

Mechanisms of Xist-Mediated Gene Silencing During the Initiation and Maintenance of X Chromosome Inactivation

Thesis by
Mackenzie Strehle

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

The logo for the California Institute of Technology (Caltech), featuring the word "Caltech" in a bold, orange, sans-serif font.

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2024
(Defended March 29, 2024)

© 2024

Mackenzie Strehle
ORCID: 0000-0003-1410-8701

ACKNOWLEDGEMENTS

I am incredibly thankful for my PhD experience at Caltech, which would not have been possible without the love and encouragement from so many individuals both inside and outside the lab. The last six years have been filled with some of the highest highs and lowest lows of my life, and to everyone who has supported me through these times, I am forever grateful.

First and foremost, I would like to thank my advisor, Mitch Guttman. Mitch was one of the first professors I met during my graduate school interviews at Caltech, and I remember leaving his office after my interview thinking “Wow, that man talks fast” and also, “How do I join his lab?”. From the beginning of my time as his student, Mitch has had a profound impact on my development as a scientist and critical thinker. During my first year in the lab, he gave me the opportunity to write a review article with him, and he leaned into what would become my thesis work from the earliest days of my PhD, providing guidance and new ideas to drive the project forward. He has also gone above and beyond to support me both personally and professionally throughout the last six years, and I am extremely thankful for his kindness and thoughtfulness. Mitch is an exceptional mentor and I feel incredibly privileged to have found my home in the Guttman lab at Caltech.

There are many people in the lab I must also thank. Most importantly, I would like to thank my mentor for the first four years of my PhD, Joanna Jachowicz. Joanna began training me during my rotation and we quickly formed a close bond, both as mentor/mentee and as friends. She has been instrumental in shaping me into the scientist I am today, and I will forever cherish the memories we made working and going on adventures together. Prashant Bhat is my best

friend in the lab and has been a constant source of love, laughter, and encouragement. He is a brilliant scientist, and I will miss having his bright smile and boundless energy sitting next to me every day. Mario Blanco has also played a fundamental role in my development as a scientist, and he has been a huge support for me both professionally and personally. I also thank Amy Chow for her help in designing experiments and bringing delicious treats for cookie hour. In all seriousness, Amy is vital to the lab and my PhD would not have been possible without her tireless work to make sure we are organized and prepared to perform our experiments. Thank you to Olivia for always making time to go on walks to the turtle pond together and chase down free food across campus. Thank you to the team who helps our manuscripts come to life: Inna-Marie Strazhnik for her beautiful scientific illustrations, and Shawna Hiley for her editorial advice and ideation. Thank you to Anthony Vasquez for his help scheduling meetings, keeping track of lab business, and making sure we are well fed. And thank you to all the lab members whom I've had the great pleasure of working with throughout the last six years: Jimmy Guo, Abhik Banerjee, Jasmine Thai, Linlin Chen, Zahra Azhar, Drew Honson, Peter Chovanec, and Jessica Mann.

I would also like to thank several faculty mentors. I thank Paul Sternberg, Magda Zernicka-Goetz, and Shasha Chong for providing valuable feedback and advice on my thesis work. Shasha has been especially helpful with her extensive knowledge of Halo imaging and the biophysics of molecular condensates. I thank Betty Hong for the opportunity to lead as head teaching assistant for Bi 8 – I learned so much about being a great mentor and teacher from our time working together. I also thank my undergraduate mentor Alan Christensen “Dr. C”, who offered me the opportunity to start working as an undergraduate researcher in his lab during my freshman year at the University of Nebraska. If not for this,

I truly don't believe I would be where I am today. And thank you to Steve Caplan and Naava Naslavsky at the University of Nebraska Medical Center for also fostering my growth as a young scientist during the summer I spent working in their lab.

