromosomen, een proces die wordt betrokken bij de initiatie van XCI, geen absolute vereiste is voor het optreden van XCI.

*Xist*, het gen verantwoordelijk voor het uitvoeren van XCI, wordt negatief gereguleerd door een tweede noncoding RNA gen, *Tsix*. *Tsix* wordt getranscribeerd in antisense richting en overlapt volledig met *Xist*. Omdat deze twee genen overlappen is het moeilijk om directe en indirecte effecten op hun regulering te onderscheiden. In **hoofdstuk 3** bedenken we daarom een strategie om hun vroege genetische regulatie onafhankelijk van invloeden van de antisense partner te bestuderen. Hiervoor werd *Xist* vervangen door een groene (EGFP) en *Tsix* door een rode (mCherry) fluorescente reporter. Het resultaat is dat de fluorescente reporters door de endogene promoters van *Xist* en *Tsix* gereguleerd worden, waardoor een cel EGFP accumuleert wanneer de *Xist* promotor en mCherry accumuleert wanneer de *Tsix* promotor actief is. Experimenten met deze cellijnen tonen dat beide genen antagonistische rollen hebben, *Tsix* onderdrukt *Xist* en vice versa. Door de fluorescerende reportereiwitten in deze cellijnen kunnen we gemakkelijk de activiteit van *Xist* en *Tsix* in levende cellen bepalen, bijvoorbeeld met behulp van op fluorescentie gebaseerde technieken in combinatie met microscopie of flow-cytometry. Wij hebben de dynamiek van *Xist* en *Tsix* regulatie tijdens ES cel differentiatie onderzocht door continue live cell imaging uit te voeren. Het blijkt dat de regulatie van *Xist* en *Tsix* tijdens differentiatie onderling is afgestemd maar niet direct afhankelijk van elkaar is. Een bep-



aald expressie niveau is ook stabiel dan eerder gedacht. Dat betekent dat zodra een “beslissing” voor zowel activering of deactivering genomen wordt, deze “beslissing” langer blijft bestaan. Bijvoorbeeld, de moedercel en beide dochtercellen van een celdeling hebben een vergelijkbaar niveau *Xist* / *Tsix* activatie. De fluorescerende reporters maken het ook mogelijk de impact die XCI activatoren en remmers, zoals RNF12 en REX1, op *Xist* en *Tsix* regulering hebben zorgvuldig te analyseren. Bovendien tonen experimenten met vrouwelijke cellen die één van de twee X-chromosomen verloren hebben aan dat niet alleen *Xist* maar ook *Tsix* door het aantal X-chromosomen in een cel gereguleerd wordt. Tot slot worden onder specifieke experimentele omstandigheden twee semi-stabiele populaties in de reporter cellijnen geobserveerd. Één populatie heeft een lager *Tsix* / mCherry niveau, de andere populatie een hoger *Tsix* / mCherry niveau. Zorgvuldige analyse van deze populaties toont aan dat ze eventueel met twee verschillende hogere orde 3D conformaties van het DNA rondom *Xist* en *Tsix* corresponderen. Omdat de verschillende populaties ook de uitkomst van het XCI proces voorspellen, dus *Tsix* / mCherry lage cellen up-reguleren *Xist* terwijl *Tsix* / mCherry hoge cellen dat niet doen, verbinden deze resultaten de 3D-conformatie van het DNA op een prachtige manier met functionele gevolgen.

Tijdens het begin van de ontwikkeling ondergaat het embryo een aantal opvallende veranderingen. Deze veranderingen worden voornamelijk gedreven door uitgebreide modulatie van gen-regulerende netwerken en gelijktijdige reorganisatie van het chromatine, het verpakte DNA template. In deze periode wordt ook XCI geïnitieerd en vastgelegd. Het tweede deel van het proefschrift richt zich daarom op het samenspel van genregulatie en de structuur van chromatine in deze vroege

