# Integrative Analysis of X-Chromosome Reactivation Kinetics in a Novel Reprogramming System

Moritz Bauer

TESI DOCTORAL UPF / 2019

Thesis Supervisor
Dr. Bernhard Payer

Gene Regulation, Stem Cells and Cancer Program
Centre for Genomic Regulation (CRG)

DEPARTMENT OF EXPERIMENTAL AND HEALTH SCIENCE


- Für meinen Opa -

*Remember, if it would be easy, everyone could do it.*
(My Dad )

# Dedication and Acknowledgments

After four years, my PhD has come to an end. Looking back at it now, I remember people's stories, their struggles, the madness, how they described it as a Tolkien-like drama, always battling forces bigger than them. However, it never felt this way to me. And the reason is simple: I was never alone in it.

I want to thank Bernhard, for trusting me so much, for letting me choose my own fight, while literally, always keeping his door open. I want to thank the whole lab, for simply being a great bunch of people, for always sticking together, for lunch together, coffee breaks and way too many sweets from around the world. I want to thank Guillaume, who took us in in time of need, for guiding us when we lost our way. It would have been impossible without you. Quique, for being our wizard, for joining this project, for giving meaning to millions of numbers. I want to thank my PhD committee, Stephan, Marcus and again Guillaume for year by year, putting us on the right track again. I want to thank the whole Hi-C crew, François, Gregoire, and Ralph, for answering my countless questions and making it seem so easy. Thank you, Isa and César, for sharing the first steps into ATAC and telling me, when my protocol was wrong. *Moltes gràcies* to the lovely people of the FACS facility Àlex, Erika, Eva, and Òscar, for allowing me to spend borderline crazy hours there, for always looking after my girls, for all the free Spanish lessons.

*Ze Gang*.
Birgit, for having been my anchor in this new world, being our real estate agent, for pushing me when I needed it. This new life was so easy with your help. Just please don't make me this drunk again…
Claudia, for being my shoulder to cry on in the hallway, for telling me when I used two fonts, for nurturing us when we would otherwise survive on chips and beer. Grazie!
Dan, for just being the most relaxed and easygoing roommate I could have ever asked for. For all the talks after work, when I just needed to let it all out. For the shared whiskeys when I needed them. Cheers mate!
Jackie, for truly knowing me. For literally always being there, for just having to turn around in my chair to have someone where no secrets where needed. To the music shared together. I guess I paid you well in chewing gums.
João, for the deep talks, for always asking when you saw me down.
Paul, for being the buddy I wish everyone could have. For being a constant on this rollercoaster. To beer and football and simple life!

Leire, for sticking with me despite my questionable volley skills. To La Croqueta y El Schnitzel!
Marta, for catching me when I was falling. For sharing my thoughts and emotions. For simply having the best food together.
Oz, for being a true friend, for sharing the deepest, for simply sharing pizza. For 35mm.

*The Molbios* (including wonderful additions): Chrissi, Birgit, Motz, Mada, Thom, Jassi, Papollo, Kathi, Michi, Meri, Laura, Jakob, David, Strebi, Niko, Ulrich. For being the most amazing group of

Uni friends possible. For all the nights shared together. This is what *home* feels like. Weil es sich immer noch so anfühlt, als ob wir uns erst gestern gesehen hätten.

Eyup, für die Musik, die innige Freundschaft, für ISIS und TOOL. Dafür dass du mit mir die Bestie zum Leben erweckt hast.
Johanna, dafür dass du schon so früh an mich geglaubt hast.
Lukas, dafür das du mich all die Jahre ertragen hast. Mit mir gelitten hast, du immer nur einen Anruf weg warst.
Miri, für das Verständnis, fürs Zuhören. Bis bald im Raimann.

Mama, Papa, Laura. Ich hoff ich hab's euch nicht zu schwer gemacht. So weit weg. Es war nicht immer leicht. Oft merkt man erst, wie wunderbar etwa ist, wenn es plötzlich fehlt. So ging es mir. Danke für alles, für all die Liebe. Dafür dass ihr mich zu dem gemacht habt der ich bin. Dafür, dass ich genau weiß, wie sich „Zu Hause" anfühlt.

# Abstract

The reactivation of the inactive X chromosome has the potential to provide a unique system to study the developmentally induced formation of euchromatin. However, insights into this process were hampered by the lack of adequate systems, which would allow the dissection of the process using high-throughput sequencing techniques. Here I describe the development of a novel induced pluripotent stem cell reprogramming system that allows the isolation of cells poised for X-reactivation, subsequently achieving near-deterministic efficiency of X-reactivation. Utilizing this novel system, we were able to reveal that the reactivation of silenced genes occurs rapidly and can be divided into distinct initiation and completion phases. Similarly, we could show that chromatin opening of the inactive X proceeds in a two-step fashion, initiating in close proximity to previously open regions, and possibly being initiated by pluripotency factors. Finally, we could show that mega-domains and TADs correspond to two different levels of three-dimensional genome organization superimposed on the Xi, independent of gene expression. We conclude that gene expression and chromatin accessibility during X-reactivation share similar kinetics, while genome organization might follow distinct principles.

# Resumen

La reactivación del cromosoma X inactivo tiene el potencial de proporcionar un sistema único para estudiar la formación de eucromatina inducida por el desarrollo. Sin embargo, la comprensión de este proceso se vio obstaculizada por la falta de sistemas adecuados, lo que permitiría la disección del proceso utilizando técnicas de secuenciación de alto rendimiento. Aquí describo el desarrollo de un nuevo sistema de reprogramación de células madre pluripotentes inducidas que permite el aislamiento de células preparadas para la reactivación de X, logrando posteriormente la eficiencia casi determinista de la reactivación de X. Utilizando este novedoso sistema, pudimos revelar que la reactivación de genes silenciados ocurre rápidamente y puede dividirse en distintas fases de iniciación y finalización. Del mismo modo, podríamos mostrar que la apertura de cromatina de la X inactiva se realiza en dos pasos, iniciando en las proximidades de regiones previamente abiertas, y posiblemente iniciada por factores de pluripotencia. Finalmente, podríamos mostrar que los mega-domains y los TADs corresponden a dos niveles diferentes de organización del genoma tridimensional superpuestos en el Xi, independientemente de la expresión génica. Llegamos a la conclusión de que la expresión génica y la accesibilidad a la cromatina durante la reactivación X comparten una cinética similar, mientras que la organización del genoma podría seguir principios distintos.

# Preface

One of the fundamental problems modern biology is trying to solve is how cellular complexity arises from a singular genetic code. Starting from a single zygote, the human body develops into a piece of complex machinery made up of 37 trillion cells. During this process, the underlying genetic code remains the same, however, what changes, it how it is packed. The DNA is wrapped around proteins called histones and together forms an assembly termed chromatin. Chromatin allows the cells to package different parts of the genome at different densities. Pieces of DNA that need to be transcribed are loosely packed to be open (euchromatin), whereas other parts are tightly packed to safeguard them from unwanted activation (heterochromatin). Chromatin accessibility changes dynamically throughout development, however, the arguably most dramatic change can be observed on the X chromosome in female mammals. Here, in a process called X chromosome inactivation, one of the two X undergoes a unique transformation from a structure that resembles the rest of the chromosomes, to be tightly packed to ensure almost complete transcriptional inactivation. For decades, this process has served as a powerful model to study the developmentally induced formation of heterochromatin. However, how could you revert it? Intriguingly, the reverse process can also be observed during development. X chromosome reactivation overturns the process of inactivation to completely restructure and reactivate the previously silenced chromosome. The reactivation of the X provides a unique system to study the developmentally induced formation of euchromatin. Nevertheless, whereas inactivation of the X has been studied in detail, much less is known about its reversal. We, therefore, aimed to describe for the first time, how the X chromosome is reactivated chromosome-wide, how the kinetics of structural and transcriptional changes go hand in hand to re-establish a whole chromosome to its former state.

# Table of Contents

# Introduction

## X Chromosome Dynamics *In Vivo* and *In Vitro*

To achieve gene dosage balance between males (XY) and females (XX), mammals transcriptionally inactivate one of the two X chromosomes in females during early embryonic development. However, this process is far from static and actually involves two rounds of both X chromosome inactivation (XCI) and X chromosome reactivation (XCR) *(Figure I1A)*. First, during pre-implantation development, imprinted XCI occurs, which from the 4-cell stage onwards, specifically inactivates the paternally inherited X (Huynh and Lee 2003) (Okamoto et al. 2004). Subsequently, imprinted XCI is erased by X-chromosome reactivation (XCR) in the inner cell mass (ICM) of mature blastocysts (Mak et al. 2004) (Okamoto et al. 2004), whereas imprinted XCI is maintained in the extraembryonic tissues of the placenta (Takagi and Sasaki 1975). XCR at this stage allows for random XCI to occur later in the differentiating postimplantation epiblast, where either the maternal or paternal X is randomly chosen for inactivation, a decision that is maintained thus throughout adult life (Lyon 1961) (Monk and Harper 1979). Finally, random XCI is erased in the germ cell lineage (Sugimoto and Abe 2007), to ensure the inheritance of an active X to the next generation. However, throughout the process, a subset of X-linked genes escapes X-inactivation and was therefore termed Escapees (Disteche 1999).
Importantly, these processes can be recapitulated *in vitro* as well *(Figure I1B).* Random XCI can be modeled using embryonic stem cell (ESC) differentiation (J. T. Lee et al. 1996) (Panning and Jaenisch 1996) (Penny et al. 1996) and XCR can be observed during induced pluripotent stem cells (iPSC) reprogramming (Maherali et al. 2007).

## X Chromosome Inactivation (XCI)

### The long non-coding RNA Xist Coats the Inactive X

The key player in X chromosome inactivation is the long non-coding RNA Xist (Borsani et al. 1991) (Brockdorff et al. 1991) (Borsani et al. 1991) (Penny et al. 1996) (Marahrens et al. 1997a). *Xist* is monoallelically upregulated from one X chromosome, starts coating the future inactive X *in cis* (A. Wutz and Jaenisch 2000) and thereby triggers a cascade of events that ultimately lock the X chromosome in a transcriptionally inactive state (Escamilla-Del-Arenal, da Rocha, and Heard 2011). Over its length of ~15 kb, *Xist* is characterized by unique repeat regions termed A-to-F repeats *(Figure I5A)*. Whereas the A-repeat region facilitates transcriptional silencing (Anton Wutz, Rasmussen, and Jaenisch 2002), other regions mediate the recruitment of a variety of chromatin remodeling factors (Anton Wutz, Rasmussen, and Jaenisch 2002) (Almeida et al. 2017).

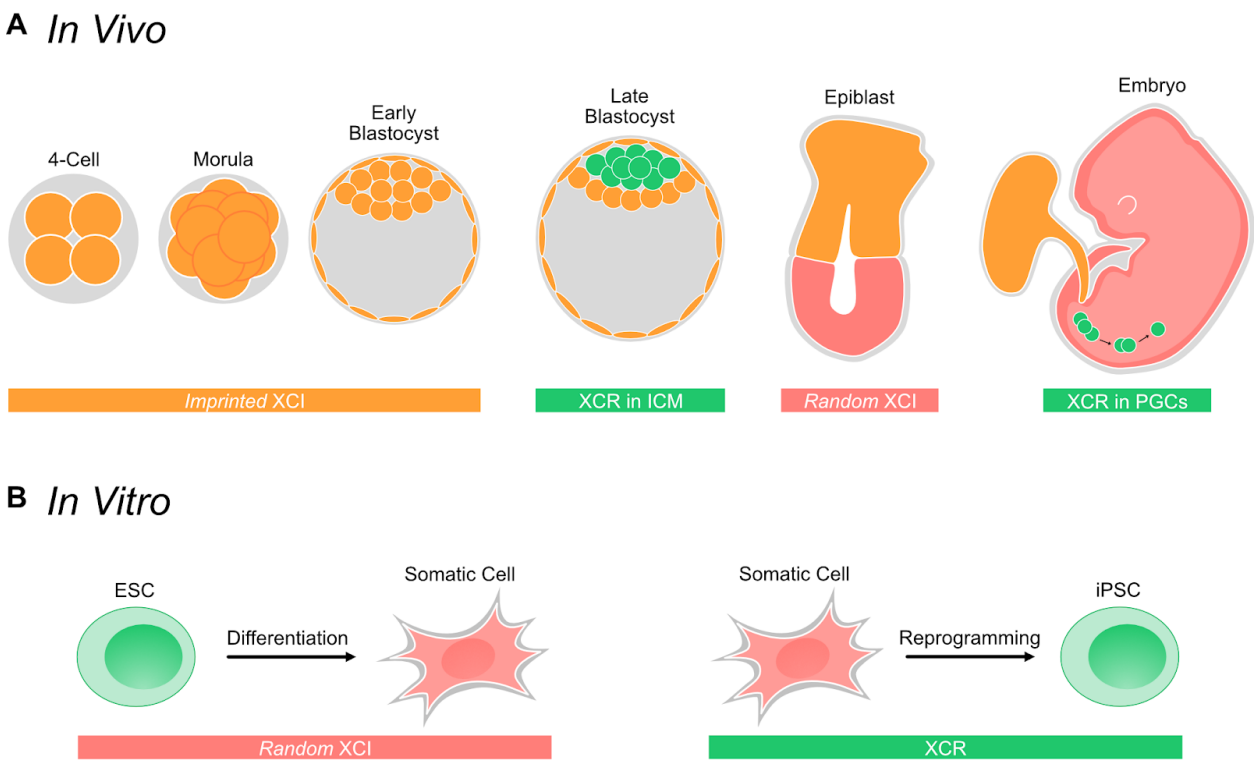



**Figure I1. Overview of X Chromosome Dynamics During Mouse Development**

***(A)*** *In vivo*, imprinted X chromosome inactivation (orange) occurs from the 4-cell stage onwards and is maintained in the extra-embryonic tissues. In the late blastocyst, cells of the inner cell mass, have undergone X chromosome reactivation (green), to later in the epiblast allow random XCI (red) to occur. Random XCI is then reversed in the germ cell lineage (green). **(B)** *In vitro*, random XCI occurs during the differentiation of embryonic stem cells. XCR can be modeled during the reprogramming of somatic cells to induced pluripotent stem cells (iPSC).

## Regulation of XCI Through the X Inactivation Center

The regulation of *Xist,* and ultimately the choice of which X to inactivate, is controlled via a single locus on the X chromosome, the X inactivation center (*Xic*) (Augui, Nora, and Heard 2011) *(Figure I2)*. The *Xic* in mice spans a region of 100kb-500kb (J. T. Lee et al. 1996), with the core region sufficient for silencing in *cis*, being only 100kb in size (J. T. Lee, Lu, and Han 1999). Strikingly, this core regions consists solely of non-protein-coding genes that code for the long non-coding RNAs (lncRNAs) *Xist*, *RepA (J. Zhao et al. 2008)*, *Tsix (J. T. Lee, Davidow, and Warshawsky 1999)* (Jeannie T. Lee and Lu 1999) (J. T. Lee 2000) (Sado et al. 2001), *Xite* (Jeannie T. Lee 2005) (Tian, Sun, and Lee 2010) and *Jpx (Tian, Sun, and Lee 2010)*. Outside of the core region, lie various other lncRNAs, such as *Ftx* (Chureau et al. 2011), *Tsx* (Simmler 1996), *Linx* (Nora et al. 2012), as well as one protein-coding gene, *Rnf12* (also known as Rlim) (Jonkers et al. 2009). As described further below, the *Xic* is spatially partitioned into two domains, TAD-D, harboring the promoter of *Tsix*, as well as, *Xite*, *Linx* and *Tsx*, and TAD-E, containing the promoter of *Xist*, as well as *Ftx*, *Jpx, RepA* and *Rnf12* (Nora et al. 2012), thereby

functionally separating cis-regulatory elements (van Bemmel et al. 2019). The arguably most well studied of these various lncRNAs besides *Xist* is *Tsix ((Jeannie T. Lee and Lu 1999)* (Jeannie T. Lee 2005) (Tian, Sun, and Lee 2010). In mice, *Tsix* is expressed antisense to *Xist* and overlaps with the majority of its gene body. *Tsix'* transcriptional pattern reflects its function as an antagonist of *Xist*, showing high expression in ESC that during differentiation remains only on the active X. However, whereas the overlapping transcriptional units of *Xist* and *Tsix* could suggest an active role of *Tsix* in preventing *Xist* expression during the initiation of XCI, its function rather seems to be, in conjunction with *Xite*, to prevent *Xist* expression from the active X, after the initial choice has already been made (Monkhorst et al. 2008) (Gayen et al. 2015). By contrast, *Jpx* has been shown to serve as a positive regulator of *Xist* (Tian, Sun, and Lee 2010). *Jpx* RNA is crucial to evict CTCF from *Xist*, which would otherwise block its upregulation (Sun et al. 2013). Furthermore, the tight relationship of the X chromosome state and pluripotency can be exemplified by the regulation of *Xist* through the *Xic* as well. The E3 ubiquitin ligase RNF12 is upregulated during differentiation and targets the pluripotency factor REX1 for degradation (Barakat et al. 2011) (Jonkers et al. 2009), which in ESCs binds the *Xist* promoter to inhibit its expression (Gontan et al. 2018). Moreover, other pluripotency factors have been shown to bind multiple binding hubs at the *Xic* as well, most prominently via the direct binding of core pluripotency factors OCT4, SOX2, NANOG, and PRDM14 to *Xist* intron 1 (Navarro et al. 2008). However, the functional consequence of their binding is debatable, as the deletion of *Xist* intron 1, did not affect its repression in ESCs or its silencing during iPSC reprogramming (Minkovsky et al. 2013). Taken together, multiple regulatory layers are in place at the *Xic* to ensure faithful regulation of *Xist*, to ultimately control XCI.

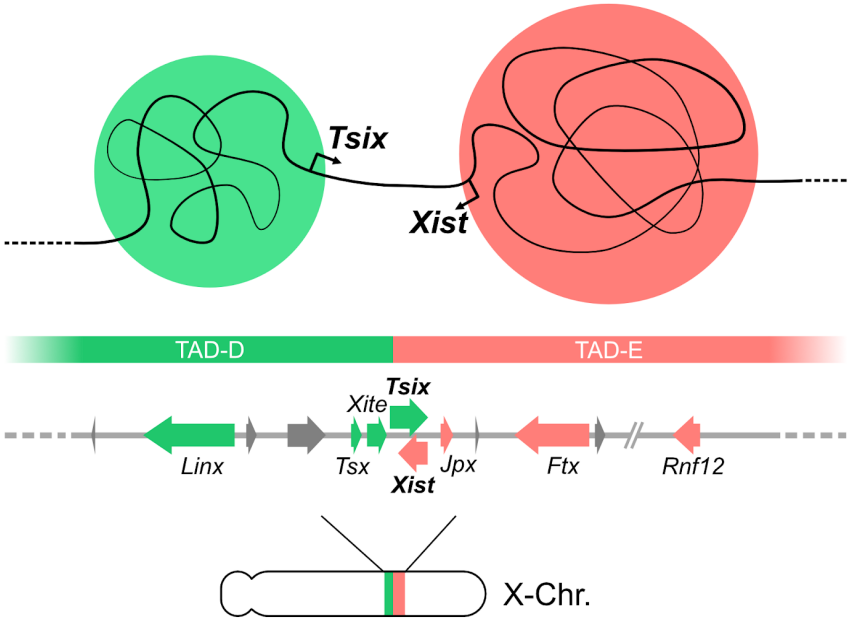



**Figure I2. X Inactivation Center**

The X inactivation center (Xic) is separated into two topologically associating domains (TAD), functionally separating *Xist* and *Tsix* regulators. TAD-D (green), harboring *Tsix*, *Xite*, *Tsx,* and *Linx*, and TAD-E (red) harboring *Xist*, *Jpx*, *Ftx,* and *Rnf12*.

## Heterochromatic Landscape of the Inactive X

X-chromosome inactivation (XCI) has served as a paradigm to study the formation of facultative heterochromatin during development. Multiple layers of repressive histone modifications are in place (Chadwick and Willard 2004) to synergistically ensure the transcriptional silencing of the inactive X (Csankovszki, Nagy, and Jaenisch 2001) (A. Minajigi et al. 2015), which ultimately define the unique heterochromatic landscape of the Xi *(Figure I3)*. The inactive X is marked by an accumulation of H3K27me3, deposited by the polycomb repressive complex 2 (PRC2), which makes it visually stand out as a nuclear spot in H3K27me3 immunostainings (J. Wang et al. 2001) (Plath 2003) (Silva et al. 2003). It further also harbors polycomb repressive complex 1 (PRC1) and its associated mark H2AK119ub (Napoles et al. 2004). Whereas polycomb-group proteins cover the majority of the Xi, unbound regions are interspersed with H3K9me3, which plays a role in the silencing of repeats (Rougeulle et al. 2004) (Keniry et al. 2016). Moreover, H4K20me1 marks the inactive X as well (Kohlmaier et al. 2004). After the initiation of XCI, macroH2A starts accumulating on the Xi (Mermoud et al. 1999) (Costanzi et al. 2000) (Nesterova et al. 2002) and silencing of the Xi is ultimately locked in via extensive CpG island DNA methylation (Gendrel et al. 2012).

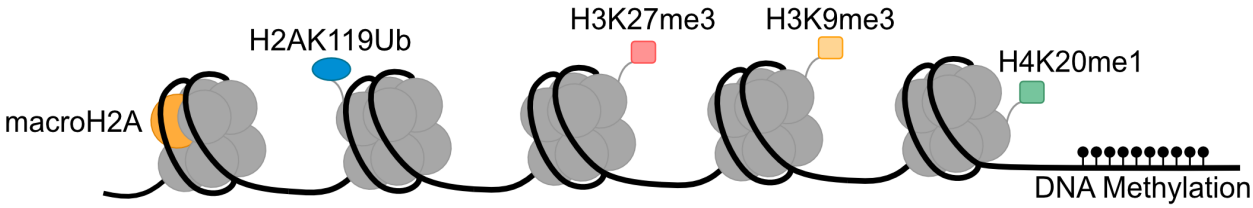



**Figure I3. Heterochromatic Landscape of the Inactive X**

The inactive X is prominently marked by several repressive DNA and histone modifications. Among others, the histone variant macroH2A. Ubiquitination of lysine 119 of histone H2A (H2AK119Ub) catalyzed by polycomb repressive complex 1. Trimethylation of lysine 27 of histone H3 (H3K27me3) catalyzed by polycomb repressive complex 2. Trimethylation of lysine 9 of histone H3 (H3K9me3) catalyzed by SETDB1. Monomethylation of lysine 20 of histone H4 (H3K9me3) catalyzed by SET8/PR-Set7.

## Spreading of Xist RNA

With a growing understanding of the regulation of *Xist* and the emergence of high-throughput sequencing techniques, it was starting to be possible to study in more detail the initiation of X chromosome inactivation. To elucidate how XCI proceeds, it was necessary to understand how Xist was able to broadly coat the X chromosome while being expressed from a single locus. Therefore, it was crucial to identify its initial binding sites and spreading mechanism. To shed light on this issue, two groups independently developed biochemical methods to assess the binding of Xist to chromatin (Engreitz et al. 2013) (Simon et al. 2013). Using capture hybridization analysis of RNA targets with deep sequencing (CHART-seq) during the *in vitro* differentiation of mouse ES cells, Simon et al. were able to show that Xist first bound gene-rich

regions and only later extended into gene-poor regions. However, when Xist was transiently stripped away using locked nucleic acids (LNA) in fully differentiated mouse embryonic fibroblasts (MEF), Xist reassociated simultaneously in early and late domains. Together, this showed that the mechanism of Xist spreading in the initiation phase is distinct from its maintenance state. However, why gene-rich regions were targeted first couldn't be explained.
Engreitz et al. used RNA Antisense Purification (RAP) in fully differentiated female mouse lung fibroblasts and showed enrichment of Xist across almost the entire X chromosome, only excluding regions that showed gene expression. To now gain insights into its initial spreading, expression of *Xist* was induced from its endogenous locus in male ESC cells. This revealed initial Xist spreading sites that were distal to its endogenous locus, arguing that its spreading didn't simply occur in a linear fashion from its site of transcription. Early sites neither showed enrichment for specific sequence motifs nor for any tested genomic annotations, including repeat elements. However, it was possible that Xist would spread first to sites in close spatial proximity in 3D. Analysis of genome-wide chromosome conformation capture data (Hi-C) (Lieberman-Aiden et al. 2009) showed a strong correlation between initial Xist binding sites and the frequency at which these sites contacted the *Xist* genomic locus. Strikingly, when the expression of *Xist* was induced from an ectopic locus 50 Mb proximal to the endogenous locus, the initial binding sites of Xist changed and now strongly correlated with proximity contacts of the ectopic site. Taken together, this revealed, that chromosome conformation plays a crucial role in determining the initial localization of the Xist RNA to gene-rich regions on the X chromosome *(Figure I4)*.

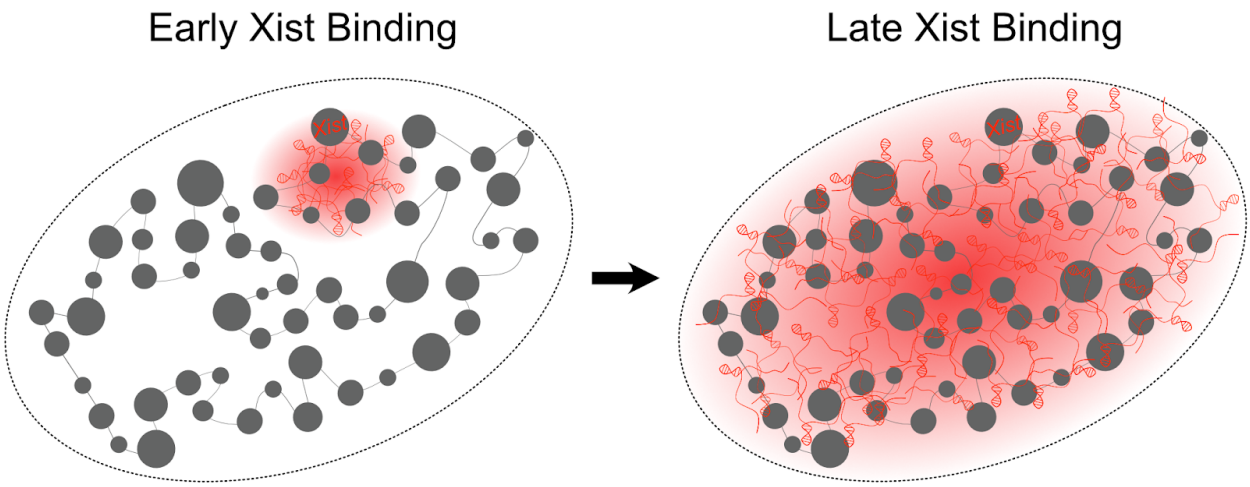



**Figure I4. Spreading of Xist RNA**

Xist RNA initially spreads to sites in close spatial proximity (Early Xist Binding) before ultimately covering the whole inactive X chromosome (Late Xist Binding).

## Xist Binding Partners Define Early Chromatin Changes

To gain mechanistic insight into how Xist facilitates XCI, several studies simultaneously aimed to systematically identify Xist binding partners (Mira-Bontenbal and Gribnau 2016). These were

either performed by pulldown of Xist followed by mass spectrometry (McHugh et al. 2015) (Chu et al. 2015) (A. Minajigi et al. 2015) or alternatively by gene knockdown (Moindrot et al. 2015) or knockout screens (Monfort et al. 2015). Whereas the number of identified binding partners differed between the studies, several hits were conserved between them.

**SPEN and HDAC3**

SPEN (also known as SHARP), a transcriptional repressor that activates the histone deacetylase HDAC3 (Guenther, Barak, and Lazar 2001) as well as the SPEN family member RBM15, were consistently identified in all five studies. Whereas Xist localization was unaffected upon knockdown of either SPEN or HDAC3, compromised transcriptional silencing, as well as the exclusion of RNA polymerase II (Pol II) from the inactive X, was observed (McHugh et al. 2015). Intriguingly, SPEN was shown to specifically interact with the Xist A-repeat region, which is required for its transcriptional silencing effect, but not for its localization or spreading (Anton Wutz, Rasmussen, and Jaenisch 2002). This suggests that the interaction of Xist with SPEN/SHARP is crucial to facilitate activation of HDAC3, thereby deacetylating histones and initiating transcriptional silencing on the X *(Figure I5B)*.

**RBM15/B and m$^6$A RNA methylation**

RBM15 and RBM15B, SPEN family members, make up the core subunits of the m$^6$A RNA methyltransferase complex (Ping et al. 2014). Similarly to SPEN, double knockdown of RBM15 and RBM15B prevented Xist-mediated gene silencing (Patil et al. 2016). This was facilitated by bridging Xist to METTL3, the major m$^6$A RNA methyltransferase (Bokar et al. 1994), and was shown to facilitate methylation of the 78 $^6$A residues, which are found over the entire length of *Xist*. m$^6$A modified Xist can further be specifically recognized by the YTH protein DC1. However, how DC1 is responsible for gene silencing is currently unknown *(Figure I5C)*.

**HnRNPU, hnRNPK, and Polycomb-group Proteins**

Furthermore, these studies confirmed the previously identified interaction of Xist with the matrix protein hnRNPU (also known as SAF-A) via its E-repeat (Hasegawa et al. 2010) and additionally identified an additional member of the hnRNP protein family, hnRNPK. Whereas hnRNPU is required for accumulation of Xist on the Xi via its binding to matrix-attachment regions (MAR) *(Figure I5D)*, hnRNPK was shown to be necessary for the accumulation of H2AK119ub and H3K27me3 on the Xi without affecting Xist RNA localization (Chu et al. 2015) *(Figure I5E)*.

Moreover, direct interactions of PRC2 with multiple regions of Xist have been suggested as well ((J. Zhao et al. 2008) (Maenner et al. 2010) (da Rocha et al. 2014) (Cifuentes-Rojas et al. 2014). However, other reports also claimed that PRC2 does not interact with Xist directly (Cerase et al. 2014). Taking into account the promiscuous RNA binding activity of PRC2, binding specificity with Xist still needs to be demonstrated *in vivo* (Davidovich et al. 2013) (Davidovich et al. 2015). This issue especially becomes apparent when taking reports into account, that suggest a different recruitment model of PRC2 to the Xi. Whereas in the canonical model of polycomb recruitment, PRC2 is deposited first and subsequently recruits PRC1 via the recognition of H3K27me3 by the canonical PRC1 subunit CBX (Min, Zhang, and Xu 2003)

(Bernstein et al. 2006), a non-canonical recruitment model seems to exist during XCI. Here, a non-canonical PRC1 complex, comprising of PCGF3/5, signals the recruitment of both PRC1 and PRC2 (Almeida et al. 2017). Binding of PRC2 is facilitated by its associated factor JARID2, which is able to recognize the PRC1 deposited mark H2AK119ub (Cooper et al. 2016). Moreover, the initial binding of PRC1 is promoted through interaction with hnRNPK, which can directly bind PCGF3/5 (Pintacuda et al. 2017) as well as Xist via its B-repeat (Chu et al. 2015), thereby bridging the two and facilitating PRC1 recruitment. Taken together, PRC1 recruitment to the Xi is facilitated through its interaction with hnRNPK and seems to precede PRC2 recruitment.

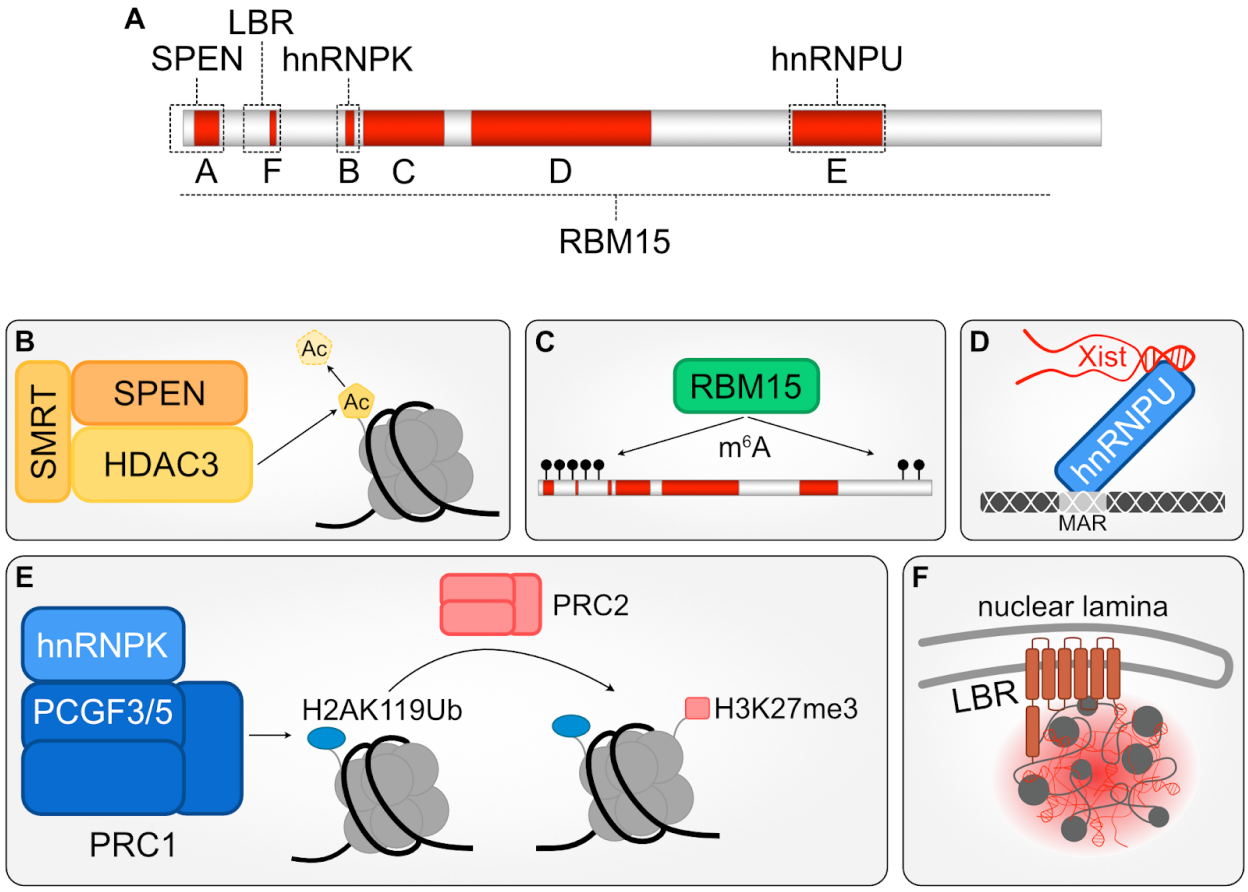



**Figure I5. Functions of Xist RNA silencing partners for XCI**

***(A)*** *Xist* harbors 6 repeat regions, A, F, B, C, D, and E. **(B)** SPEN binds repeat A of Xist via SMRT and bridges it to HDAC3, to promote deacetylation of histones. **(C)** RBM15 and RBM15B binding to Xist is spread along the whole RNA and catalyzes the methylation of 78 $^{6}$A residues along Xist. **(D)** HnRNPU binds Xist' E-repeat and is required for the accumulation of Xist on the Xi via its binding to matrix-attachment regions (MAR). **(E)** HnRNPK binds repeat B of Xist and via binding to PCGF3/5 bridges it polycomb repressive complex 1 (PRC1), which catalyzes H2AK119Ub. H2AK119Ub is recognized by polycomb repressive complex 2 (PRC2), which catalyzes H3K27me3. **(F)** LBR binds Xist' repeat F and thereby tethers the Xi to the nuclear lamina.

**Lamina Association**

Finally, it was shown that Xist interacts specifically with the Lamin B Receptor (LBR) (McHugh et al. 2015) (A. Minajigi et al. 2015), a transmembrane protein that is required for anchoring chromatin to the nuclear lamina (Gruenbaum et al. 2005). As XCI leads to the recruitment of the Xi to the nuclear lamina, it was an intriguing possibility that this is facilitated by the direct interaction of LBR with Xist. Indeed, upon knockdown of LBR, as well as mutation of an LBR binding site of *Xist*, recruitment of the Xi to the nuclear lamina was disrupted (C.-K. Chen et al. 2017). Moreover, gene silencing couldn't spread to actively transcribed genes, suggesting that the nuclear lamina interaction is necessary to shape the three-dimensional nuclear structure of the inactive X to promote chromosome-wide silencing *(Figure I5F)*.

However, a systematic analysis of these various factors found that while SPEN was critical for the silencing of the majority of X-linked genes, interactions with LBR and RBM15 only made minor contributions (Nesterova et al. 2019). Taken together, Xist binds multiple proteins, whose distinct functions show that X chromosomal silencing, Xist recruitment, and spreading are mechanistically segregated and reveals that most histone modifiers are not directly recruited by Xist RNA, but rather, through its binding partners.

**Firre**

Notably, Xist might not be the only lncRNA involved in facilitating chromatin changes on the X. The lncRNA *Firre* is expressed from the active X (Fang et al. 2019) and was shown to interact, similar to Xist, with the nuclear-matrix protein hnRNPU (Hacisuleyman et al. 2014). This interaction appeared to be critical for the anchoring of the Xi to the nucleolus, as well as H3K27me3 deposition on the Xi (F. Yang et al. 2015) (Fang et al. 2019).

## Spreading of Repressive Histone Modifications

XCI initiates with the binding of Xist to defined early entry sites (Engreitz et al. 2013) (Simon et al. 2013), followed by the recruitment of multiple repressive histone modifiers. However, their timing, as well as their initial binding sites, are not merely randomly defined. Addressing the binding of PRC2 during XCI, Pinter and colleagues were able to show that it initially localized to ~150 strong sites, coinciding with bivalent domains and CpG island (Pinter et al. 2012), that highly correlate with early Xist binding sites (Simon et al. 2013). With proceeding XCI, PRC2 then spreads locally from its initial sites into intergenic regions. However, whereas this would propose a uni-directional relationship, where PRC2 simply follows Xist, a recent study rather points towards an interdependent spreading mechanism, as ablation of PRC1 or PRC2 impaired Xist spreading, leading to a failure of de novo Xi silencing (Colognori et al. 2019).