There are several other people whom I've crossed paths with on my journey to this current moment I would like to thank. Thank you to Emma Purfeerst, who was my first mentor in the lab as an undergraduate and taught me so much about molecular biology experimental techniques, which I've used throughout my PhD. I thank Kriti Bahl for her mentorship during the summer I spent working at the University of Nebraska Medical Center. Thank you also to Helen McBride, who I had the privilege of working under as an intern with the Office of Technology Transfer & Corporate Partnerships at Caltech. Since I met Helen, she has been an amazing mentor and source of insight as I begin to plan my future. And thank you to Liz Ayala for working tirelessly to help run the division and always having an open door to chat.

Thank you to all my friends who have stood by me throughout my PhD. Thanks to Meghan Matt for being my best friend for the last thirteen years. Thank you to Michaela Mapes and Kurt Cronican for always making time to get together. And thank you to Tara Chari, Ben Riviere, Kishan Patel, Mate Borsos, Mark Zhang, Morgan Schwartz, Matthew Langley, Fayth Tan, Lex Colon, Maddie Vera-Colon, Gio Tomaleri, and Terry Kim – I have enjoyed immensely getting to know and spend time with each of you.

I would not be where I am today without my family, and words cannot begin to describe how grateful I am for their eternal love and support. To my parents, thank you for fostering my love of math and science since I was a young girl and teaching me that I can do anything I put my mind to.

My parents made immense sacrifices to get me to where I am today, and I cannot thank them enough for all they've done for me. Thank you to my sister, Allie, for all the laughs and adventures, and for inspiring me by your own drive and determination. And thank you to my grandparents, Dean and Judy, who have always stood by my side and reminded me to give my best at everything I do.

Finally, thank you to my best friend and life partner, Andrew Schacht. Thank you for supporting me through the highs and lows of the last eight years and for being my biggest cheerleader in all aspects of my life. I cannot imagine having completed this journey without Andrew, and I am forever grateful for his endless love, trust, and encouragement.

ABSTRACT

X chromosome inactivation (XCI) is a critical development process during which one of the two X chromosomes in female mammals is silenced to balance gene expression with males. XCI is initiated by upregulation of the long noncoding RNA (lncRNA) Xist from the future inactive X chromosome (Xi), which recruits a variety of proteins *in cis* to mediate transcriptional repression that is maintained throughout the lifetime of the organism. Recent studies have demonstrated that silencing following Xist expression is dependent on direct recruitment of the transcriptional silencing protein SHARP (also known as SPEN); however, the mechanism underlying formation of the Xi silencing compartment has remained poorly defined. Similarly, it has long been thought that maintenance of XCI occurs independently of Xist and depends on differential DNA methylation enrichment on the Xi, but the evidence in support of these views is lacking. Here, we show how low copy numbers of Xist can recruit SHARP in superstoichiometric excess to initiate gene silencing on the X and mediate formation of the silent Xi compartment. We also provide preliminary evidence suggesting that maintenance of XCI is Xist independent, but dependent on DNA methylation and histone deacetylation. Together, these results offer a more holistic view of the molecular mechanisms underlying both initiation and maintenance XCI, as well as provide a framework for further investigation into lncRNA biology and epigenetic regulation more broadly.

PUBLISHED CONTENT AND CONTRIBUTIONS

1. Strehle, M. & Guttman, M. Xist drives spatial compartmentalization of DNA and protein to orchestrate initiation and maintenance of X inactivation. *Curr. Opin. Cell. Biol.* **64**, 139–147 (2020).
<https://doi.org/10.1016/j.ceb.2020.04.009>

M.S. wrote the review with M.G.

2. Jachowicz, J.W.*, Strehle, M.*, Banerjee, A.K., Blanco, M.R., Thai, J. & Guttman, M. Xist spatially amplifies SHARP/SPEN recruitment to balance chromosome-wide silencing and specificity to the X chromosome. *Nat. Struct. Mol. Biol.* **29**, 239–249 (2022).
<https://doi.org/10.1038/s41594-022-00739-1>
*Equal Contribution

M.S. performed experiments, analyzed and interpreted data, generated figures and schematics, and wrote the paper with J.W.J. and M.G.