fase van de embryonale ontwikkeling. In **appendix A** wordt een techniek gebruikt die huidfibroblasten in zogenaamde geïnduceerde pluripotente stamcellen (iPS cellen) transformeert. IPS cellen lijken op ES cellen, ze kunnen bijvoorbeeld in verschillende celtypes differentiëren. Dus door het transformeren van fibroblasten in iPS cellen worden deze cellen in de omgekeerde volgorde terug door differentiatie gedwongen, totdat ze opnieuw pluripotent zijn. De fibroblasten die in **appendix A** gebruikt worden, zijn gemaakt van een drager van een *FMRI* mutatie. Normaal gesproken wordt DNA rond de *FMRI* promotor in dragers van deze mutatie tijdens de vroege ontwikkeling gemethyleerd. Daardoor wordt dit gen uitgeschakeld en veroorzaakt het de ziekte fragile X syndroom. In dit specifieke geval blijft de *FMRI* promotor zonder DNA methylering en is het gen actief. Verrassenderwijs wordt de *FMRI* promotor bij transformatie van de fibroblasten toch nog gemethyleerd. Dat betekent dat de cellen die getransformeerd kunnen worden door een specifieke ontwikkelingsfase gingen waarin de mutatie “juist” herkend wordt. Dit leidt tot DNA methylering en uitschakeling van *FMRI*. Daarom het *FMRI* gen in eerste instantie, tijdens de vroege embryonale ontwikkeling van de drager van de mutatie, niet wordt gemethyleerd en uitgeschakeld blijft een onderwerp voor toekomstig onderzoek.

Ook in **hoofdstuk 4** wordt een speciale reguleringsfase tijdens de vroege embryonale ontwikkeling geanalyseerd. Omdat veel genen, bijvoorbeeld *Xist* en *Tsix*, als sense-antisense gen paren in elkaar gezet zijn, werd een reporter systeem gecreëerd dat deze situatie imiteert. Een sequentie die voor het groen fluorescent eiwit EGFP codeert werd in antisense oriëntatie geplaatst van een induceerbare promotor, een promotor die door toevoeging of verwijdering van het geneesmiddel doxycycline aan- of uitgeschakeld





wordt. Dit construct werd in ES cellen geïntroduceerd zodat antisense transcriptie plaatsvindt door de EGFP coderende sequentie / promoter na toevoeging van doxycycline. In ongedifferentieerde ES cellen wordt het EGFP door inductie van antisense transcriptie door het EGFP reversibel uitgeschakeld. Dat wil zeggen, zodra doxycycline uit het medium verwijderd wordt en antisense transcriptie stopt, komt EGFP opnieuw tot expressie. Dit effect is niet toe te schrijven aan transcriptionele interferentie, de directe botsing van transcriptie machines die uit tegengestelde richting op elkaar lopen. In plaats daarvan blijkt dat het effect gemedieerd wordt door chromatine modificaties die door de doorlopende transcriptiemachinerie in de EGFP promoter neergezet worden. In tegenstelling tot de reversibele uitschakeling in ongedifferentieerde ES cellen, leidt antisense transcriptie door EGFP tijdens differentiatie tot stabiele uitschakeling van het EGFP. In deze situatie worden aanvullende kenmerken van gene repressie zoals DNA methylering gedetecteerd. Deze resultaten geven nog maar eens aan dat cellen door een speciale fase tijdens de vroege ontwikkeling gaan, waarin specifieke kenmerken van chromatine herkend worden en een blijvend effect op de structuur van chromatine en genregulatie hebben.

Tenslotte worden de resultaten van dit proefschrift in de algemene discussie in **hoofdstuk 5** in het licht van de huidige ontwikkelingen op het gebied van X-chromosoom inactivatie en genregulatie besproken.



## Zusammenfassung

Genregulation ist der Prozess, der bestimmt in welchem Maß und in welchem Raum-Zeit Kontext Gene in einer Zelle an- oder ausgeschaltet werden. Wenn verschiedene Kombinationen von Genen in zwei verschiedenen Zellen aktiv sind, ist auch die molekulare Zusammensetzung dieser Zellen verschieden und die Zellen können dementsprechend auch verschiedene Aufgaben ausführen. Darum ist Genregulation von fundamentaler Bedeutung um eine Vielzahl verschiedener Zelltypen und Gewebe in mehrzelligen Organismen zu generieren und aufrecht zu erhalten. Dieser Zusammenhang zeigt sich auch sehr deutlich in den vielen Fällen, in denen eine fehlerhafte Regulierung von Genexpression zu Krankheiten wie zum Beispiel Krebs führt.