To gain a more detailed view of the spatiotemporal dynamics of early chromatin changes in XCI, Żylicz and colleagues, performed native ChIP-seq for seven histone modifications in a synchronized *Xist*-induction system (Żylicz et al. 2019). This revealed that loss of H3K27ac marks one of the first events following Xist RNA accumulation. The rapid deacetylation is achieved via already pre-loaded HDAC3 on putative enhancers on the X chromosome to ensure efficient transcriptional silencing, as in its absence, silencing is delayed, however not fully

prevented. Concurrently, PRC1-dependent H2AK119ub also rapidly accumulates at intergenic regions that are pre-marked Polycomb (PcG) marks and lie in close spatial proximity to Xist RNA entry sites (Engreitz et al. 2013). However, the deposition of H3K27me3 by PRC2 is delayed in comparison to H2AK119ub, favoring a non-canonical PcG protein recruitment model as mentioned above. Nevertheless, while the accumulation of H2AK119ub marks a very early event during XCI, it might not initiate gene repression, as analysis of HDAC3 mutants revealed that its spreading required, at least partially, prior histone deacetylation and transcriptional silencing. However, as both processes occur around the same time, PRC1 might be necessary to facilitate efficient XCI.

## Kinetics of Gene Silencing

Nevertheless, while gene-rich regions seemed to be targeted first by Xist, followed by the stepwise accumulation of repressive histone modifications, the kinetics of gene silencing on the X remained to be defined. However, a prerequisite for the analysis of X-chromosomal expression data is a hybrid mouse strain background. As the two X chromosomes would otherwise be indistinguishable by sequencing-based assays, crosses of genetically distant mouse strains, such as *Mus musculus musculus* and *Mus musculus castaneus*, are utilized. The high rate of polymorphisms between them allows for allele-specific quantification using mostly single-nucleotide polymorphisms (SNP).

***In Vivo***

Analysis of allele-specific single-cell RNA-sequencing (scRNA-seq) data of early mouse embryos showed gradual imprinted X inactivation beyond the 4-cell stage until the early blastocyst (Q. Deng et al. 2014). However, the spreading of silencing wasn't a simple function of the distance to the X-inactivation center in two dimensions, as some early silencing regions were positioned far away from it. A more detailed analysis of the process revealed that already at the 8-cell stage, few genes completed XCI (Borensztein, Syx, et al. 2017). Nevertheless, the majority of genes only completed silencing at the blastocyst stage. Moreover, whereas early and intermediate silencing genes did tend to lie closer to the *Xic*, some escapees did lie close to it as well. Furthermore, substantiating observations on early Xist RNA binding sites (Engreitz et al. 2013), it was found that genes located within or close to these showed the earliest silencing. Of important note, significant differences were observed when comparing different mouse crosses, pointing to a high degree of variability in silencing kinetics.

***In Vitro***

Whereas the above-mentioned studies using scRNA-seq allowed to take into account the heterogeneity of the inactivation process, they failed to paint a complete picture of the silencing along the X. Due to technical limitations such as low capture efficiency and sequencing coverage of scRNA-seq, only 125 of 580 expressed X-linked genes have been analyzed with confidence by Borensztein, Syx, and colleagues. *In vitro* systems are therefore useful tools to circumvent this problem by being able to use a higher number of cells that combined with more conventional RNA-seq techniques allow for increased coverage of transcripts.

During a time course of ESC to embryoid body differentiation, X-linked genes showed a gradual downregulation from the future inactive X (259 genes on the X chromosome analyzed out of a total of 590 genes with expression > 0.5 RPKM) (Marks et al. 2015). After 8 days of differentiation, the active X showed on average a four-fold higher expression than the inactive one. Curiously, lowly expressed genes showed faster XCI dynamics. Clustering analysis of the time course revealed four different classes of genes that shared similar dynamics. “Early” ones, that were silenced after 2 days and accounted for 46% of the analyzed genes. “Intermediate”, which mainly showed silencing between 4 and 8 days. “Late” ones that only showed mild downregulation after 8 days and “not silenced” that included known escapee genes. Confirming previous observations, “late” silencing genes were significantly higher expressed than genes in the other clusters. Moreover, when plotting the four clusters over the linear X chromosome, it became apparent that genes of the “early” cluster on average tended to lie closer to the *Xic* than the other clusters. Following this trend, “intermediate” genes lied closer to the *Xic* than “late” and “not silenced” genes.

**Predictors of Silencing Kinetics**

Building on these results, it remained to be answered what predisposed genes on the X chromosome to be silenced early, late or not at all. To be able to predict their silencing behavior, Sousa and colleagues first measured silencing dynamics using allele-specific Precision nuclear Run-On sequencing (PRO-seq) during a time course of ectopic *Xist* induction in mESC (Sousa et al. 2019). This system allows for more direct quantification of silencing and circumvents the issue of heterogeneity in XCI states of *in vitro* differentiation systems (G. Chen et al. 2016). To determine silencing factors in an unbiased manner, a Random Forest machine learning approach was used to predict the silencing behavior of X-linked genes from a total of 77 epigenetic and genomic features, including high-throughput ChIP-seq and Bisulfite-seq datasets, as well as features such as gene density, 3D chromatin interaction frequencies and the linear distance to other genomic features. This revealed, that in agreement with a previous study, the most important feature determining early silencing, was a genes' proximity to the *Xist* locus (Marks et al. 2015). Moreover, enrichment for polycomb-group proteins was associated with early silencing as well. However, contrasting previous studies, rapidly silenced genes were also found to lie in regions of low gene density, and in close proximity to LINE elements, a discrepancy that will need to be addressed (Engreitz et al. 2013) (Simon et al. 2013). Finally, arguing for a role of 3D chromatin organization in Xist mediated gene silencing, early silencing was observed to preferentially occur at genes that are close to lamina-associated domains (LADs), which generally contain repressed genes, while genes close to the boundaries of topologically associating domains (TADs) tended to be silenced late *(Figure I6)*.

To summarize, silencing kinetics seem to vary across the X chromosome, with the most important determinant, being close linear and 3D proximity to the *Xist* locus.

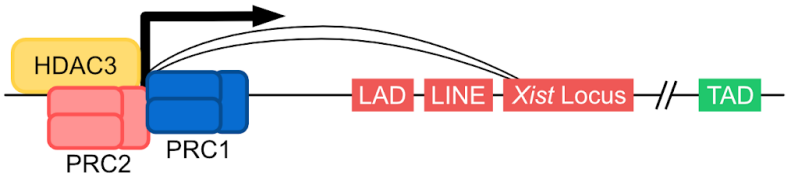



**Figure I6. Predictors of Early Gene Silencing**
Promoters of genes predicted to become silenced early during XCI are bound by HDAC3, as well as by polycomb repressive complex 1 (PRC1) and 2 (PRC2). Early silencing genes tend to lie in close proximity to lamina-associated domains (LAD), long interspersed nuclear elements (LINE) and the *Xist* locus. Moreover, 3D proximity to the *Xist* locus is predictive as well. Moreover, genes further away from TAD boundaries tend to become silenced early.

# X-Chromosome Reactivation (XCR)

## Linking Pluripotency to Two Active X Chromosomes

The study of X-chromosome reactivation raises the question, how the activity of two X chromosomes instead of one influences cells. One of the earliest findings in mESCs showed that DNA methylation is globally reduced in XX compared to XY and X0 cells (Zvetkova et al. 2005). This hypomethylation state is associated with a reduction in protein levels of the *de novo* DNA methyltransferases DNMT3A/B. A more detailed analysis in a large variety of pluripotent cells revealed that the ratio of X chromosomes to autosomes directly correlates with global methylation levels, rather than the absolute number of X chromosomes inside a pluripotent cell (Choi et al. 2017). Notably, expression levels of known DNA methylation regulators were not significantly different between male and female cells. However, at the protein level, female cells showed markedly reduced levels of DNMT1 and DNMT3A/B, as well as elevated levels of TET2, compared to male cell lines. Mechanistically, this seems to be regulated by the X-linked phosphatase *Dusp9*, which dosage-dependently regulates MAPK/GSK/AKT pathways, pluripotency, and epigenetic factors. Overexpression of DUSP9 in male cells led to hypomethylation, whereas the heterozygous deletion of *Dusp9* in female cells increased methylation levels. However, hypermethylation of XX ESCs was not sufficient to cause a transition to a male-like transcriptional signature, arguing for the existence of other pathways and X-linked genes involved in mediating the effects of X-dosage (Song et al. 2019). Adding to this, it is important to note, that X dosage mediated DNA methylation differences between male and female cells, seem to be restricted to *in vitro* cultured cells, as the inner cell mass of male and female embryos were shown to be hypomethylated similarly (Choi et al. 2017). This further suggests that female ESCs might be molecularly closer to the ground state of pluripotency than male ESCs.

Moreover, these *in vitro* differences manifest themselves in the cells' ability to differentiate. During the differentiation into epiblast-like cells (EpiLC), female XX ESCs were shown to be delayed and lag behind their male XY counterparts by roughly one day (Schulz et al. 2014). This delay was accompanied by sustained expression of pluripotency-associated factors. Importantly, female XO cells showed a behavior similar to male cells, further adding to the notion that the X chromosome dosage is directly responsible for this effect. However, when cells were forced to undergo XCI by inducing expression of *Xist* using a doxycycline-inducible promoter, the delay in differentiation was eliminated, and they were able to differentiate at speeds comparable to male cells. Finally, to circumvent the asynchronicity of the differentiation process, single-cell qRT-PCR was performed and was able to directly show that expression of pluripotency-factors and *Xist* is reciprocal.

## XCR in the Blastocyst

*In vivo*, X-chromosome reactivation was first described in the blastocyst *(Figure I1A)*. As outlined previously, early in development, cells first undergo paternal X-chromosome

inactivation, where the X gets prominently marked by Xist and members of the polycomb repressive complex 2 (PRC2). However, in the inner cell mass (ICM) of mature blastocysts, a rapid loss of Xist and of the PRC2 complex member EED, accompanied by a gradual disappearance of its associated mark H3K27me3, were observed on the X (Okamoto et al. 2004) (Mak et al. 2004). Moreover, linking X reactivation to pluripotency, loss of EED from the inactive X was specifically observed in *Nanog* expressing cells, that demarcate the nascent epiblast (Silva et al. 2009). Mechanistically, this link could be explained by the cooperative binding of key pluripotency factors OCT4, SOX2, NANOG, and PRDM14 within intron 1 of *Xist,* to repress its expression in mESCs (Navarro et al. 2008). Indeed, both *Nanog* knockout (Silva et al. 2009) and *Prdm14* knockout (Payer et al. 2013) blastocysts failed to erase the X-specific EED and H3K27me3 accumulation, respectively.

However, reactivation of X-linked genes did not necessarily seem to require the loss of Xist coating, as RNA FISH analysis of a small set of genes in the ICM showed that X-linked gene transcription could already occur before the loss of Xist (Williams et al. 2011). This finding was further supported by allele-specific scRNA-seq data, which showed that 26 out of 116 analyzed X-linked genes underwent reactivation independently of Xist RNA and H3K27me3 loss, already before lineage segregation at E3.5 (Borensztein, Okamoto, et al. 2017). Nevertheless, the majority of X-linked genes only reactivated later when Xist enrichment was lost and pluripotency markers such as *Nanog* were expressed.

However, it remained to be answered, if XCR was a mere reversal of the inactivation process. Comparing the kinetics of XCR (Borensztein, Okamoto, et al. 2017) to XCI (Borensztein, Syx, et al. 2017) revealed that the former did not simply mirror the latter. Further arguing for different mechanisms of XCI and XCR, only a slight tendency for late and very late reactivating genes to lie in close proximity to *Xist* was observed, with some early reactivating genes even lying in close proximity to *Xist*. Therefore, X reactivation timings couldn't be predicted from the distance of a gene to *Xist*. Furthermore, neither Xist entry sites (Engreitz et al. 2013), nor gene expression levels were found to be predictors of early or late reactivation.

However, when analyzing the enrichment of histone modifications, early reactivating genes were found to be enriched for H3K4me3, whereas late ones showed enrichment for H3K27me3. The latter was further proven to be directly linked to late reactivating genes, as the homozygous knockout of the H3K27me3 demethylase UTX, significantly impaired the efficiency of X-linked gene reactivation for late-reactivated genes. Nevertheless, this couldn’t explain the different timing of X-linked gene reactivation, especially not why some genes were able to reactivate even before Xist downregulation occurred. Due to the strong link of XCR and pluripotency, as explained above, the finding that X chromosome reactivation was correlated with the expression of pluripotency factors was not surprising. Yet, half of the X-linked genes, irrespective of their reactivation timing, presented at least one binding site for core pluripotency factors, ruling them out as singular initiators of early reactivation. Apart from pluripotency factors, *Myc* expression was also found to be associated with reactivation, albeit to a lesser degree. However, in contrast to pluripotency factors, early reactivating genes, as well as escapees, showed significant enrichment for Myc-factor binding sites compared to late reactivating genes, suggesting a possible role for MYC in the early reactivation of X-linked genes.

Therefore, whereas the correlation of pluripotency factors and X reactivation is clear, the mechanistic link still isn't. Nevertheless, a study by Payer and colleagues was able to show that the homozygous deletion of the pluripotency factor *Prdm14*, led to a reduction in XCR, visualized by retention of H3K27me3 on the Xi in E4.5 epiblast cells (Payer et al. 2013). This effect was further observed to a similar extent upon homozygous deletion of *Tsix*. However, double mutants of *Prdm14* and *Tsix* did not show a significantly stronger phenotype, arguing that they act through a common pathway during X reactivation

Taken together, XCR in the blastocyst occurs rapidly, however, it is not a mere reversal of the inactivation process, with distinct timings of gene reactivation along the X, possibly being controlled through pluripotency-factor expression, binding of Myc-factors and the gene-specific enrichment of histone modifications.

## XCR in Primordial Germ Cells

*In vivo*, XCR was not only described in the blastocyst but as well later in development, during the migration of primordial germ cells (PGC) *(Figure I1A)*. Already at E7.75, it was observed that some PGCs showed *Xist* downregulation, as well as biallelic expression of some X-linked genes (Sugimoto and Abe 2007). However, in contrast to the blastocyst, reactivation seems to occur gradually in PGCs, as even at 14.5, not a single PGC was detected that had fully undergone XCR. This gradual reactivation is concomitant with the PRDM14-dependent progressive removal of H3K27me3 from the inactive X-chromosome along the PGC migration path (Mallol, Guirola, and Payer 2019).

Intriguingly, reactivation in PGCs seems to depend on signals from the surrounding somatic cells (Chuva de Sousa Lopes et al. 2008). Culturing X inactive PGCs in a transwell system with XX somatic cells on the bottom was sufficient to trigger XCR, observed by the upregulation of an X-linked GFP transgene. Strikingly, XY somatic cells were not able to induce reactivation.

Furthermore, XCR in PGCs has been recapitulated *in vitro* as well. Reflecting the developmental progression, primordial germ cell-like cells were generated from embryonic stem cells, through epiblast-like cells and aggregation with somatic gonadal cells, reaching a pre-meiotic stage similar to E12.5 (Hayashi et al. 2012). At this stage, cells had lost the H3K27me3 spot demarcating the former inactive X, indicative of X-chromosome reactivation.

## XCR during iPSC Reprogramming

The advent of induced pluripotent stem cells (iPSC) (Takahashi and Yamanaka 2006), reprogramming of somatic cells into pluripotent ESC-like cells by the overexpression of the four factors *Oct4*, *Sox2*, *c-Myc*, and *Klf4*, held the promise for a new kind of regenerative medicine. However, whereas the iPSC revolution might still be years away from truly changing medicine, iPSC did revolutionize biological research already (Scudellari 2016). iPS cells have proven to be a valuable tool for disease modeling, and moreover, provided a system to dissect the process of cell fate decision making, how transcription factors and chromosome architecture guide cellular remodeling (Stadhouders, Filion, and Graf 2019). Along these lines, it enabled researchers to examine the reactivation of the X chromosome *in vitro*.

**Reprogramming Stages and Timing of XCR**

Shortly after iPSC had been first described, it was shown that female mouse embryonic fibroblasts (MEF), harboring one inactive X, had successfully reactivated it upon conversion into iPSC, as shown by biallelic expression of *Tsix* and *Pgk1*, as well as the loss of Xist coating (Maherali et al. 2007). Moreover, it became evident that XCR marks a late event of the reprogramming process, with upregulation of the pluripotency surface marker SSEA-1 preceding it, and coinciding with the upregulation of the pluripotency transcription factor *Sox2* (Stadtfeld et al. 2008). However, the sequential order of events, how the multiple layers of repressive epigenetic features are removed from the inactive X, remained to be shown. A study by Pasque and colleagues set out to address this question, by assessing changes in X-linked chromatin marks and the associated changes in pluripotency factor expression by microscopy (Pasque et al. 2014). First evaluating H3K27me3, a hallmark of the inactive X deposited by EZH2, it became apparent that it was exclusively lost in cells already expressing the late pluripotency marker NANOG. Moreover, the first NANOG+ cells to appear in culture, still exhibited Xist RNA coating, indicating that late pluripotency activation is a prerequisite for Xist downregulation. Subsequently, only after Xist coating of the inactive X was lost, biallelic expression of X-linked genes, as shown by RNA FISH of 4 genes, was restored. Furthermore, mirroring the XCI events *in vivo*, monoallelic expression of the *Xist* regulator *Tsix*, was first detected in NANOG+ cells from the active X, and only later switched to biallelic expression during the loss of Xist coating. Two of the latest events of XCI during differentiation are the incorporation of macroH2A1 (Mermoud et al. 1999) and the establishment of DNA methylation (Gendrel et al. 2012). Conversely, during reprogramming, the loss of macroH2A.1 occurred after the loss of EZH2 and therefore doesn't represent the reverse sequence X inactivation dynamics. Loss of DNA methylation during reprogramming marked one of the last events of XCR. Using reduced representation bisulfite sequencing, it was observed that even in SSEA1+ reprogramming intermediates, CpG islands across the X chromosome still showed DNA methylation levels comparable to MEFs, whereas promoters of pluripotency genes had mostly lost it already. Moreover, assessing the levels of 5-hydroxymethylcytosine (5hmC), which marks a prominent intermediate step during DNA demethylation, it became apparent that while globally 5hmC levels gradually increased, it was excluded from the former Xi until late stages of reprogramming.

Taken together, this detailed analysis of XCR during reprogramming reveals a defined sequence of chromatin changes, that however doesn't necessarily reflect a mere reversal of the inactivation process. Moreover, as shown *in vivo* in the blastocyst, reactivation specifically initiates in cells that underwent upregulation of pluripotency genes, with the subsequent downregulation of *Xist*, followed by biallelic expression of X-linked genes. Yet, the exact timing of pluripotency activation and *Xist* downregulation is still controversial, as single-cell RNA-Seq analysis during reprogramming revealed a different order of events, with *Xist* downregulation preceding late pluripotency activation (Schiebinger et al. 2019) *(Figure I7)*.

**Regulation of *Xist***

Successful XCR in female cells was shown to be dependent on the downregulation of *Xist*. Constitutive expression of *Xist* during iPSC formation resulted in a decrease of XCR in

NANOG+ cells, detected by an approximately 50% reduction in biallelic expression of the X-linked gene *Atrx* (Pasque et al. 2014). Consequently, knockdown of *Xist* late during reprogramming improved reprogramming efficiency by promoting the pre-iPSC to iPSC transition (Q. Chen et al. 2014). However, knockdown of *Xist* early during reprogramming had the opposite effect and decreased reprogramming efficiency by interfering with the mesenchymal-to-epithelial transition (MET), showing the dualistic nature of *Xist* and arguing for a cautious evaluation of results based on constitutive ablation of factors during reprogramming. Moreover, a direct link between *Xist* and pluripotency factors was previously proposed via the direct binding of core pluripotency factors to *Xist* intron 1 (Navarro et al. 2008). However, the deletion of intron 1, did not Interfere, among others, with the loss of Xist RNA coating upon reprogramming of MEFs to iPSCs (Minkovsky et al. 2013). This argues for a more indirect *Xist* repression mechanism, possibly via the silencing of the *Xist* activator RNF12 (Gontan, Achame, Demmers, Barakat, Rentmeester, van IJcken, Anton Grootegoed, et al. 2012) (Payer et al. 2013).

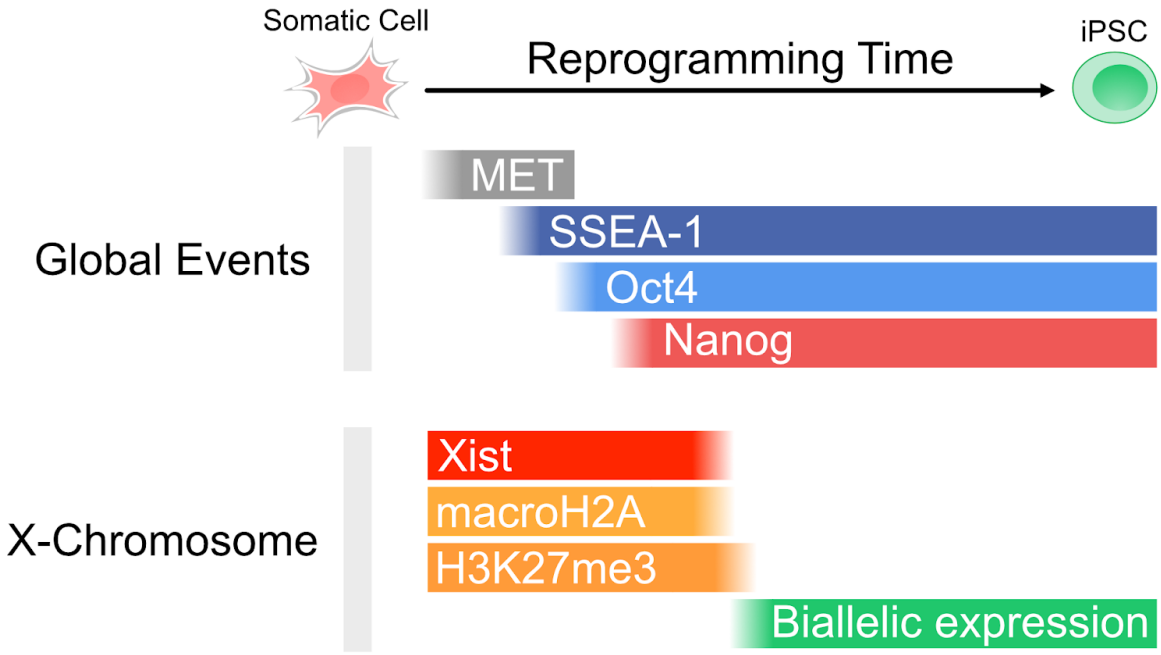



**Figure I7. Stages of iPSC Reprogramming**

During reprogramming, cells first undergo mesenchymal-to-epithelial transition (MET), followed by upregulation of the cell surface marker SSEA-1, and subsequent expression of the pluripotency markers *Oct4* and then *Nanog*. When the expression of Nanog is initiated, cells start down-regulating *Xist*, evicting macroH2A from the Xi, and with a slight delay, removing H3K27me3. Concomitantly, biallelic expression of X-linked genes is beginning to be restored.

**DNA Demethylation**

As mentioned previously, DNA demethylation marks one of the last steps of iPSC reprogramming as well as XCR. Strikingly, double knockout of the DNA demethylases Tet1 and Tet2, which catalyze the conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine

(5hmC), while preventing induction of 5hmC during reprogramming, did not interfere with the formation of pluripotent colonies (Pasque et al. 2014). Nevertheless, while the authors claim that Tet1, Tet2, and global 5hmC are dispensable for XCR, it is important to point out that X-linked gene reactivation was only tested by RNA FISH for a single gene, *Atrx*. Furthermore, in contrast, a study investigating the dynamics of DNA methylation and hydroxymethylation in a highly efficient reprogramming system revealed a requirement for Tet2, observed by a dramatic reduction in pluripotent colonies formed (Sardina et al. 2018). It is, therefore, possible that in the context of conventional MEF to iPSC systems as deployed by Pasque and colleagues, which are usually slow and inefficient, passive DNA demethylation is sufficient, whereas fast and efficient systems might need active DNA demethylation by TET enzymes. Therefore, DNA demethylation might play a context-dependent role in X-chromosome reactivation.

**Kinetics of Gene Reactivation**

Finally, in contrast to X inactivation (Borensztein, Syx, et al. 2017) (Marks et al. 2015), the kinetics of X-linked gene reactivation during iPSC reprogramming, until very recently, had not been described yet. A study by Janiszewski and colleagues aimed to address this question (Janiszewski et al. 2019). The authors deployed a conventional MEF to iPSC system with highly polymorphic MEF from a *mus m. musculus* and *mus m. castaneus* cross, to allow for allele-specific quantification of transcription dynamics. Furthermore, to isolate reprogramming intermediates, cells were FACS-sorted for expression of the early reprogramming marker SSEA-1. Cells further harbored an X-GFP transgene on the inactive X (Pasque et al. 2011), to allow visualization of X-chromosome activation. This showed a gradual increase in the number of X-GFP+ cells in the SSEA1+ population over a number of 5 days. Time-resolved maps of X-linked allelic expression ratios revealed that genes along the X reactivated with highly different kinetics, with very late reactivating genes being delayed up to 5 days, compared to early ones. This is in contrast to the reactivation kinetics observed *in vivo* in the epiblast, where XCR is completed in one day (Borensztein, Syx, et al. 2017). Moreover, early reactivation preceded the activation of late pluripotency genes as well as the full downregulation of *Xist*, a stark contrast to previous studies, which observed XCR only after both were completed (Pasque et al. 2014) (Schiebinger et al. 2019). Correlation with various genetic and epigenetic features showed that genes of early and intermediate reactivation timings were significantly closer to escapees compared to late ones. Moreover, binding of the 4 Yamanaka factors, OCT4, SOX2, KLF4, and c-MYC was significantly enriched in early, intermediate and late, compared to very late reactivating genes, however, no difference was observed between these first three categories.

During reprogramming, several populations of cells are often superposed on a single day. Therefore, especially in inefficient reprogramming systems where the majority of cells don't undergo full reprogramming, careful attention has to be given to the isolation of pure cell populations. Unfortunately, cell populations isolated in this study occur to be highly mixed. First, whereas SSEA-1 is a suitable marker to isolate cells at very early stages of reprogramming, it fails to distinguish later stages (Schwarz et al. 2018). Second, SSEA1+ populations and X-GFP+ populations were not isolated separately, therefore SSEA1+ cells showed a gradual enrichment of X-GFP+ cells, possibly masking early reactivation events in X-GFP+ cells. These

shortcomings in the experimental design of the study, therefore argue for a careful interpretation of the results, and observations will have to be validated in optimized experimental systems. However, slow reactivation kinetics compared to *in vivo* might as well be explained by the increased difficulty to revert random XCI, which includes macroH2A deposition and increased DNA methylation on the Xi, compared to imprinted XCI. Nonetheless, this would conversely suggest that inefficient reprogramming systems are an insufficient proxy for XCR in the inner cell mass.

To summarize, as observed in *vivo*, XCR strongly correlates with the activation of pluripotency genes, as well as the downregulation of *Xist* during iPSC reprogramming. However, it remains to be answered, if the reactivation kinetics of X-linked genes mirror those observed in the epiblast, and how much differences in the experimental design contribute to these.

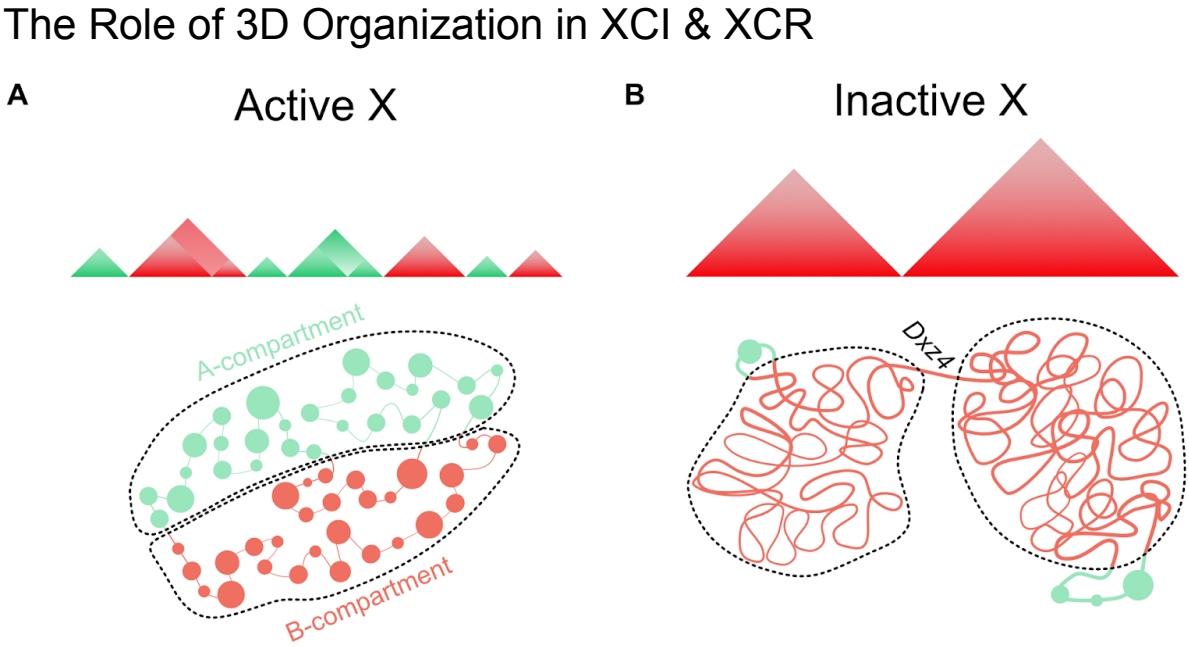



**Figure I8. Distinct Topology of the Inactive X versus the Active X**

Typical Hi-C profiles are depicted on the top and a schematic representation of genome organization is shown at the bottom. **(A)** 3D organization of the active X shows the typical TAD organization into active (A-compartment, green) and repressive (B-compartment, red) compartments. **(B)** The inactive X is organized into two large repressive mega-domains, divided by the *Dxz4* locus. Few TAD-like structures with escapees loop out from the repressive domain.

## Chromosome Conformation Capture

Eukaryotic interphase chromosomes are not randomly distributed in the nucleus, but instead were shown to be confined to distinct spaces, so-called chromosome territories, revealing, among others, separate territories for the active and the inactive X (Cremer and Cremer 2010). Moreover, the introduction of chromosome conformation capture (3C) technologies allowed to resolve in further detail their structural organization (Grob and Cavalli 2018). Briefly, 3C is based on formaldehyde crosslinking of chromatin, followed by enzymatic or physical chromatin fragmentation and subsequent proximity ligation (Dekker 2002). The resulting DNA hybrids are indicative of three-dimensional chromatin contacts. However, 3C and their adaptations 4C (Z. Zhao et al. 2006) (Simonis et al. 2006) and 5C (Dostie et al. 2006) are limited to a set of target loci, not allowing unbiased genome-wide analysis. Therefore, the introduction of Hi-C, which now allowed to assess chromosome conformation genome-wide, revolutionized and jump-started the field (Lieberman-Aiden et al. 2009). Hi-C did not only confirm the existence of chromosome territories but it, more importantly, revealed that chromatin is spatially segregated into two compartments: A, corresponding to open chromatin and high mRNA expression and B, corresponding to closed chromatin and low expression. Moreover, the genome was shown to be partitioned into megabase-sized local chromatin interaction domains, termed topologically

associating domains (TADs), whose boundaries were enriched for the insulator binding protein CTCF (Dixon et al. 2012) (Nora et al. 2012). These domains not only cover most of the genome but are also stable across various cell types and even conserved between mice and humans. Furthermore, *in situ* Hi-C, an adaptation of the protocol where proximity ligation is performed in intact nuclei, improved its efficiency and resolution, and now allowed to identify loops genome-wide, which frequently link promoters and enhancers (Rao et al. 2014).

## Structural Principles of the Inactive X

The nature of the inactive X and its unique necessity to be transcriptionally inactive, not only manifests itself in a discrete epigenetic landscape as described earlier, but also in a distinct three-dimensional organization.

Early studies on the X chromosome, already showed that the inactive X is not only more compact than its active counterpart (Naughton et al. 2010), but also that its territory is characterized by a spherical shape and smooth surface, in contrast to the flatter shape and more irregular surface of the active X (Eils et al. 1996). However, the advent of chromosome conformation capture technologies now allowed for a more detailed and comprehensive view of chromatin folding on the X. Using allele-specific 4C of 4 viewpoints, it became apparent that while on the active X, loci formed multiple long-range interactions, the inactive X was mostly devoid of such, with silenced loci lacking preferred interactions and only escapees engaging in long-range contacts (Splinter et al. 2011). Strikingly, upon deletion of *Xist,* a refolding into a structure resembling the active X was observed, while gene silencing wasn't affected, possibly due to the persistence of DNA-methylation. However, how Xist maintained the unique structure of the Xi wasn't known. Mechanistic insights into this process came from a study on Xist binding partners (A. Minajigi et al. 2015). This revealed that cohesins, integral structural components of the 3D genome (Baranello, Kouzine, and Levens 2014), were among the most enriched factors. Notably, deletion of *Xist* led to a restoration of cohesin binding along the inactive X. This shows that Xist actively repels cohesins in cis, thereby preventing the establishment of a chromosome structure resembling the active X. To provide a chromosome-wide view of the inactive X, Deng and colleagues utilized an *in situ* DNase Hi-C protocol in mouse brain and Patski cells (X. Deng et al. 2015). This revealed a bipartite structure of the mouse Xi, similar to what has been observed in humans (Rao et al. 2014). The two resulting domains (here termed superdomains) are separated by a ~200kb large boundary element and are of roughly similar size (72 Mb and 94 Mb respectively). Notably, this unique structural organization was not only observed using Hi-C but as well by using a multiplexed FISH method, which showed a spatial partitioning of the Xi into two large domains (S. Wang et al. 2016). However, whereas the superdomains differ in size between mice and humans, the boundary region seemed to be conserved and to lie near the *DXZ4/Dxz4* macrosatellite locus in both species. Yet, allelic TAD calling was not possible due to insufficient sequence depth (X. Deng et al. 2015). This issue was overcome by a later study performing allele-specific Hi-C in a clonal neural progenitor cell (NPC) line (Giorgetti et al. 2016). This confirmed the organization of the Xi in two large domains (here called mega-domains) and further revealed a notable lack of A/B compartments, as well as a global lack of TADs (Later studies argued for attenuation of TADs, rather than a complete loss, a

discrepancy that could be explained by the increased sensitivity of the *in situ* Hi-C protocol of the later study (C.-Y. Wang et al. 2018)). However, few residual TAD-like structures were still to be observed. Integration of Hi-C with RNA-seq and ATAC-seq data revealed that residual TADs corresponded to transcribed and accessible regions where genes escape X inactivation (Giorgetti et al. 2016). Furthermore, whereas the deletion of the mega-domain boundary did lead to a loss of the bipartite structure, it did not affect the efficiency of gene silencing. Nevertheless, one clone showed a substantial reduction in escapees, an effect that however wasn't uniformly observed in multiple clones. Consistent with these observations in the mouse, deletion of *DXZ4* from the human Xi chromosome resulted in loss of the mega-domain structures as well (Darrow et al. 2016) Remarkably, the bipartite structure of the Xi does not depend on a specific orientation of *Dxz4*, as upon its inversion, the overall structure of the Xi persisted (Bonora et al. 2018). Strikingly, analysis of a *Xist* mutant that results in a failure of Xi silencing, properly formed mega-domains while at the same time having visibly less attenuated TADs, showing that mega-domain formation is neither sufficient nor a consequence of gene silencing and that TADs and mega-domains can co-exist (Colognori et al. 2019).
Taken together, the inactive X shows a unique structural organization into two mega-domains that is further characterized by a lack of A/B compartments and an attenuation of TADs *(Figure I8)*.

## Spatial Regulation of the X Inactivation Center

The role of genome architecture in XCI can be exemplified by its regulatory function on the X inactivation center (*Xic*). The two lncRNAs *Xist* and *Tsix*, while adopting opposite transcriptional patterns during XCI, share overlapping transcriptional units. Therefore, to achieve accurate expression timings, their corresponding regulatory elements need to be confined to their associated regulatory landscape to prevent uncoordinated expression. To ensure this, it was discovered that a chromatin transitional zone exists between the two domains (Tsai et al. 2008). This harbored a conserved element, *RS14*, that presented strong binding for CTCF, serving as a boundary between the two domains to facilitate coordinated expression (Spencer et al. 2011). In accordance, the deletion of *RS14* resulted in aberrant XCI in female cells, through a failure to upregulate *Xist*, arguing for a functionally important spatial separation of the two domains. Indeed, 5C across the *Xic*, revealed its partitioning into TADs, that separate the promoters of *Xist* (TAD-E) and *Tsix* (TAD-D) (Nora et al. 2012). Moreover, the functional separation of the two TADs became apparent, as each showed a high correlation in transcriptional dynamics within. Furthermore, whereas the position of the TADs was conserved during differentiation, differences in their internal organization were observed. The functional importance of this spatial separation was later shown by genomic inversions of the *Xist/Tsix* locus in male mESCs, which placed their respective promoters in each other's TADs (van Bemmel et al. 2019). Capture-C (Hughes et al. 2014) (Davies et al. 2016) was deployed to first assess the topological environment of the *Xist* and *Tsix* promoters in the wild-type situation. As expected, both promoters preferentially interacted with sequences within their respective TADs. However, a 40kb, as well as a 70kb inversion of the locus, resulted in a switch of their interaction profiles, with the *Tsix* promoter now preferentially interacting with sequences within the *Xist* TAD, and vice versa for the *Xist*

promoter. Moreover, whereas the expression of *Tsix* was not affected by the inversions, the placement in a new interaction environment did lead to a significant upregulation of *Xist*, which manifested itself in the formation of Xist clouds in a significant percentage of cells. A highly similar phenotype was observed when the inversions were generated in female mESCs. Nonetheless, upon differentiation into EpiLCs, differences between wild-type and mutant cells decreased, with the majority of cells exhibiting normal differentiation hallmarks. However, some cells showed reduced expression of *Xist* and a concordant impairment of *Tsix* silencing. In summary, this shows how the three-dimensional organization of the genome not only influences the global structure of the inactive X but also how it partitions its regulatory landscape to ensure proper timing of XCI kinetics *(Figure I2)*.

## Dynamics of Structural Changes During XCI and XCR

As described previously, the inactive X is characterized by a unique structure, that partitions it into two mega-domains. However, how this structure is formed during differentiation and how structural changes and XCI kinetics relate to each other, remained elusive. To answer the latter, Froberg and colleagues used *in situ* Hi-C to assess mega-domain formation during a time course of mESC embryoid body differentiation (Froberg et al. 2018). Whereas after 3 days of differentiation, cells already showed upregulation of *Xist* and coating of the Xi, no mega-domains were observed yet. Only later, after 7 days, mega-domains had formed, showing that the initiation of XCI precedes mega-domain formation. However, silencing of X-linked genes could only be detected later on day 7, when mega-domains had already formed. Unfortunately, no time points in between were sampled, therefore it remains to be seen if mega-domain formation or X-linked gene silencing occurs first.