3. Markaki, Y., Gan Chong, J., Wang, Y., Jacobsen, E.C., Luong, C., Tan, S.Y.X., Jachowicz, J.W., Strehle, M., Maestrini, D., Banerjee, A.K., Mistry, B.A., Dror, I., Dossin, F., Schoneberg, J., Heard, E., Guttman, M., Chou, T. & Plath, K. Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **25**, 6174–6192.e32 (2021).
<https://doi.org/10.1016/j.cell.2021.11.028>

M.S. contributed the initial observation that the intrinsically disordered region of SPEN is critical for silencing on the X chromosome.

In loving memory of my father, Mike Strehle, who always encouraged me to do my best and reach for the stars. His love, patience, and curiosity have shaped me into the person I am today, and I hope in the future I am able to make the same lasting impact he made on everyone with whom he crossed paths. I miss you every day.

TABLE OF CONTENTS

Acknowledgements	iii
Abstract.....	vii
Published Content and Contributions	viii
Table of Contents	x

Chapter 1: Xist drives spatial compartmentalization of DNA and protein to orchestrate initiation and maintenance of X inactivation

1.1 Abstract	2
1.2 Overview.....	3
1.3 The Xist lncRNA orchestrates XCI.....	4
1.4 How Xist initiates transcriptional silencing on the X chromosome	5
1.5 How Xist and its silencing complex spread across the inactive X chromosome	8
1.6 How Xist establishes an epigenetically heritable state that maintains inactivation.....	10
1.7 A possible role for protein compartmentalization and phase separation in the maintenance of X chromosome inactivation	12
1.8 References.....	16

Chapter 2: Xist spatially amplifies SHARP/SPEN recruitment to balance chromosome-wide silencing and specificity to the X chromosome.....

2.1 Abstract	26
2.2 Introduction.....	27

2.3 Results.....	29
2.3.1 SHARP accumulates on the Xi in stoichiometric excess to Xist	29
2.3.2 SHARP forms concentration-dependent assemblies in the nucleus	33
2.3.3. SHARP recruitment to the Xi depends on IDR-mediated assembly	37
2.3.4 SHARP assemblies are required for gene silencing on the X	40
2.3.5 Low Xist expression levels limit spreading to autosomes.....	44
2.4 Discussion	48
2.5 Author Note.....	51
2.6 Methods.....	52
2.7 Extended Data	73
2.8 Supplementary Information	81
2.8.1 Supplementary Note.....	81
2.8.2 Supplementary Tables	82
2.9 References.....	85
Chapter 3: Investigating the mechanisms underlying maintenance of X chromosome inactivation.....	93
3.1 Abstract	94
3.2 Introduction.....	95
3.3 Results.....	98
3.3.1 Maintenance of XCI is Xist-independent, but sensitive to epigenetic inhibition	98

3.3.2 A fluorescent-based reporter system enables more efficient measurement of gene silencing during the transition to XCI maintenance	102
3.4 Discussion	106
3.5 Methods.....	109
3.6 Extended Data	114
3.7 References.....	115
Chapter 4: Conclusion and future directions	120
4.1 Conclusion and future directions	121

Chapter 1

Xist drives spatial compartmentalization of DNA and protein to orchestrate initiation and maintenance of X inactivation

Mackenzie Strehle and Mitchell Guttman

A modified version of this chapter was published as: Strehle, M. & Guttman, M. Xist drives spatial compartmentalization of DNA and protein to orchestrate initiation and maintenance of X inactivation. *Curr. Opin. Cell. Biol.* **64**, 139–147 (2020). <https://doi.org/10.1016/j.ceb.2020.04.009>

1.1 | Abstract

X chromosome inactivation is the process whereby one of the X chromosomes in female mammalian cells is silenced to equalize X-linked gene expression with males. XCI depends on the long noncoding RNA Xist, which coats the inactive X chromosome in *cis* and triggers a cascade of events that ultimately lead to chromosome-wide transcriptional silencing that is stable for the lifetime of an organism. In recent years, the discovery of proteins that interact with Xist have led to new insights into how the initiation of XCI occurs. Nevertheless, there are still various unknowns about the mechanisms by which Xist orchestrates and maintains stable X-linked silencing. Here, we review recent work elucidating the role of Xist and its protein partners in mediating chromosome-wide transcriptional repression, as well as discuss a model by which Xist may compartmentalize proteins across the inactive X chromosome to enable both the initiation and maintenance of XCI.