Genregulation wird auf verschiedenen Ebene durch Kontrolle der Gentranskription gesteuert. Diese Ebenen umfassen: i) Proteine, die Gene direkt aktivieren oder hemmen, sogenannte Transkriptionsfaktoren; ii) nicht-kodierende RNAs, RNAs die Genexpression beeinflussen aber nicht in Proteine translatiert werden; iii) antisense Transkription, also Transkription von beiden DNA-Strängen, die daher in gegenläufige Richtungen läuft und so auch überlappen kann; iv) epigenetische Prozesse, also die Eigenschaften der als Chromatin verpackten DNA, die nicht auf der Basenzusammensetzung der DNA beruhen, zum Beispiel DNA-Methylierung, Histonmodifikationen und die Organisation von DNA im 3-dimensionalen Zellkern. Zusammen sind diese Ebenen für die korrekte Ausführung von Genregulations- und Entwicklungsprogrammen verantwortlich.

Ein weiteres Phänomen, das zeigt

wie wichtig korrekte Genregulation ist, ist ein genetischer Regulationsmechanismus, der als Dosiskompensation (dosage compensation) bezeichnet wird. Dosiskompensation ist durch Evolution in Spezies entstanden, in denen männliche und weibliche Individuen unterschiedliche Sexchromosomen besitzen. In Säugetieren haben Weibchen zum Beispiel zwei X Chromosomen, während Männchen ein X und ein Y Chromosom tragen. Das bedeutet, dass Weibchen theoretisch zweimal mehr Genprodukte von ihren X Chromosomen produzieren als Männchen. Um diesen Unterschied auszugleichen, ist ein Dosiskompensationsmechanismus entstanden, der X-Chromosom-Inaktivierung (XCI) genannt wird und fast alle Gene auf einem der beiden X-Chromosomen während der frühen Embryonalentwicklung von weiblichen Säugetieren ausschaltet.

XCI ist ein schrittweiser Prozess bei dem zuerst die Anzahl X-Chromosomen in einer Zelle „gezählt“ wird um sicherzugehen, dass nur in Weibchen und nur ein X-Chromosom ausgeschaltet wird. Dieses „Zählen“ („counting“) wird über den XCI-Aktivator RNF12 geregelt, der auf dem X-Chromosom kodiert ist und XCI nur aktiviert wenn in ausreichend hoher Konzentration vorhanden. In Weibchen, die ja zwei X-Chromosomen haben, ist das der Fall. Die Inaktivierung eines X-Chromosoms in Weibchen führt also auch zur Repression von einer *Rnf12*-Kopie und dadurch sinkt die RNF12-Konzentration in diesen Zellen und XCI des zweiten X-Chromosoms findet nicht statt. In XY Männchen ist die RNF12-Konzentration nicht hoch genug um XCI zu initiieren. RNF12 aktiviert XCI, indem es zum Abbau von REX1, einem Transkriptionsfaktor der das *Xist*-Gen unterdrückt, führt. Die nicht-kodierende RNA *Xist* wird also durch RNF12 hochreguliert. In einem „Auswahl“ („choice“) genannten Prozess wird *Xist* Expression daraufhin



auf nur einem X-Chromosom beibehalten, was schließlich dazu führt, dass sich Xist RNA ausbreitet und das gesamte X-Chromosom bedeckt. Die nicht-kodierende RNA Xist rekrutiert in der Folge Proteinkomplexe, die die Struktur des X-Chromosoms verändern und zur Repression von Genen auf dem X-Chromosom führen. Zu einem späteren Zeitpunkt von XCI tritt auch DNA-Methylierung an den ausgeschalteten Genen auf und stabilisiert so den Aus-Zustand.

XCI gilt als ein Paradebeispiel für Genregulation insgesamt, da es praktisch alle Aspekte der Genregulation im Kontext der Embryonalentwicklung rekapituliert. Hinzu kommt praktischerweise, dass embryonale Stammzellen (ES-Zellen) von Mäusen und ES Zelldifferenzierung als *in vitro* Modell zur Erforschung von XCI benutzt werden können. Diese Doktorarbeit versucht ein neues Licht auf die vielen Ebenen der Genregulation zu werfen, die die Initiierung und Aufrechterhaltung von XCI steuern. Besonders der dynamische Aspekt der Regulierung von Genen, die am XCI Prozess beteiligt sind, und der entwicklungsbiologische Kontext werden dabei mit Hilfe von ES-Zellen betrachtet.