More detailed insights into how the structural rearrangement of the X chromosome proceeds, came from studies on the architectural protein SMCHD1 (structural maintenance of chromosomes hinge domain containing 1) (C.-Y. Wang et al. 2018) (Gdula et al. 2019) (C.-Y. Wang et al. 2019). SMCHD1 was shown to be enriched on the inactive X (Blewitt et al. 2008), an interaction that was dependent on the polycomb repressive complex 1 deposited mark H2AK119ub (Jansz et al. 2018). Deletion of *Smchd1* not only led to a derepression of X-linked genes but further correlated with a loss of H3K27me3 (Sakakibara et al. 2018), a gain of H3K4me3 and a loss of Xist binding from the derepressed genes. Strikingly, this also revealed a reappearance of compartments, which are normally absent from the Xi. However, these compartments, termed S1/S2, were distinct from the A/B compartments of the Xa, not only in their position on the X but also in their size, being larger than their A/B counterparts. Moreover, binding of SMCHD1 was shown to be rather uniform along the Xi, thereby bridging S1/S2 compartments, arguing for a role of SMCHD1 in merging these. This unique intermediate structure was not only observed in an *Smchd1* ablated state. Inspection of *in situ* Hi-C contact heatmaps of early stages of embryoid body differentiation, when SMCHD1 had not yet been recruited to the Xi yet, revealed the formation of transient compartments that highly correlated with the S1/ S2 structures observed in *Smchd1* knockout cells. Finally, the initial partitioning of the Xi into S1/S2 compartments seems to depend on PRC1 (C.-Y. Wang et al. 2019). Whereas deletion of *Smchd1* in MEFs led to the re-appearance of S1/S2 compartments, these were

markedly weakened if PRC1 was abolished, as well by double knockdown of RING1A/B. Together with the finding that knockdown of hnRNPK, which as previously mentioned bridges PRC1 and Xist (Chu et al. 2015) (Pintacuda et al. 2017), had a highly similar effect, and the PRC1 dependent recruitment of SMCHD1 to the Xi (Jansz et al. 2018), a model can be proposed where Xist starts coating the Xi, subsequently recruiting PRC1 to form S1/S2 compartments, which are ultimately merged by recruitment of SMCHD1 via PRC1 to form the two mega-domains.
Taken together, this reveals the stepwise folding mechanism of the inactive X that is characterized by a unique compartmental structure *(Figure I9)*.

Nonetheless, much less is known about the structural changes occurring during XCR. Stadhouders and colleagues utilized in *situ* Hi-C to characterize genome topology changes in a highly efficient iPSC reprogramming system and found that the TAD structure on the X was reestablished, concordant with the upregulation of naive pluripotency factors and downregulation of *Xist (Stadhouders et al. 2018)*. However, due to the lack of allele-specific information in this study, limited conclusions could be drawn regarding the X and it remains to be answered if XCR involves a unique intermediate folding state as observed during XCI.

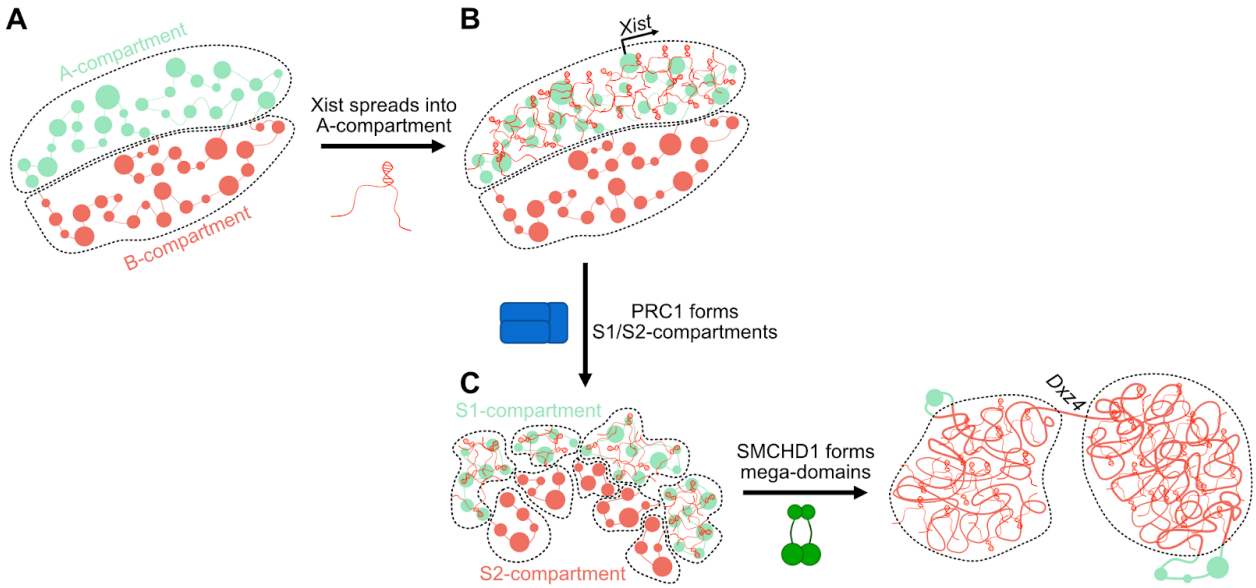



**Figure I9. Stepwise folding of the Inactive X**

**(A)** Xist initially spreads into A-compartmental chromatin. **(B)** Xist then recruits polycomb repressive complex 1 (PRC1) to reconfigure the Xi into S1 (green) and S2 (red) compartments, with itself being restricted to the S1-compartment. **(C)** Finally, SMCHD1 is recruited to merge S1/S2 compartments and form a compartment-less structure, organized into two mega-domains.

# Chapter 1 - Development of *Nano*X

## Aim

Previous studies on X chromosome reactivation during iPSC reprogramming were hindered by the lack of a complete toolset that would allow its study without compromising on either cell purity and/or possible read-out techniques that require large cell numbers. We, therefore, set out to develop a novel reprogramming system that would address common shortcomings of conventional, mostly mouse embryonic fibroblast (MEF)-based, reprogramming systems (Summary of cell line features and comparison to conventional systems see *Figure R1*). First, a hybrid mouse strain background was required, as sequences from the X would otherwise be indistinguishable by sequencing-based assays in cells from inbred strain backgrounds. Second, an easy read-out of the X activity of cells was necessary, to be able to isolate pure cell populations of Xa/Xa and Xi/Xa. Third, cells needed to be able to be reprogrammed into iPSC by the simple addition of doxycycline. Fourth, as the efficiency of iPSC reprogramming is inherently low, a marker was required to isolate prospective iPSC intermediates poised for reprogramming and XCR from a pool of incompletely reprogramming cells. Fifth, high scalability of the system was crucial to be able to generate a large number of cells to allow for sequencing-based high-throughput experiments.

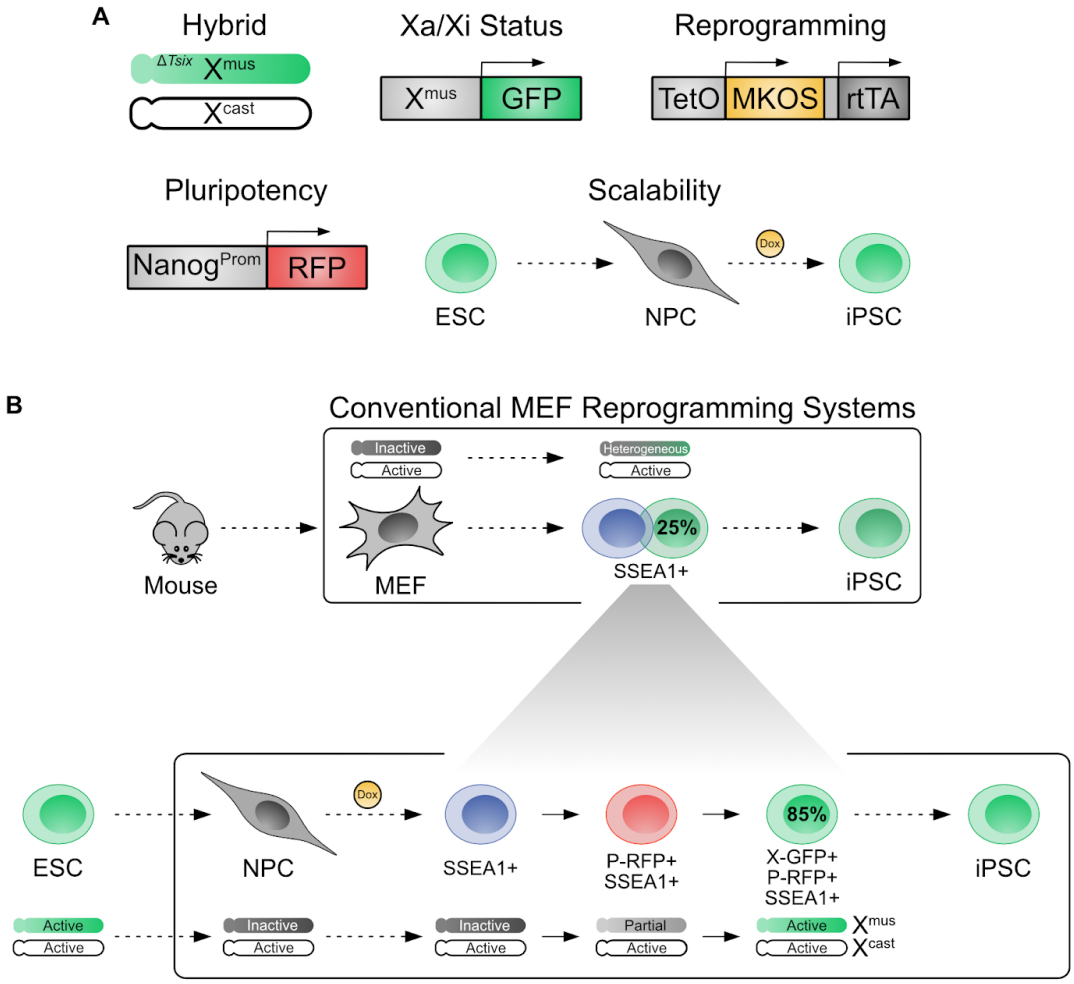



**Figure R1. Features of the Cell Line**

**(A)** Schematic of the 5 features implemented in the cell line. **(B)** Comparison of our novel iPSC system with conventional MEF reprogramming systems.

# Results

## Hybrid Mouse Strain Background

To be able to distinguish sequence reads from the two X-chromosomes, we decided to utilize the female F2 ESC line EL16.7 TST, which was derived from a cross of the genetically distant mouse strains *Mus musculus musculus* (mus) with *Mus musculus castaneus* (cas) *(Ogawa, Sun, and Lee 2008)*. As a result, cells contain one X chromosome from *M.m musculus* ($X^{mus}$) and one from *M.m castaneus* ($X^{cas}$), allowing us to distinguish them via sequence polymorphisms that occur on average every 300 bp (Marks et al. 2015). Chromosome 13 is fully heterozygous as well, while the rest of the genome consists of intermingled blocks of *M.m musculus* and *M.m castaneus*. Moreover, EL16.7 TST contains a truncation of *Tsix* on $X^{mus}$

(*Tsix*$^{TST/+}$), which abrogates *Tsix* expression and leads to the non-random inactivation of X$^{mus}$ upon differentiation (Luikenhuis, Wutz, and Jaenisch 2001).

As EL16.7 TST cells were used to be grown on feeder cells, they were adapted to feeder-free conditions. Cells were seeded at low density (1,000 cells per 10 cm plate) on 0.2% gelatin in standard ESC medium and then colonies with proper ESC morphology (round, shiny edges) were picked *(Figure R2A)*. Two additional rounds of picking followed to ensure a stable ESC state until cells could be expanded. Quantitative RT-PCR showed proper expression of the pluripotency marker *Nanog*, as well as of *Xist (Figure R2B)*. Karyotyping was performed to guarantee diploidy, and especially critical for this study the presence of two X chromosomes. Whereas the majority of cells harbored a diploid chromosome set *(Figure R2D)* with always two X chromosomes, cells as well displayed a trisomy of chromosome 8, which was already present in control cells before the gelatin adaption *(Figure R2C)*. Trisomy 8 is considered to be a fairly common chromosomal abnormality in mouse ESCs, as it conveys a growth advantage compared to karyotypically normal cells (Liu et al. 1997) (Gaztelumendi and Nogués 2015). While it correlates with low efficiency of germline transmission, it doesn't impact the cell's differentiation potential (Gaztelumendi and Nogués 2015).

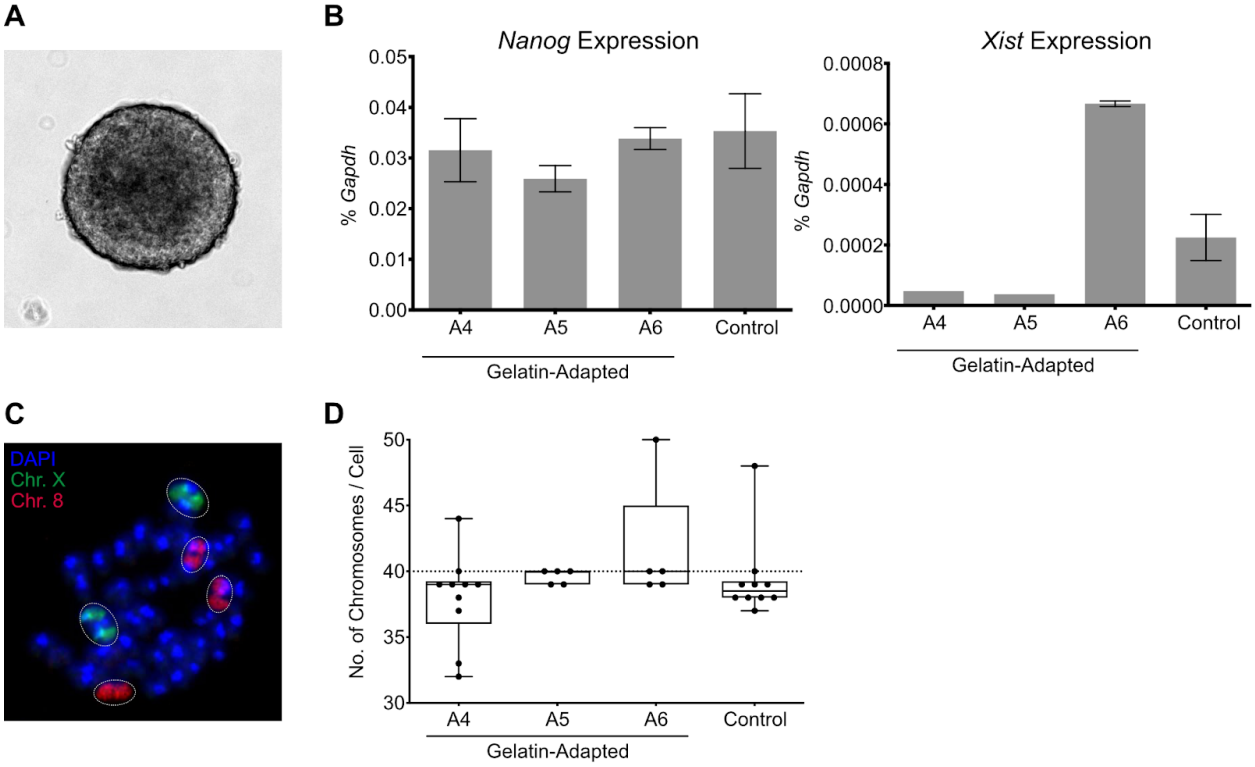



**Figure R2. Gelatin Adaption of EL16.7 TST**

**(A)** Example of an ESC colony with proper round morphology and shiny edges. **(B)** Gene expression (qRT-PCR) of *Nanog* and *Xist* after gelatin-adaption. Values were normalized against *Gapdh* expression. Error bars indicate SD (n = 3 technical replicates). **(C)** Chromosome painting shows two Chr. X (green) and trisomy 8 (red). **(D)** Chromosome counting shows on average 40 chromosomes. N = 10 for A4 and control. N = 5 for A5 and A6. Control, parental cell line before gelatin adaption.

## X Activity Reporters

In order to visualize the activity of the X chromosomes at cellular resolution, we decided to integrate fluorescent reporters on either the $X^{mus}$ (X-GFP) or the $X^{mus}$ and the $X^{cas}$ (X-Dual Color / X-DC) *(Figure R3A)*. We selected the *Hprt* locus for targeted integration, as it has been shown that reporters inserted there, obey XCI (Ciavatta et al. 2006). For this, we utilized previously used reporter constructs consisting of a CMV/beta-actin enhancer/promoter (CAG), driving nuclear-localized GFP or tdTomato reporters (Wu et al. 2014). These constructs are additionally flanked in tandem by Chicken hypersensitive site 4 (cHS4) DNA insulators that prevent repression of the transgene, but not its random X inactivation (Ciavatta et al. 2006). We modified these reporters to facilitate targeted integration of the GFP reporter into $X^{mus}$ and the tdTomato reporter into $X^{cas}$ by replacing the homology arms with allele-specific variants. We furthermore reduced the length of the homology arms from around 5kb to 2kb, as CRISPR/Cas9 assisted targeted integration doesn't necessitate long homology regions (H. Yang et al. 2013). The modified homology arms target the second exon of *Hprt*. Targeted integration would lead to the deletion of exon 2, which abrogates the gene's function and confers resistance to 6-thioguanine (6-TG) (Gallagher et al. 1984). Targeted cells can, therefore, be selected by the addition of 6-TG (Liao, Tammaro, and Yan 2015).

We nucleofected EL16.7 TST ES cells with a mix of circularized GFP and tdTomato reporter vectors and a single guide RNA vector targeting exon 2 of *Hprt*. We found, that after selection with 6-TG, the vast majority of the cells did not show a fluorescent signal, showing that Cas9-mediated genome editing via nonhomologous end joining (NHEJ) alone was sufficient to delete Hprt's function *(Figure R3B)*. However, we still found around 1.4% of the cells to be either GFP+ or tdTomato+, with a considerably smaller amount of 0.18% of cells being double-positive, indicating that homozygous integrations are a rare event. We then went on to confirm the correct integration of single clones via Southern blot, using an external probe binding upstream of the right homology arm and two internal probes for GFP and tdTomato. We found that approximately 15% of X-DC *(Figure R3C)*, as well as 30% of X-GFP clones *(Figure R3D)*, correctly integrated the reporters. Importantly, previous attempts using linearized targeting vectors were unsuccessful as clones consistently showed multiple integrations of the vectors (data not shown). As our Southern blot strategy couldn't distinguish between integrations of the vectors in the mus or the cas allele, we subsequently functionally tested correctly targeted clones by differentiating ESC and monitoring the downregulation of the reporters, illustrating XCI. We performed retinoic acid as well as embryoid body differentiation for 6-8 days of X-DC lines and analyzed their outgrowth and found that 5 out of 5 tested clones, correctly inactivated the X-GFP reporter, demonstrating the integration of the reporter in the *mus* allele *(Figure R4B)*. Similarly, embryoid body differentiation of X-GFP lines revealed the correct inactivation of the reporter upon differentiation in 5 out of 7 tested clones *(Figure R4C)*.

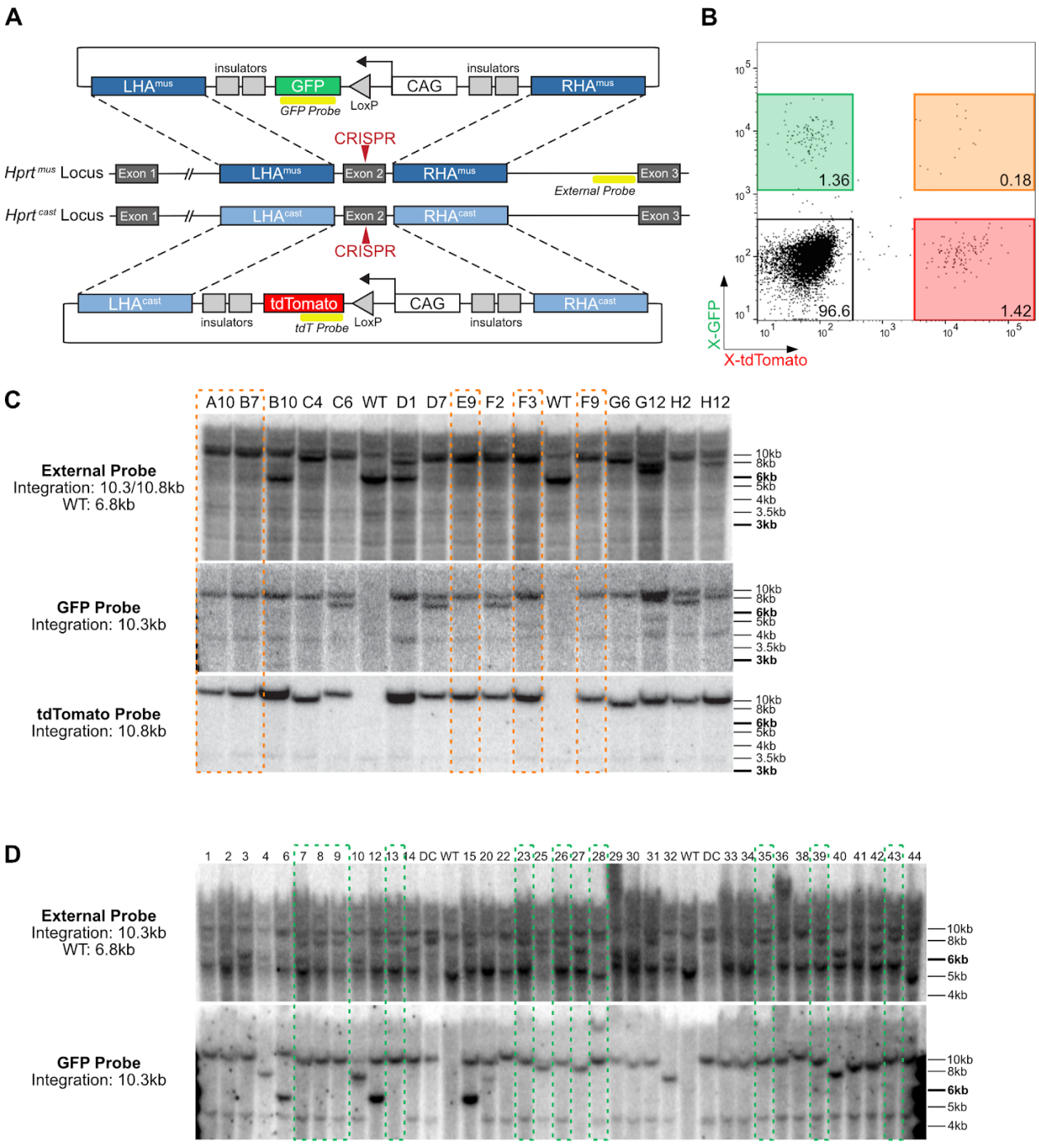



**Figure R3. X-Activity Reporter Integration**

**(A)** Schematic of the reporter constructs and targeting strategy. **(B)** FACS gating strategy for the isolation of GFP and tdTomato reporter expressing clones. Numbers indicate the percentage of cells.
*(Figure legend continues on next page)*

**(C)** Southern blot analysis of *Hprt* X-DC targeted alleles. Genomic DNA was digested with BamHI. Expected fragment sizes: WT (wild-type) = 6.8kb, GFP targeted = 10.3kb, tdTomato targeted = 10.8kb. Clones with correct integration are indicated by an orange outline. **(D)** Southern blot analysis of the *Hprt* X-GFP targeted allele. Genomic DNA was digested with BamHI. Expected fragment sizes: WT (wild-type) = 6.8kb, GFP targeted = 10.3kb. Clones with correct integration are indicated by a green outline.

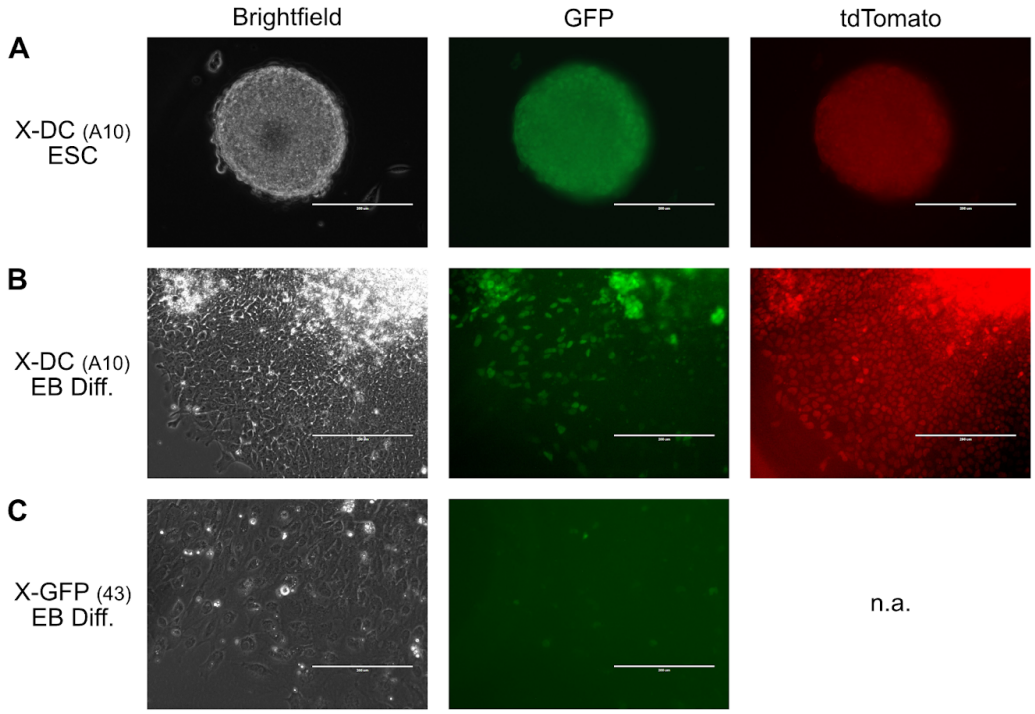



**Figure R4. X-Activity Reporter Verification by Embryoid Body Differentiation**

**(A)** X-DC targeted ESC show expression of both X-GFP and X-tdTomato. **(B)** X-DC targeted cells show downregulation of X-GFP after embryoid body differentiation. **(C)** X-GFP targeted cells show downregulation of X-GFP after embryoid body differentiation.

## All-in-one Reprogramming Cassette

To be able to easily and efficiently reprogram cells, we utilized a previously published optimized polycistronic reprogramming cassette, termed MKOSimO, for the targeted integration into the *Sp3* locus (Chantzoura et al. 2015) *(Figure R5A)*. Cells reprogrammed using MKOSimO showed a significant reduction in heterogeneity, a critical parameter when trying to reveal molecular cornerstones of iPSC formation, such as in our case XCR. Moreover, the cassette already included a rtTA transgene, allowing for doxycycline-inducible cassette expression, and thereby also reducing additional gene targeting steps, and furthermore a mOrange fluorescent reporter to follow cassette expression. We nucleofected X-GFP (clone 43) ESC with a mix of circularized Sp3 MKOSimO vector and a single guide RNA vector targeting the third intron of the *Sp3* gene for CRISPR/Cas9 assisted targeted integration. After selection with neomycin, clones were picked and doxycycline added for 24 hours to monitor cassette expression using mOrange. 4

out of 12 (1, 9, 10 and 11) tested clones expressed mOrange, indicating proper cassette expression. We then further analyzed the clones by Southern blot using a probe binding the 3' homology arm and found that 2 out of the 4 clones (9 and 11), correctly integrated the reprogramming cassette, giving a targeting efficiency of around 16% *(Figure R5B)*.

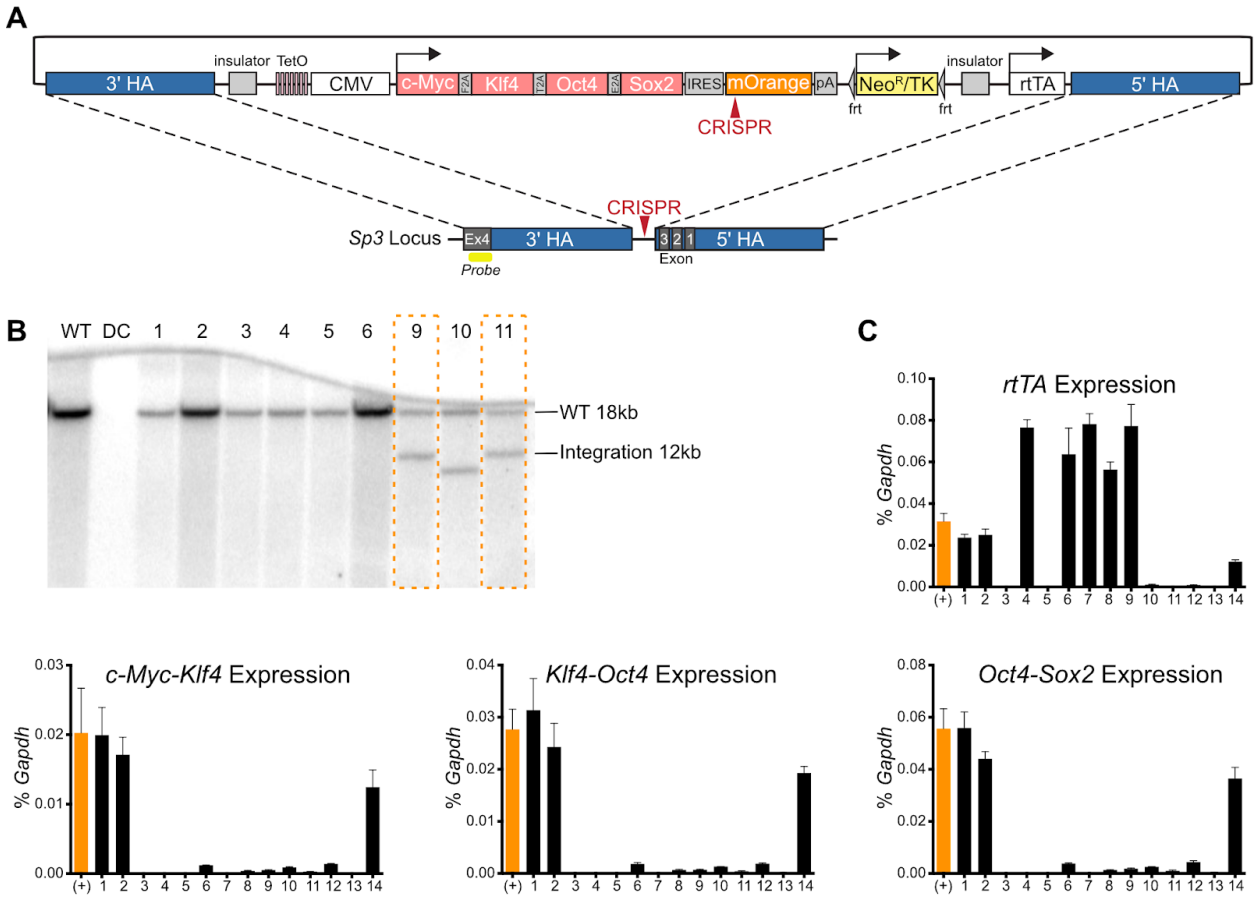



**Figure R5. Reprogramming Cassette Integration**

**(A)** Schematic of the reprogramming cassette, targeting strategy and CRISPR/Cas9 cut sites. **(B)** Southern blot analysis of the *Sp3* MKOSimO targeted allele. Genomic DNA was digested with SacI. Expected fragment sizes: WT (wild-type) = 18kb, targeted = 12kb. Clones with correct integration are indicated by an orange outline. **(C)** Gene expression (qRT-PCR) of reprogramming cassette specific transcripts of *rtTA*, *c-Myc-Klf4*, *Klf4-Oct4*, and *Oct4-Sox2* after mOrange knockout. Values were normalized against *Gapdh* expression. Error bars indicate SD (n = 3 technical replicates).

## Knockout of mOrange

As downstream applications of the X-GFP Sp3 MKOSimO cell line necessitated the use of an RFP reporter, which spectrally overlaps with mOrange, we decided to knockout mOrange in the reprogramming cassette using CRISPR-Cas9. We nucleofected X-GFP Sp3 MKOSimO (clone 11) ESC with a single guide RNA vector targeting the 5'-end of the mOrange reporter *(Figure R5A)*, induced cassette expression for 24h by addition of doxycycline and then isolated mOrange-negative cells by FACS. We screened single clones for reprogramming cassette

expression by qRT-PCR and found that 3 out of 14 (1, 2 and 14) tested clones maintained expression of the cassette, as well as of rtTA *(Figure R5C)*.

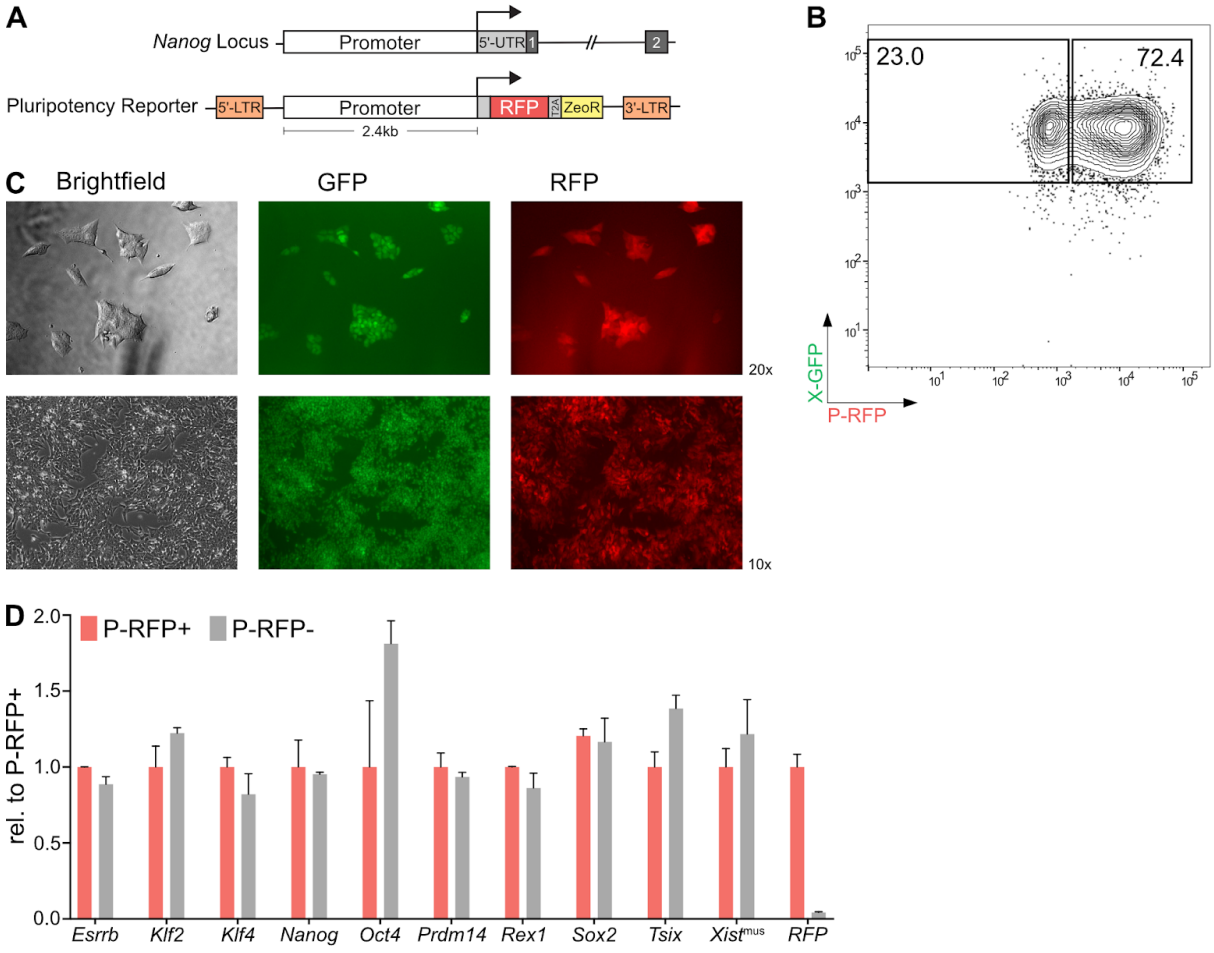



**Figure R6. Pluripotency Reporter**

**(A)** Schematic of the pluripotency reporter. **(B)** FACS analysis of P-RFP reporter expression of ESC cultured in serum/LIF conditions. Shown is a representative contour plot. Numbers indicate the percentage of cells. **(C)** P-RFP targeted ESC show expression of both X-GFP and P-RFP. **(D)** Gene expression (qRT-PCR) of pluripotency genes, as well as *Xist*$^{mus}$ and *RFP* in P-RFP- and P-RFP+ sorted cells. Values were normalized against *Gapdh* expression. Error bars indicate SD (n = 3 technical replicates).

## Pluripotency Reporter (P-RFP)

To isolate prospective iPSC intermediates poised for complete reprogramming and XCR from a pool of incompletely reprogramming cells, we decided to utilize a commercially available *Nanog* reporter. The vector harbors an RFP reporter driven by a promoter fragment of *Nanog* in a lentiviral backbone (P-RFP) *(Figure R6A)*. We infected X-GFP Sp3 MKOS ESC with the vector and isolated single clones by FACS based on RFP expression. Comparably to some *Nanog*

knockin reporter cell lines (Chambers et al. 2007) (Kalmar et al. 2009) (Abranches, Bekman, and Henrique 2013), we found a dynamic expression of the reporter when ESC were cultured in serum/LIF conditions, with around 70% of cells being P-RFP+ *(Figure R6B)*. However, P-RFP- cells shared the same pluripotency factor expression pattern, including *Nanog*, as P-RFP+ cells *(Figure R6D)*, arguing that this population does not describe a differentiated sub-population.