1.2 | Overview

X chromosome inactivation (XCI) is a developmental process during which one of the two X chromosomes in female placental mammals is randomly silenced to ensure dosage balance in gene expression between males and females. XCI has emerged as a mechanistic paradigm for studying epigenetics, gene regulation, three-dimensional (3D) nuclear structure, and long noncoding RNA (lncRNA) biology.

The Xist lncRNA orchestrates random XCI by spreading across one of the two X chromosomes^{1,2} and recruiting dozens of chromatin-modifying proteins, DNA methylation enzymes, transcriptional repressors, and RNA-binding proteins (RBPs) to the future inactive X chromosome (Xi)³⁻⁷. During this process, Xist triggers large-scale remodeling of the Xi, including tethering the entire chromosome to the nuclear periphery⁸⁻¹⁰, chromosome-wide compaction^{11,12}, loss of topologically associating domains¹³, and establishment of a unique 3D silencing compartment^{11,13-16}. Ultimately, Xist induces stable and heritable chromosome-wide transcriptional silencing that is maintained in an Xist-independent manner^{17,18}.

The mechanisms underlying the functions mediated by Xist have been at the center of intensive research efforts over the last several decades. In this review, we discuss recent mechanistic insights into how Xist initiates transcriptional silencing, spreads across the X chromosome, and establishes epigenetic maintenance of the silent state. We describe an emerging model for Xist RNA in driving spatial compartmentalization of various proteins on the Xi and the potential roles of this process in the initiation and maintenance of transcriptional silencing.

1.3 | The Xist lncRNA orchestrates XCI

Xist was initially identified as the only gene that is expressed exclusively from the Xi^{19,20}. Based on this unusual expression pattern, it was immediately suspected to be an essential regulator controlling XCI, a hypothesis that has since been confirmed by several lines of evidence, including: (i) deletion of Xist before random XCI disrupts initiation of XCI *in vivo*^{21,22} and (ii) induction of Xist expression is sufficient to initiate XCI even in contexts where this process does not normally occur, such as in mouse embryonic stem cells, in male cells, or when it is expressed on an autosome^{18,23}.

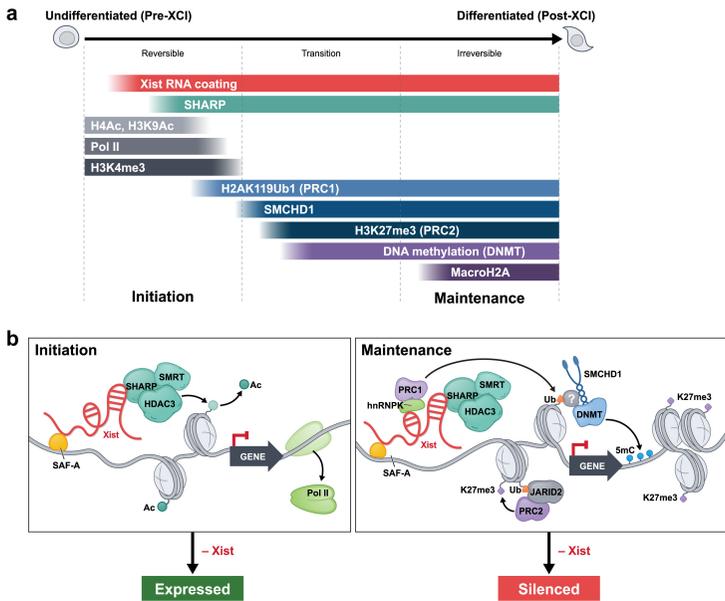


Figure 1. Protein recruitment and chromatin modifications during the initiation and maintenance of XCI. *a*, X chromosome inactivation can be separated into two phases – initiation and maintenance. These stages correspond to precise timepoints during differentiation and have specific molecular