Eine grundlegende Einleitung in **Kapitel 1** erstellt eine Übersicht des Standes der Wissenschaft in den Bereichen Genregulation, nicht-kodierende RNAs und XCI. **Kapitel 2** beschreibt eine Reihe von Experimenten, die versuchen die Gennetzwerke die XCI kontrollieren und die Rolle von RNF12, dem einzigen zum jetzigen Zeitpunkt bekannten XCI-Aktivator, besser zu verstehen. Unsere Ergebnisse mit mehreren *Rnf12*-Deletionen in verschiedenen Situationen zeigen, dass RNF12 XCI auf zwei Ebenen reguliert. Die im Vergleich zu männlichen XY Zellen zweifach höhere RNF12-Konzentration in weiblichen XX Zellen ist essentiell um den XCI Prozess zu starten.

Darüber hinaus ist eine aktive Kopie von *Rnf12* nötig, um auch zu einem späteren Zeitpunkt die Genrepression auf dem inaktivierten X-Chromosom aufrecht zu erhalten, und XCI korrekt abzuschließen. Eine zweite Serie von Experimenten, in denen die Regionen um das *Xist*-Gen deletiert wurden, deutet darauf hin, dass viele *Xist*-regulierende Gene in nächster Nachbarschaft zu *Xist* ihre Aufgabe hauptsächlich über Co-Aktivierung *in cis* erfüllen. Das bedeutet, dass diese Gene *Xist* nur aktivieren können, wenn sie auf demselben Chromosom und direkt neben *Xist* liegen. Es ist nicht abschließend geklärt wie ein solcher Mechanismus funktionieren könnte, eine Möglichkeit wäre jedoch, dass diese aktiven Gene andere aktivierende Faktoren rekrutieren und so auch benachbarte Gene, zum Beispiel *Xist*, kumulativ aktivieren. Daraus folgt auch, dass das Wiedereinsetzen von diesen Genen in Zellen, in denen die ursprünglichen Gene deletiert wurden, als sogenanntes Transgen, nicht zur Aktivierung von *Xist* führt. Im Gegensatz dazu kann ein *Rnf12*-Transgen *Xist* unabhängig von seiner Position im Genom aktivieren. Dies bestätigt RNF12 als einzigen bis dato bekannten XCI *trans*-Aktivator. Außerdem zeigen wir, dass X-Chromosom-„Paarung“ („pairing“), ein Prozess bei dem sich beide X-Chromosomen vorübergehend sehr nahe kommen und der in der Vergangenheit mit der Regulation von XCI in Verbindung gebracht wurde, für den korrekten Ablauf von XCI nicht benötigt wird.

*Xist*, das Gene, das für XCI verantwortlich ist, wird negativ von einem zweiten nicht-kodierendem RNA Gen reguliert. Dieses Gen, *Tsix*, wird in antisense Richtung zu *Xist* transkribiert und überlappt vollständig mit diesem. Auf Grund dieser Anordnung ist es schwierig zwischen direkten und indirekten Effekten auf die Regulation dieser beiden Gene zu unterscheiden. In **Kapitel 3** konzipieren wir daher