## Neural Precursor Cell Differentiation

As high scalability of the system was crucial to be able to generate a large number of cells to allow for sequencing-based high-throughput experiments, as well as to circumvent the lengthy process of producing transgenic mice to derive MEFs, we decided to generate neural precursor cells (NPC) *in vitro* by differentiation from ESC. We adapted a previously published protocol (Abranches et al. 2009) with several critical modifications. As most differentiation protocols are optimized for male ESC, we first increased the duration of the protocol and accordingly adapted the starting cell number, to account for the delay in differentiation observed in female ESC, as two active X delay the exit from pluripotency (Schulz et al. 2014). Moreover, as we observed a large heterogeneity between different plates of the same differentiation, we intended to synchronize cells, as well as switch them to a more naive like state, by the addition of a 6-24h incubation in 2i/LIF medium (Ying et al. 2008) before the start of the differentiation. Using this optimized protocol, we observed that after 4 days of differentiation, small colonies had formed, from which after an additional 2-4 days, cells with the characteristic NPC morphology appeared in the outgrowth *(Figure R7A)*. The differentiation of the cells was accompanied by the downregulation of the pluripotency factors *Nanog* and *Oct4*, as well as the upregulation of the NPC marker *Nestin (Figure R7B)*. Moreover, cells underwent XCI, defined by the upregulation of Xist from $X^{mus}$, as well as visualized using our X-DC line, by downregulation of GFP from $X^{mus}$, whereas tdTomato expression from the $X^{cas}$ stayed consistent *(Figure R7C)*.
We then went on to assess the dynamics of our P-RFP reporter cell line during NPC differentiation. We found that the majority of cells had downregulated the P-RFP reporter by day 6, preceding the downregulation of SSEA-1 at day 8 *(Figure R8A)*. This showed that our reporter cell line could recapitulate the correct order of events, with naive pluripotency genes being downregulated first, and general pluripotency markers being downregulated later during differentiation. Furthermore, we found that downregulation of P-RFP, occurred concurrently with the downregulation of the X-GFP reporter at day 6 (*Figure R8B)*, again indicating the link between pluripotency downregulation and XCI.

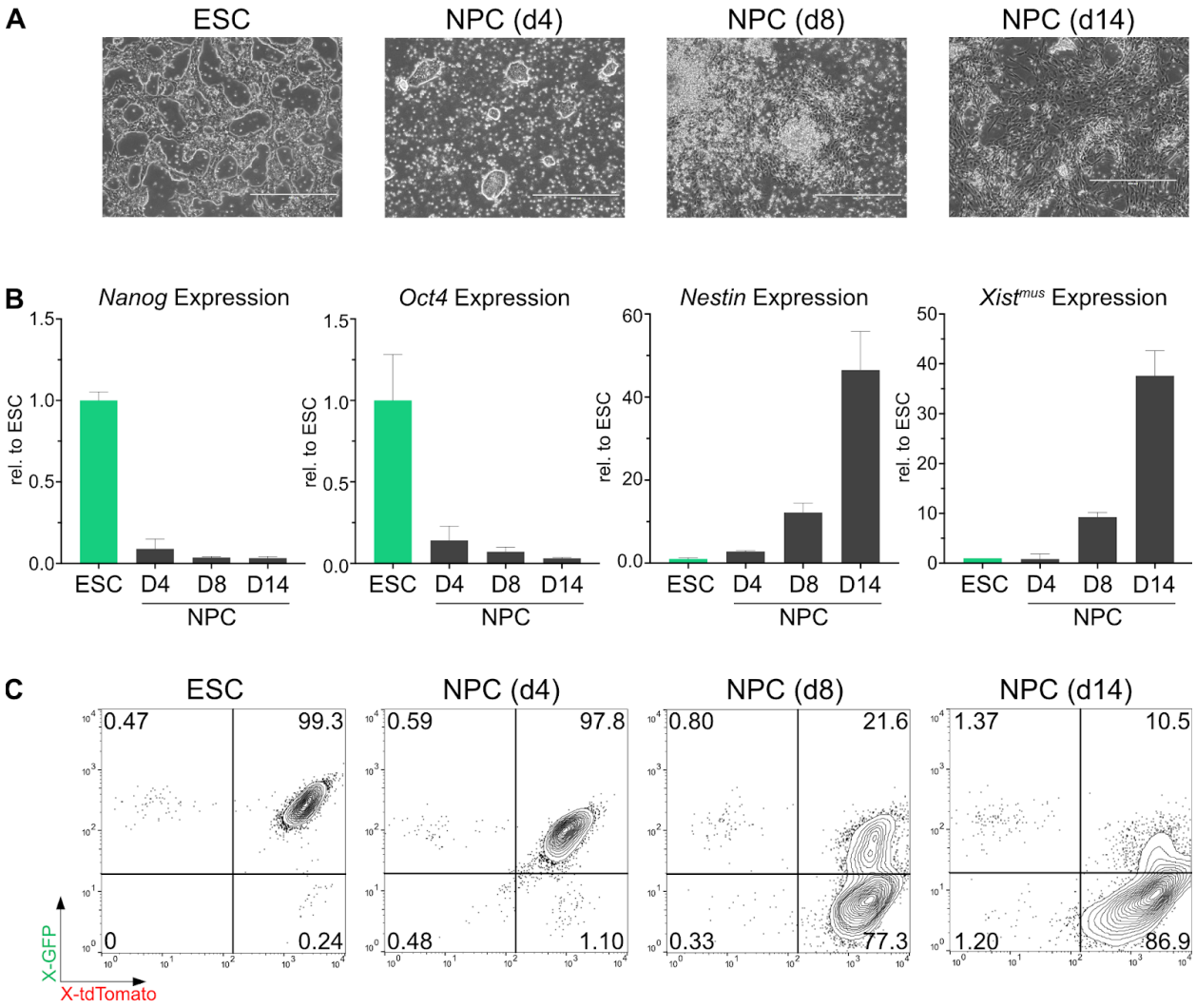
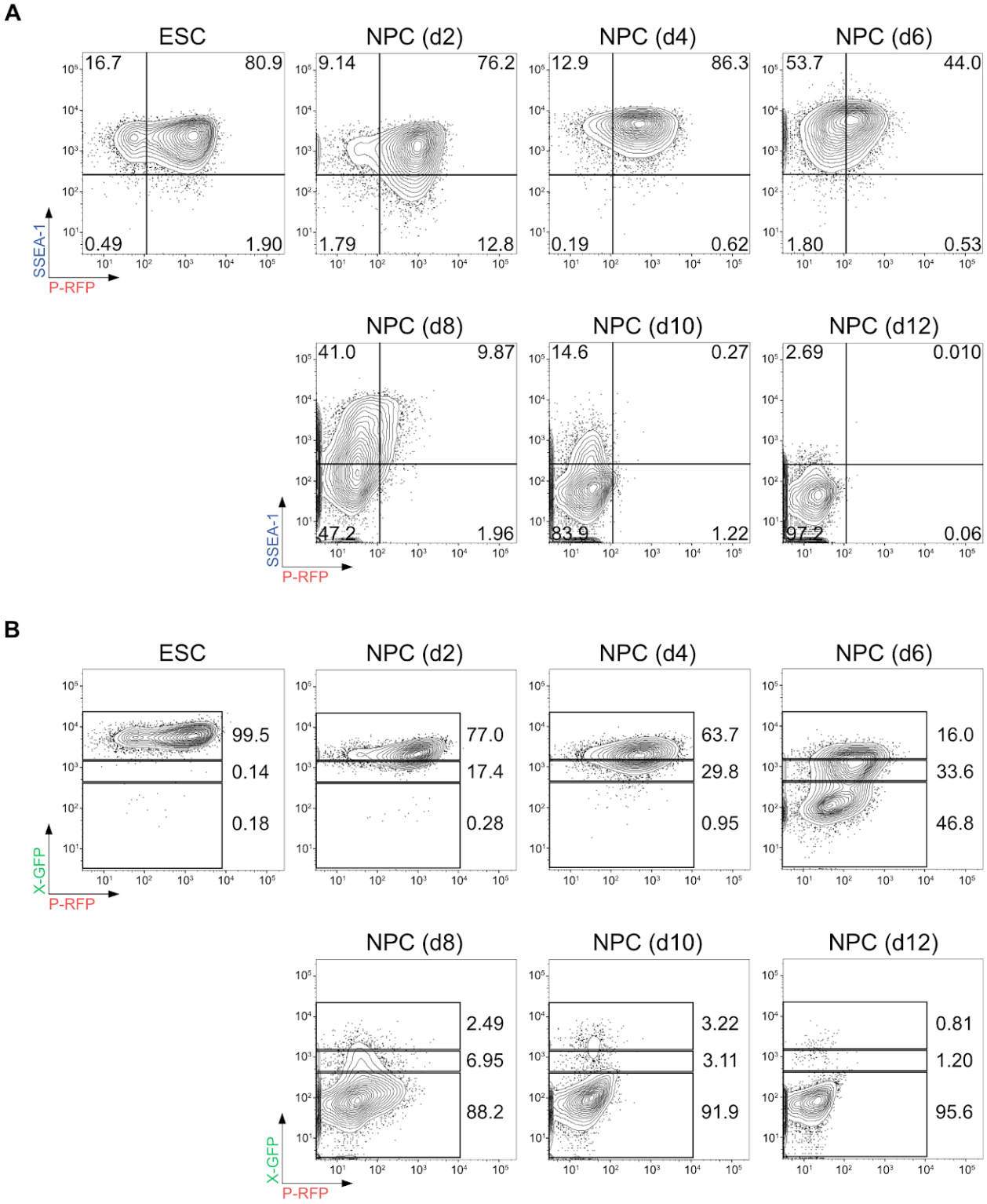



**Figure R7. Neural Precursor Cell Differentiation of the X-DC Line**

**(A)** Representative brightfield images of ESC cultured in 2i/LIF conditions and neural precursor cells after 4, 8 and 14 days of differentiation respectively. **(B)** Gene expression (qRT-PCR) of pluripotency markers *Nanog* and *Oct4*, NPC marker *Nestin* and *Xist*$^{mus}$ during a time course of NPC differentiation. Values were normalized against *Gapdh* expression. Error bars indicate SD (n = 3 technical replicates). **(C)** FACS analysis of X-GFP and X-tdTomato reporter expression during a time course of NPC differentiation. Shown are representative contour plots. Numbers indicate the percentage of cells.




*(Figure legend on next page)*

**Figure R8. Neural Precursor Cell Differentiation of the P-RFP / X-GFP line**

**(A)** FACS analysis of SSEA-1 and P-RFP reporter expression during a time course of NPC differentiation. Shown are representative contour plots gated on live cells. Numbers indicate the percentage of cells. **(B)** FACS analysis of X-GFP and P-RFP reporter expression during a time course of NPC differentiation. Shown are representative contour plots gated on live cells. Numbers indicate the percentage of cells.

## iPSC Reprogramming

We then set out to test the NPC's ability to reprogram. We found that upon addition of doxycycline, colonies formed when cells were grown either on irradiated mouse embryonic fibroblasts (iMEF) or gelatin-coated plates. However, the colony-forming ability was drastically reduced on gelatin, being around 10% of the observed numbers on iMEF (data not shown). To increase reprogramming efficiency, we tested combinations of small molecules, which have been previously shown to enhance reprogramming. Combinatorial treatment with a TGF-b inhibitor (iAlk5) together with an activator of Wnt signaling (CHIR) in the presence of L-ascorbic acid (L-AA), previously shown to achieve efficiencies of up to 100% in MEF-based systems, did increase efficiency, albeit only to about 8% *(Figure R8A)* (Vidal et al. 2014). However, the single addition of L-AA showed even higher efficiencies of over 10% (Esteban et al. 2010). Importantly, 50% of colonies showed upregulation of X-GFP as well *(Figure R9A and R9B)*. Of note, we also observed that a very high cell density of the starting NPC *(Figure R9C)* was crucial to achieving a high reprogramming efficiency (data not shown). Moreover, we found that the dissociation of NPC decreased efficiencies as well, independent of the cell density. We, therefore, limited the dissociation steps during the NPC's differentiation to a single one at day 9 when cells were FACS purified (data not shown).

Furthermore, having established our main reprogramming conditions, we set out to assess how the addition of L-AA improved its efficiency and if perturbation of the system with inhibitors targeting epigenetic modifiers was feasible. We reprogrammed P-RFP reporter cells and first analyzed early pluripotency by SSEA-1. We found that SSEA1 levels increased until day 6, however, we couldn't detect significant differences between the different conditions, apart from a slight increase in SSEA1+ cells upon treatment with the EZH2 inhibitor GSK126. In contrast, we consistently found P-RFP+ levels to be higher throughout the reprogramming time course upon addition of L-AA, peaking around 20% at day 8, compared to just ~7% without. Moreover, we found that a high dose of 5-aza further increased the percentage of P-RFP+ cells to up to 50%. These general trends were also observable for the X-GFP reporter. Whereas cells without L-AA only showed upregulation of X-GFP in ~12% of cells, we found X-GFP levels of >40% upon addition of L-AA. Moreover, again a high dose of 5-aza improved this to 75%. However, it's important to note, that while the percentages of X-GFP cells increased, the timing of X reactivation did not change and was consistently observed at day 6.

We, therefore, conclude that the addition of L-AA does not influence early stages of reprogramming, but rather later stages, positively influencing the acquisition of naive pluripotency and XCR. The addition of 5-aza has previously been shown to considerably enhance reprogramming efficiency (Mikkelsen et al. 2008) (Huangfu et al. 2008). However,

cautious interpretation of our results obtained from high doses of 5-aza is necessary. While the results shown here could argue for an increase in efficiency of late pluripotency acquisition and XCR, we did observe a considerable amount of cell death (data not shown) in this condition, arguing that the perceived increase in efficiency, is rather accounted for by a decrease in the number of non-reprogramming cells.

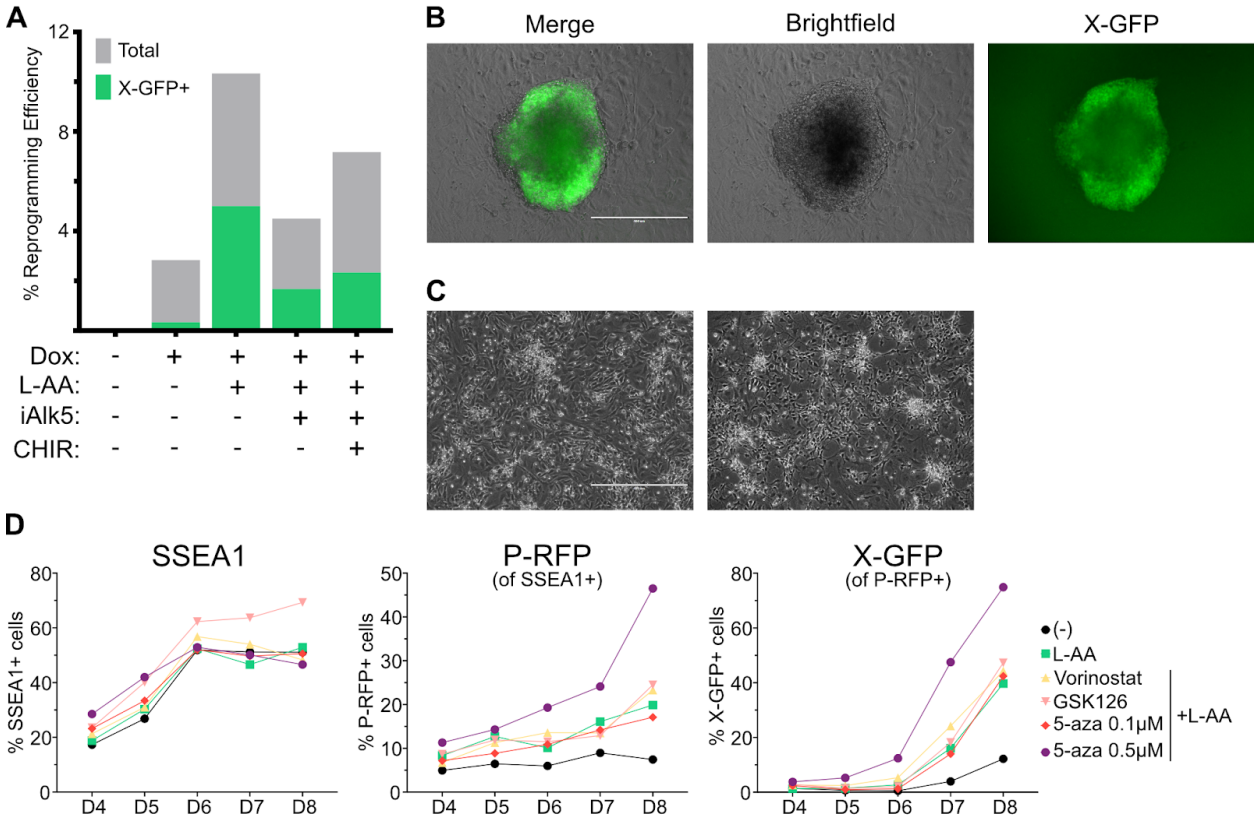



**Figure R9. iPSC Reprogramming**

**(A)** Reprogramming efficiency of cells upon addition of doxycycline (Dox), L-ascorbic acid (L-AA), TGF-b RI Kinase Inhibitor II (iAlk5), GSK3β antagonist/Wnt activator CHIR99021 (CHIR). Reprogramming efficiency was calculated by dividing the number of colonies by the number of cells plated after 10 days of reprogramming. Colonies were counted as X-GFP+ if ~50% of cells showed X-GFP expression. **(B)** Representative brightfield images of iPSC after 10 days of reprogramming. **(C)** Optimal density of NPCs for efficient iPSC reprogramming (left). Suboptimal density of NPCs (right). **(D)** FACS analysis of reprogramming stages. SSEA-1 expression gated on live cells. P-RFP expression gated on SSEA1+ cells. X-GFP expression gated on P-RFP+ cells. (-), without L-AA.

# Materials and Methods

## Cell Lines Conceived

| Name | Unique Feature(s) | Location |
|---|---|---|
| EL16.7 TST | Gelatin adapted | / |
| EL16.7 TST X-GFP | X-GFP reporter | *Hprt*$^{mus}$ locus |
| EL16.7 TST X-DC | X-GFP & X-tdTomato reporter | *Hprt*$^{mus}$ & *Hprt*$^{cas}$ loci |
| EL16.7 TST X-GFP MKOSimO | MKOSimOrange reprogramming cassette with mOrange expression | *Sp3* locus |
| EL16.7 TST X-GFP MKOS | Knockout of mOrange | *Sp3* locus |
| EL16.7 TST X-GFP MKOS P-RFP | Nanog promoter fragment driving RFP | Viral delivery, random integration, not mapped |

## Embryonic Stem Cell Culture

Mouse embryonic stem cells (ESCs) were cultured on 0.2% gelatin-coated dishes in DMEM (Thermo Fisher Scientific, 31966021), supplemented with 10% FBS (ES-qualified, Thermo Fisher Scientific, 16141079) 1,000 U/ml LIF (ORF Genetics, 01-A1140-0100), 1 mM Sodium Pyruvate (Thermo Fisher Scientific, 11360070), 1x MEM Non-Essential Amino Acids Solution (Thermo Fisher Scientific, 11140050), 50U/ml penicillin/streptomycin (Ibian Tech, P06-07100) and 0.1 mM 2-mercaptoethanol (Thermo Fisher Scientific, 31350010). Cells were incubated at 37°C with 5% $CO_2$. The medium was changed every day and cells were passaged using 0.05% Trypsin-EDTA (Thermo Fisher Scientific, 25300054). Cells were monthly tested for mycoplasma contamination using PCR.

## Karyotyping

ESC were grown on 10cm plates until they reached 70-80% confluency and then 20 µg/ml colcemid (Merck Millipore, 234109) were added and incubated o/n. Cells were washed twice with PBS, detached using 0.05% Trypsin-EDTA and again washed with PBS and pelleted. The cell pellet was loosened by tapping the tube and then dropwise 2 ml of warm (37°C) hypotonic solution (KCl 0.075M in $H_2O$) were added with a 1ml pipette while vortexing at low speed. The tube was flicked to resuspend the pellet, as pipetting would lead to cells getting stuck in the tip. The cells were kept vortexed and 8 ml more of hypotonic solution were added. Cells were then incubated at 37°C in a water bath for 10 min without agitation. Then dropwise 1 ml of cold (-20°C) Carnoy fixative (methanol, acetic acid, 3:1) were added while vortexing at low speed.

Cells were centrifuged at 1800 rpm for 10 min at 4°C, the supernatant was removed, and the pellet loosened. As above, 2 ml Carnoy fixative were added dropwise and then an additional 8 ml Carnoy fixative was added.
Microscopy slides were washed in Carnoy fixative and stored in ice-cold $H_2O$ until use. Samples were centrifuged at 1800 rpm for 8 minutes at 4°C and resuspended in 200 µl freshly prepared room temperature Carnoy fixative. Slides were then dried with cellulose paper until they were completely clean and then placed shortly in distilled $H_2O$, to leave a thin layer of $H_2O$ on it. The sample was then collected using a glass Pasteur pipette and exactly one drop was released onto the slide from a height of about 50 cm. The slides were then left to dry at room temperature.
5 µl of each XMP 8 orange (MetaSystems Probes, D-1408-050-OR), as well as XMP X green (MetaSystems Probes, D-1420-05-FI), were added to each spot, covered with a 22x22mm$^2$ coverslip and sealed with rubber cement. Sample and probes were then denatured at 75°C for 2 min and then hybridized in a humidified chamber at 37°C o/n.
The coverslip was removed, and slides were first washed in 0.4x SSC at 72°C for 2 min and then in 2x SSC, 0.05% Tween-20 at room temperature for 30 sec. Slides were rinsed briefly in distilled $H_2O$ to avoid crystal formation and let air dry. 10µl Vectashield (Vector Laboratories, H1200) were added and overlaid with a 22x60mm$^2$ coverslip, left for 10 min and then sealed with nail polish. Slides were stored at 4°C. Images were acquired using a Leica DM6000 B.

## X Activity Reporter

A GFP and a tdTomato reporter construct were targeted into the second exon of *Hprt* on $X^{mus}$ and $X^{cas}$ as follows.

### Cloning of Reporter Constructs

Homology arms flanking the target site were amplified from genomic DNA of male tail-tip fibroblasts from *musculus* and *castaneous* origin, to give allele-specific targeting constructs. The TaKaRa Ex Taq DNA Polymerase (Clontech, RR001C) was used with primers adding restriction enzyme cut sites on the 5'-end for subsequent cloning steps. The amplified DNA was separated by agarose gel electrophoresis and bands at expected sizes of 2.2kb for the left homology arm (LHA) and 1.5kb for the right homology arm (RHA) were cut out and purified using the QIAEX II Gel Extraction Kit (Qiagen, 20021). To facilitate homology arm insertion into a vector backbone, restriction enzyme double digests using PstI and Cfr42I for the left arm and PstI and XhoI for the right arm were then performed and DNA was subsequently purified using the QIAEX II Gel Extraction Kit. Digested left homology arms were then ligated into pBluescript II SK(+) (Addgene, 212205) using the T4 DNA Ligase (New England Biolabs, M0202) and then electroporated into ElectroMAX Stbl4 cells (Invitrogen, 11635-018) according to manufacturer's instructions to give the construct pB LHA. Correct integration was verified by digest with DpnI. pB LHA was then double digested using PstI and XhoI and the right homology arm was ligated in and transformed as described above for the left arm to give pB LHA-RHA. Correct integration was verified by digest with DpnI as well as HindIII/XhoI. Next, the reporter construct cHS4-CAG-nlsGFP-cHS4 (Wu et al. 2014), was to be cloned between the two homology arms.

The integration was performed in two steps. First, a 2.5kb region downstream of the GFP/tdTomato reporter, containing one of the two cHS4 tandem repeats was excised from the original targeting vector by restriction digest using EcoRI and MluI. This fragment was then ligated into pB LHA-RHA and ligation products electroporated into Stbl4 cells to give pB HA-cHS4. Correct integration was verified by digest with EcoRI/HindIII. Second, the residual 5.5kb parts of the reporter constructs, containing the GFP/tdT reporter as well as the second cHS4 insulator tandem, were excised from the original targeting vector by restriction digest using EcoRI. This fragment was then ligated into pB HA-cHS4 and ligation products electroporated into Stbl4 cells to give the final constructs. Correct integration was verified by digest with HindIII as well as XbaI.

**Table 1. Primers for Homology Arm Amplification**

Homology arm binding regions are shown in bold**.** Restriction enzyme recognition sites are underlined. The following recognition sites were inserted: PstI, CTGCAG; Cfr42I, CCGCGG; XhoI, CTCGAG; MluI, ACGCGT; EcoRI, GAATTC.

| Amplicon | Forward (5'-3') | Reverse (5'-3') |
|---|---|---|
| Left Homology Arm | 5'-TGCTTACCGCGG**TATCCCTGCCTTTGTACCCG**-3' | 5'-GAATTCTGCAGACGCGT**AAGAAAGACAGCTGTTTCAAAGTGG**-3' |
| Right Homology Arm | 5'-ACGCGTCTCGACTGCAGGGAATTC**AGTAAGACCTCGATTGAAGTTATT**T-3' | 5'-TAAGCACTCGAG**CTTTCAGGGGGACTGCTGAG**-3' |

## Integration of Reporter Constructs

$5x10^6$ EL16.7 TST A4 ES cells were mixed with 1.6 µg circularized GFP vector, 1.6 µg circularized tdTomato vector and 5 µg single guide RNA vector PX459 (Addgene, 48139) (5'-TATACCTAATCATTATGCCG-3'), to achieve an optimal ratio of Cas9 to targeting vector equal to 5:1 (Pinder, Salsman, and Dellaire 2015). Cells were nucleofected with the AMAXA Mouse Embryonic Stem Cell Nucleofector Kit (LONZA, VPH-1001) using program A-30 and 7.5 µM RS-1 (Merck, 553510) was added to enhance homology-directed repair. To select for the disruption of *Hprt*, 4 days after the nucleofection 10 µM 6-thioguanine (Sigma-Aldrich, A4882-250MG) was added for 3 days. GFP+ and GFP+/tdT+ cells were then isolated by FACS using a BD Influx (BD Biosciences) and a day later subjected to an additional 3 days of 6-TG selection. Cells were then grown 2 more days without selection when single clones could be picked. Single clones were screened by Southern blot hybridization. Inactivation of the X-GFP construct upon differentiation was confirmed using retinoic acid as well as embryoid body differentiation.

## Verification of Reporter Constructs Using Southern Blot

The external southern blot probe was PCR amplified from genomic DNA (Forward: 5'-ACTGGTCAAGGAAATGGTGCT-3', Reverse 5'-CCAACACACCAGCTCAACCA-3') using DreamTaq Green PCR Master Mix (2X) (Thermo Fisher Scientific, K1082) to give a 402bp

fragment. The tdTomato probe was PCR amplified from the targeting construct (Forward: 5'-GGGCGAGGAGGTCATCAAAG-3', Reverse 5'-TGATGACGGCCATGTTGTTG-3') to give a 743bp fragment. The GFP probe was generated by excision from the targeting construct using restriction enzymes NotI and BamHI to give a 778bp fragment.

10µg genomic DNA was digested with 40U BamHI (Thermo Fisher Scientific, ER0051) overnight. Subsequently, genomic DNA was separated on a 0.8% agarose gel for 5h at 120V and 120mA. After the run, the DNA was assessed visually, with a smear indicating proper quality for the further procedures. The gel was then depurinated for 20 min using 0.25M HCl, denaturated for 30 min using 0.5M NaOH/1.5M NaCl and finally neutralized for 30 min using 0.5M Tris pH 7.4/1.5M NaCl. Each incubation step was performed on a rocking platform and the gel was washed with $H_2O$ between every step. The DNA was then transferred onto an Amersham Hybond-XL membrane (GE Healthcare, RPN303S). The membrane was equilibrated for 10 minutes in 20x SSC. The blotting took place overnight at room temperature with 20x SSC as a buffer and with three layers of Whatman chromatography paper on top of the membrane. The next day, the membrane was dried for 30 min at 80°C and the DNA then cross-linked to the membrane using a *Stratagene* UV Stratalinker 1800, using the auto cross-link program. 25ng probes were radioactively labeled with dCTP, [α-32P] (Perkin Elmer, NEG513H250UC) using High Prime (Roche, 11585592001) and then purified using ProbeQuant G-50 Micro Columns (GE Healthcare, GE28-9034-08) according to the manufacturer's instruction. The membranes were prehybridized with Church buffer ("Sodium Phosphate Buffer for Church and Gilbert Hybridization:" 2015a) for 1h at 65°C and then the radioactively labeled probes were added to the solution and incubated overnight at 65°C. The next day, the membranes were washed with low stringency buffer (2xSSC, 0.1% SDS in $H_2O$) and then with high stringency buffer (0.5xSSC, 0.1% SDS in $H_2O$), each for 20 minutes at 65°C. Membranes incubated with the external probe were additionally washed for 20 min at even higher stringency (0.2xSSC, 0.1% SDS in $H_2O$). The membranes were then wrapped into a plastic film and put into an exposure cassette overnight. Visualization was done using a Typhoon Trio phosphorimager (Amersham).

### Embryoid Body Differentiation

Cells were seeded at $4.7x10^4$ cells/$cm^2$ in uncoated 10cm bacterial dishes in DMEM (Thermo Fisher Scientific, 31966021), supplemented with 10% FBS (Thermo Fisher Scientific, 10270-106), 1 mM Sodium Pyruvate (Thermo Fisher Scientific, 11360070), 1x MEM Non-Essential Amino Acids Solution (Thermo Fisher Scientific, 11140050), 50U/ml penicillin/streptomycin (Ibian Tech, P06-07100) and 0.1 mM 2-mercaptoethanol (Thermo Fisher Scientific, 31350010). The medium was changed every other day. After 6 days, embryoid bodies were transferred to 0.2% gelatin-coated 10cm plates and after two additional days, XCI was assessed in the outgrowth.

### Retinoic Acid Differentiation

Cells were differentiated as described previously (Rohwedel, Guan, and Wobus 1999). Briefly, cells were seeded at $5x10^3$ cells/$cm^2$ on 0.2% gelatin-coated 6-well plates in embryoid body

differentiation medium with 1µM retinoic acid (Enzo Life Sciences, BML-GR100-0500) added to induce differentiation. The medium was changed daily, and cells were cultured up to 12 days.

## Reprogramming Cassette

### Integration of the Cassette

An all-in-one gene targeting vector with doxycycline-inducible reprogramming factors, MKOSimO neotk rtTA Sp3 (Chantzoura et al. 2015), was targeted into the third intron of the *Sp3* gene in the ESC line EL16.7 TST X-GFP using CRISPR-Cas9. $5x10^6$ ESC were mixed with 3.8 µg circularized targeting vector and 2.5 µg single guide RNA vector PX459 (5'-GTGACAATCTCCGGAAAGCG-3') and nucleofected with the AMAXA Mouse Embryonic Stem Cell Nucleofector Kit (LONZA, VPH-1001) using program A-24. 7.5 µM RS-1 (Merck, 553510) was added to enhance homology-directed repair. Cells were selected with 300 µg/ml G418 for 5 days. Clones were selected for expression of mOrange upon the addition of 1 mg/ml doxycycline for 24 hours and then screened by Southern blot hybridization.

### Verification of Cassette Integration Using Southern Blot

The external southern blot probe was PCR amplified from genomic DNA (Forward: 5'-ACAGTTAGGAGGAGCACCAA-3', Reverse 5'-GGCTGCTGAGAATCTTGAAG-3') to give a 918bp fragment. Southern blot was performed as described for X activity reporters except for the following modifications. Genomic DNA was digested with 20U SacI-HF (New England Biolabs, R3156S) and agarose gel separation was performed for 12h.

### mOrange Knockout

Knockout of mOrange was generated using CRISPR-Cas9. $5x10^6$ ESC were mixed with 1.8 µg single guide RNA vector PX459 V2 (Addgene, 62988) (5'-CAACGAGGACTACACCATCG-3') and nucleofected with the AMAXA Mouse Embryonic Stem Cell Nucleofector Kit (LONZA, VPH-1001) using program A-30. The next day, 2 µg/ml Puromycin was added for 24h to select for transfected cells. mOrange-negative cells were isolated by FACS using a BD Influx and single clones were screened for maintenance of proper cassette expression by quantitative RT-PCR.

## P-RFP Pluripotency Reporter

Lentivirus, encoding the mouse *Nanog* promoter driving RFP expression, was purchased from System Biosciences Inc (SR10044VA-1). EL16.7 TST X-GFP MKOS ESCs were seeded at $1x10^4$ cells/$cm^2$ on 6-well plates. 5h later, cells were infected at an MOI of 30 in the presence of 10µl/ml polybrene (Sigma-Aldrich, TR-1003-G). Three days later, RFP-positive cells were FACS purified using a BD Influx. Single clones were picked a week later and selected based on proper RFP expression using FACS analysis on a BD LSRFortessa.

## Neural Precursor Cell Differentiation

ESCs were differentiated to neural precursor cells (NPCs) as described previously (Abranches et al. 2009). ESCs were seeded at a density of $2.75x10^5$ cells/$cm^2$ in 2i/LIF medium, N2B27 (50% DMEM/F12 (Thermo Fisher Scientific, 21041025), 50% Neurobasal medium (Thermo Fisher Scientific, 12348017), 1x N2 (Thermo Fisher Scientific, 17502048), 1x B27 (Thermo Fisher Scientific, 12587001)) supplemented with 0.4 µM PD0325901 (Selleck Chemicals, S1036-5mg), 3 µM CHIR99021 (Sigma-Aldrich, SML1046-5MG) and 1,000 U/ml LIF (ORF Genetics, 01-A1140-0100). 24 hours later, cells were dissociated using Accutase (Thermo Fisher Scientific, 00-4555-56), diluted in wash buffer (DMEM/F-12, no phenol red (Thermo Fisher Scientific, 21041025), 10mM HEPES (Thermo Fisher Scientific, 15630049), 0.5% BSA (Sigma-Aldrich, A7979-50ML) and plated at $2.95x10^4$ cells/$cm^2$ in RHB-A (Takara Bio, Y40001) on 0.2% gelatin-coated T75 flasks, changing media every other day. On days 6 and 8, media was supplemented with 10 ng/ml EGF (R&D Systems, 236-EG-200) and 10 ng/ml bFGF (Thermo Fisher Scientific, 13256029) and additionally with 10µM ROCK inhibitor (Sellekchem, S1049) on day 8. On day 9, cells were dissociated using 2 ml Accutase per flask (Thermo Fisher Scientific, 00-4555-56), diluted in wash buffer, passed through a 40µM filter and centrifuged for 5 min at 900 rpm. SSEA-1 expressing cells were then removed by MACS sorting using Anti-SSEA-1 (CD15) MicroBeads (Miltenyi Biotech, 130-094-530). Cells were incubated for 15 min in the fridge with MicroBeads, followed by an additional 15 min of staining with SSEA-1 eFluor 660 (Thermo Fisher Scientific, 50-8813-42). Cells were then washed once with MACS buffer (0.5% BSA, 2mM EDTA in PBS), passed through a 40µM filter and centrifuged for 5 min at 900 rpm. Cells were then passed through an LS column (Miltenyi Biotec, 130-042-401), washed three times with MACS buffer and the flow-through collected. Cells were then resuspended at $4x10^6$ cells/ml in NPC FACS buffer (50U/ml penicillin/streptomycin, 0.5% BSA, 10mM HEPES in RHB-A), which for critical samples was additionally added with EGF, FGF, and ROCKi. SSEA1-/P-RFP-/X-GFP- cells were then FACS purified using a BD FACSAria II SORP or a BD Influx (BD Biosciences) at a maximum flow rate of 4,000 ev/s to improve survival. FACS sorted cells were seeded at $3.5x10^5$ cells/$cm^2$ on 0.2% gelatin-coated dishes in RHB-A, supplemented with EGF, FGF, and ROCKi. The medium was changed daily until day 12 when cells reached 100% confluency.

## Reprogramming of Neural Precursor Cells

Reprogramming of day 12 neural precursor cells was induced by the addition of 1 mg/ml doxycycline (Tocris, 4090/50) and 25 mg/ml L-ascorbic acid (Sigma-Aldrich, A7506-25G) to the NPC medium (RHB-A supplemented with EGF and FGF). 24 hours later, cells were dissociated using Accutase and seeded on irradiated mouse embryonic fibroblasts (iMEF) in ESC medium containing 15% FBS and supplemented with 1 mg/ml doxycycline and 25 mg/ml L-ascorbic acid. The optimal seeding density of NPCs for reprogramming was $1x10^5$ cells/$cm^2$ for assessing day 4 of reprogramming, with a 50% reduction in cell numbers for every consecutive day. The medium was changed every other day. To isolate iPSC, SSEA1+/P-RFP+/X-GFP+ cells were

isolated using FACS at day 10 of reprogramming and re-plated on 0.2% gelatin-coated plates in ESC medium and kept in doxycycline free conditions for 5-6 days.
Additional experiments, comparing the effect of small molecules on reprogramming efficiency were performed by addition of 250nM TGF-b RI Kinase Inhibitor II (also known as ALK5 Inhibitor II, SB431542) (Enzo Life Sciences, ALX-270-445-M001) and 3μM GSK3β antagonist/Wnt activator CHIR99021 (Sigma-Aldrich, SML1046-5MG) from day 0 of reprogramming.
Experiments inhibiting epigenetic modifiers were performed by addition of 1μM Vorinostat (non-selective HDAC inhibitor) (Sigma-Aldrich, SML0061-5MG), 2μM GSK126 (EZH2 inhibitor) (Cayman Chemical, 15415) and 0.1μM or 0.5μM 5-Aza-2′-deoxycytidine (DNA methylation inhibitor) (Sigma-Aldrich, A3656-5MG) from day 2 of reprogramming.

## RNA-Fluorescent *In Situ* Hybridization

Strand-specific RNA FISH was performed with fluorescently labeled oligonucleotides (IDT) as described previously (Del Rosario et al. 2017). Briefly, cells were fixed with 4% paraformaldehyde for 10 minutes at room temperature and then permeabilized for 5 minutes on ice in 0.5% Triton-X. 10ng/ml equimolar amounts of Cy5 labeled Xist probes BD384-Xist-Cy5-3'-AM (5'-ATG ACT CTG GAA GTC AGT ATG GAG /3Cy5Sp/ -3') and BD417-5'Cy5-Xist-Cy5-3'-AM (5'- /5Cy5/ATG GGC ACT GCA TTT TAG CAA TA /3Cy5Sp/ -3') were hybridized in 40% formamide, 10% dextran sulfate, 2xSSC pH 7 at room temperature overnight. Slides were then washed in 30% formamide 2xSSC pH 7 at room temperature, followed by washes in 2xSSC pH 7 and then mounted with Vectashield (Vector Laboratories, H1200). Images were acquired using an EVOS and a Cy5 light cube (Thermo Fisher Scientific).