characteristics, including chromatin modifications and structural changes on the Xi. **b**, Xist recruits a variety of regulatory complexes to the Xi. Initiation (left): Xist is tethered to the X chromosome through SAF-A, a DNA and RNA-binding protein. Xist interacts with SHARP, which recruits SMRT and HDAC3, to evict Pol II and silence transcription. Removal of Xist during initiation leads to reactivation of gene expression from the Xi. Maintenance (right): Several architectural and chromatin modifying proteins are recruited to the Xi. For example, Xist interacts with hnRNPK to recruit PRC1 and deposit H2AK119ub1, which in turn recruits PRC2 and its associated H3K27me3 histone modification. H2AK119ub1 is also required for recruiting SMCHD1 and DNMT enzymes through an unknown intermediate (indicated with a question mark). During this stage, Xist is dispensable for continued transcriptional silencing.

Xist expression is required within a critical developmental period for proper establishment of XCI. During this window, removal of the Xist RNA leads to reversal of the silencing phenotype¹⁸, yet following this initiation phase, the X chromosome becomes stably repressed and the loss of Xist expression does not lead to dramatic reactivation of X chromosome expression^{17,18,24}. Accordingly, XCI can be divided into two stages – initiation of silencing and maintenance of silencing – each of which is associated with distinct chromatin modifications (**Figure 1**).

1.4 | How Xist initiates transcriptional silencing on the X chromosome

To mediate XCI, Xist acts as a molecular scaffold to recruit proteins to the Xi. Recent studies have identified approximately a dozen RBPs that bind directly to Xist^{3,5} and interact with many more auxiliary proteins that are also involved in XCI^{5,7}. Notably, among these RBPs is SHARP (SMRT/HDAC-associated repressor protein; also known as

Spn)³⁻⁶ – a critical repressive protein that binds to Xist and interacts with the SMRT co-repressor²⁵, which is known to activate the histone deacetylase activity of the HDAC3 complex²⁶. HDAC3 requires interaction with the deacetylation activation domain of the NCoR/SMRT complex to initiate a conformational change that enables its catalytic activity²⁷. Genetic knockout²⁸, perturbations³, and drug treatment²⁹ have all shown that the histone deacetylase activity of HDAC3 is required for silencing on the Xi. In contrast, other HDACs, including other members of the class I HDAC family, appear to be dispensable for this function^{3,28}. Importantly, histone deacetylation has long been noted as one of the earliest chromatin modifications on the Xi^{30,31}. Nevertheless, it remains unclear whether the interaction between Xist and SHARP directly recruits HDAC3 to the Xi or if the Xist-SHARP interaction activates HDAC3 that is already prebound to the X chromosome^{3,28,32}.

SHARP interacts with a highly conserved region at the 5' end of Xist called the A-repeat, a tandem repeat containing 7-8 copies of a GC-rich sequence³³. This region forms a multivalent interaction with SHARP, such that a single copy of Xist binds to multiple SHARP proteins. Although the exact stoichiometry of the SHARP complex on a single molecule of Xist *in vivo* is unknown, recent biochemical studies have suggested that there might be four proteins bound per Xist molecule³⁴.

Several studies have shown that the interaction between Xist and the SHARP/SMRT/HDAC3 complex (Xist-SHARP) is required to evict RNA polymerase II (Pol II) from the Xi³⁻⁶. Nevertheless, this molecular interaction alone cannot explain how the Xist-SHARP complex leads to deterministic silencing of the entire X chromosome during mammalian development as its stoichiometry is not high enough to silence each gene individually. Specifically, there are ~200