ein Strategie um die frühe genetische Regulation von *Xist* und *Tsix* unabhängig von Einflüssen des jeweiligen antisense-Partners zu untersuchen. Zu diesem Zweck wurde *Xist* mit einem grünen (EGFP) und *Tsix* mit einem roten (mCherry) fluoreszierendem Reporter, einer Art Marker, ersetzt. Als Resultat werden diese Reporter nun durch die endogenen Promotoren von *Xist* und *Tsix* gesteuert, was bedeutet, dass Zellen mit aktivem *Xist*-Promotor EGFP und Zellen mit aktivem *Tsix*-Promotor mCherry akkumulieren. Experimente mit diesen Zelllinien zeigen, dass beide Gene antagonistische Rollen haben, *Tsix* inhibiert *Xist* und umgekehrt. Die fluoreszierenden Reporter in diesen Zelllinien erlauben uns die Aktivität von *Xist* und *Tsix* mit Hilfe von fluoreszenzbasierten Techniken wie Mikroskopie oder Durchflusszytometrie einfach zu bestimmen. Aus diesem Grund konnten wir die Dynamik der Regulation von *Xist* und *Tsix* während der ES Zelldifferenzierung mit kontinuierlicher live-Mikroskopie von lebenden Zellen verfolgen. Anscheinend ist die Regulation von *Xist* und *Tsix* zwar im Differenzierungskontext aufeinander abgestimmt, die beiden Gene hängen aber nicht notwendigerweise direkt voneinander ab. Expressionslevel erscheinen außerdem stabiler zu sein als bisher vermutet. Das heißt, sobald eine „Entscheidung“ für Aktivierung oder Ausschalten getroffen wurde, wird diese Entscheidung erst einmal beibehalten. Zum Beispiel haben die Mutterzelle und beide Tochterzellen einer Zellteilung ein meist sehr ähnliches Niveau an *Xist*/*Tsix*-Aktivierung. Die Fluoreszenzreporter ermöglichen auch eine sorgfältige Analyse des Einflusses von XCI-Aktivatoren und -Inhibitoren, zum Beispiel RNF12 und REX1, auf die Regulation von *Xist* und *Tsix*. Darüber hinaus zeigen Experimente mit weiblichen Zellen, die ein X-Chromosom verloren hatten, dass nicht nur *Xist* sondern auch *Tsix* von der Anzahl X-Chromosomen in einer

Zelle beeinflusst wird. Schließlich werden unter bestimmten experimentellen Bedingungen zwei semi-stabile Populationen in den Reporterzelllinien beobachtet. Die eine Population hat ein niedrigeres, die andere ein höheres Level an *Tsix*/mCherry. Die sorgfältige Analyse dieser Population suggeriert, dass sie möglicherweise zwei verschiedene 3D-Konformationen der DNA um *Xist* und *Tsix* herum darstellen. Da die beiden Populationen auch das Ergebnis des XCI-Prozesses voraussagen, Zellen mit niedrigem Level *Tsix*/mCherry aktivieren nämlich *Xist* während die Zellen mit hohem *Tsix*/mCherry Level das nicht tun, verbinden diese zwei Population auf elegante Weise die 3D-Konformation der DNA mit funktionellen Konsequenzen.

In den frühesten Phasen seiner Entwicklung durchläuft der Embryo eine Reihe von ausgeprägten Wandlungen. Diese Veränderungen werden hauptsächlich durch die umfangreiche Modulation von Genregulationsnetzwerken und gleichzeitiger Remodellierung der Chromatinstruktur angetrieben. In diesem Zeitraum findet auf XCI statt. Der zweite Teil der Doktorarbeit befasst sich daher mit der Wechselwirkung zwischen Genregulation und Chromatinstruktur in dieser frühen Entwicklungsphase. In **Appendix A** wird eine spezielle Technik benutzt um Hautfibroblasten künstlich in sogenannte induzierte pluripotente Stammzellen (iPS-Zellen) zu reprogrammieren. IPS-Zellen ähneln ES-Zellen, auch sie können in verschiedene Zelltypen differenziert werden. Die Fibroblasten werden also sozusagen in umgekehrter Reihenfolge rückwärts durch den Differenzierungsprozess gezwungen bis sie schließlich als iPS-Zellen wieder pluripotent sind. Die Fibroblasten, die in **Appendix A** benutzt wurden, stammen vom Träger einer *FMRI*-Mutation. Normalerweise wird der Promotor des *FMRI*-Gens in Trägern dieser Mutation in frühen Stadien der