## RNA Isolation and Quantitative RT-PCR

RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, 74136) or RNeasy Micro Kit (Qiagen, 74004) and quantified by Nanodrop. cDNA was produced with a High-Capacity RNA-to-cDNA Kit (Thermo Fisher Scientific, 4387406) and was used for qRT-PCR analysis in triplicate reactions with Power SYBR Green PCR Master Mix (Thermo Fisher Scientific, 4367659).

**Table 2. Quantitative RT-PCR Primers Used in this Study**

| Target Transcript | Forward (5'-3') | Reverse (5'-3') |
|---|---|---|
| *c-Myc-Klf4* | AGGAAACGACGAGAACAGTTGA | GACGCAGTGTCTTCTCCCTTC |
| *Esrrb* | CAAGAGAACCATTCAAGGCAACA | CATCCCCACTTTGAGGCATTT |
| *Gapdh* | ATGAATACGGCTACAGCAACAGG | CTCTTGCTCAGTGTCCTTGCTG |
| *Klf2* | TCGAGGCTAGATGCCTTGTGA | AAACGAAGCAGGCGGCAGA |
| *Klf4* | TGGTGCTTGGTGAGTTGTGG | GCTCCCCCGTTTGGTACCTT |
| *Klf4-Oct4* | GGACCACCTTGCCTTACACAT | GAAGCTTAGCCAGGTTCGAGA |
| *Nanog* | CTTTCACCTATTAAGGTGCTTGC | TGGCATCGGTTCATCATGGTAC |
| *Nestin* | TAGCCCTACCACTTCCTGCT | GAGGTGACCCTTGGGTTAGA |
| *Oct4* | GATGCTGTGAGCCAAGGCAAG | GGCTCCTGATCAACAGCATCAC |
| *Oct4-Sox2* | ACCACACTCTACTCAGTCCCT | AGCTCCGTCTCCATCATGTT |
| *Prdm14* | AGCACCCAACCGACTTACAG | AAGTGTGGCACATCACCAAA |
| *Rex1* | CCCTCGACAGACTGACCCTAA | TCGGGGCTAATCTCACTTTCAT |
| *rtTA* | CTACCACCGATTCTATGCCCC | CGCTTTCGCACTTTAGCTGTT |
| *RFP* | GTCCCTCGTCAGGGAATCTTG | GGGATCGGACTTTGCCTTGTA |
| *Sox2* | CATGAGAGCAAGTACTGGCAAG | CCAACGATATCAACCTGCATGG |
| *Tsix* | TGGGTCATTGGCATCTTAGTC | CCCAGGGTGTCTGATCTCTT |
| *Xist (mus)* | ATCATACTAAAGGCCACACAAAGAAT | ATTTGGATTGCAAGGTGGAT |

# Chapter 2 - Dynamics of X-Reactivation during iPSC Reprogramming

*The following chapter contains a soon to be submitted manuscript.*

## Integrative Analysis of X-Chromosome Reactivation Kinetics

Moritz Bauer[1, 2]*, Enrique Vidal[1, 2*], Stefan Pinter[3], Guillaume Filion[1, 2]† & Bernhard Payer[1,2]†

[1]Gene Regulation, Stem Cells and Cancer Program, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain
[2]Universitat Pompeu Fabra (UPF), Barcelona, Spain
[3]Department of Genetics and Genome Sciences, Institute for Systems Genomics, University of Connecticut Health Center, USA

* These authors contributed equally
†Lead contacts and correspondence

# Introduction

In mammals, females have two X chromosomes but males have only one. The chromosomal imbalance is resolved by inactivating one of the two X chromosomes specifically in females. While the active X chromosome resembles in many aspects an autosome, the inactive X has a unique repressive configuration, which sets it apart from other chromosomes (Payer 2017; Galupa and Heard 2018; Jégu, Aeby, and Lee 2017). This repressed state integrates multiple layers of epigenetic gene regulation under the control of the long non-coding RNA Xist. The Xist RNA coats the X from which it is expressed and sets up a heterochromatic environment by recruiting repressor complexes. Genes are silenced as they acquire repressive histone marks and their promoters are DNA-methylated, meanwhile the chromosome folds into a unique conformation consisting of two mega-domains, while lacking topologically associating domains (TADs) (Giorgetti et al. 2016; Rao et al. 2014; Anand Minajigi et al. 2015; X. Deng et al. 2015). Therefore, the inactive X is safeguarded from accidental reactivation on multiple levels (Csankovszki, Nagy, and Jaenisch 2001; Anand Minajigi et al. 2015) in order to avoid gene dosage imbalances, which can lead to lethality or cancer (Yildirim et al. 2013; Chaligné et al. 2015; Pageau et al. 2007; Marahrens et al. 1997b).

In order to reset the epigenetic memory of the inactive state, the X chromosome needs to be reactivated (Payer 2016). While X-inactivation has been the subject of intense study for many decades as a classic example of epigenetic gene regulation (Morey and Avner 2011; Jeannie T. Lee 2011; Lyon 1961), X-reactivation has been more elusive and has only been investigated much more recently (Payer and Lee 2014; Pasque and Plath 2015; Cantone 2017). In mice, X-reactivation occurs twice during early development. The first round takes place at the blastocyst stage within the pluripotent epiblast of the inner cell mass (Mak et al. 2004; Okamoto et al. 2004; Borensztein, Okamoto, et al. 2017; Payer et al. 2013). This allows the female embryo to switch from an imprinted form of X-inactivation – whereby the X inherited from the father's sperm is inactivated – to a random form where either X can be inactivated. This leads to genetic mosaicism in females, allowing them to compensate for X-linked mutations and to reduce the severity of X-linked disorders (Migeon 2006). The second round of X-reactivation takes place in primordial germ cells during their migration and colonization of the gonads, ensuring that an active X can be transmitted to the next generation (de Napoles, Nesterova, and Brockdorff 2007; Sugimoto and Abe 2007; Chuva de Sousa Lopes et al. 2008). Mechanistically, X-reactivation in the epiblast and in the germ cell lineage have both common and distinct features. While in both cases, Xist RNA is downregulated, silencing marks like H3K27me3 are erased and genes are reactivated, the kinetics differ greatly, as X-reactivation in the blastocyst epiblast occurs within a day, while it takes several days during germ cell development. This could be due to differences between imprinted and random X-inactivation, as the imprinted form is considered to be less stable (Dubois et al. 2014; Corbel et al. 2013; Hadjantonakis et al. 2001).

X-reactivation also takes place *in vitro*: female embryonic stem cells (ESCs) derived from the blastocyst epiblast and induced pluripotent stem cells (iPSCs) generated from somatic cells through reprogramming have two active X chromosomes (Maherali et al. 2007). Therefore, the

X-reactivation process is linked to de-differentiation into the naïve pluripotent stem cell state (Payer and Lee 2014), but the mechanisms that govern the transition have not yet been elucidated. Although pluripotency factors like PRDM14 (Payer et al. 2013) and chromatin modifiers like UTX/KDM6A (Borensztein, Okamoto, et al. 2017) are known to play a role in X-reactivation, the elicitors of X-reactivation during de-differentiation are still mostly unknown. It is unclear how the interplay of X-chromosome sequence, 3D-structure, chromatin status, and *trans*-acting factors affect the reactivation of X-linked genes. In particular, it is unknown, if the dramatic topological rearrangement of the X chromosome from an inactive state with two mega-domains into an autosome-like active state with TADs (Giorgetti et al. 2016; Rao et al. 2014; Anand Minajigi et al. 2015; X. Deng et al. 2015) occurs before or after X-linked genes become reactivated. This is especially relevant, as cause and effect between chromosome topology and transcriptional activity have been under debate (Stadhouders et al. 2018; Rowley et al. 2017; Hug et al. 2017). Furthermore, it has been proposed that early reactivation of X-chromosomal genes might be driven by transcription factors like Myc (Borensztein, Okamoto, et al. 2017). A temporal analysis of chromatin-opening during X-reactivation would reveal if pioneering transcription factors contribute to gene reactivation. Finally, X-inactivation kinetics are determined by Xist RNA spreading along the X chromosome, which occurs first to actively transcribed gene-rich regions, which are in three-dimensional proximity to the *Xist* locus, followed by more distant, gene-poor regions (Simon et al. 2013; Engreitz et al. 2013). Single-cell RNA-seq analysis of mouse blastocysts suggested that X-reactivation is not simply a reversal of X-inactivation kinetics, but is rather determined by other parameters like the erasure of the H3K27me3 mark (Borensztein, Okamoto, et al. 2017). However, a fine-scale assessment of the exact order of early versus late reactivated regions along the X chromosome during an X-reactivation time-course will be required to gain further insight. In summary, a detailed kinetic epigenomic analysis of the X-reactivation process, which has been missing so far, is critically needed to clarify, which epigenetic events are instructive and which are secondary for X-linked gene-reactivation.

ESC differentiation has been the workhorse of X-inactivation research for decades (Jeannie T. Lee 2011; Morey and Avner 2011). More recently, iPSC reprogramming has proven to be a useful tool to investigate the mechanisms of X-reactivation (Payer et al. 2013; Pasque et al. 2014). Further *in vitro* systems for studying X-reactivation are cell fusion between somatic cells and pluripotent stem cells (Cantone et al. 2017, 2016; Takagi et al. 1983; Tada et al. 2001) and *in vitro* models for germ cell differentiation from ESCs (Hayashi et al. 2012; Ohta et al. 2017; Yamashiro et al. 2018). Overall, their scalability and amenability to genetic manipulation and perturbations make these *in vitro* systems ideally suited to interrogate the X-reactivation process by epigenomic analysis methods.

Here we developed an *in vitro* model to study the reactivation of the X chromosome during reprogramming, with millions of cells transitioning to iPSCs. The unparalleled efficiency and purity of this system allowed us to study for the first time the conformation of the inactive X chromosome during the process. We observe that the inactive X chromosome undergoes large-scale but gradual changes of conformation. The X chromosome adopts an intermediate structure that was not observed during the inactivation process. The reconfiguration of the inactive X chromosome mirrors gradual changes of transcription and chromatin composition

where remodeling proceeds inwards from the chromosome tips to the XIC. Overall, these results provide the first dynamics of the molecular events operating during the reactivation of the X chromosome.

# Results

## Creation of the *Nano*X Reporter Model for X Chromosome Reactivation

Studying the kinetics of X-chromosome reactivation during iPSC reprogramming using conventional mouse embryonic fibroblast (MEF)-based reprogramming systems (Maherali et al. 2007; Payer et al. 2013; Pasque et al. 2014; Janiszewski et al. 2019) has been challenging for three main reasons: First, the fraction of cells that undergo successful reprogramming and X-reactivation is usually very low, so that it has been so far impossible to follow the process using assays that require a large number of input cells, such as Hi-C or mass spectrometry. Second, the conversion is usually asynchronous and cells at different stages are superposed at every time point. Finally, in inbred strain backgrounds, the two X chromosomes are indistinguishable by sequencing-based assays, to the effect that previous knowledge of the process was mostly based on microscopy (Maherali et al. 2007; Payer et al. 2013; Pasque et al. 2014).

We thus set out to create an *in vitro* model that would address all those difficulties. For this, we used the female mouse ESC line EL16.7 TST, which has been derived from an F2 cross of *Mus musculus musculus* with *Mus musculus castaneus (J. T. Lee and Lu 1999; Ogawa, Sun, and Lee 2008)*. As a result, cells contain one X chromosome from *M.m musculus* ($X^{mus}$) and one from *M.m castaneus* ($X^{cas}$), allowing us to distinguish them via sequence polymorphisms that occur on average every 300 bp (Marks et al. 2015). Chromosome 13 is fully heterozygous as well, while the rest of the genome consists of intermingled blocks of *M.m musculus* and *M.m castaneus*. Furthermore, the cell line harbors a *Tsix* truncation (TST) on the $X^{mus}$ chromosome, which leads to nonrandom choice of the $X^{mus}$ chromosome to be inactivated during differentiation (Luikenhuis, Wutz, and Jaenisch 2001; Ogawa, Sun, and Lee 2008). This allowed us to specifically follow the $X^{mus}$ chromosome through the X-inactivation and X-reactivation process, while the $X^{cas}$ remained active throughout and served as our control.

Next, we circumvented the low reprogramming efficiency and X-reactivation asynchrony by engineering the cell line to allow us to isolate a large number of prospective iPSC intermediates that will undergo X-reactivation. In order to isolate cells poised for X-reactivation, we inserted a lentiviral RFP reporter driven by a promoter fragment of Nanog (P-RFP) (System Biosciences). Moreover, to monitor the activity of the X chromosome, we integrated a GFP reporter into the *Hprt* -locus on $X^{mus}$ (X-GFP) (Wu et al. 2014). Furthermore, in order to make our cell line reprogrammable to iPSC fate, we incorporated an optimized all-in-one doxycycline-inducible reprogramming cassette into the *Sp3* locus (Chantzoura et al. 2015). Finally, to consistently obtain a large and homogeneous population of somatic cells for reprogramming, we optimized differentiation of our ESCs into neural precursor cells (NPCs) (Abranches et al. 2009), during which they would undergo X-chromosome inactivation. This would then provide the starting

material for iPSC reprogramming, during which we would be able to study the X-reactivation process. The unique combination of these features is key to the unprecedented purity of cells with ongoing X-reactivation in our model system. The features of this reporter cell line called *Nano*X (Nanog and X Chromosome reporter) are presented in *Figure 1A*.

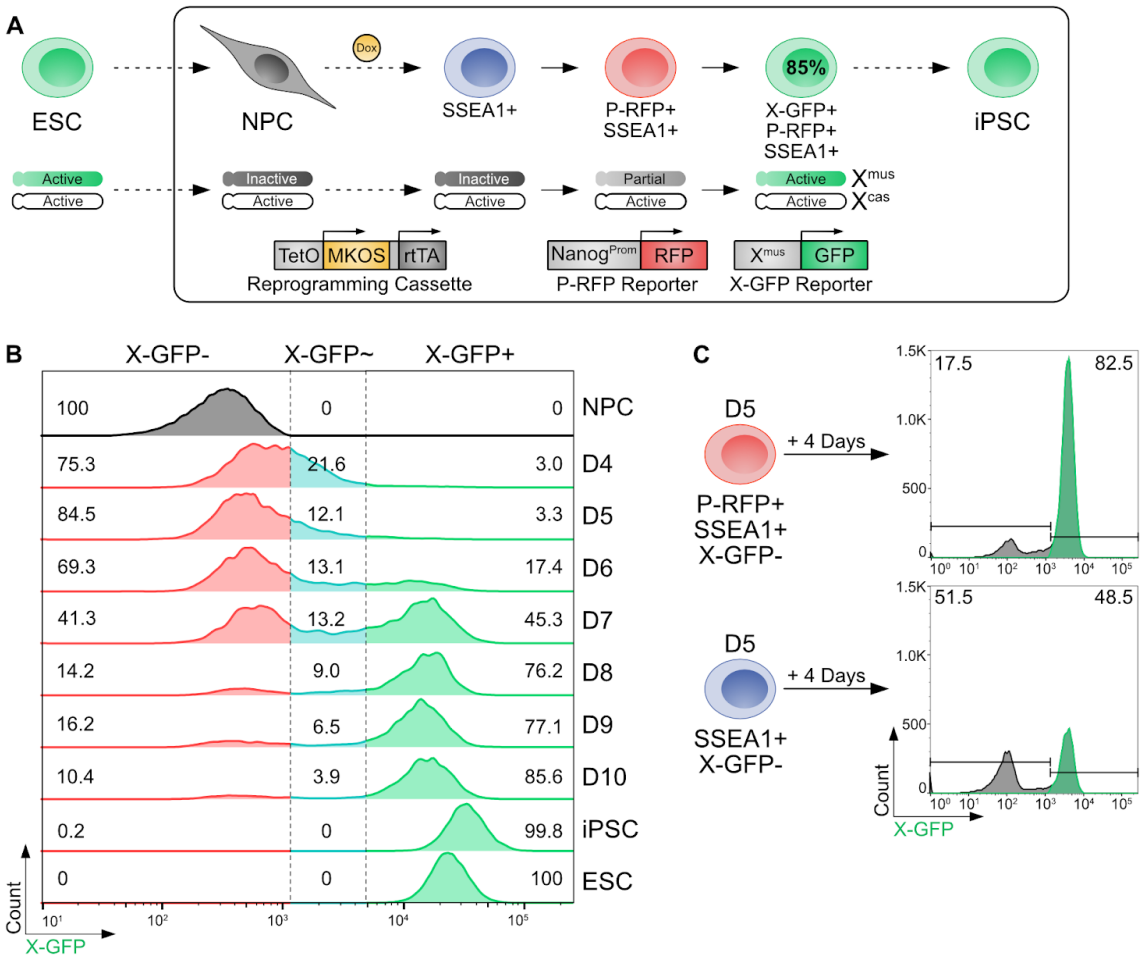



**Figure 1. A Novel Reprogramming System to Trace X-Chromosome Reactivation**

**(A)** Schematic representation of the *Nano*X reprogramming system. **(B)** FACS analysis of X-GFP reporter expression during a reprogramming time course. Shown are representative histograms gated on P-RFP+ cells. Numbers indicate the percentage of cells. **(C)** X-reactivation efficiency of indicated reprogramming intermediates isolated on day 5 and then reprogrammed for an additional 4 days. Shown are representative histograms gated on SSEA-1+ cells. Numbers indicate the percentage of cells.

We tested the capacity of the *Nano*X reporter cell line to efficiently reprogram into iPSCs and found that it faithfully recapitulated the correct order of events during X-reactivation (Schiebinger et al. 2019), with around 15-25% of SSEA1+ cells upregulating P-RFP first (*Figure S1B*) and subsequently up to 85% of P-RFP+ cells upregulating X-GFP (*Figure 1B and Figure S1C*). To test, if P-RFP+ cells represent a population primed for X-reactivation, we isolated SSEA1+/P-RFP-/X-GFP- and SSEA1+/P-RFP+/X-GFP- cells on day 5 and continued

reprogramming. FACS analysis 4 days later showed that while only half of the SSEA1+ cells were able to reactivate the X chromosome, over 80% of P-RFP+ cells successfully reactivated the X (*Figure 1C*). We conclude that our system holds unique advantages to conventional MEF-based reprogramming systems, by allowing us to isolate cells poised for X-reactivation and subsequently achieving near-deterministic efficiency of X-reactivation. Therefore our *Nano*X system allowed us to follow the kinetics of X-reactivation by sorting homogenous cell populations, which is a prerequisite for reliable epigenomic analysis of this process.

## Characterization of NanoX iPSC Reprogramming Shows Rapid X-Reactivation

Utilizing this specialized reprogramming system, we set out to obtain a high-resolution map of X-linked gene reactivation and chromatin opening in relation to the iPSC-reprogramming process. We performed reprogramming of *Nano*X NPCs and employed fluorescence-activated cell sorting (FACS) to isolate iPSC intermediates from day 4 to 10 of reprogramming in a 24h interval. To this end, we performed RNA-sequencing and ATAC-sequencing (Assay for Transposase-Accessible Chromatin with High Throughput Sequencing, see section further below) on a total of 12 distinct subpopulations, capturing 7 consecutive days of reprogramming, as well as fully reprogrammed iPSC clones (*Figure S1A*).

To define the trajectory towards X-reactivation, we performed principal component analysis (PCA) of our RNA-Seq data. As we found that already at D4 (*Figure 2A and 2B*), cells clustered far away from NPCs, we repeated the analysis excluding NPCs in order to get a better resolution along the reprogramming time course (*Figure 2C and 2D*). This allowed us to reveal a more fine-scaled trajectory, starting from D4 SSEA1+ cells, following a path from D4 to D5 and D6 P-RFP+ cells, and leading towards X-GFP+ cells, which closely clustered together.

When we focused our analysis specifically on the X chromosome, we calculated the average allelic expression ratios between the mus and cas alleles of X-linked genes (see methods) (*Figure 2B and 2D*). This revealed a very similar trajectory to the one observed genome-wide (*Figure 2A and 2C*), illustrating the tight correlation of pluripotency and X-chromosome status. Furthermore, as in our reprogramming time course samples from multiple populations would have been superposed on single days, we decided to utilize the inherent asynchronicity of iPSC reprogramming systems, to resolve the key time point of X-reactivation at day 6 in finer detail. We thereby followed the pseudotime-trajectory revealed by the PCA and therefore sorted in our further analysis the samples in order of first D6 X-GFP-, followed by D6 X-GFP~ (intermediate) and then D6 X-GFP+.

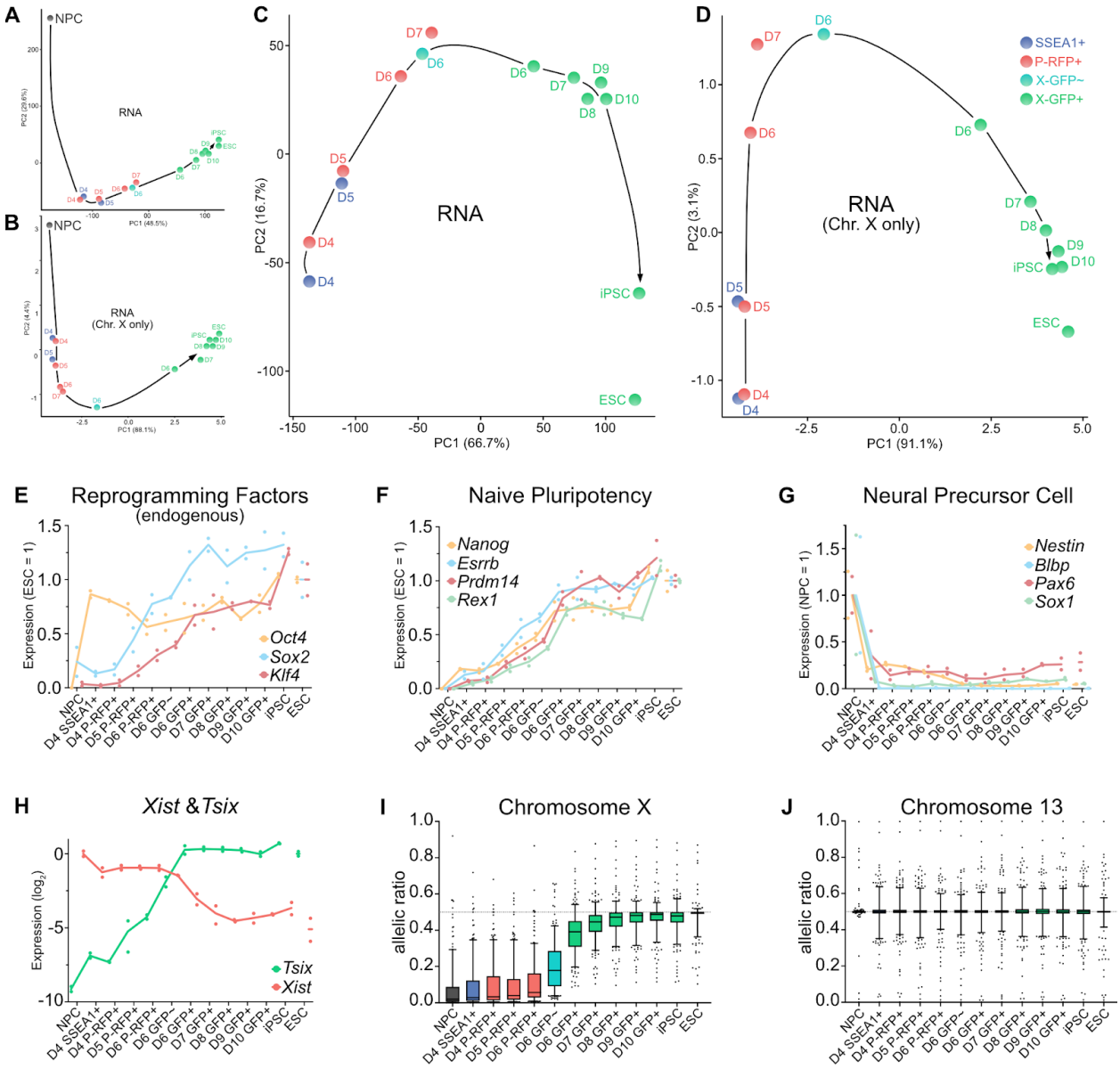



**Figure 2. Characterization of iPSC Reprogramming Dynamics**

**(A)** PCA of dynamics of gene expression during reprogramming (n = 21,314 genes). Black arrow, hypothetical trajectory. **(B)** PCA of dynamics of allele-specific gene expression changes of chromosome X during reprogramming using the allelic ratio (= mus/(mus+cas)) (n = 586 genes). Black arrow, hypothetical trajectory. **(C)** As (A) excluding NPC. **(D)** As (B) excluding NPC. **(E)** Average endogenous gene expression kinetics of *Oct4*, *Sox2,* and *Klf4* during reprogramming (n = 2, relative to the levels in ESC). Endogenous expression calculated via the genes 3'-UTR. **(F)** Average gene expression kinetics of *Nanog*, *Esrrb*, *Prdm14*, and *Rex1* during reprogramming (n = 2, relative to the levels in ESC). **(G)** Average gene expression kinetics of *Nestin*, Blbp, *Map2*, and *Sox1* during reprogramming (n = 2, relative to the levels in NPC). **(H)** Average gene expression kinetics of *Tsix*$^{mus}$ (relative to the levels in NPC) and *Xist*$^{mus}$ (relative to the levels in ESC) during reprogramming (n = 2).

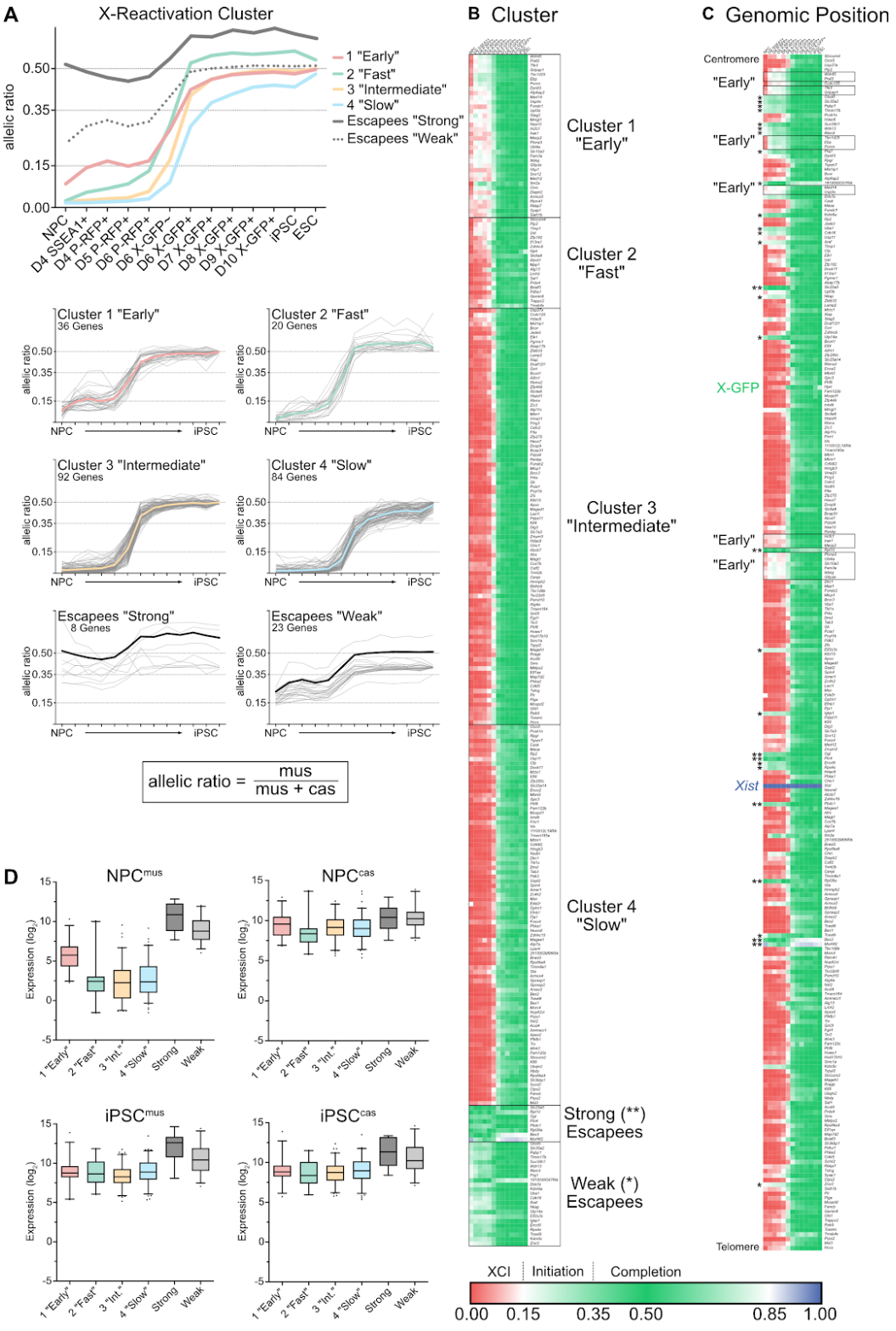
*(Figure legend continues on next page)*

**(I)** Allele-specific expression of genes expressed from chromosome X (n = 308, min. >40 counts per gene). Allelic-ratio = mus/(mus+cas). All box plots depict the 25th to 75th percentiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 5 and 95 percentiles range. **(J)** Allele-specific expression of genes expressed from chromosome 13 (n = 395, min. >40 counts per gene). Allelic ratio = mus/(mus+cas). All box plots depict the 25th to 75th percentiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 5 and 95 percentiles range.

Previous reports have outlined the sequential activation of core pluripotency factors during reprogramming (Stadhouders et al. 2018; Rowley et al. 2017; Hug et al. 2017; Polo et al. 2012). When we followed the trajectory of events identified by the PCA, we found that similarly, we could already detect high endogenous expression levels of *Oct4* at D4, preceding full activation of *Sox2* in D6 P-RFP+ cells and of *Klf4* in D6 X-GFP+ cells (*Figure 2E*). Similar to *Klf4*, when we assessed the expression of key naive pluripotency markers *Nanog*, *Esrrb*, *Prdm14*, and *Rex1*, we found these to be upregulated gradually, already reaching expression levels similar to ESCs in day 6 X-GFP+ cells (*Figure 2F*). Concomitantly, markers of neural precursor cells *Nestin*, *Blbp*, *Pax6*, and *Sox1,* were already strongly downregulated in D4 cells, showing the rapid loss of the somatic gene expression signature (*Figure 2G*). Furthermore, coinciding with the initiation of X-GFP expression, we observed downregulation of *Xist* to about 10% of NPC levels on D6 (*Figure 2H*). *Tsix*, positively correlating with X-reactivation, showed the reciprocal behavior, with a strong upregulation in D6 X-GFP+ cells. As we have been using a *Tsix* truncation (TST) cell line with non-functional *Tsix* on the $X^{mus}$ chromosome, we thereby confirmed our previous observation that *Tsix* is not essential for *Xist*-downregulation during X-reactivation in iPSCs (Payer et al. 2013). *Xist* downregulation during iPSC reprogramming therefore might depend on additional repressive mechanisms including the binding of both core and naive pluripotency factors along their binding hubs at the *X-chromosome inactivation center* (*Xic*) (Payer et al. 2013; Navarro et al. 2011; Gontan, Achame, Demmers, Barakat, Rentmeester, van IJcken, Grootegoed, et al. 2012; Navarro et al. 2008; Payer and Lee 2014).

We next set out to investigate if X-linked gene expression dynamics, followed upregulation of our X-GFP reporter. In order to assess X-linked gene reactivation, we assessed the allelic ratio for 286 genes, which passed our quality control criteria throughout the reprogramming time course (see methods). Indeed, whereas D6 P-RFP+ cells still portrayed levels similar to NPCs, we found a clear switch in D6 X-GFP+ cells, giving an allelic ratio of close to 0.5 (*Figure 2I*), showing reactivation of the inactive X in this population. In contrast, allelic ratios of chromosome 13 remained consistent throughout reprogramming (*Figure 2J*), confirming that the changes observed were X-chromosome specific. To exclude that variation in expression of the $X^{cas}$ is confounding allelic ratio measurements, we separately assessed the expression of $X^{mus}$ and $X^{cas}$ throughout reprogramming and found $X^{mus}$ genes being upregulated whereas no significant change occurred in expression of $X^{cas}$ (*Figure S2A*).

Taken together, while naive pluripotency is acquired gradually during iPSC reprogramming, X-chromosome reactivation occurs rapidly, coinciding with full naive pluripotency gene expression and *Xist* downregulation.




*(Figure legend on next page)*

**Figure 3. Differential Timing of X-Reactivation During iPSC Reprogramming**
**(A)** Allelic ratio (= mus/(mus+cas)) of clusters of X-linked genes during iPSC reprogramming (line shows mean value). **(B)** X-linked genes ordered by reactivation timing categories as in (A). The color gradient represents allelic expression ratio. Red (<0.15), inactive X. White to light green (>0.15, <0.35), reactivation initiation. Green (>0.35), reactivation completion. Blue (>0.85), mono-allelic expression from $X^{mus}$. **(C)** X-linked genes ordered by genomic location on the X chromosome (n = 321). Genes that do not fit any of the clusters are shown as well. Escapees are marked by an asterisk (*), strong**, weak *. Color gradient as for (B). **(D)** Expression of clusters of X-linked genes from $X^{mus}$ and $X^{cas}$. All box plots depict the 25th to 75th percentiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 5 and 95 percentiles range.

## Kinetic Cluster Analysis Reveals Distinct Temporal Waves of Gene Reactivation

Having observed rapid reactivation of the inactive X chromosome, we wanted to assess if this was true for all X-linked genes, or if different reactivation kinetics for clusters of genes were in place. We focussed on a list of 286 X-linked genes, which passed our quality control criteria (see methods) From a list of 558 protein-coding genes, where sufficient sequence polymorphisms were present to allow allele-specific calling, we focussed on 286 genes, which throughout our reprogramming time course, consistently contained over 40 RNA-Seq reads, and were biallelically expressed in iPSC and ESCs (allelic ratio >0.35). To distinguish genes undergoing X-reactivation from genes escaping X-inactivation, we selected an allelic ratio of >0.15 in NPCs to define escapees. We found 31 genes escaping X-inactivation in our system, out of which 21 were previously identified in NPCs as well (Giorgetti et al. 2016; Rao et al. 2014; Anand Minajigi et al. 2015; X. Deng et al. 2015)). However, only 8 out of these 31 genes reached an allelic ratio of >0.35 in NPCs as well, being close to biallelic expression. We, therefore, divided escapees into two groups of strong (allelic ratio >0.35) and weak (allelic ratio >0.15 and <0.35) (*Figure 3A*).
Focussing on genes undergoing X-inactivation during NPC differentiation, we found 232 genes with an allelic ratio of <0.15 in NPCs. Corroborating our previous data (*Figure 2I*), we found that all genes to at least have initiated reactivation (ratio >0.15) in D6 X-GFP+ cells (*Figure 3A*). To distinguish different patterns of reactivation, we performed hierarchical K-means clustering on the allelic ratio of genes undergoing reactivation over the reprogramming time course. This enabled us to reveal four different clusters of XCR genes (*Figure 3A/B*). Cluster 1, termed “early”, was characterized by early initiation of XCR already in D4 P-RFP+ cells (36 genes). Genes in cluster 2 reactivated rapidly, showing allelic ratios of >0.35 in D6 X-GFP+ cells, and was therefore termed “fast” (20 genes). However, the majority of genes were found in clusters 3 and 4. Cluster 3, “intermediate”, featured delayed kinetics compared to “fast” genes, nevertheless with completion being similar to cluster 1 (92 genes). Finally, genes of cluster 4 showed the slowest kinetics of reactivation (“slow”, 84 genes).
When we then assessed heat maps of allelic expression, with genes ordered according to their location along the X chromosome (*Figure 3C*), we found that early initiating genes were not distributed randomly, but rather clustered together, in perceived vicinity to escapees. Interestingly, “early” genes already showed increased expression from the mus allele in NPCs

(*Figure 3D*), suggesting that these might be poised for early initiation. Thereby they showed very similar kinetics in upregulation compared to weak escapees during iPSC reprogramming (*Figure 3A*), suggesting that they might belong to a related category with similar regulatory principles. Alternatively, early genes could be poised for reactivation by the special chromatin environment due to their vicinity to escapee genes. Moreover, considering that we observed early initiation of reactivation for this subset of genes already at D4 when *Xist* was still being highly expressed, we asked if Xist RNA was still coating the inactive X, or if it has already detached from the chromosome. We, therefore, performed Xist RNA FISH on D5 P-RFP+ cells and found that the majority of cells still harbored a single Xist cloud (*Figure S3*). This suggests that early genes might escape Xist based silencing, even when the majority of the X chromosome is still coated by Xist.

In summary, in agreement with X-chromosome reactivation in the mouse blastocyst *in vivo (Borensztein, Okamoto, et al. 2017)*, X-reactivation during iPSC reprogramming in our system occurs rapidly. However, the timing of initiation and completion of reactivation can differ, with early initiation not priming for faster completion of the process. Moreover, whereas completion of reactivation is solely observed in populations showing downregulation of *Xist*, initiation of reactivation of a subset of genes can occur independently of loss of Xist RNA expression and coating, arguing for a two-step process of reactivation.

## Accessibility to the Inactive X Chromosome is Gained in a Stepwise Fashion Around Pluripotency Factor Binding Hubs

A major event during the X-reactivation process is the gain of chromatin accessibility linked to the binding of transcription factors. While the active X chromosome is comprised of open and closed regions with similar proportion to autosomes, the inactive X chromosome is in a closed heterochromatic state almost devoid of open regions ((Giorgetti et al. 2016; Rao et al. 2014; Anand Minajigi et al. 2015; X. Deng et al. 2015)). In order to interrogate the changes in accessibility over time during X-reactivation, we have performed ATAC-Seq on our sorted cell populations during reprogramming. We thereby wanted to gain information on the spatial distribution and kinetics of chromatin opening along the X chromosome to identify potential hubs of early chromatin opening. Specifically, we wanted to assess the enrichment of transcription factor binding motifs within early accessible regions in order to gain mechanistic insights into the chromatin opening process associated with X-reactivation and identify candidate trans-factors potentially important for it.