copies of Xist within each cell³⁵ and the complex would need to localize across the >167 million base pairs of the X chromosome (~1 copy/megabase of DNA) to achieve complete silencing. One possible explanation is that the complex can sample multiple sites and activate HDAC3 that is already present across the chromosome. Another possibility is that multivalent interactions formed between Xist and SHARP increase the spatial concentration of the silencing complex and allow it to act preferentially on the X chromosome, even when not directly bound to an Xist molecule. This process may be facilitated by cooperative interactions between the large intrinsically disordered regions (IDRs) contained within SHARP (**Figure 2**). Although a recent study showed that the Spen paralog and ortholog C-terminal (SPOC) domain of SHARP is sufficient to repress transcription when tethered to Xist, the level of silencing achieved by this domain alone is lower than that observed for full-length SHARP³². Moreover, we note that this SPOC domain is also responsible for recruiting SMRT to the Xi, and SMRT itself contains many IDRs, which may similarly promote the formation of a high concentration compartment.

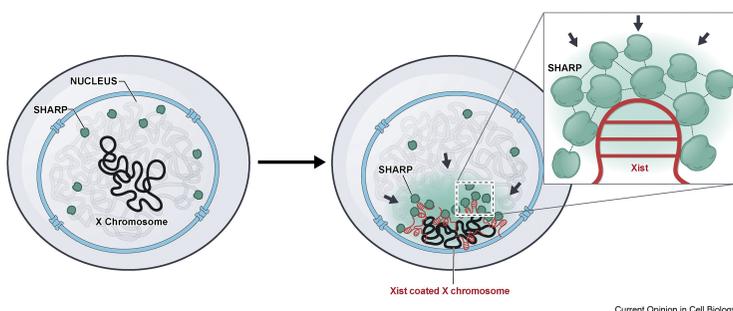


Figure 2. Multivalent recruitment of SHARP to the Xi. SHARP, and its associated SMRT/HDAC3 complex, are recruited to the Xi by direct binding of SHARP to the A-repeat of Xist. The repetitive nature of this region enables a single copy of Xist to bind to

multiple copies of the SHARP protein. In this way, induction of Xist expression leads to dynamic assembly of a high local concentration territory of SHARP over the Xi. As a result, SHARP may be able to silence the entire X chromosome despite the Xist RNA being expressed at sub-stoichiometric levels relative to its target sites on the Xi. Such preferential recruitment of SHARP to the Xi may be reinforced by self-association of the IDRs contained within the SHARP protein.

1.5 | How Xist and its silencing complex spread across the inactive X chromosome

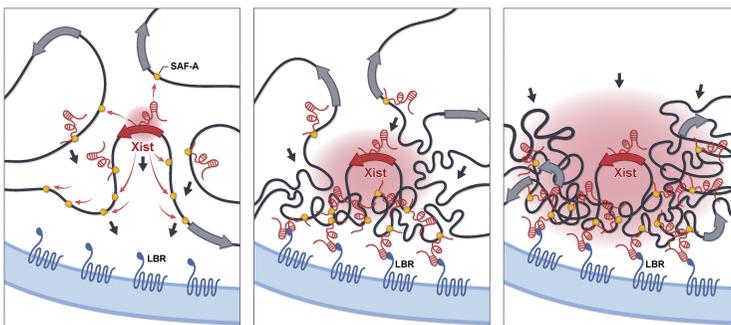
High-resolution studies mapping Xist localization over time in both inducible Xist systems and endogenous female differentiation have shown that Xist initially localizes to sites on the X chromosome that are in 3D spatial proximity to its transcription locus and are largely enriched for inactive genes^{1,2,36}. Because these inactive regions tend to be abundant in L1 elements, there is also a correlation between early Xist spreading and L1 sites. Indeed, these L1 elements appear to be among the earliest silenced features on the X chromosome¹¹. Nevertheless, L1 concentration does not appear to be required for initial spreading as Xist still accumulates at regions that do not possess L1 elements¹.

Several studies have identified an RNA and DNA-binding protein called SAF-A (scaffold attachment factor A; also known as hnRNPU) as a critical factor for tethering Xist to chromosomal DNA³⁷. While SAF-A appears to be localized broadly across chromatin throughout the nucleus (rather than being enriched specifically on the Xi)³⁷, it may act to specifically localize Xist on the X chromosome by sequestering the RNA on chromatin close to its transcriptional locus through 3D diffusion. Recent studies have also shown that binding of the nuclear matrix protein

CIZ1 (Cip1-interacting zinc finger protein 1) to the E-repeat of Xist is crucial for its proper localization to the Xi^{38,39}.