Embryonalentwicklung methyliert, was in der Repression von *FMRI* und der Krankheit Fragiles-X-Syndrom resultiert. In diesem speziellen Fall jedoch blieb der *FMRI*-Promotor des Trägers unmethyliert und das Gen aktiv. Überraschenderweise wurde der *FMRI*-Promotor während der Reprogrammierung zu iPS-Zellen dann doch noch methyliert. Diese Beobachtung impliziert, dass die reprogrammierten Zellen eine spezielle Entwicklungsphase durchlaufen haben, in der die Mutation „korrekt“ erkannt wurde, was daraufhin zu DNA-Methylierung und Genrepression führte. Warum das *FMRI*-Gen nicht von vornherein, also während der frühen Embryonalentwicklung des Trägers der Mutation, ausgeschaltet wurde, wird in zukünftigen Studien zu klären sein. Auch in **Kapitel 4** wird eine spezielle Phase der Genregulation in frühen Entwicklungsstadien untersucht. Da viele Gene wie zum Beispiel *Xist* und *Tsix* als sense-antisense Gen-Paar angeordnet sind, wurde ein Reporter-System generiert um diese Situation nachzuahmen. Eine Sequenz die für das grün fluoreszierende Protein EGFP codiert wurde in antisense, also gegenläufiger, Richtung zu einem induzierbaren Promotor eingesetzt. Der induzierbare Promotor ist ein Promotor, der durch Hinzugabe oder Entfernen des Stoffes Doxycyclin an- oder ausgeschaltet werden kann. Dieses Konstrukt wurde in so einer Weise in ES-Zellen eingebaut, dass also antisense Transkription durch die für EGFP kodierende Sequenz und deren Promoter durch Hinzugabe oder Entfernen von Doxycyclin einfach an- oder ausgeschaltet werden konnte. In undifferenzierten ES-Zellen führte die Induktion von antisense Transkription durch das EGFP zu reversibler Repression von EGFP. Das bedeutet, dass EGFP erneut zur Expression kam, nachdem Doxycyclin aus dem Medium entfernt und so die antisense Transkription durch das EGFP gestoppt wurde. Dieser Effekt beruht nicht

auf direkter transkriptioneller Interferenz, also zum Beispiel der Kollision von Transkriptionskomplexen, die aus gegenläufigen Richtungen aufeinander zu laufen. Vielmehr scheint dieser Effekt durch Chromatinmodifikationen herbeigeführt zu werden, die durch die Transkriptionsmaschinerie im Promoter von EGFP erzeugt werden. Im Gegensatz zu der reversiblen Repression, die in undifferenzierten ES-Zellen beobachtet wurde, verursacht während der Differenzierung induzierte antisense Transkription eine stabile Repression des EGFP. Das heißt, auch nach dem Auswaschen von Doxycyclin und dem Stopp von antisense Transkription bleibt das EGFP weiterhin ausgeschaltet. In dieser Situation wurden darüber hinaus zusätzliche Merkmale von Genrepression, nämlich DNA-Methylierung, entdeckt. Diese Ergebnisse verdeutlichen noch einmal, dass Zellen während der frühen Entwicklung eine spezielle Phase durchlaufen, in der spezifische Merkmale des Chromatins erkannt werden und einen bleibenden Einfluss auf die Chromatinstruktur und Genregulation haben.

In **Kapitel 5**, der Generellen Diskussion, werden die Ergebnisse dieser Doktorarbeit schließlich im Licht von gegenwärtigen Entwicklungen in den Bereichen X Chromosom Inaktivierung und Genregulation diskutiert.

## Curriculum Vitae

**Name** Friedemann Loos  
**Date of Birth** September 22, 1983, Münster, Germany  
**Nationality** German  
**Phone** +31(0)6-24963142  
**Email** friedemannloos@gmail.com



## Work Experience

**Oct 2010** PhD candidate at Erasmus MC, Rotterdam, The Netherlands

**to present** Prospective degree: PhD (March 2015)

Projects in the field of X chromosome inactivation integrating stem cell research, epigenetics and molecular biology. Cloning/targeting strategies in stem cells, generation of mouse models, (live time-lapse) microscopy, cell sorting/flow cytometry, (non-coding) RNA techniques and basic molecular biology / biochemistry / protein methods.

## Education

**Oct 2005** Ludwig-Maximilians-University Munich, Germany

**to Jun 2010** Degree: Diploma ("Diplom")

Major: Genetics

Minors: Immunology, Cell Biology, Microbiology

## Internships & Thesis

**Oct 2009** Diploma Thesis: The Role of ACF1 in DNA Break Repair (Becker lab)

**to Jun 2010** Project involving a chromatin remodeler and DNA repair. Methods include tissue culture, biochemistry, DNA repair assays and molecular biology.