First, we observed, as expected, that globally the inactive $X^{mus}$ chromosome in NPCs had only around 15% of the number of accessible peaks of the active $X^{cas}$ (*Figure 4A*). The percentage of open sites changed rapidly on day 6 of reprogramming in X-GFP+ cells when the previously inactive $X^{mus}$ reached a similar percentage of open sites to the active $X^{cas}$. In comparison, chromosome 13 did not show allelic differences in chromatin openness, with only minor changes in the number of accessible sites throughout the time course of the experiment (*Figure 4A*). To further substantiate this finding, we assessed the correlation of populations throughout the time course for $X^{mus}$ and $X^{cas}$ (*Figure 4B*). We found that whereas the majority of changes on the active $X^{cas}$ occurred already between NPCs and D4 SSEA1+ cells, the inactive $X^{mus}$

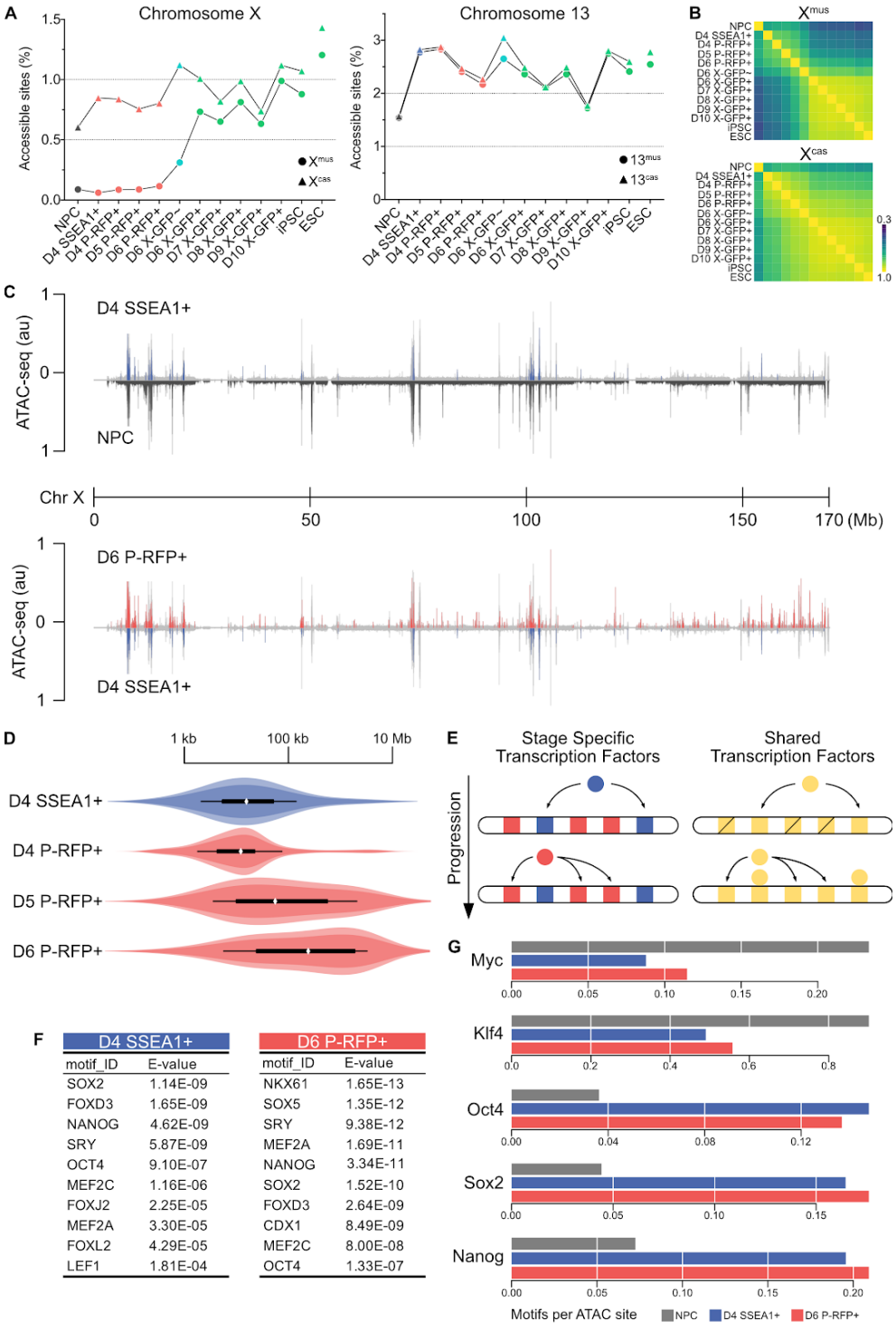
displayed distinct kinetics, gradually changing between NPCs and D6 X-GFP~, followed by an abrupt shift in the D6 X-GFP+ population marking a breakpoint, from which on little change was observed. This confirms that the opening of the X chromosome during reprogramming involves small but gradual changes during its initiation period that culminated in a rapid switch, which is not generally observed on autosomes.

We then wanted to look into more detail where the X chromosome gained accessibility first (*Figure 4C*). Interestingly, we found already in NPCs clustered regions of openness, which correlated in location with escapee genes from our gene expression data. Already on day 4 of reprogramming in the SSEA1+ subpopulation, we observed the gain of newly opened regions, which were consistently in the vicinity of the open escapee regions in NPCs (*Figure 4C and 4D*). On day 6, the time when we observed dramatic changes in chromatin opening within our X-reactivating subpopulations (*Figure 4A and 4B*), we found that in the P-RFP+ population newly opening regions not only appeared next to already open sites but did appear as well at a further distance from them (*Figure 4C and 4D*). This indicates that chromatin accessibility on the inactive X is regained in a two-step manner: Chromatin opening initiates on day 4 near escapee regions, which prime their surroundings to become accessible. This is followed by later regions on day 6, which either appear near existing open regions, but can also form at a distance independently from previously open chromatin.

Next, we wanted to see, if the presence of specific transcription factor binding sites could explain this two-step model of chromatin opening. We thereby considered two alternative hypotheses (*Figure 4E*):

In the first, stage-specific transcription factors become expressed at different time-points during reprogramming and bind to all their binding sites once they become expressed. Thus, different stage-specific transcription factors open different regions at different time points. In the second hypothesis, shared transcription factors, which are expressed throughout different stages of reprogramming could gain accessibility to different sites at different time-points, depending on other parameters like for example chromatin status or cooperative binding with other co-factors. In order to test these competing hypotheses, we looked for enrichment of transcription factor binding motifs at newly gained sites at different time points (day 4 and day 6, *Figure 4F*). Through this analysis, it became apparent that on both time points pluripotency factors were highly enriched, including OCT4, SOX2, and NANOG. When comparing the enrichment of pluripotency factor motifs on newly gained sites on day 4 and day 6, we did not detect apparent differences (*Figure 4G*). Neither did we detect any apparent motifs, which were exclusively enriched on early (day 4) versus later (day 6) sites. Although we cannot exclude the binding of stage-specific pioneer factors with undefined binding motifs, this led us to conclude that based on known binding motifs, both early and late sites shared the same transcription factor binding sites. In particular, as we expressed the pluripotency factors OCT4, SOX2, KLF4, and C-MYC from our reprogramming cassette throughout the reprogramming time-course, which all showed the same enrichment at early and late sites (*Figure 4G*), this speaks in favor of the shared transcription factor model (*Figure 4E*). Not the stage-specific expression of pioneering factors, but rather other parameters like chromatin status seem to determine which regions become accessible first and which later during X-reactivation. Nevertheless, core pluripotency factors, which have been described to have pioneering activity (Soufi, Donahue, and Zaret 2012; Soufi

et al. 2015) and bind the genome in a cooperative manner during different stages of iPSC reprogramming (Chronis et al. 2017; Knaupp et al. 2017; Li et al. 2017), might be key in the chromatin-opening process during X-reactivation due to their enrichment at opening sites at different stages, even if they don't control its exact timing.




F

| D4 SSEA1+ | | D6 P-RFP+ | |
|---|---|---|---|
| motif_ID | E-value | motif_ID | E-value |
| SOX2 | 1.14E-09 | NKX61 | 1.65E-13 |
| FOXD3 | 1.65E-09 | SOX5 | 1.35E-12 |
| NANOG | 4.62E-09 | SRY | 9.38E-12 |
| SRY | 5.87E-09 | MEF2A | 1.69E-11 |
| OCT4 | 9.10E-07 | NANOG | 3.34E-11 |
| MEF2C | 1.16E-06 | SOX2 | 1.52E-10 |
| FOXJ2 | 2.25E-05 | FOXD3 | 2.64E-09 |
| MEF2A | 3.30E-05 | CDX1 | 8.49E-09 |
| FOXL2 | 4.29E-05 | MEF2C | 8.00E-08 |
| LEF1 | 1.81E-04 | OCT4 | 1.33E-07 |

*(Figure legend on next page)*

**Figure 4. Dynamics of Chromatin Opening**

**(A)** Percentage of allele-specific accessible sites of Chromosome X and 13 as determined by ATAC-seq. **(B)** Correlation heat maps of chromatin accessibility changes. **(C)** ATAC-seq peaks along the entire length of the $X^{mus}$ chromosome are shown. Colored peaks represent peaks forming for the first time in the specified population compared to the population on the bottom. Light grey depicts all peaks. **(D)** Violin plots showing the average distance of novel peaks to already accessible regions in NPC. **(E)** Two models for transcription factor-induced chromatin opening. **(F)** Transcription factor motif enrichment at novel peaks are shown. **(G)** Number of motifs of indicated transcription factors per ATAC site are shown.

## A Unique Structural Intermediate Forms During X Reactivation

The inactive X adopts a unique 3D conformation, without evident compartments or TADs (Giorgetti et al. 2016; Rao et al. 2014; Anand Minajigi et al. 2015; X. Deng et al. 2015). However, how the restructuring of the inactive X correlates with gene reactivation and chromatin opening is currently unknown. We, therefore, wanted to interrogate the structure of the inactive X during the process of X-reactivation. We performed cell cycle-specific *in situ* Hi-C on NPCs, ESCs as well as D5 P-RFP+ iPSCs (*Figure S4*). We decided on this specific iPSC time point, as it marked a cell population we have shown to be poised for X-reactivation (*Figure 1C*), with the initiation of reactivation having commenced already.

When we assessed Hi-C matrices from NPCs, we could recapitulate the previously described mega-domain structure of the inactive X *(Figure 5A/B*) (Giorgetti et al. 2016; Rao et al. 2014; Anand Minajigi et al. 2015; X. Deng et al. 2015). Moreover, TADs were visibly absent as expected for the inactive X. When we compared matrices obtained for D5 iPSCs to NPCs and ESCs, it became apparent, that they still strongly resembled the inactive X of NPCs. Nevertheless, upon closer inspection, we noticed that while the inactive X was still partitioned into two mega-domains, we observed regions where TADs started to appear (*Figure 6A/B*). Intriguingly, whereas we could observe TAD formation in regions we have shown to harbor early initiators of reactivation (*Figure 6A*), we were able to additionally identify regions of early TAD formation, which lacked clusters of early initiation of fast reactivating genes (*Figure 6B*).

While a more detailed analysis of this unique intermediate structure will be necessary, we can conclude that mega-domains and TADs, similarly to during X-inactivation (Colognori et al. 2019), are able to coexist during X-reactivation, arguing that these correspond to two different levels of three-dimensional genome organization superimposed on the Xi. Moreover, early TAD formation wasn't exclusively observed in regions of early initiation of gene reactivation, arguing that chromosomal restructuring and gene expression might be separate events.

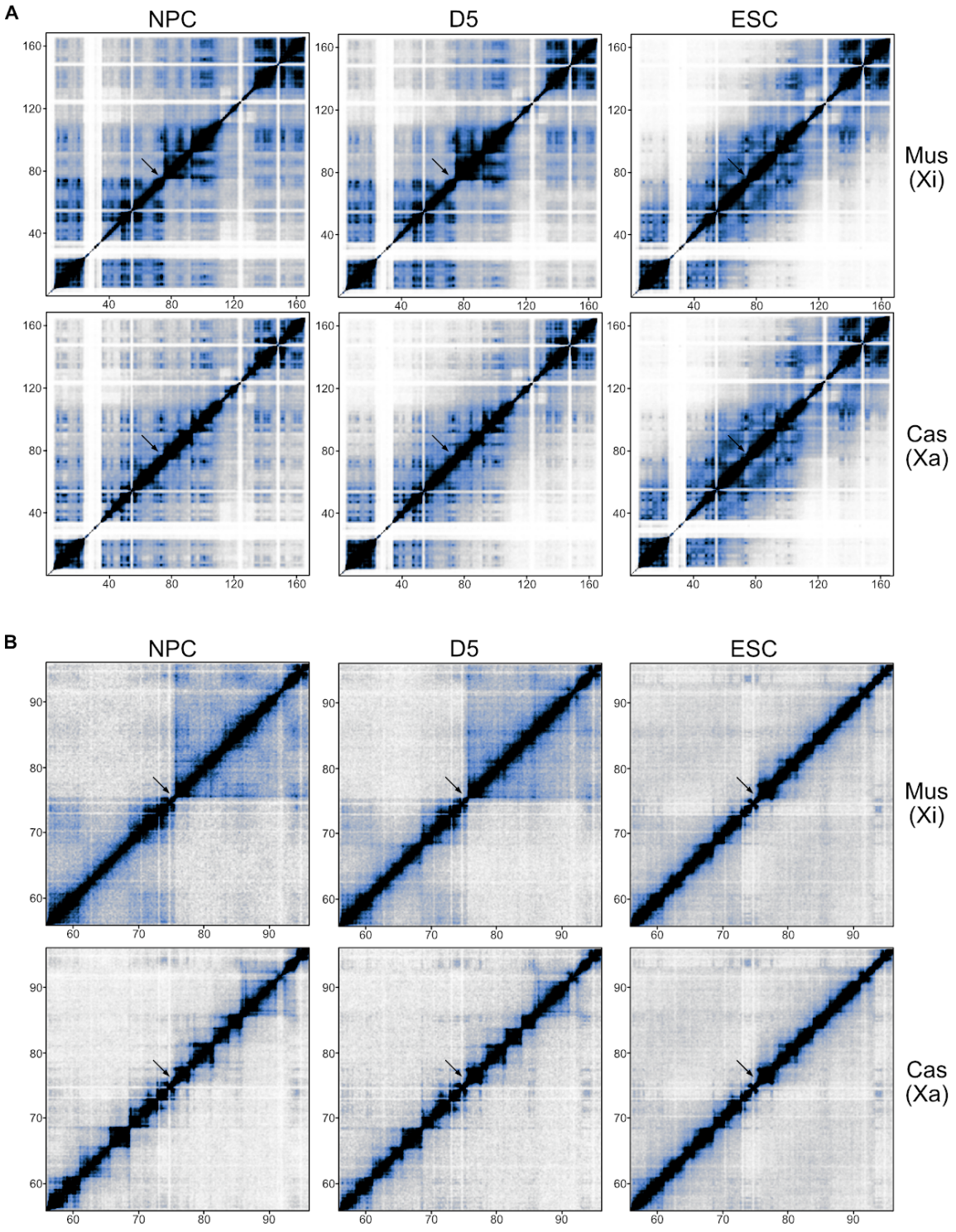



**Figure 5. Overview of Chromosome Structure.**

**(A)** Allele-specific Hi-C maps of the Xi (mus) and Xa (cas) chromosomes. Entire chromosome is shown. **(B)** As (A). Zoom-in of the mega-domain boundary is shown. Scale is shown in mega-bases (Mb). Mega-domain border is indicated by a black arrow.

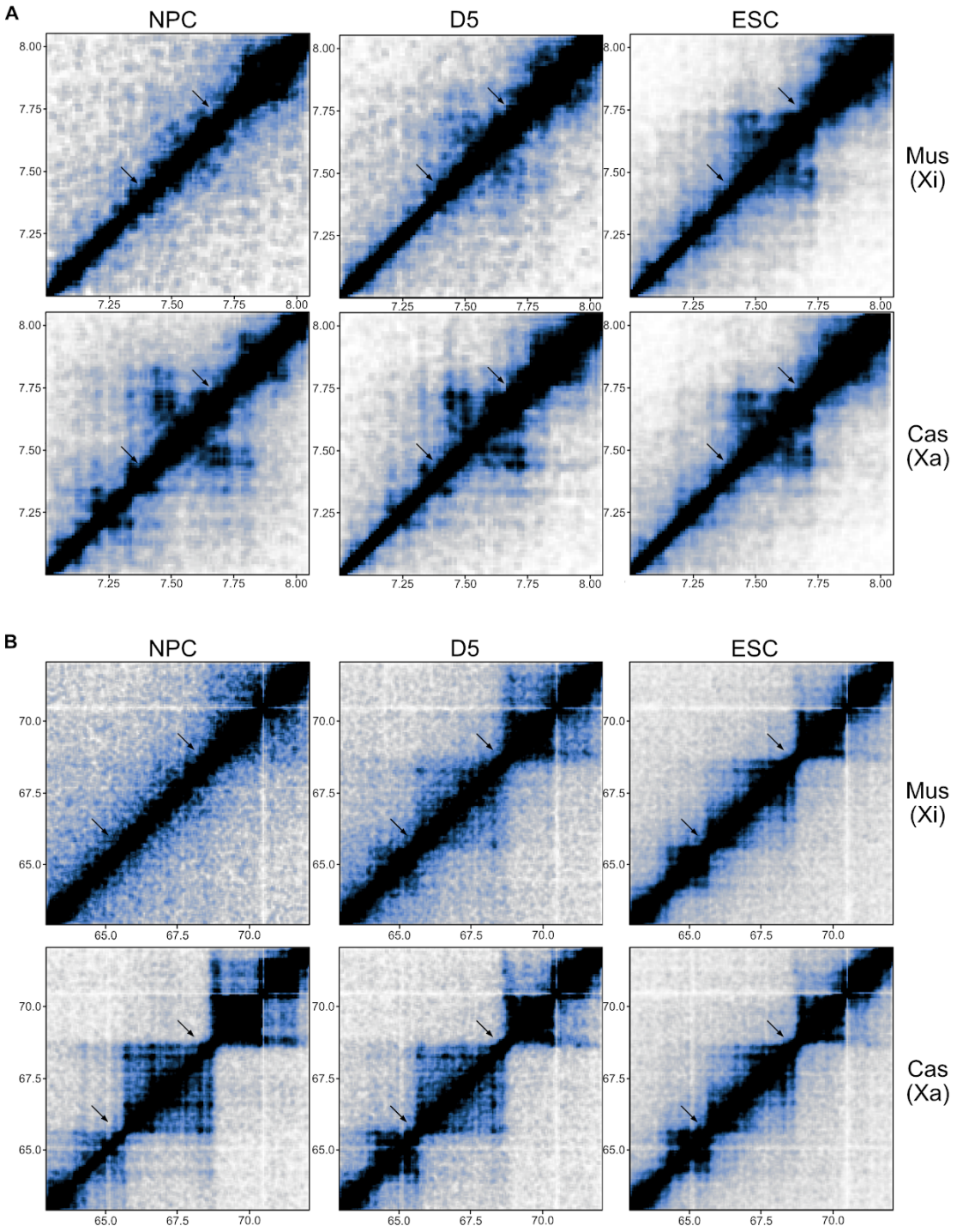



**Figure 6. Early TAD Formation in the Presence of Mega-Domains**

**(A)** Allele-specific Hi-C maps of the Xi (mus) and Xa (cas) chromosomes. Zoom-in of an early reactivating cluster is shown. **(B)** As (A). Zoom-in of an Early forming TAD is shown. Scale is shown in mega-bases (Mb). TAD borders are indicated by black arrows.

# Discussion

Here we have shown that utilizing a novel iPSC system, we were able to obtain a high-resolution map of gene reactivation and chromatin opening of the inactive X and relate it to an intermediate 3D conformation. It has been shown previously that X-reactivation *in vivo* occurs rapidly (Borensztein, Okamoto, et al. 2017). Similarly, we found that genic reactivation during iPSC reprogramming occurs rapidly, taking approximately 24 hours. However, in contrast to previous studies, we were able to distinguish the initiation and completion of reactivation. In combination with clustering analysis, this revealed that the majority of X-linked genes initiates and completes reactivation synchronously. Nevertheless, subsets of genes were either able to initiate reactivation earlier or complete the process faster. Importantly, early initiation didn't prime for fast reactivation, nor did higher expression. However, early initiating genes did show increased expression in the X-inactive state. This might argue that higher expression primes for earlier initiation. Nevertheless, why this doesn't affect reactivation completion remains to be explained. As multiple levels of repressive chromatin modifications, most notably DNA methylation and H3K27me3 methylation and the absence of histone acetylation, are in place to prevent spurious reactivation of X-linked genes in somatic cells (Csankovszki, Nagy, and Jaenisch 2001), we could speculate that higher expression levels are sufficient to remove a subset of these layers, which would allow for genes to initiate reactivation early. However, completion of the process might require the removal of additional repressive layers. Yet, which repressive modifications would possibly correspond to early or late removed layers remains to be shown. Moreover, when X-reactivation initiated, we still observed a cloud of Xist marking the inactive X. Whereas XCR initiation can, therefore, commence in the presence of Xist, or if Xist RNA is lost locally similar to escapees, remains to be shown.

Gene expression and chromatin opening are two tightly intertwined events. However, it is currently unknown if gene reactivation and chromatin opening follow similar trajectories during X-reactivation. Our analysis of the dynamic landscape of accessible chromatin of the X revealed that similar to gene expression, chromatin opening occurs rapidly. Moreover, we were able to show that the initial chromatin opening proceeds from primary sites in close proximity to regions already accessible in NPCs. Only later during our time course, secondary sites further from already open regions became accessible. Intriguingly, early initiating genes were found to lie in close proximity to genes escaping X-inactivation, which overlapped with regions of early chromatin opening. This argues that gene reactivation and chromatin opening are tightly linked during X-reactivation and proceed from hubs of higher chromatin accessibility.

*In vivo*, X-reactivation occurs specifically in cells of the inner cell mass, where pluripotency factors are expressed. However, the requirement of pluripotency factors for the different stages of reactivation is not well understood yet. Analyzing the expression of pluripotency factors, we found that in contrast to X-reactivation that occurs rapidly, pluripotency factor expression increased gradually. We observed that whereas endogenous *Oct4* expression already reached peak levels at D4 of reprogramming, naive pluripotency factor expression like *Nanog* or *Prdm14,* only reached full expression levels in cells having completed X-reactivation at D6. Nonetheless, when we analyzed transcription factor motifs of early accessible regions on the Xi,

we found significant enrichment for pluripotency factors OCT4, SOX2, and NANOG. Moreover, all three were not only enriched in primary accessible sites close to escapee regions but as well in secondary regions further away. Notably, OCT4 and SOX2 have previously been described to cooperatively bind and open closed chromatin, a mechanism termed pioneering activity (Peñalosa-Ruiz et al. 2019). As we overexpress OCT4 and SOX2 from the beginning from the reprogramming cassette, these factors might, therefore, be involved in chromatin opening and gene reactivation at all stages, however, they might rely on the removal of currently unknown layers of repressive chromatin modifications first.

Whereas the initiation of chromatin opening and gene reactivation seem to be tightly linked, the link to genome structure might be less pronounced. Recent evidence suggests that transcription, as well as architectural protein occupancy, determines 3D chromatin organization in most eukaryotes (Rowley et al. 2017). However, if transcription is truly causative in the formation of higher-order 3D genome structures (Rowley and Corces 2018) and especially if common principles extrapolate to the X chromosome, remains to be shown. Whereas we found that one hub of early transcriptional initiation showed the early formation of TADs as well, we were able to identify regions of early TAD formation that didn't overlap with any of our described early initiating regions. Therefore, how genome structure shapes gene expression and vice versa on the X, remains to be explained. Moreover, we were able to show that mega-domains and TADs are able to coexist during the X-reactivation process, arguing that these correspond to two different levels of three-dimensional genome organization superimposed on the Xi.

In conclusion, our study provides unique insights into the dynamics of X-reactivation and establishes it as a powerful system to investigate the formation of euchromatin, and the relationship between gene expression, chromatin accessibility, and genome conformation during the process.

# Supplementary Figures

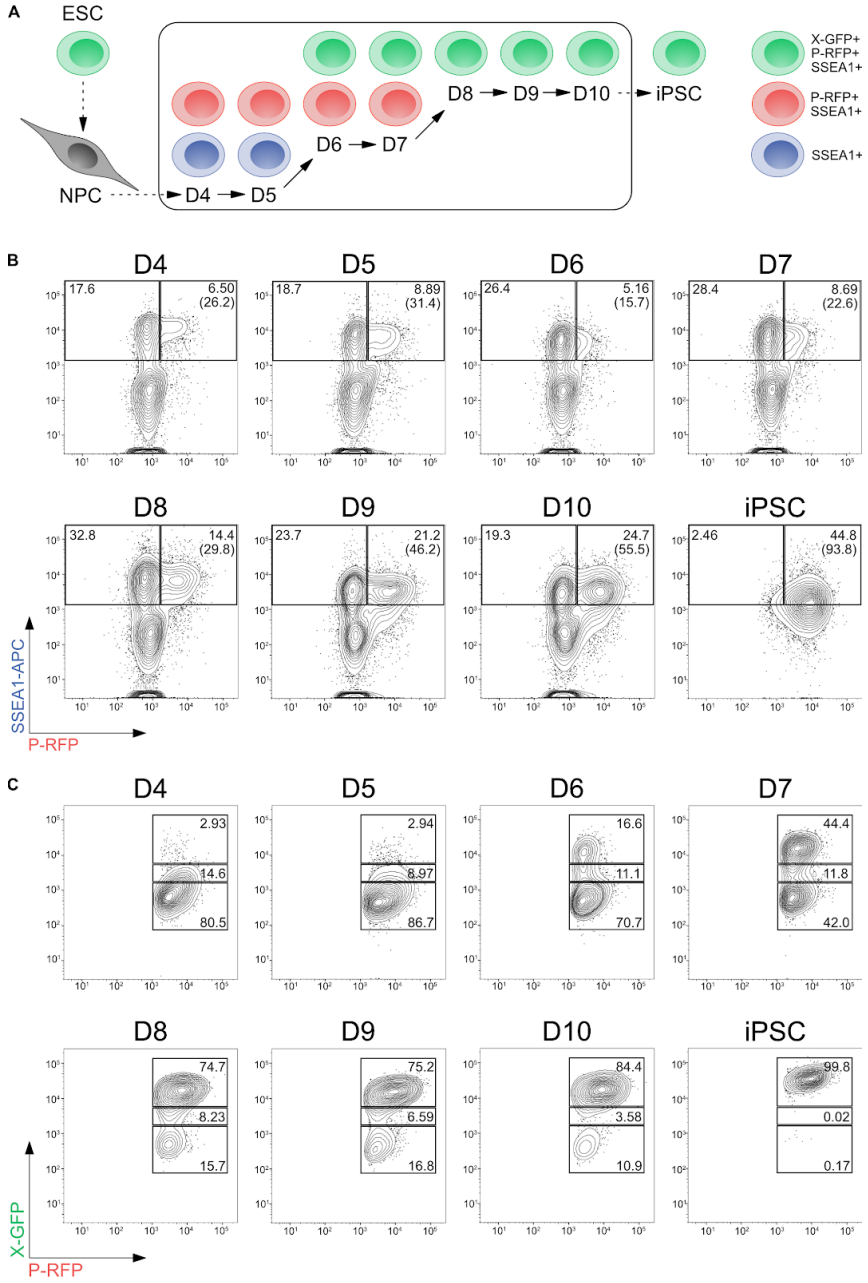



**Figure S1. Cell Populations Sorted During Reprogramming**

**(A)** Schematic representation of the populations sorted during reprogramming. **(B)** FACS gating strategy of P-RFP reporter expression during a reprogramming time course. Shown are representative contour plots gated on live cells. Numbers indicate the percentage of cells. Numbers in brackets indicate the percentage of P-RFP+ cells out of SSEA1+ cells. **(C)** FACS gating strategy of X-GFP reporter expression during a reprogramming time course. Shown are representative contour plots gated on P-RFP+ cells. Numbers indicate the percentage of cells.

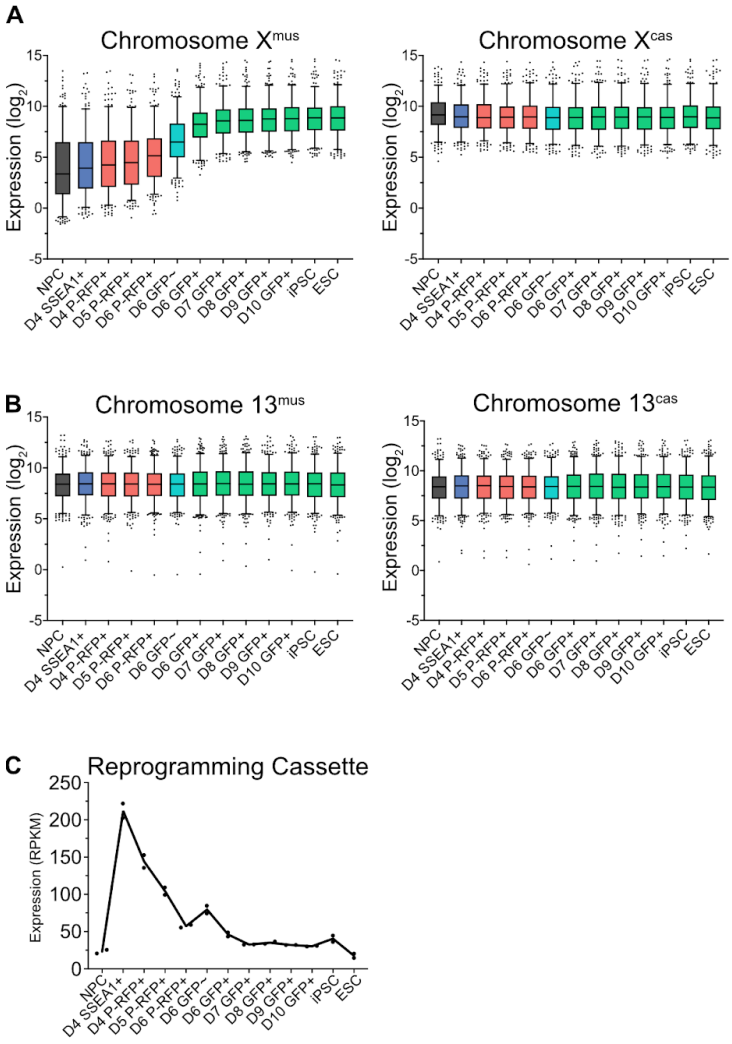



**Figure S2. RNA Expression during Reprogramming**

**(A)** Gene expression from chromosome $X^{mus}$ and $X^{cas}$ (n = 308). All box plots depict the 25th to 75th percentiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 5 and 95 percentiles range. **(B)** Gene expression from chromosome $13^{mus}$ and $13^{cas}$ (n = 395). All box plots depict the 25th to 75th percentiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 5 and 95 percentiles range. **(C)** Average gene expression kinetics of the reprogramming cassette calculated using F2A and T2A containing reads (n = 2).

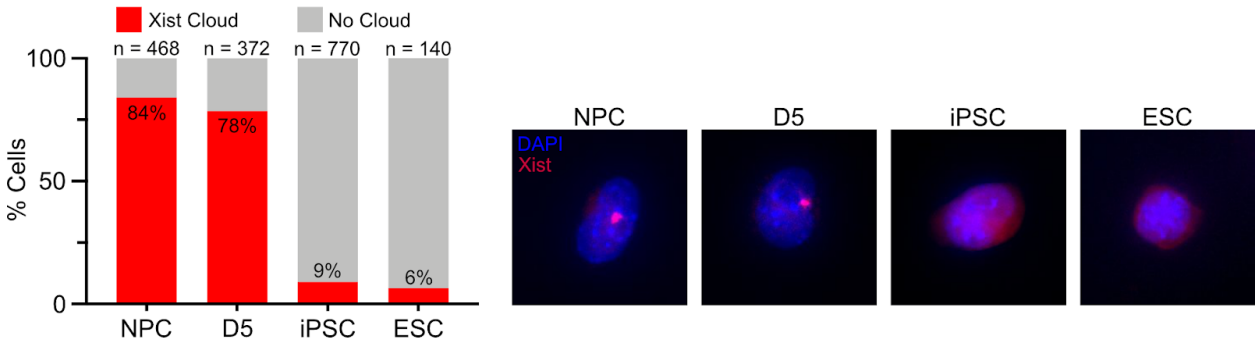



**Figure S3. Xist RNA FISH**

Number of cells containing Xist RNA FISH clouds are shown.

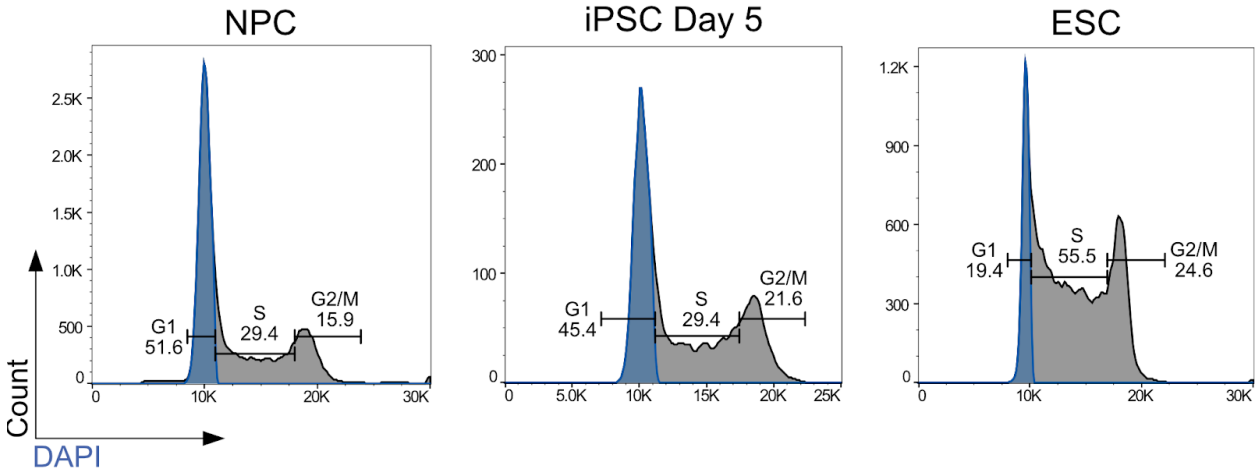



**Figure S4. Cell Cycle Sorting of Hi-C samples**

FACS analysis of cell cycle using DAPI. Blue indicates sorted G1 population. Numbers indicate the percentage of cells.

# Methods

## Embryonic Stem Cell Culture

Mouse embryonic stem cells (ESCs) were cultured on 0.2% gelatin-coated dishes in DMEM (Thermo Fisher Scientific, 31966021), supplemented with 10% FBS (ES-qualified, Thermo Fisher Scientific, 16141079) 1,000 U/ml LIF (ORF Genetics, 01-A1140-0100), 1 mM Sodium Pyruvate (Thermo Fisher Scientific, 11360070), 1x MEM Non-Essential Amino Acids Solution (Thermo Fisher Scientific, 11140050), 50U/ml penicillin/streptomycin (Ibian Tech, P06-07100) and 0.1 mM 2-mercaptoethanol (Thermo Fisher Scientific, 31350010). Cells were incubated at 37°C with 5% $CO_2$. The medium was changed every day and cells were passaged using 0.05% Trypsin-EDTA (Thermo Fisher Scientific, 25300054). Cells were monthly tested for mycoplasma contamination using PCR.

## Generation of Cell Lines

### X-GFP Reporter

We used the female F2 ESC line EL16.7 TST, that was derived from a cross of *Mus musculus musculus* with *Mus musculus castaneus* (Ogawa, Sun, and Lee 2008). As a result, cells contain one X chromosome from *M.m musculus* ($X^{mus}$) and one from *M.m castaneus* ($X^{cas}$). Moreover, EL16.7 TST contains a truncation of *Tsix* on $X^{mus}$ ($Tsix^{TST/+}$), which abrogates *Tsix* expression and leads to the non-random inactivation of $X^{mus}$ upon differentiation.

A GFP reporter construct was targeted into the second exon of *Hprt* on $X^{mus}$ as follows: Homology arms flanking the target site were amplified from genomic DNA and cloned into pBluescript II SK(+) (Addgene, 212205) by restriction-enzyme based cloning and the cHS4-CAG-nlsGFP-cHS4 construct, kindly provided by J. Nathans (Wu et al. 2014), was cloned between the two homology arms.

$5x10^6$ cells were mixed with 1.6 µg circularised targeting vector and 5 µg single guide RNA vector PX459 (Addgene, 48139) (5'-TATACCTAATCATTATGCCG-3'), to achieve an optimal ratio of Cas9 to targeting vector equal to 5:1 (Pinder, Salsman, and Dellaire 2015). Cells were nucleofected with the AMAXA Mouse Embryonic Stem Cell Nucleofector Kit (LONZA, VPH-1001) using program A-30 and 7.5 µM RS-1 (Merck, 553510) was added to enhance homology-directed repair. To select for the disruption of *Hprt*, cells were grown in the presence of 10 µM 6-thioguanine (Sigma-Aldrich, A4882-250MG) for 6 days, and GFP+ cells were isolated by FACS using a BD Influx (BD Biosciences). Single clones were screened by Southern blot hybridization. Inactivation of the X-GFP construct upon differentiation was confirmed using embryoid body differentiation.

pSpCas9(BB)-2A-Puro (PX459) was a gift from Feng Zhang (Addgene plasmid # 48139 ; http://n2t.net/addgene:48139 ; RRID:Addgene_48139)

### Reprogramming Cassette

An all-in-one gene targeting vector with doxycycline-inducible reprogramming factors, MKOSimO neotk rtTA Sp3, kindly provided by K. Kaji (Chantzoura et al. 2015), was targeted into the third intron of the *Sp3* gene in the ESC line EL16.7 TST X-GFP using CRISPR-Cas9. $5x10^6$ cells were mixed with 3.8 µg circularised targeting vector and 2.5 µg single guide RNA vector PX459 (5'-GTGACAATCTCCGGAAAGCG-3') and nucleofected with the AMAXA Mouse Embryonic Stem Cell Nucleofector Kit (LONZA, VPH-1001) using program A-24. 7.5 µM RS-1 (Merck, 553510) was added to enhance homology-directed repair. Cells were selected with 300 µg/ml G418 for 5 days. Clones were selected for expression of mOrange upon the addition of 1 mg/ml doxycycline for 24 hours and then screened by Southern blot hybridization.