Since early Xist accumulation occurs at transcriptionally inactive regions, the initial Xi compartment that forms following Xist spreading is depleted of Pol II. Although actively transcribed genes initially appear to loop out of this region, over time they are relocated into the Xi compartment and undergo silencing^{1,11}. Importantly, repositioning of active genes into this silencing compartment is dependent on the A-repeat region of Xist^{1,11}.

The A-repeat of Xist has also been implicated in positioning the Xi at the nuclear periphery through its interaction with Lamin B Receptor (LBR)^{3,10} – a transmembrane protein that acts to anchor chromatin to the nuclear lamina⁴⁰. One of the earliest observations about the Xi was that it is tethered at the nuclear lamina^{8,9}. However, it was unclear what, if any, functional significance this association might play until it was recently shown that the deletion of LBR disrupts Xi association with the nuclear lamina, and by doing so precludes Xist spreading to active genes across the X chromosome¹⁰. Together, these observations suggest a model for how Xist localizes specifically on the future Xi and spreads across the chromosome over time (**Figure 3**).



Current Opinion in Cell Biology

Figure 3. Xist coating of the Xi. *Xist spreads across the Xi by initially diffusing to DNA sites that are in close 3D proximity to its transcription locus (red arrows), where the RNA is sequestered by binding to SAF-A on chromatin (left). Xist-bound DNA regions are repositioned to the nuclear lamina by binding to the Lamin B receptor (LBR, black arrows), leading to structural changes on DNA that enable other sites to come into close proximity to the Xist transcriptional locus (middle). This alters the accessibility of DNA such that Xist is able to spread across the entire chromosome in later stages of XCI (right).*

1.6 | How Xist establishes an epigenetically heritable state that maintains inactivation

Xist triggers a cascade of histone modifications and DNA methylation on the Xi that are deposited throughout the process of XCI. For example, Polycomb repressive complex 1 (PRC1) is recruited to the Xi through a direct interaction between Xist and hnRNPk, which interacts with the noncanonical Pcgf3/5 components of the PRC1 complex⁴¹. PRC1 deposits H2A ubiquitylation marks (H2AK119ub1) across the Xi, leading to subsequent recruitment of the PRC2 complex and its associated H3K27me3 modifications⁴².

The PRC1 complex has been implicated as a key regulator of XCI because Pcgf3/5 loss of function leads to a significant reduction in X-linked silencing⁴³. Pcgf3/5 may contribute to maintenance of silencing because Pcgf3/5 knockout embryos show lethality in females at mid-gestation stages – several days after initiation of XCI has already occurred⁴². Moreover, it has been shown that deletion of other components of the PRC1 complex do not impact proper initiation of silencing on the X chromosome^{44,45}.

PRC1 might play a role in maintenance of XCI, at least in part through its ability to recruit SMCHD1 to the X

chromosome. Specifically, PRC1-mediated deposition of H2AK119ub1 has been shown to be critical for recruiting SMCHD1 – a cohesin family protein recently shown to be essential for mediating structural changes on the Xi^{14,15}. Independently, mutations in SMCHD1 have been shown to cause defective X-linked silencing and female embryonic lethality at mid-gestation^{46,47}. Loss of SMCHD1 also reduces DNA methylation levels across the X chromosome⁴⁸. Nevertheless, it remains unclear the full extent of how PRC1 might contribute to X-linked silencing mechanisms and whether SMCHD1 mediates its functional role in XCI through structural changes, recruitment of DNA methylation, or other possible mechanisms.

DNA methylation (methylation of cytosines within CpG dinucleotides) has long been suspected to play a key role in the maintenance of XCI⁴⁹ as it is generally associated with transcriptional repression, especially when clustered within promoter regions, and is known to be recognized by several silencing proteins⁵⁰. DNA methylation is deposited by *de novo* methyltransferases (DNMT3A/B) and propagated after cell division through a maintenance methyltransferase (DNMT1)⁵⁰. In this way, DNA methylation is maintained across the cell cycle through a self-reinforcing epigenetic mechanism⁵⁰.