**Sep 2008** Visiting Graduate Student at University of Washington, Seattle, USA

**to Jun 2009** (Fulbright Scholarship, Henikoff lab at FHCRC)

Project in the field of transcriptional regulation by RNA polymerase 2. Methodology includes chromatin immuno-precipitation, Western blotting and molecular cloning.



- May 2008**      Lab aide in B. Jungnickel's lab, Helmholtz Centre, Munich, Germany
- to Jul 2008**    Design and cloning of a targeting vector to generate a knock-out cell line.
- 
- Feb 2008**      Practical in A. Brachmann's / M. Parniske's lab, Munich, Germany
- to Apr 2008**    Yeast genetics, cloning of fusion proteins, metabolic assays in yeast.

**Further activities and work**

- Oct 2003**      Social work, school for handicapped children, Cajamarca, Peru
- to Mar 2005**    Education/caretaking of children with diverse degrees of mental/physical disability
- 
- since**          Instructor of Seven Star Mantis Kung Fu
- Jul 2003**      Teaching groups of all sizes (private training to >50 ) and ages (6 to 50 years), giving seminars, involved in organization of tournaments

**Interests**

Climbing/Mountaineering, Seven Star Mantis Kung Fu, Nature, Literature



## List of Publications

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. ISBN: 978-87-93237-07-0 (Print). River Publishers. 2015 Jan.

de Esch CEF, Ghazvini M, **Loos F**, Schelling-Kazaryan N, Widagdo W, Munshi ST, van der Wal E, Douben H, et al., Epigenetic characterization of the FMR1 promoter after reprogramming of a human fibroblast line carrying an unmethylated full mutation into pluripotent stem cells, Stem Cell Reports. 2014 Oct 14;3(4):548-55.

Tomi NS, Davari K, Grotzky D, **Loos F**, Böttcher K, Frankenberger S, Jungnickel B, Analysis of SHPRH functions in DNA repair and immunoglobulin diversification, DNA Repair (Amst). 2014 Oct 10. pii: S1568-7864(14)00241-9.

Barakat TS, **Loos F**, van Staveren S, Myronova E, Ghazvini M, Grootegoed JA, Gribnau J, The *Trans*-Activator RNF12 and *Cis*-Acting Elements Effectuate X Chromosome Inactivation Independent of X-Pairing, Molecular Cell. 2014 Mar 20;53(6):965-78.

Ellerbeck M, Schüßler A, Brucker D, Dafinger C, **Loos F**, Brachmann A, Characterization of three ammonium transporters of the glomeromycotan fungus *Geosiphon pyriformis*, Eukaryot Cell. 2013 Nov;12(11):1554-62.

**Loos F**, van Wijk L, Grootegoed JA, Gribnau J, Chromatin mediated reversible silencing of sense-antisense gene pairs in ESCs is consolidated upon differentiation. *Submitted*.

**Loos F**, Grootegoed JA, Gribnau J, *Xist* and *Tsix* transcription dynamics are regulated by the X-to-autosome ratio and predicted by semi-stable transcriptional states. *Manuscript in preparation*.