Knockout of mOrange was generated using CRISPR-Cas9. $5x10^6$ cells were mixed with 1.8 µg single guide RNA vector PX459 V2 (Addgene, 62988) (5'-CAACGAGGACTACACCATCG-3') and nucleofected with the AMAXA Mouse Embryonic Stem Cell Nucleofector Kit (LONZA, VPH-1001) using program A-30. mOrange-negative cells were isolated by FACS using a BD Influx and single clones were screened for maintenance of proper cassette expression by quantitative RT-PCR.

pSpCas9(BB)-2A-Puro (PX459) V2.0 was a gift from Feng Zhang (Addgene plasmid # 62988 ; http://n2t.net/addgene:62988 ; RRID:Addgene_62988)

### Southern blot

Genomic DNA (10 µg) was digested with appropriate restriction enzymes overnight. Subsequently, genomic DNA was separated on a 0.8% agarose gel and transferred to an Amersham Hybond-XL membrane (GE Healthcare, RPN303S). Probes were synthesized by PCR amplification and labeled with dCTP, [α-32P] (Perkin Elmer, NEG513H250UC) using High Prime (Roche, 11585592001) and hybridization performed in Church buffer (“Sodium Phosphate Buffer for Church and Gilbert Hybridization:” 2015b).

### P-RFP Pluripotency Reporter

Lentivirus, encoding the mouse *Nanog* promoter driving RFP expression, was purchased from System Biosciences (SR10044VA-1). EL16.7 TST X-GFP MKOS ESCs were infected at an MOI of 30 and RFP-positive cells FACS purified using a BD Influx. Single clones were isolated and selected based on proper RFP expression using FACS analysis on a BD LSRFortessa.

## Neural Precursor Cell Differentiation

ESCs were differentiated to neural precursor cells (NPCs) as described previously (Abranches et al. 2009). ESCs were seeded at a density of $2.75x10^5$ cells/$cm^2$ in N2B27 (50% DMEM/F12

(Thermo Fisher Scientific, 21041025), 50% Neurobasal medium (Thermo Fisher Scientific, 12348017), 1x N2 (Thermo Fisher Scientific, 17502048), 1x B27 (Thermo Fisher Scientific, 12587001)) supplemented with 0.4 µM PD0325901 (Selleck Chemicals, S1036-5mg), 3 µM CHIR99021 (Sigma-Aldrich, SML1046-5MG) and 1,000 U/ml LIF (ORF Genetics, 01-A1140-0100). 24 hours later, cells were dissociated using Accutase (Thermo Fisher Scientific, 00-4555-56) and plated at $2.95x10^4$ cells/$cm^2$ in RHB-A (Takara Bio, Y40001) on 0.2% gelatin-coated T75 flasks, changing media every other day. On days 6 and 8, media was supplemented with 10 ng/ml EGF (R&D Systems, 236-EG-200) and 10 ng/ml bFGF (Thermo Fisher Scientific, 13256029) and additionally with 10µM ROCK inhibitor (Sellekchem, S1049) on day 8. On day 9, cells were dissociated using Accutase (Thermo Fisher Scientific, 00-4555-56) and SSEA1 expressing cells were removed by MACS sorting using Anti-SSEA-1 (CD15) MicroBeads (Miltenyi Biotech, 130-094-530). To completely remove cells that hadn't undergone XCI, cells were stained with SSEA-1 eFluor 660 (Thermo Fisher Scientific, 50-8813-42) for 15 min at 4°C, washed once with 0.5% BSA in PBS and then SSEA1-/Nanog-RFP-/X-GFP- cells were FACS purified using a BD FACSAria II SORP or a BD Influx (BD Biosciences) at a maximum flow rate of 4,000 ev/s to improve survival. FACS sorted cells were seeded at $3.5x10^5$ cells/$cm^2$ on 0.2% gelatin-coated dishes in RHB-A, supplemented with EGF, FGF, and ROCKi. The medium was changed daily until day 12 when cells reached 100% confluency.

## Reprogramming of Neural Precursor Cells

Reprogramming of day 12 neural precursor cells was induced by the addition of 1 mg/ml doxycycline (Tocris, 4090/50) and 25 mg/ml L-ascorbic acid (Sigma-Aldrich, A7506-25G) to the NPC medium (RHB-A supplemented with EGF and FGF). 24 hours later, cells were dissociated using Accutase and seeded on irradiated mouse embryonic fibroblasts (iMEF) in ESC medium containing 15% FBS and supplemented with 1 mg/ml doxycycline and 25 mg/ml L-ascorbic acid. The medium was changed every other day. To isolate iPSC, SSEA1+/Nanog-RFP+/X-GFP+ cells were isolated using FACS at day 10 of reprogramming and re-plated on 0.2% gelatin-coated plates in ESC medium and kept in doxycycline free conditions for 5-6 days.

## RNA Isolation, Quantitative RT-PCR and RNA-Sequencing

RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, 74136) or RNeasy Micro Kit (Qiagen, 74004) and quantified by Nanodrop. cDNA was produced with a High-Capacity RNA-to-cDNA Kit (Thermo Fisher Scientific, 4387406) and was used for qRT-PCR analysis in triplicate reactions with Power SYBR Green PCR Master Mix (Thermo Fisher Scientific, 4367659). Libraries were prepared using the TruSeq Stranded Total RNA Library Preparation Kit (Illumina, 20020597) followed by paired-end sequencing (2x125bp) on an Illumina HiSeq 2500.

## Assay for Transposase-Accessible Chromatin with High Throughput Sequencing (ATAC-Seq)

ATAC-seq was performed as described previously (Corces et al. 2017) with minor modifications. 50,000 FACS purified cells were resuspended in 50 µl cold lysis buffer (10 mM Tris-HCl pH 7.4,

10 mM NaCl, 3 mM $MgCl_2$, 0.01% Digitonin, 0.1% Tween-20, 0.1% IGEPAL CA-630). After 3 min the lysis was washed out using 1 ml cold lysis buffer containing Tween-20, but no Digitonin or IGEPAL CA-630. Cells were centrifuged for 10 min at 500 rcf and 4°C, supernatant was removed and nuclei were resuspended in 50 µl transposition reaction mix (25 µl Tn5 Transposase buffer (Illumina, 15027866), 2.5 µl Tn5 transposase (Illumina, 15027865), 16.5 µl PBS, 0.01% Digitonin, 0.1% Tween-20, 5 µl nuclease-free water) and incubated at 37°C for 45 min with 1000 RPM mixing. DNA was isolated using the MinElute PCR Purification Kit (Qiagen, 28004). Library amplification was performed by two sequential PCR reactions (8 and 4-7 cycles, respectively) using the NEBNext High Fidelity PCR Master Mix (New England Biolabs, M0541S). DNA was then double-size selected using 0.5x and 1.5x Agencourt AMPure XP beads (Beckman, A63880) in order to isolate fragments between 100bp and 1kb. Library quality was assessed on a Bioanalyzer, followed by paired-end sequencing (2x125bp) on an Illumina HiSeq 2500.

## Cell Isolation and Purification for ATAC-Seq and RNA-Seq

Cells were dissociated using Accutase (Thermo Fisher Scientific, 00-4555-56) (for NPCs), 0.05% Trypsin-EDTA (Thermo Fisher Scientific, 25300054) (for ESCs) or 0.25% Trypsin-EDTA (Thermo Fisher Scientific, 25200056) (for iPSCs) and then stained with SSEA-1 eFluor 660 (Thermo Fisher Scientific, 50-8813-42) for 15 min at 4°C. Cells were washed once with 0.5% BSA in PBS and then FACS sorted using a BD FACSAria II SORP or a BD Influx.

## Cell Isolation and Purification for Hi-C

Purification of $G_0G_1$ cells based on DNA content was performed as described previously (Bonev et al. 2017) with minor modifications. Briefly, cells were dissociated using Accutase (Thermo Fisher Scientific, 00-4555-56) (for NPCs), 0.05% Trypsin-EDTA (Thermo Fisher Scientific, 25300054) (for ESCs) or 0.25% Trypsin-EDTA (Thermo Fisher Scientific, 25200056) (for iPSCs) and then stained with SSEA-1 eFluor 660 (Thermo Fisher Scientific, 50-8813-42). iPSCs were additionally MACS sorted using Anti-SSEA-1 (CD15) MicroBeads (Miltenyi Biotech, 130-094-530). Cells were then fixed for 10 min at room temperature with freshly prepared 1% formaldehyde in PBS (Sigma-Aldrich, F8775-4X25ML) and the reaction then quenched by addition of 0.2M glycine (NZYTech, MB01401). $1x10^6$ cells/ml were permeabilized using 0.1% saponin (Sigma-Aldrich, 47036-50G-F). 10 µg/ml DAPI (Thermo Fisher Scientific, D1306) and 100 µg/ml RNase A (Thermo Fisher Scientific, EN0531) were added and samples incubated for 30 min at room temperature protected from light with slight agitation. After washing once with cold PBS, samples were resuspended in cold 0.5% BSA in PBS at a concentration of $1x10^7$ cells/ml and immediately FACS purified using a BD FACSAria II SORP or a BD Influx. After FACS sorting, dry cell pellets were snap-frozen in dry ice and stored at -80°C.

## *In situ* Hi-C Library Preparation

*In situ* Hi-C was performed as described previously (Stadhouders et al. 2018) with minor modifications. One million cells purified for $G_0G_1$ were used as starting materials. Cells were

lysed using 250 µl cold lysis buffer (10 mM Tris-HCl pH 8, 10 mM NaCl, 0.2% IGEPAL CA-630) supplemented with 50 µl protease inhibitor cocktail (Sigma-Aldrich, P8340-1ML). Cells were digested with 100 U MboI (New England Biolabs) and incubated for 2 hours at 37°C under rotation, followed by the addition of another 100U for 2 hours and another 100U before overnight incubation. The next day a final 100U were added and incubated for 3 hours. After fill-in with biotin-14-dATP (Thermo Fisher Scientific, 19524016), ligation was performed with 10,000 U T4 DNA Ligase (New England Biolabs, M0202M) overnight at 24°C under rotation. After de-crosslinking, DNA was purified using ethanol precipitation and sonicated to an average size of 300–700 bp with a Bioruptor Pico (Diagenode; seven cycles of 20 s on and 60 s off). Ligation products containing biotin-14-dATP were pulled-down using Dynabeads MyOne Streptavidin T1 beads (Thermo Fisher Scientific, 65601) and end-repaired and A-tailed using the NEBNext End Repair/dA-Tailing Module (New England Biolabs, E6060S and E6053S). Libraries were amplified using the NEBNext High Fidelity PCR Master Mix and NEBNext Multiplex Oligos for Illumina (New England Biolabs, M0541S and E7335S) for 8 cycles and size-selected with 0.9x Agencourt AMPure XP beads. Library quality was assessed on a Bioanalyzer and by low-coverage sequencing on an Illumina NextSeq 500, followed by high-coverage paired-end sequencing (2x125bp) on an Illumina HiSeq 2500.

## RNA-Fluorescent *In Situ* Hybridization

Strand-specific RNA FISH was performed with fluorescently labeled oligonucleotides (IDT) as described previously (Del Rosario et al. 2017). Briefly, cells were fixed with 4% paraformaldehyde for 10 minutes at room temperature and then permeabilized for 5 minutes on ice in 0.5% Triton-X. 10ng/ml equimolar amounts of Cy5 labeled Xist probes BD384-Xist-Cy5-3'-AM (5'-ATG ACT CTG GAA GTC AGT ATG GAG /3Cy5Sp/ -3') and BD417-5'Cy5-Xist-Cy5-3'-AM (5'- /5Cy5/ATG GGC ACT GCA TTT TAG CAA TA /3Cy5Sp/ -3') were hybridized in 40% formamide, 10% dextran sulfate, 2xSSC pH 7 at room temperature overnight. Slides were then washed in 30% formamide 2xSSC pH 7 at room temperature, followed by washes in 2xSSC pH 7 and then mounted with Vectashield (Vector Laboratories, H1200). Images were acquired using an EVOS and a Cy5 light cube (Thermo Fisher Scientific).

## Bioinformatic Analysis

### Allele-specific Alignment

Allele-specific alignment of reads for RNA-seq, ATAC-seq and Hi-C was performed as described previously (Pinter et al. 2012). Briefly, mouse genome sequencing data for 129S1/SvImJ (mus) and CAST/EiJ (cas) were aligned to the C57BL/6J reference genome (NCBI mm10) and screened for high quality single nucleotide polymorphisms (SNPs) and insertions/deletions indels. Reads that didn't uniquely align to either mus or cas were discarded.

### Allele-Specific Expression

From a list of 806 protein-coding genes on the X-chromosome, we masked out conditions and gene combinations that didn't have enough reads to be reliable, either due to low expression or low amount of sequence polymorphisms, leaving 586 genes that passed these criteria.
The allelic ratio of these genes was then calculated by dividing mus reads by the sum of mus and cas reads ((mus/mus+cas)).
The absolute allele-specific expression for mus and cas alleles was calculated by multiplying the bulk counts by the allelic ratio to correct for biases introduced by SNP density variations.

### Reactivation Cluster

For the identification of X-reactivation clusters, we applied two additional filters. First, the sum of reads from both alleles had to be over 40 throughout the time course. Second, genes had to have an allelic-ratio over 0.35 in iPSCs and ESCs, to ensure biallelic expression. From this list of 286 genes, we divided genes into two groups: Genes undergoing X-reactivation by an allelic-ratio under 0.15 in NPCs and genes escaping X-inactivation by a ratio of over 0.15. Escapees were then additionally filtered by a confidence interval over 0.1 throughout the time course and were then divided into "strong" and "weak" by an allelic-ratio of over or under 0.35 respectively. Using these criteria, 232 genes underwent reactivation, 31 escaped inactivation, with 23 genes that couldn't be assigned to any of the two groups.
Hierarchical k-means clustering was performed on 232 genes undergoing X-reactivation using the R package Factoextra (R 3.6.1.). The optimal number of clusters was determined using the elbow method. The heatmaps were generated in the online software Morpheus (Morpheus, https://software.broadinstitute.org/morpheus).

# Discussion

The partitioning of genomes into active and inactive regions, euchromatic and heterochromatic, serves as an important separator to ensure precise gene regulation during mammalian development. Historically, the inactive X has served as an important model for developmentally induced gene silencing and the formation of heterochromatin. However equivalently, reactivation of the inactive X could provide a powerful model to study the formation of euchromatin and the initiation of gene expression. Yet, experimental models allowing the study of X-reactivation in the context of gene expression, chromatin accessibility, and genome conformation were crucially missing. Reprogramming to iPSC cells has served as an important system to gain insights into the acquisition of pluripotency and more recently into the reactivation of the inactive X. However, until now experimental systems have been hampered by low reprogramming efficiency, the impurity of cell populations and have been restricted to microscopy-based assays (Maherali et al. 2007; Payer et al. 2013; Pasque et al. 2014; Janiszewski et al. 2019). Here we describe a novel iPSC reprogramming system that overcomes these common shortcomings and allows for the first time the large-scale interrogation of the X-reactivation process using epigenomic analysis methods. Combining the reprogramming of *in vitro* derived neural precursor cells with fluorescent reporters for pluripotency and X-activity, allowed us to obtain highly purified cell populations transitioning to iPSCs. Previously, people have relied on the early marker SSEA-1 for the isolation of iPSC. However, we could show that only a subset of SSEA1+ cells was able to ultimately reactivate the X chromosome. Utilizing a marker for later stages of reprogramming has proven to be invaluable in our system, allowing the isolation of a subpopulation, poised for near-deterministic X-reactivation. Furthermore, induction of reprogramming in conventional systems was often shown to be variable, as systems were handicapped by either sub-optimal reprogramming cassettes, or lentiviral delivered cassettes, which depict integration site-dependent expression. We overcame these issues by using an optimized reprogramming cassette and utilized CRISPR/Cas9, to precisely target a defined genomic locus, to circumvent variability in expression induced by position effects. The utility of CRISPR/Cas9 has further proven advantageous for the integration of an X-GFP reporter, showing stronger expression compared to previously used reporters, allowing for more precise isolation of populations throughout the reprogramming time course. Moreover, mouse embryonic fibroblasts, the primary source of somatic cells for iPSC reprogramming, display extensive variability (Singhal et al. 2016). Utilizing *in vitro* derived neural precursor cells, enabled us to further reduce the variation in the system and additionally allowed to significantly upscale cell numbers for large scale experiments. The unique combination of these factors, combined with a hybrid mouse strain background to allow allele-specific calling of sequencing data, enabled us to obtain a high-resolution map of X-linked gene reactivation and chromatin opening in relation to the iPSC-reprogramming process.

Using this system, we were able to show that X-reactivation during reprogramming occurs rapidly and mostly synchronously. Moreover, chromatin accessibility and gene reactivation follow highly similar patterns, with regions of early initiation of gene reactivation being in close proximity to hubs of chromatin openness, coinciding with binding sites for pluripotency factors.

However, this tight link wasn't observed when assessing genome conformation. Whereas some regions of early chromatin opening overlapped with early TAD formation, we could as well observe the formation of TADs in regions with no detectable early gene reactivation or chromatin opening.

In summary, we have established a novel system to interrogate the reactivation of the inactive X chromosome, which beyond, might provide important insights into the formation of euchromatin.

# Conclusions

- We describe a novel iPSC system that allows the isolation of cells poised for X-reactivation, subsequently achieving near-deterministic efficiency of X-reactivation.
- X-reactivation in our iPSC system occurs rapidly, coinciding with full naive pluripotency gene expression and *Xist* downregulation.
- Timing of initiation and completion of gene reactivation can differ, with early initiation not priming for faster completion of the process.
- Completion of reactivation is solely observed in populations showing downregulation of *Xist*, whereas initiation of reactivation of a subset of genes can occur independently of loss of Xist RNA expression and coating, arguing for a two-step process of reactivation.
- Chromatin opening of the X chromosome during reprogramming involves small but gradual changes during its initiation period that culminate in a rapid switch, which is not generally observed on autosomes.
- Chromatin accessibility on the inactive X is regained in a two-step manner: Primary chromatin opening initiates near escapee regions, which prime their surroundings to become accessible. Only later, secondary chromatin opening occurs at distances independently from previously open chromatin.
- Transcription factor binding motif analysis of primary and secondary chromatin opening sites shows enrichment for pluripotency factors, including OCT4, SOX2, and NANOG.
- No apparent motifs were exclusively enriched on primary versus secondary sites, arguing for a shared transcription factor model, where factors are expressed throughout different stages of reprogramming and could gain accessibility to different sites at different time-points, depending on other parameters like chromatin status or cooperative binding with other co-factors.
- Mega-domains and TADs are able to coexist during the X-reactivation process, arguing that these correspond to two different levels of three-dimensional genome organization superimposed on the Xi.
- Early TAD formation wasn’t exclusively observed in regions of early initiation of gene reactivation, arguing that chromosomal restructuring and gene expression might be separate events.
- Gene expression and chromatin accessibility during X-reactivation share similar kinetics, while genome organization might follow distinct principles.

# Bibliography


Abranches, Elsa, Evguenia Bekman, and Domingos Henrique. 2013. “Generation and Characterization of a Novel Mouse Embryonic Stem Cell Line with a Dynamic Reporter of Nanog Expression.” *PloS One* 8 (3): e59928.

Abranches, Elsa, Margarida Silva, Laurent Pradier, Herbert Schulz, Oliver Hummel, Domingos Henrique, and Evguenia Bekman. 2009. “Neural Differentiation of Embryonic Stem Cells in Vitro: A Road Map to Neurogenesis in the Embryo.” *PloS One* 4 (7): e6286.

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

Augui, Sandrine, Elphège P. Nora, and Edith Heard. 2011. “Regulation of X-Chromosome Inactivation by the X-Inactivation Centre.” *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2987.

Barakat, Tahsin Stefan, Nilhan Gunhanlar, Cristina Gontan Pardo, Eskeatnaf Mulugeta Achame, Mehrnaz Ghazvini, Ruben Boers, Annegien Kenter, Eveline Rentmeester, J. Anton Grootegoed, and Joost Gribnau. 2011. “RNF12 Activates Xist and Is Essential for X Chromosome Inactivation.” *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.1002001.

Baranello, Laura, Fedor Kouzine, and David Levens. 2014. “CTCF and Cohesin Cooperate to Organize the 3D Structure of the Mammalian Genome.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (3): 889–90.

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

Bernstein, Emily, Elizabeth M. Duncan, Osamu Masui, Jesus Gil, Edith Heard, and C. David Allis. 2006. “Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin.” *Molecular and Cellular Biology* 26 (7): 2560–69.

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

Bokar, J. A., M. E. Rath-Shambaugh, R. Ludwiczak, P. Narayan, and F. Rottman. 1994. “Characterization and Partial Purification of mRNA N6-Adenosine Methyltransferase from HeLa Cell Nuclei. Internal mRNA Methylation Requires a Multisubunit Complex.” *The Journal of Biological Chemistry* 269 (26): 17697–704.

Bonev, Boyan, Netta Mendelson Cohen, Quentin Szabo, Lauriane Fritsch, Giorgio L. Papadopoulos, Yaniv Lubling, Xiaole Xu, et al. 2017. “Multiscale 3D Genome Rewiring during Mouse Neural Development.” *Cell* 171 (3): 557–72.e24.

Bonora, G., X. Deng, H. Fang, V. Ramani, R. Qiu, J. B. Berletch, G. N. Filippova, et al. 2018. “Orientation-Dependent Dxz4 Contacts Shape the 3D Structure of the Inactive X Chromosome.” *Nature Communications* 9 (1): 1445.

Borensztein, Maud, Ikuhiro Okamoto, Laurène Syx, Guillaume Guilbaud, Christel Picard, Katia Ancelin, Rafael Galupa, et al. 2017. “Contribution of Epigenetic Landscapes and Transcription Factors to X-Chromosome Reactivation in the Inner Cell Mass.” *Nature Communications* 8 (1): 1297.

Borensztein, Maud, Laurène Syx, Katia Ancelin, Patricia Diabangouaya, Christel Picard, Tao Liu, Jun-Bin Liang, et al. 2017. “Xist-Dependent Imprinted X Inactivation and the Early Developmental Consequences of Its Failure.” *Nature Structural & Molecular Biology* 24 (3): 226–33.

Borsani, G., R. Tonlorenzi, M. C. Simmler, L. Dandolo, D. Arnaud, V. Capra, M. Grompe, et al. 1991. “Characterization of a Murine Gene Expressed from the Inactive X Chromosome.” *Nature* 351 (6324): 325–29.

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

Cantone, Irene. 2017. “Reversal of X Chromosome Inactivation: Lessons from Pluripotent Reprogramming of Mouse and Human Somatic Cells.” *Journal of Translational Genetics and Genomics*. https://doi.org/10.20517/jtgg.2017.19.

Cantone, Irene, Hakan Bagci, Dirk Dormann, Gopuraja Dharmalingam, Tatyana Nesterova, Neil Brockdorff, Claire Rougeulle, et al. 2016. “Ordered Chromatin Changes and Human X Chromosome Reactivation by Cell Fusion-Mediated Pluripotent Reprogramming.” *Nature Communications* 7 (August): 12354.

Cantone, Irene, Gopuraja Dharmalingam, Yi-Wah Chan, Anne-Celine Kohler, Boris Lenhard, Matthias Merkenschlager, and Amanda G. Fisher. 2017. “Allele-Specific Analysis of Cell Fusion-Mediated Pluripotent Reprograming Reveals Distinct and Predictive Susceptibilities of Human X-Linked Genes to Reactivation.” *Genome Biology* 18 (1): 2.

Cerase, Andrea, Daniel Smeets, Y. Amy Tang, Michal Gdula, Felix Kraus, Mikhail Spivakov, Benoit Moindrot, et al. 2014. “Spatial Separation of Xist RNA and Polycomb Proteins Revealed by Superresolution Microscopy.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (6): 2235–40.

Chadwick, Brian P., and Huntington F. Willard. 2004. “Multiple Spatially Distinct Types of Facultative Heterochromatin on the Human Inactive X Chromosome.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (50): 17450–55.

Chaligné, Ronan, Tatiana Popova, Marco-Antonio Mendoza-Parra, Mohamed-Ashick M. Saleem, David Gentien, Kristen Ban, Tristan Piolot, et al. 2015. “The Inactive X Chromosome Is Epigenetically Unstable and Transcriptionally Labile in Breast Cancer.” *Genome Research* 25 (4): 488–503.

Chambers, Ian, Jose Silva, Douglas Colby, Jennifer Nichols, Bianca Nijmeijer, Morag Robertson, Jan Vrana, Ken Jones, Lars Grotewold, and Austin Smith. 2007. “Nanog Safeguards Pluripotency and Mediates Germline Development.” *Nature* 450 (7173): 1230–34.

Chantzoura, Eleni, Stavroula Skylaki, Sergio Menendez, Shin-Il Kim, Anna Johnsson, Sten Linnarsson, Knut Woltjen, Ian Chambers, and Keisuke Kaji. 2015. “Reprogramming Roadblocks Are System Dependent.” *Stem Cell Reports* 5 (3): 350–64.

Chen, Chun-Kan, Amy Chow, Mason Lai, and Mitchell Guttman. 2017. “Response to Comment on ‘Xist Recruits the X Chromosome to the Nuclear Lamina to Enable Chromosome-Wide Silencing.’” *Science*. https://doi.org/10.1126/science.aam5439.

Chen, Geng, John Paul Schell, Julio Aguila Benitez, Sophie Petropoulos, Marlene Yilmaz, Björn Reinius, Zhanna Alekseenko, et al. 2016. “Single-Cell Analyses of X Chromosome Inactivation Dynamics and Pluripotency during Differentiation.” *Genome Research* 26 (10): 1342–54.

Chen, Qi, Shuai Gao, Wenteng He, Xiaochen Kou, Yanhong Zhao, Hong Wang, and Shaorong

Gao. 2014. “Xist Repression Shows Time-Dependent Effects on the Reprogramming of Female Somatic Cells to Induced Pluripotent Stem Cells.” *Stem Cells* 32 (10): 2642–56.
Choi, Jiho, Kendell Clement, Aaron J. Huebner, Jamie Webster, Christopher M. Rose, Justin Brumbaugh, Ryan M. Walsh, et al. 2017. “DUSP9 Modulates DNA Hypomethylation in Female Mouse Pluripotent Stem Cells.” *Cell Stem Cell* 20 (5): 706–19.e7.
Chronis, Constantinos, Petko Fiziev, Bernadett Papp, Stefan Butz, Giancarlo Bonora, Shan Sabri, Jason Ernst, and Kathrin Plath. 2017. “Cooperative Binding of Transcription Factors Orchestrates Reprogramming.” *Cell* 168 (3): 442–59.e20.
Chu, Ci, Qiangfeng Cliff Zhang, Simão Teixeira da Rocha, Ryan A. Flynn, Maheetha Bharadwaj, J. Mauro Calabrese, Terry Magnuson, Edith Heard, and Howard Y. Chang. 2015. “Systematic Discovery of Xist RNA Binding Proteins.” *Cell* 161 (2): 404–16.
Chureau, Corinne, Sophie Chantalat, Antonio Romito, Angélique Galvani, Laurent Duret, Philip Avner, and Claire Rougeulle. 2011. “Ftx Is a Non-Coding RNA Which Affects Xist Expression and Chromatin Structure within the X-Inactivation Center Region.” *Human Molecular Genetics* 20 (4): 705–18.
Chuva de Sousa Lopes, Susana M., Katsuhiko Hayashi, Tanya C. Shovlin, Will Mifsud, M. Azim Surani, and Anne McLaren. 2008. “X Chromosome Activity in Mouse XX Primordial Germ Cells.” *PLoS Genetics* 4 (2): e30.
Ciavatta, D., S. Kalantry, T. Magnuson, and O. Smithies. 2006. “A DNA Insulator Prevents Repression of a Targeted X-Linked Transgene but Not Its Random or Imprinted X Inactivation.” *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0603754103.
Cifuentes-Rojas, Catherine, Alfredo J. Hernandez, Kavitha Sarma, and Jeannie T. Lee. 2014. “Regulatory Interactions between RNA and Polycomb Repressive Complex 2.” *Molecular Cell* 55 (2): 171–85.
Colognori, David, Hongjae Sunwoo, Andrea J. Kriz, Chen-Yu Wang, and Jeannie T. Lee. 2019. “Xist Deletional Analysis Reveals an Interdependency between Xist RNA and Polycomb Complexes for Spreading along the Inactive X.” *Molecular Cell* 74 (1): 101–17.e10.
Corbel, Catherine, Patricia Diabangouaya, Anne-Valerie Gendrel, Jennifer C. Chow, and Edith Heard. 2013. “Unusual Chromatin Status and Organization of the Inactive X Chromosome in Murine Trophoblast Giant Cells.” *Development* 140 (4): 861–72.
Corces, M. Ryan, Alexandro E. Trevino, Emily G. Hamilton, Peyton G. Greenside, Nicholas A. Sinnott-Armstrong, Sam Vesuna, Ansuman T. Satpathy, et al. 2017. “An Improved ATAC-Seq Protocol Reduces Background and Enables Interrogation of Frozen Tissues.” *Nature Methods* 14 (10): 959–62.
Costanzi, C., P. Stein, D. M. Worrad, R. M. Schultz, and J. R. Pehrson. 2000. “Histone macroH2A1 Is Concentrated in the Inactive X Chromosome of Female Preimplantation Mouse Embryos.” *Development* 127 (11): 2283–89.
Cremer, Thomas, and Marion Cremer. 2010. “Chromosome Territories.” *Cold Spring Harbor Perspectives in Biology* 2 (3): a003889.
Csankovszki, G., A. Nagy, and R. Jaenisch. 2001. “Synergism of Xist RNA, DNA Methylation, and Histone Hypoacetylation in Maintaining X Chromosome Inactivation.” *The Journal of Cell Biology* 153 (4): 773–84.
Darrow, Emily M., Miriam H. Huntley, Olga Dudchenko, Elena K. Stamenova, Neva C. Durand, Zhuo Sun, Su-Chen Huang, et al. 2016. “Deletion of DXZ4 on the Human Inactive X Chromosome Alters Higher-Order Genome Architecture.” *Proceedings of the National Academy of Sciences of the United States of America* 113 (31): E4504–12.
Davidovich, Chen, Xueyin Wang, Catherine Cifuentes-Rojas, Karen J. Goodrich, Anne R.

Gooding, Jeannie T. Lee, and Thomas R. Cech. 2015. “Toward a Consensus on the Binding Specificity and Promiscuity of PRC2 for RNA.” *Molecular Cell*. https://doi.org/10.1016/j.molcel.2014.12.017.

Davidovich, Chen, Leon Zheng, Karen J. Goodrich, and Thomas R. Cech. 2013. “Promiscuous RNA Binding by Polycomb Repressive Complex 2.” *Nature Structural & Molecular Biology* 20 (11): 1250–57.

Davies, James O. J., Jelena M. Telenius, Simon J. McGowan, Nigel A. Roberts, Stephen Taylor, Douglas R. Higgs, and Jim R. Hughes. 2016. “Multiplexed Analysis of Chromosome Conformation at Vastly Improved Sensitivity.” *Nature Methods* 13 (1): 74–80.

Dekker, J. 2002. “Capturing Chromosome Conformation.” *Science*. https://doi.org/10.1126/science.1067799.

Deng, Q., D. Ramskold, B. Reinius, and R. Sandberg. 2014. “Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells.” *Science*. https://doi.org/10.1126/science.1245316.

Deng, Xinxian, Wenxiu Ma, Vijay Ramani, Andrew Hill, Fan Yang, Ferhat Ay, Joel B. Berletch, et al. 2015. “Bipartite Structure of the Inactive Mouse X Chromosome.” *Genome Biology* 16 (August): 152.

Disteche, C. M. 1999. “Escapees on the X Chromosome.” *Proceedings of the National Academy of Sciences of the United States of America*.

Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. 2012. “Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions.” *Nature* 485 (7398): 376–80.

Dostie, J., T. A. Richmond, R. A. Arnaout, R. R. Selzer, W. L. Lee, T. A. Honan, E. D. Rubio, et al. 2006. “Chromosome Conformation Capture Carbon Copy (5C): A Massively Parallel Solution for Mapping Interactions between Genomic Elements.” *Genome Research*. https://doi.org/10.1101/gr.5571506.

Dubois, Agnès, Jane Lynda Deuve, Pablo Navarro, Sarra Merzouk, Sylvain Pichard, Pierre-Henri Commere, Anne Louise, Danielle Arnaud, Philip Avner, and Céline Morey. 2014. “Spontaneous Reactivation of Clusters of X-Linked Genes Is Associated with the Plasticity of X-Inactivation in Mouse Trophoblast Stem Cells.” *Stem Cells* 32 (2): 377–90.

Eils, R., S. Dietzel, E. Bertin, E. Schröck, M. R. Speicher, T. Ried, M. Robert-Nicoud, C. Cremer, and T. Cremer. 1996. “Three-Dimensional Reconstruction of Painted Human Interphase Chromosomes: Active and Inactive X Chromosome Territories Have Similar Volumes but Differ in Shape and Surface Structure.” *The Journal of Cell Biology* 135 (6 Pt 1): 1427–40.

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

Escamilla-Del-Arenal, Martin, Simao Teixeira da Rocha, and Edith Heard. 2011. “Evolutionary Diversity and Developmental Regulation of X-Chromosome Inactivation.” *Human Genetics* 130 (2): 307–27.

Esteban, Miguel Angel, Tao Wang, Baoming Qin, Jiayin Yang, Dajiang Qin, Jinglei Cai, Wen Li, et al. 2010. “Vitamin C Enhances the Generation of Mouse and Human Induced Pluripotent Stem Cells.” *Cell Stem Cell* 6 (1): 71–79.

Fang, H. E., Giancarlo Bonora, Jordan Lewandowski, Jitendra Thakur, Galina N. Filippova, Steven Henikoff, Jay Shendure, et al. 2019. “Trans− and Cis−acting Effects of the lncRNA Firre on Epigenetic and Structural Features of the Inactive X Chromosome.” https://doi.org/10.1101/687236.

Froberg, John E., Stefan F. Pinter, Andrea J. Kriz, Teddy Jégu, and Jeannie T. Lee. 2018.

"Megadomains and Superloops Form Dynamically but Are Dispensable for X-Chromosome Inactivation and Gene Escape." *Nature Communications* 9 (1): 5004.
Gallagher, R. E., A. C. Ferrari, A. W. Zulich, R. W. Yen, and J. R. Testa. 1984. "Cytotoxic and Cytodifferentiative Components of 6-Thioguanine Resistance in HL-60 Cells Containing Acquired Double Minute Chromosomes." *Cancer Research* 44 (6): 2642–53.
Galupa, Rafael, and Edith Heard. 2018. "X-Chromosome Inactivation: A Crossroads Between Chromosome Architecture and Gene Regulation." *Annual Review of Genetics* 52 (November): 535–66.
Gayen, Srimonta, Emily Maclary, Emily Buttigieg, Michael Hinten, and Sundeep Kalantry. 2015. "A Primary Role for the Tsix lncRNA in Maintaining Random X-Chromosome Inactivation." *Cell Reports*. https://doi.org/10.1016/j.celrep.2015.04.039.
Gaztelumendi, Nerea, and Carme Nogués. 2015. "Chromosome Instability in Mouse Embryonic Stem Cells." *Scientific Reports*. https://doi.org/10.1038/srep05324.
Gdula, Michal R., Tatyana B. Nesterova, Greta Pintacuda, Jonathan Godwin, Ye Zhan, Hakan Ozadam, Michael McClellan, et al. 2019. "The Non-Canonical SMC Protein SmcHD1 Antagonises TAD Formation and Compartmentalisation on the Inactive X Chromosome." *Nature Communications* 10 (1): 30.
Gendrel, Anne-Valerie, Anwyn Apedaile, Heather Coker, Ausma Termanis, Ilona Zvetkova, Jonathan Godwin, Y. Amy Tang, et al. 2012. "Smchd1-Dependent and -Independent Pathways Determine Developmental Dynamics of CpG Island Methylation on the Inactive X Chromosome." *Developmental Cell*. https://doi.org/10.1016/j.devcel.2012.06.011.
Giorgetti, Luca, Bryan R. Lajoie, Ava C. Carter, Mikael Attia, Ye Zhan, Jin Xu, Chong Jian Chen, et al. 2016. "Structural Organization of the Inactive X Chromosome in the Mouse." *Nature* 535 (7613): 575–79.
Gontan, Cristina, Eskeatnaf Mulugeta Achame, Jeroen Demmers, Tahsin Stefan Barakat, Eveline Rentmeester, Wilfred van IJcken, J. Anton Grootegoed, and Joost Gribnau. 2012. "RNF12 Initiates X-Chromosome Inactivation by Targeting REX1 for Degradation." *Nature*. https://doi.org/10.1038/nature11070.
Gontan, Cristina, Eskeatnaf Mulugeta Achame, Jeroen Demmers, Tahsin Stefan Barakat, Eveline Rentmeester, Wilfred van IJcken, J. Anton Grootegoed, and Joost Gribnau. 2012. "RNF12 Initiates X-Chromosome Inactivation by Targeting REX1 for Degradation." *Nature* 485 (7398): 386–90.
Gontan, Cristina, Hegias Mira-Bontenbal, Aristea Magaraki, Catherine Dupont, Tahsin Stefan Barakat, Eveline Rentmeester, Jeroen Demmers, and Joost Gribnau. 2018. "REX1 Is the Critical Target of RNF12 in Imprinted X Chromosome Inactivation in Mice." *Nature Communications*. https://doi.org/10.1038/s41467-018-07060-w.
Grob, Stefan, and Giacomo Cavalli. 2018. "Technical Review: A Hitchhiker's Guide to Chromosome Conformation Capture." *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-4939-7318-7_14.
Gruenbaum, Yosef, Ayelet Margalit, Robert D. Goldman, Dale K. Shumaker, and Katherine L. Wilson. 2005. "The Nuclear Lamina Comes of Age." *Nature Reviews. Molecular Cell Biology* 6 (1): 21–31.
Guenther, M. G., O. Barak, and M. A. Lazar. 2001. "The SMRT and N-CoR Corepressors Are Activating Cofactors for Histone Deacetylase 3." *Molecular and Cellular Biology* 21 (18): 6091–6101.
Hacisuleyman, Ezgi, Loyal A. Goff, Cole Trapnell, Adam Williams, Jorge Henao-Mejia, Lei Sun, Patrick McClanahan, et al. 2014. "Topological Organization of Multichromosomal Regions by the Long Intergenic Noncoding RNA Firre." *Nature Structural & Molecular Biology*.

https://doi.org/10.1038/nsmb.2764.
Hadjantonakis, A. K., L. L. Cox, P. P. Tam, and A. Nagy. 2001. “An X-Linked GFP Transgene Reveals Unexpected Paternal X-Chromosome Activity in Trophoblastic Giant Cells of the Mouse Placenta.” *Genesis* 29 (3): 133–40.
Hasegawa, Yuko, Neil Brockdorff, Shinji Kawano, Kimiko Tsutui, Ken Tsutui, and Shinichi Nakagawa. 2010. “The Matrix Protein hnRNP U Is Required for Chromosomal Localization of Xist RNA.” *Developmental Cell*. https://doi.org/10.1016/j.devcel.2010.08.006.
Hayashi, Katsuhiko, Sugako Ogushi, Kazuki Kurimoto, So Shimamoto, Hiroshi Ohta, and Mitinori Saitou. 2012. “Offspring from Oocytes Derived from in Vitro Primordial Germ Cell-like Cells in Mice.” *Science* 338 (6109): 971–75.
Huangfu, Danwei, Kenji Osafune, René Maehr, Wenjun Guo, Astrid Eijkelenboom, Shuibing Chen, Whitney Muhlestein, and Douglas A. Melton. 2008. “Induction of Pluripotent Stem Cells from Primary Human Fibroblasts with Only Oct4 and Sox2.” *Nature Biotechnology* 26 (11): 1269–75.
Hug, Clemens B., Alexis G. Grimaldi, Kai Kruse, and Juan M. Vaquerizas. 2017. “Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription.” *Cell* 169 (2): 216–28.e19.
Hughes, Jim R., Nigel Roberts, Simon McGowan, Deborah Hay, Eleni Giannoulatou, Magnus Lynch, Marco De Gobbi, Stephen Taylor, Richard Gibbons, and Douglas R. Higgs. 2014. “Analysis of Hundreds of Cis-Regulatory Landscapes at High Resolution in a Single, High-Throughput Experiment.” *Nature Genetics*. https://doi.org/10.1038/ng.2871.
Huynh, Khanh D., and Jeannie T. Lee. 2003. “Inheritance of a Pre-Inactivated Paternal X Chromosome in Early Mouse Embryos.” *Nature* 426 (6968): 857–62.
Janiszewski, Adrian, Irene Talon, Joel Chappell, Samuel Collombet, Juan Song, Natalie De Geest, San Kit To, et al. 2019. “Dynamic Reversal of Random X-Chromosome Inactivation during iPSC Reprogramming.” *Genome Research*. https://doi.org/10.1101/gr.249706.119.
Jansz, Natasha, Tatyana Nesterova, Andrew Keniry, Megan Iminitoff, Peter F. Hickey, Greta Pintacuda, Osamu Masui, et al. 2018. “Smchd1 Targeting to the Inactive X Is Dependent on the Xist-HnrnpK-PRC1 Pathway.” *Cell Reports* 25 (7): 1912–23.e9.
Jégu, Teddy, Eric Aeby, and Jeannie T. Lee. 2017. “The X Chromosome in Space.” *Nature Reviews. Genetics* 18 (6): 377–89.
Jonkers, Iris, Tahsin Stefan Barakat, Eskeatnaf Mulugeta Achame, Kim Monkhorst, Annegien Kenter, Eveline Rentmeester, Frank Grosveld, J. Anton Grootegoed, and Joost Gribnau. 2009. “RNF12 Is an X-Encoded Dose-Dependent Activator of X Chromosome Inactivation.” *Cell* 139 (5): 999–1011.
Kalmar, Tibor, Chea Lim, Penelope Hayward, Silvia Muñoz-Descalzo, Jennifer Nichols, Jordi Garcia-Ojalvo, and Alfonso Martinez Arias. 2009. “Regulated Fluctuations in Nanog Expression Mediate Cell Fate Decisions in Embryonic Stem Cells.” *PLoS Biology* 7 (7): e1000149.
Keniry, Andrew, Linden J. Gearing, Natasha Jansz, Joy Liu, Aliaksei Z. Holik, Peter F. Hickey, Sarah A. Kinkel, et al. 2016. “Setdb1-Mediated H3K9 Methylation Is Enriched on the Inactive X and Plays a Role in Its Epigenetic Silencing.” *Epigenetics & Chromatin*. https://doi.org/10.1186/s13072-016-0064-6.
Knaupp, Anja S., Sam Buckberry, Jahnvi Pflueger, Sue Mei Lim, Ethan Ford, Michael R. Larcombe, Fernando J. Rossello, et al. 2017. “Transient and Permanent Reconfiguration of Chromatin and Transcription Factor Occupancy Drive Reprogramming.” *Cell Stem Cell* 21 (6): 834–45.e6.
Kohlmaier, Alexander, Fabio Savarese, Monika Lachner, Joost Martens, Thomas Jenuwein, and

Anton Wutz. 2004. “A Chromosomal Memory Triggered by Xist Regulates Histone Methylation in X Inactivation.” *PLoS Biology*. https://doi.org/10.1371/journal.pbio.0020171.
Lee, Jeannie T. 2005. “Regulation of X-Chromosome Counting by Tsix and Xite Sequences.” *Science* 309 (5735): 768–71.
———. 2011. “Gracefully Ageing at 50, X-Chromosome Inactivation Becomes a Paradigm for RNA and Chromatin Control.” *Nature Reviews. Molecular Cell Biology* 12 (12): 815–26.
Lee, Jeannie T., and Naifang Lu. 1999. “Targeted Mutagenesis of Tsix Leads to Nonrandom X Inactivation.” *Cell*. https://doi.org/10.1016/s0092-8674(00)80061-6.
Lee, J. T. 2000. “Disruption of Imprinted X Inactivation by Parent-of-Origin Effects at Tsix.” *Cell* 103 (1): 17–27.
Lee, J. T., L. S. Davidow, and D. Warshawsky. 1999. “Tsix, a Gene Antisense to Xist at the X-Inactivation Centre.” *Nature Genetics* 21 (4): 400–404.
Lee, J. T., and N. Lu. 1999. “Targeted Mutagenesis of Tsix Leads to Nonrandom X Inactivation.” *Cell* 99 (1): 47–57.
Lee, J. T., N. Lu, and Y. Han. 1999. “Genetic Analysis of the Mouse X Inactivation Center Defines an 80-Kb Multifunction Domain.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (7): 3836–41.
Lee, J. T., W. M. Strauss, J. A. Dausman, and R. Jaenisch. 1996. “A 450 Kb Transgene Displays Properties of the Mammalian X-Inactivation Center.” *Cell* 86 (1): 83–94.
Liao, Shuren, Margaret Tammaro, and Hong Yan. 2015. “Enriching CRISPR-Cas9 Targeted Cells by Co-Targeting the HPRT Gene.” *Nucleic Acids Research* 43 (20): e134.
Li, Dongwei, Jing Liu, Xuejie Yang, Chunhua Zhou, Jing Guo, Chuman Wu, Yue Qin, et al. 2017. “Chromatin Accessibility Dynamics during iPSC Reprogramming.” *Cell Stem Cell* 21 (6): 819–33.e6.
Lieberman-Aiden, Erez, Nynke L. van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, et al. 2009. “Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome.” *Science* 326 (5950): 289–93.
Liu, X., H. Wu, J. Loring, S. Hormuzdi, C. M. Disteche, P. Bornstein, and R. Jaenisch. 1997. “Trisomy Eight in ES Cells Is a Common Potential Problem in Gene Targeting and Interferes with Germ Line Transmission.” *Developmental Dynamics: An Official Publication of the American Association of Anatomists* 209 (1): 85–91.
Luikenhuis, S., A. Wutz, and R. Jaenisch. 2001. “Antisense Transcription through the Xist Locus Mediates Tsix Function in Embryonic Stem Cells.” *Molecular and Cellular Biology* 21 (24): 8512–20.
Lyon, Mary F. 1961. “Gene Action in the X-Chromosome of the Mouse (Mus Musculus L.).” *Nature*. https://doi.org/10.1038/190372a0.
Maenner, Sylvain, Magali Blaud, Laetitia Fouillen, Anne Savoye, Virginie Marchand, Agnès Dubois, Sarah Sanglier-Cianférani, et al. 2010. “2-D Structure of the A Region of Xist RNA and Its Implication for PRC2 Association.” *PLoS Biology* 8 (1): e1000276.
Maherali, Nimet, Rupa Sridharan, Wei Xie, Jochen Utikal, Sarah Eminli, Katrin Arnold, Matthias Stadtfeld, et al. 2007. “Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution.” *Cell Stem Cell* 1 (1): 55–70.
Mak, Winifred, Tatyana B. Nesterova, Mariana de Napoles, Ruth Appanah, Shinya Yamanaka, Arie P. Otte, and Neil Brockdorff. 2004. “Reactivation of the Paternal X Chromosome in Early Mouse Embryos.” *Science* 303 (5658): 666–69.
Mallol, Anna, Maria Guirola, and Bernhard Payer. 2019. “PRDM14 Controls X-Chromosomal and Global Epigenetic Reprogramming of H3K27me3 in Migrating Mouse Primordial Germ

Cells.” *Epigenetics & Chromatin* 12 (1): 38.
Marahrens, Y., B. Panning, J. Dausman, W. Strauss, and R. Jaenisch. 1997a. “Xist-Deficient Mice Are Defective in Dosage Compensation but Not Spermatogenesis.” *Genes & Development*. https://doi.org/10.1101/gad.11.2.156.
———. 1997b. “Xist-Deficient Mice Are Defective in Dosage Compensation but Not Spermatogenesis.” *Genes & Development* 11 (2): 156–66.
Marks, Hendrik, Hindrik H. D. Kerstens, Tahsin Stefan Barakat, Erik Splinter, René A. M. Dirks, Guido van Mierlo, Onkar Joshi, et al. 2015. “Dynamics of Gene Silencing during X Inactivation Using Allele-Specific RNA-Seq.” *Genome Biology* 16 (August): 149.
McHugh, Colleen A., Chun-Kan Chen, Amy Chow, Christine F. Surka, Christina Tran, Patrick McDonel, Amy Pandya-Jones, et al. 2015. “The Xist lncRNA Interacts Directly with SHARP to Silence Transcription through HDAC3.” *Nature* 521 (7551): 232–36.
Mermoud, J. E., C. Costanzi, J. R. Pehrson, and N. Brockdorff. 1999. “Histone macroH2A1.2 Relocates to the Inactive X Chromosome after Initiation and Propagation of X-Inactivation.” *The Journal of Cell Biology* 147 (7): 1399–1408.
Migeon, Barbara R. 2006. “The Role of X Inactivation and Cellular Mosaicism in Women’s Health and Sex-Specific Diseases.” *JAMA: The Journal of the American Medical Association* 295 (12): 1428–33.
Mikkelsen, Tarjei S., Jacob Hanna, Xiaolan Zhang, Manching Ku, Marius Wernig, Patrick Schorderet, Bradley E. Bernstein, Rudolf Jaenisch, Eric S. Lander, and Alexander Meissner. 2008. “Dissecting Direct Reprogramming through Integrative Genomic Analysis.” *Nature* 454 (7200): 49–55.
Minajigi, A., J. E. Froberg, C. Wei, H. Sunwoo, B. Kesner, D. Colognori, D. Lessing, et al. 2015. “A Comprehensive Xist Interactome Reveals Cohesin Repulsion and an RNA-Directed Chromosome Conformation.” *Science*. https://doi.org/10.1126/science.aab2276.
Minajigi, Anand, John Froberg, Chunyao Wei, Hongjae Sunwoo, Barry Kesner, David Colognori, Derek Lessing, et al. 2015. “Chromosomes. A Comprehensive Xist Interactome Reveals Cohesin Repulsion and an RNA-Directed Chromosome Conformation.” *Science* 349 (6245). https://doi.org/10.1126/science.aab2276.
Min, Jinrong, Yi Zhang, and Rui-Ming Xu. 2003. “Structural Basis for Specific Binding of Polycomb Chromodomain to Histone H3 Methylated at Lys 27.” *Genes & Development* 17 (15): 1823–28.
Minkovsky, Alissa, Tahsin Stefan Barakat, Nadia Sellami, Mark Henry Chin, Nilhan Gunhanlar, Joost Gribnau, and Kathrin Plath. 2013. “The Pluripotency Factor-Bound Intron 1 of Xist Is Dispensable for X Chromosome Inactivation and Reactivation in Vitro and in Vivo.” *Cell Reports* 3 (3): 905–18.
Mira-Bontenbal, Hegias, and Joost Gribnau. 2016. “New Xist-Interacting Proteins in X-Chromosome Inactivation.” *Current Biology: CB* 26 (10): 1383.
Moindrot, Benoit, Andrea Cerase, Heather Coker, Osamu Masui, Anne Grijzenhout, Greta Pintacuda, Lothar Schermelleh, Tatyana B. Nesterova, and Neil Brockdorff. 2015. “A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing.” *Cell Reports* 12 (4): 562–72.
Monfort, Asun, Giulio Di Minin, Andreas Postlmayr, Remo Freimann, Fabiana Arieti, Stéphane Thore, and Anton Wutz. 2015. “Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells.” *Cell Reports* 12 (4): 554–61.
Monkhorst, Kim, Iris Jonkers, Eveline Rentmeester, Frank Grosveld, and Joost Gribnau. 2008. “X Inactivation Counting and Choice Is a Stochastic Process: Evidence for Involvement of

an X-Linked Activator." *Cell*. https://doi.org/10.1016/j.cell.2007.12.036.
Monk, Marilyn, and Mary I. Harper. 1979. "Sequential X Chromosome Inactivation Coupled with Cellular Differentiation in Early Mouse Embryos." *Nature*. https://doi.org/10.1038/281311a0.
Morey, Céline, and Philip Avner. 2011. "The Demoiselle of X-Inactivation: 50 Years Old and as Trendy and Mesmerising as Ever." *PLoS Genetics* 7 (7): e1002212.
Napoles, Mariana de, Mariana de Napoles, Jacqueline E. Mermoud, Rika Wakao, Y. Amy Tang, Mitusuhiro Endoh, Ruth Appanah, et al. 2004. "Polycomb Group Proteins Ring1A/B Link Ubiquitylation of Histone H2A to Heritable Gene Silencing and X Inactivation." *Developmental Cell*. https://doi.org/10.1016/j.devcel.2004.10.005.
Napoles, Mariana de, Tatyana Nesterova, and Neil Brockdorff. 2007. "Early Loss of Xist RNA Expression and Inactive X Chromosome Associated Chromatin Modification in Developing Primordial Germ Cells." *PloS One* 2 (9): e860.
Naughton, Catherine, Duncan Sproul, Charlotte Hamilton, and Nick Gilbert. 2010. "Analysis of Active and Inactive X Chromosome Architecture Reveals the Independent Organization of 30 Nm and Large-Scale Chromatin Structures." *Molecular Cell* 40 (3): 397–409.
Navarro, Pablo, Ian Chambers, Violetta Karwacki-Neisius, Corinne Chureau, Céline Morey, Claire Rougeulle, and Philip Avner. 2008. "Molecular Coupling of Xist Regulation and Pluripotency." *Science* 321 (5896): 1693–95.
Navarro, Pablo, Michael Moffat, Nicholas P. Mullin, and Ian Chambers. 2011. "The X-Inactivation Trans-Activator Rnf12 Is Negatively Regulated by Pluripotency Factors in Embryonic Stem Cells." *Human Genetics* 130 (2): 255–64.
Nesterova, Tatyana B., Jacqueline E. Mermoud, Kathy Hilton, John Pehrson, M. Azim Surani, Anne McLaren, and Neil Brockdorff. 2002. "Xist Expression and macroH2A1.2 Localisation in Mouse Primordial and Pluripotent Embryonic Germ Cells." *Differentiation; Research in Biological Diversity* 69 (4-5): 216–25.
Nesterova, Tatyana B., Guifeng Wei, Heather Coker, Greta Pintacuda, Joseph S. Bowness, Tianyi Zhang, Mafalda Almeida, et al. 2019. "Systematic Allelic Analysis Defines the Interplay of Key Pathways in X Chromosome Inactivation." https://doi.org/10.1101/477232.
Nora, Elphège P., Bryan R. Lajoie, Edda G. Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Piolot, et al. 2012. "Spatial Partitioning of the Regulatory Landscape of the X-Inactivation Centre." *Nature* 485 (7398): 381–85.
Ogawa, Y., B. K. Sun, and J. T. Lee. 2008. "Intersection of the RNA Interference and X-Inactivation Pathways." *Science* 320 (5881): 1336–41.
Ohta, Hiroshi, Kazuki Kurimoto, Ikuhiro Okamoto, Tomonori Nakamura, Yukihiro Yabuta, Hidetaka Miyauchi, Takuya Yamamoto, et al. 2017. "Expansion of Mouse Primordial Germ Cell-like Cells Recapitulates an Epigenetic Blank Slate." *The EMBO Journal* 36 (13): 1888–1907.
Okamoto, Ikuhiro, Arie P. Otte, C. David Allis, Danny Reinberg, and Edith Heard. 2004. "Epigenetic Dynamics of Imprinted X Inactivation during Early Mouse Development." *Science* 303 (5658): 644–49.
Pageau, Gayle J., Lisa L. Hall, Shridar Ganesan, David M. Livingston, and Jeanne B. Lawrence. 2007. "The Disappearing Barr Body in Breast and Ovarian Cancers." *Nature Reviews. Cancer* 7 (8): 628–33.
Panning, B., and R. Jaenisch. 1996. "DNA Hypomethylation Can Activate Xist Expression and Silence X-Linked Genes." *Genes & Development* 10 (16): 1991–2002.
Pasque, Vincent, Astrid Gillich, Nigel Garrett, and John B. Gurdon. 2011. "Histone Variant macroH2A Confers Resistance to Nuclear Reprogramming." *The EMBO Journal* 30 (12): 2373–87.

Pasque, Vincent, and Kathrin Plath. 2015. “X Chromosome Reactivation in Reprogramming and in Development.” *Current Opinion in Cell Biology* 37 (December): 75–83.
Pasque, Vincent, Jason Tchieu, Rahul Karnik, Molly Uyeda, Anupama Sadhu Dimashkie, Dana Case, Bernadett Papp, et al. 2014. “X Chromosome Reactivation Dynamics Reveal Stages of Reprogramming to Pluripotency.” *Cell* 159 (7): 1681–97.
Patil, Deepak P., Chun-Kan Chen, Brian F. Pickering, Amy Chow, Constanza Jackson, Mitchell Guttman, and Samie R. Jaffrey. 2016. “m6A RNA Methylation Promotes XIST-Mediated Transcriptional Repression.” *Nature*. https://doi.org/10.1038/nature19342.
Payer, Bernhard. 2016. “Developmental Regulation of X-Chromosome Inactivation.” *Seminars in Cell & Developmental Biology* 56 (August): 88–99.
———. 2017. “Epigenetic Regulation of X-Chromosome Inactivation.” *Epigenetics in Human Reproduction and Development*. https://doi.org/10.1142/9789813144279_0005.
Payer, Bernhard, and Jeannie T. Lee. 2014. “Coupling of X-Chromosome Reactivation with the Pluripotent Stem Cell State.” *RNA Biology* 11 (7): 798–807.
Payer, Bernhard, Michael Rosenberg, Masashi Yamaji, Yukihiro Yabuta, Michiyo Koyanagi-Aoi, Katsuhiko Hayashi, Shinya Yamanaka, Mitinori Saitou, and Jeannie T. Lee. 2013. “Tsix RNA and the Germline Factor, PRDM14, Link X Reactivation and Stem Cell Reprogramming.” *Molecular Cell* 52 (6): 805–18.
Peñalosa-Ruiz, Georgina, Ann Rose Bright, Klaas W. Mulder, and Gert Jan C. Veenstra. 2019. “The Interplay of Chromatin and Transcription Factors during Cell Fate Transitions in Development and Reprogramming.” *Biochimica et Biophysica Acta, Gene Regulatory Mechanisms* 1862 (9): 194407.
Penny, Graeme D., Graham F. Kay, Steven A. Sheardown, Sohaila Rastan, and Neil Brockdorff. 1996. “Requirement for Xist in X Chromosome Inactivation.” *Nature*. https://doi.org/10.1038/379131a0.
Pinder, Jordan, Jayme Salsman, and Graham Dellaire. 2015. “Nuclear Domain ‘knock-In’ Screen for the Evaluation and Identification of Small Molecule Enhancers of CRISPR-Based Genome Editing.” *Nucleic Acids Research* 43 (19): 9379–92.
Ping, Xiao-Li, Bao-Fa Sun, Lu Wang, Wen Xiao, Xin Yang, Wen-Jia Wang, Samir Adhikari, et al. 2014. “Mammalian WTAP Is a Regulatory Subunit of the RNA N6-Methyladenosine Methyltransferase.” *Cell Research*. https://doi.org/10.1038/cr.2014.3.
Pintacuda, Greta, Guifeng Wei, Chloë Roustan, Burcu Anil Kirmizitas, Nicolae Solcan, Andrea Cerase, Alfredo Castello, et al. 2017. “hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing.” *Molecular Cell* 68 (5): 955–69.e10.
Pinter, S. F., R. I. Sadreyev, E. Yildirim, Y. Jeon, T. K. Ohsumi, M. Borowsky, and J. T. Lee. 2012. “Spreading of X Chromosome Inactivation via a Hierarchy of Defined Polycomb Stations.” *Genome Research*. https://doi.org/10.1101/gr.133751.111.
Plath, K. 2003. “Role of Histone H3 Lysine 27 Methylation in X Inactivation.” *Science*. https://doi.org/10.1126/science.1084274.
Polo, Jose M., Endre Anderssen, Ryan M. Walsh, Benjamin A. Schwarz, Christian M. Nefzger, Sue Mei Lim, Marti Borkent, et al. 2012. “A Molecular Roadmap of Reprogramming Somatic Cells into iPS Cells.” *Cell* 151 (7): 1617–32.
Rao, Suhas S. P., Miriam H. Huntley, Neva C. Durand, Elena K. Stamenova, Ivan D. Bochkov, James T. Robinson, Adrian L. Sanborn, et al. 2014. “A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping.” *Cell* 159 (7): 1665–80.
Rocha, Simão Teixeira da, Valentina Boeva, Martin Escamilla-Del-Arenal, Katia Ancelin, Camille Granier, Neuza Reis Matias, Serena Sanulli, et al. 2014. “Jarid2 Is Implicated in the Initial

Xist-Induced Targeting of PRC2 to the Inactive X Chromosome." *Molecular Cell* 53 (2): 301–16.
Rohwedel, Jürgen, Kaomei Guan, and Anna M. Wobus. 1999. "Induction of Cellular Differentiation by Retinoic Acid in Vitro." *Cells Tissues Organs*. https://doi.org/10.1159/000016699.
Rougeulle, C., J. Chaumeil, K. Sarma, C. D. Allis, D. Reinberg, P. Avner, and E. Heard. 2004. "Differential Histone H3 Lys-9 and Lys-27 Methylation Profiles on the X Chromosome." *Molecular and Cellular Biology*. https://doi.org/10.1128/mcb.24.12.5475-5484.2004.
Rowley, M. Jordan, and Victor G. Corces. 2018. "Organizational Principles of 3D Genome Architecture." *Nature Reviews. Genetics* 19 (12): 789–800.
Rowley, M. Jordan, Michael H. Nichols, Xiaowen Lyu, Masami Ando-Kuri, I. Sarahi M. Rivera, Karen Hermetz, Ping Wang, Yijun Ruan, and Victor G. Corces. 2017. "Evolutionarily Conserved Principles Predict 3D Chromatin Organization." *Molecular Cell* 67 (5): 837–52.e7.
Sado, T., Z. Wang, H. Sasaki, and E. Li. 2001. "Regulation of Imprinted X-Chromosome Inactivation in Mice by Tsix." *Development* 128 (8): 1275–86.
Sakakibara, Yuki, Koji Nagao, Marnie Blewitt, Hiroyuki Sasaki, Chikashi Obuse, and Takashi Sado. 2018. "Role of SmcHD1 in Establishment of Epigenetic States Required for the Maintenance of the X-Inactivated State in Mice." *Development*. https://doi.org/10.1242/dev.166462.
Sardina, Jose Luis, Samuel Collombet, Tian V. Tian, Antonio Gómez, Bruno Di Stefano, Clara Berenguer, Justin Brumbaugh, et al. 2018. "Transcription Factors Drive Tet2-Mediated Enhancer Demethylation to Reprogram Cell Fate." *Cell Stem Cell* 23 (6): 905–6.
Schiebinger, Geoffrey, Jian Shu, Marcin Tabaka, Brian Cleary, Vidya Subramanian, Aryeh Solomon, Joshua Gould, et al. 2019. "Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming." *Cell* 176 (6): 1517.
Schulz, Edda G., Johannes Meisig, Tomonori Nakamura, Ikuhiro Okamoto, Anja Sieber, Christel Picard, Maud Borensztein, Mitinori Saitou, Nils Blüthgen, and Edith Heard. 2014. "The Two Active X Chromosomes in Female ESCs Block Exit from the Pluripotent State by Modulating the ESC Signaling Network." *Cell Stem Cell* 14 (2): 203–16.
Schwarz, Benjamin A., Murat Cetinbas, Kendell Clement, Ryan M. Walsh, Sihem Cheloufi, Hongcang Gu, Jan Langkabel, et al. 2018. "Prospective Isolation of Poised iPSC Intermediates Reveals Principles of Cellular Reprogramming." *Cell Stem Cell* 23 (2): 289–305.e5.
Scudellari, Megan. 2016. "How iPS Cells Changed the World." *Nature News* 534 (7607): 310.
Silva, Jose, Winifred Mak, Ilona Zvetkova, Ruth Appanah, Tatyana B. Nesterova, Zoe Webster, Antoine H. F. M. Peters, Thomas Jenuwein, Arie P. Otte, and Neil Brockdorff. 2003. "Establishment of Histone h3 Methylation on the Inactive X Chromosome Requires Transient Recruitment of Eed-Enx1 Polycomb Group Complexes." *Developmental Cell* 4 (4): 481–95.
Silva, Jose, Jennifer Nichols, Thorold W. Theunissen, Ge Guo, Anouk L. van Oosten, Ornella Barrandon, Jason Wray, Shinya Yamanaka, Ian Chambers, and Austin Smith. 2009. "Nanog Is the Gateway to the Pluripotent Ground State." *Cell* 138 (4): 722–37.
Simmler, M. 1996. "A 94 Kb Genomic Sequence 3' to the Murine Xist Gene Reveals an AT Rich Region Containing a New Testis Specific Gene Tsx." *Human Molecular Genetics*. https://doi.org/10.1093/hmg/5.11.1713.
Simonis, Marieke, Petra Klous, Erik Splinter, Yuri Moshkin, Rob Willemsen, Elzo de Wit, Bas van Steensel, and Wouter de Laat. 2006. "Nuclear Organization of Active and Inactive

Chromatin Domains Uncovered by Chromosome Conformation Capture–on-Chip (4C)." *Nature Genetics*. https://doi.org/10.1038/ng1896.

Simon, Matthew D., Stefan F. Pinter, Rui Fang, Kavitha Sarma, Michael Rutenberg-Schoenberg, Sarah K. Bowman, Barry A. Kesner, Verena K. Maier, Robert E. Kingston, and Jeannie T. Lee. 2013. "High-Resolution Xist Binding Maps Reveal Two-Step Spreading during X-Chromosome Inactivation." *Nature* 504 (7480): 465–69.

Singhal, Prabhat K., Slim Sassi, Lan Lan, Patrick Au, Stefan C. Halvorsen, Dai Fukumura, Rakesh K. Jain, and Brian Seed. 2016. "Mouse Embryonic Fibroblasts Exhibit Extensive Developmental and Phenotypic Diversity." *Proceedings of the National Academy of Sciences of the United States of America* 113 (1): 122–27.

"Sodium Phosphate Buffer for Church and Gilbert Hybridization:" 2015a. *Cold Spring Harbor Protocols*. https://doi.org/10.1101/pdb.rec086579.

———. 2015b. *Cold Spring Harbor Protocols*. https://doi.org/10.1101/pdb.rec086579.

Song, Juan, Adrian Janiszewski, Natalie De Geest, Lotte Vanheer, Irene Talon, Mouna El Bakkali, Taeho Oh, and Vincent Pasque. 2019. "X-Chromosome Dosage Modulates Multiple Molecular and Cellular Properties of Mouse Pluripotent Stem Cells Independently of Global DNA Methylation Levels." *Stem Cell Reports* 12 (2): 333–50.

Soufi, Abdenour, Greg Donahue, and Kenneth S. Zaret. 2012. "Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome." *Cell* 151 (5): 994–1004.

Soufi, Abdenour, Meilin Fernandez Garcia, Artur Jaroszewicz, Nebiyu Osman, Matteo Pellegrini, and Kenneth S. Zaret. 2015. "Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming." *Cell* 161 (3): 555–68.

Sousa, Lisa Barros De Andrade e., Lisa Barros De Andrade e Sousa, Iris Jonkers, Laurene Syx, Julie Chaumell, Christel Picard, Benjamin Foret, et al. 2019. "Kinetics of Xist-Induced Gene Silencing Can Be Predicted from Combinations of Epigenetic and Genomic Features." *Genome Research*, no. 29 (June): 1087–99.

Spencer, Rebecca J., Brian C. del Rosario, Stefan F. Pinter, Derek Lessing, Ruslan I. Sadreyev, and Jeannie T. Lee. 2011. "A Boundary Element between Tsix and Xist Binds the Chromatin Insulator Ctcf and Contributes to Initiation of X-Chromosome Inactivation." *Genetics* 189 (2): 441–54.

Splinter, Erik, Elzo de Wit, Elphège P. Nora, Petra Klous, Harmen J. G. van de Werken, Yun Zhu, Lucas J. T. Kaaij, et al. 2011. "The Inactive X Chromosome Adopts a Unique Three-Dimensional Conformation That Is Dependent on Xist RNA." *Genes & Development* 25 (13): 1371–83.

Stadhouders, Ralph, Guillaume J. Filion, and Thomas Graf. 2019. "Transcription Factors and 3D Genome Conformation in Cell-Fate Decisions." *Nature* 569 (7756): 345–54.

Stadhouders, Ralph, Enrique Vidal, François Serra, Bruno Di Stefano, François Le Dily, Javier Quilez, Antonio Gomez, et al. 2018. "Transcription Factors Orchestrate Dynamic Interplay between Genome Topology and Gene Regulation during Cell Reprogramming." *Nature Genetics* 50 (2): 238–49.

Stadtfeld, Matthias, Nimet Maherali, David T. Breault, and Konrad Hochedlinger. 2008. "Defining Molecular Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse." *Cell Stem Cell* 2 (3): 230–40.

Sugimoto, Michihiko, and Kuniya Abe. 2007. "X Chromosome Reactivation Initiates in Nascent Primordial Germ Cells in Mice." *PLoS Genetics* 3 (7): e116.

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

Tada, M., Y. Takahama, K. Abe, N. Nakatsuji, and T. Tada. 2001. “Nuclear Reprogramming of Somatic Cells by in Vitro Hybridization with ES Cells.” *Current Biology: CB* 11 (19): 1553–58.
Takagi, N., and M. Sasaki. 1975. “Preferential Inactivation of the Paternally Derived X Chromosome in the Extraembryonic Membranes of the Mouse.” *Nature* 256 (5519): 640–42.
Takagi, N., M. A. Yoshida, O. Sugawara, and M. Sasaki. 1983. “Reversal of X-Inactivation in Female Mouse Somatic Cells Hybridized with Murine Teratocarcinoma Stem Cells in Vitro.” *Cell* 34 (3): 1053–62.
Takahashi, Kazutoshi, and Shinya Yamanaka. 2006. “Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors.” *Cell* 126 (4): 663–76.
Tian, Di, Sha Sun, and Jeannie T. Lee. 2010. “The Long Noncoding RNA, Jpx, Is a Molecular Switch for X Chromosome Inactivation.” *Cell*. https://doi.org/10.1016/j.cell.2010.09.049.
Tsai, Chia-Lun, Rebecca K. Rowntree, Dena E. Cohen, and Jeannie T. Lee. 2008. “Higher Order Chromatin Structure at the X-Inactivation Center via Looping DNA.” *Developmental Biology* 319 (2): 416–25.
Vidal, Simon E., Bhishma Amlani, Taotao Chen, Aristotelis Tsirigos, and Matthias Stadtfeld. 2014. “Combinatorial Modulation of Signaling Pathways Reveals Cell-Type-Specific Requirements for Highly Efficient and Synchronous iPSC Reprogramming.” *Stem Cell Reports* 3 (4): 574–84.
Wang, Chen-Yu, David Colognori, Hongjae Sunwoo, Danni Wang, and Jeannie T. Lee. 2019. “PRC1 Collaborates with SMCHD1 to Fold the X-Chromosome and Spread Xist RNA between Chromosome Compartments.” *Nature Communications*. https://doi.org/10.1038/s41467-019-10755-3.
Wang, Chen-Yu, Teddy Jégu, Hsueh-Ping Chu, Hyun Jung Oh, and Jeannie T. Lee. 2018. “SMCHD1 Merges Chromosome Compartments and Assists Formation of Super-Structures on the Inactive X.” *Cell* 174 (2): 406–21.e25.
Wang, J., J. Mager, Y. Chen, E. Schneider, J. C. Cross, A. Nagy, and T. Magnuson. 2001. “Imprinted X Inactivation Maintained by a Mouse Polycomb Group Gene.” *Nature Genetics* 28 (4): 371–75.
Wang, Siyuan, Jun-Han Su, Brian J. Beliveau, Bogdan Bintu, Jeffrey R. Moffitt, Chao-Ting Wu, and Xiaowei Zhuang. 2016. “Spatial Organization of Chromatin Domains and Compartments in Single Chromosomes.” *Science* 353 (6299): 598–602.
Williams, Lucy H., Sundeep Kalantry, Joshua Starmer, and Terry Magnuson. 2011. “Transcription Precedes Loss of Xist Coating and Depletion of H3K27me3 during X-Chromosome Reprogramming in the Mouse Inner Cell Mass.” *Development* 138 (10): 2049–57.
Wu, Hao, Junjie Luo, Huimin Yu, Amir Rattner, Alisa Mo, Yanshu Wang, Philip M. Smallwood, Bracha Erlanger, Sarah J. Wheelan, and Jeremy Nathans. 2014. “Cellular Resolution Maps of X Chromosome Inactivation: Implications for Neural Development, Function, and Disease.” *Neuron* 81 (1): 103–19.
Wutz, A., and R. Jaenisch. 2000. “A Shift from Reversible to Irreversible X Inactivation Is Triggered during ES Cell Differentiation.” *Molecular Cell* 5 (4): 695–705.
Wutz, Anton, Theodore P. Rasmussen, and Rudolf Jaenisch. 2002. “Chromosomal Silencing and Localization Are Mediated by Different Domains of Xist RNA.” *Nature Genetics* 30 (2): 167–74.
Yamashiro, Chika, Kotaro Sasaki, Yukihiro Yabuta, Yoji Kojima, Tomonori Nakamura, Ikuhiro Okamoto, Shihori Yokobayashi, et al. 2018. “Generation of Human Oogonia from Induced

Pluripotent Stem Cells in Vitro." *Science*. https://doi.org/10.1126/science.aat1674.
Yang, Fan, Xinxian Deng, Wenxiu Ma, Joel B. Berletch, Natalia Rabaia, Gengze Wei, James M. Moore, et al. 2015. "The lncRNA Firre Anchors the Inactive X Chromosome to the Nucleolus by Binding CTCF and Maintains H3K27me3 Methylation." *Genome Biology*. https://doi.org/10.1186/s13059-015-0618-0.
Yang, Hui, Haoyi Wang, Chikdu S. Shivalila, Albert W. Cheng, Linyu Shi, and Rudolf Jaenisch. 2013. "One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering." *Cell* 154 (6): 1370–79.
Yildirim, Eda, James E. Kirby, Diane E. Brown, Francois E. Mercier, Ruslan I. Sadreyev, David T. Scadden, and Jeannie T. Lee. 2013. "Xist RNA Is a Potent Suppressor of Hematologic Cancer in Mice." *Cell* 152 (4): 727–42.
Ying, Qi-Long, Jason Wray, Jennifer Nichols, Laura Batlle-Morera, Bradley Doble, James Woodgett, Philip Cohen, and Austin Smith. 2008. "The Ground State of Embryonic Stem Cell Self-Renewal." *Nature*. https://doi.org/10.1038/nature06968.
Zhao, Jing, Bryan K. Sun, Jennifer A. Erwin, Ji-Joon Song, and Jeannie T. Lee. 2008. "Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome." *Science* 322 (5902): 750–56.
Zhao, Zhihu, Gholamreza Tavoosidana, Mikael Sjölinder, Anita Göndör, Piero Mariano, Sha Wang, Chandrasekhar Kanduri, et al. 2006. "Circular Chromosome Conformation Capture (4C) Uncovers Extensive Networks of Epigenetically Regulated Intra- and Interchromosomal Interactions." *Nature Genetics*. https://doi.org/10.1038/ng1891.
Zvetkova, Ilona, Anwyn Apedaile, Bernard Ramsahoye, Jacqueline E. Mermoud, Lucy A. Crompton, Rosalind John, Robert Feil, and Neil Brockdorff. 2005. "Global Hypomethylation of the Genome in XX Embryonic Stem Cells." *Nature Genetics* 37 (11): 1274–79.
Żylicz, Jan Jakub, Aurélie Bousard, Kristina Žumer, Francois Dossin, Eusra Mohammad, Simão Teixeira da Rocha, Björn Schwalb, et al. 2019. "The Implication of Early Chromatin Changes in X Chromosome Inactivation." *Cell*. https://doi.org/10.1016/j.cell.2018.11.041.

# Abbreviations / Glossary

| | |
|---|---|
| **Cas** | *Mus musculus castaneus* |
| **ESC** | Embryonic stem cell |
| **Escapee** | Gene with biallelic expression in somatic cells, escaping X-chromosome inactivation |
| **FACS** | Fluorescence-activated cell sorting |
| **iMEF** | Irradiated mouse embryonic fibroblast / Feeder cell |
| **iPSC** | Induced pluripotent stem cell |
| **MEF** | Mouse embryonic fibroblast |
| **Mus** | *Mus musculus musculus* |
| **NPC** | Neural precursor cell |
| **PRC1/2** | Polycomb repressive complex 1/2 |
| **P-RFP** | RFP pluripotency reporter |
| **TAD** | Topologically associating domain |
| **Xa** | Active X-chromosome |
| **$X^{cas}$** | X-chromosome *M.m castaneus* |
| **XCI** | X-chromosome inactivation |
| **XCR** | X-chromosome reactivation |
| **X-DC** | X-Dual color GFP and tdTomato X-chromosome activity reporters |
| **X-GFP** | GFP X-chromosome activity reporter |
| **Xi** | Inactive X-chromosome |
| **Xic** | X-chromosome inactivation center |
| **$X^{mus}$** | X-chromosome *M.m musculus* |

This thesis was funded by a ”la Caixa” fellowship.