# PhD Portfolio

## Summary of PhD training and teaching

Name PhD student: Friedemann Loos	PhD period: October 2010 – March 2015
Erasmus MC Department: Developmental Biology	Promotor(s): Joost Gribnau
Research School: Biomedical Sciences	Supervisor: Joost Gribnau
<b>1. PhD training</b>	
	<b>Year</b>
<b>General courses</b>	
- Laboratory animal science (art.9/FELASA-C)	2012
- Handling of laboratory animals (IVC)	2013
<b>Specific courses (e.g. Research school, Medical Training)</b>	
- Biochemistry and Biophysics	2010
- Cell and Developmental Biology	2011
- Genetics	2011
- Literature course	2011
<b>Seminars and workshops</b>	
- (Confocal) Microscopy introductory courses (Zeiss LSM700, Leica SP5, Nikon Ti Eclipse Spinning Disc)	2011, 2013, 2014
- MGC PhD Workshops	2011-2014
- Adobe Photoshop and Illustrator workshop	2011
- Joint XCI Retreat (with Heard laboratory), Spetses, Greece	2012
- BD FACS Aria III operator training, Erebodegem, Belgium	2013
- Joint XCI Retreat (with Heard laboratory), Rotterdam, The Netherlands	2013
- 23 <sup>rd</sup> MGC Symposium, Rotterdam, The Netherlands	2013
- 24 <sup>th</sup> MGC Symposium, Rotterdam, The Netherlands	2014
- Erasmus MC PhD day	2014
- Plantum Matchmaking Day	2014
<b>Presentations</b>	
- Reproduction and Development/Biochemistry Work Discussions	2010-2015
- MGC PhD Workshops	2011-2014
- TRR81 International Research Consortium Meetings, Hirschegg, Austria	2011, 2012
- Joint X Chromosome Inactivation Retreats (with Heard laboratory)	2012, 2013
<b>(Inter)national conferences</b>	
- 4 <sup>th</sup> International NIRM symposium, Amsterdam, The Netherlands	2010
- EMBO Workshop 50 Years of X-inactivation Research, Oxford, UK (poster)	2011
- TRR81 International Chromatin Symposium, Giessen, Germany (poster)	2011
- 5 <sup>th</sup> International NIRM symposium, Amsterdam, The Netherlands (poster)	2012
- 11 <sup>th</sup> Annual ISSCR Meeting, Boston, USA (poster)	2013
- Max Planck Epigenetics Meeting, Freiburg, Germany (poster)	2014
<b>2. Teaching</b>	
	<b>Year</b>
<b>Supervising Master's theses/internships</b>	
- Supervision of Bioinformatics Msc. internship student	2012
- Supervision of Biology Msc. internship student	2014



## Acknowledgements

With my time in Rotterdam coming to an end, I want to say thanks to all the people that made it happen.

Dear Joost, from the first time I talked to you there was no doubt that I wanted to work with you. In the last years I have not come to regret that choice a single time. Rather, I realized to the full extent what kind of a great personality you are. Be it scientifically, with your exuberantly exploratory mind, be it as a dancer, a birdwatcher or just another nice person to talk to about the world and all, an objective and fair observant. Thank you for all that!

Dear Anton, I also want to thank you for giving me the opportunity to start working in the Department of Reproduction and Development. I am very glad for all the scientific and non-science-related input I got from you and Willy over these years.

I am very grateful that Frank Grosveld, Peter Verrijzer and Adriaan Houtsmuller took the time to be part of my inner committee and read my thesis. Actually, thanks also for approving the thesis! Thank you Rob Willemsen and Raymond Poot to complete my committee and be there for my thesis defense.

Initially, this curious Dutch tradition-construct called paranymph made me mostly smile...By now though I am very happy that good Ruben, Michael and Agnese take on that task. I am proud to have you standing there! Especially Agnese has been a cornerstone for my fascination for science, my thesis and far beyond that.

From a technical point of view I thank all the great facilities of the Erasmus MC and all the people who taught

me how to hold the tweezers here and how to make it stick there. However, I would like to single out Gert-Jan and Gert from the OIC, and Reinier from the Flow Cytometry facility. Without their help this thesis would be deprived of most of the more interesting pages.

I will not start to list all the great, kind and fantastic people I met during my time at the Erasmus MC. The risk of forgetting somebody by accident is just too big. People that have left, people that are still there. Even without great words I am sure the things we share are pretty clear to each one of you. To a big part, you were the reason, I would almost say the essence, of having a good time in Rotterdam. I don't know if you are my friends, but this one truth is plain: I kinda like you!

A somewhat far-reaching but still justifiable extension of the group of people that made this thesis happen includes my pre-Erasmus friends, most of which by now live in Berlin, even though there is also some in Münster and actually all around Germany and the globe. It is great to know that in spite of my pathologic sluggishness in picking up the phone to give you a call you still talk to me. It has been an honor for all these years!

Now, the last group I want to thank was absolutely required to finish this thesis, they were even absolutely required to start it. Not surprisingly, I am talking about my family. I am grateful that every single one of you, from littlest cousin to oldest grandpa, exists. Dear Mama and Papa, I guess some might object, but in my opinion you have done a tremendous job in bringing me, Swaantje, Josina and Jakob, who are the gold standard for siblings, up. I have seen quite a bit since then, and still have never met any parents cooler than you!