The proposed role for DNA methylation in maintenance of XCI is based on three key lines of evidence. First, DNA methylation is known to be enriched on the Xi^{51,52} and is specifically deposited on the Xi in an Xist-dependent manner by the DNMT3B *de novo* methyltransferase⁴⁸. Second, genetic and pharmacological disruptions of DNA methylation in somatic cells have shown that specific genes on the X chromosome can undergo variable levels of reactivation from the Xi^{29,53-56}. Finally, the mechanism of propagation of DNA methylation through cell division

provides a simple epigenetic memory mechanism that could explain how silencing is propagated in an Xist-independent manner⁵⁰.

Still, there are several lines of evidence that raise doubts about whether DNA methylation is the only component required for maintenance of XCI. Treatment of various human somatic cell lines with drugs that lead to demethylation (e.g. 5-azacytidine) does not cause reactivation of most genes on the Xi^{17,57}. This is true even when coupled with genetic deletion of Xist¹⁷. Furthermore, genetic deletion of DNMT3A/B does not impair the propagation of the silenced state in mouse embryos⁵⁸, and human patients with immunodeficiency, centromeric region instability or facial anomalies (ICF) syndrome (caused by a loss of function mutation in both copies of the DNMT3B gene) maintain normal XCI despite global hypomethylation across the Xi⁵⁹. Finally, the transition to the maintenance state of XCI appears to occur before establishment of DNA methylation⁴⁸. Together these results indicate that other factors are likely to be important for maintaining transcriptional silencing on the Xi.

1.7 | A possible role for protein compartmentalization and phase separation in the maintenance of X chromosome inactivation

Recent studies have shown that proteins involved in many critical transcriptional processes, including chromatin regulation and transcription, undergo liquid-liquid phase separation (LLPS)⁶⁰. These observations have led to a model whereby LLPS in the nucleus could explain the dynamic formation of membrane-less compartments that coordinately regulate different aspects of gene expression. A central tenant of this model is that cooperative interactions between

high local concentrations of nucleic acids and proteins – particularly those with intrinsically disordered regions (IDRs) – can lead to the formation of phase-separated bodies that act to compartmentalize molecular components and biochemical functions in cells⁶⁰. For example, many RBPs contain low-complexity domains that facilitate self-aggregation, and in several cases, RNA has been shown to facilitate this phase separation process by increasing local concentrations of such RBPs^{61,62} or by directly participating in heterotypic multivalent interactions through RNA sequence repeats⁶³.

Several observations suggest that such concentration-dependent phase transition may lead to compartmentalization of silencing components on the Xi. (i) Many of the RBPs that interact with Xist are known to contain IDRs (e.g. Ptbp1, SAF-A, SHARP)⁶⁴, and several of these have been shown to individually undergo RNA-mediated phase-separation *in vitro*⁶⁵⁻⁶⁷. (ii) Xist contains repetitive sequences that are known to form direct multivalent interactions with these RBPs^{34,68}, suggesting a molecular mechanism for how Xist can initiate compartmentalization and phase transitions *in vivo*. (iii) Molecules within the inactive compartment diffuse rapidly, with properties that are consistent with diffusion of liquids⁶⁹. (iv) The Xist-coated nuclear compartment appears to be stable even on removal of DNA, suggesting that it consists of RNA-protein interactions⁷⁰. (v) There is a critical developmental time point at which Xist-mediated silencing transitions from a reversible to an irreversible process¹⁸. Such sharp transitions are a physical characteristic of molecular phase-transition events⁷¹.

In such a model, induction of Xist expression leads to high concentration of the RNA over the Xi. The various RBPs that interact with Xist then preferentially localize within this

compartment to achieve a high local concentration. Upon achieving high local concentration, these RBPs can interact with each other to undergo LLPS and, in this way, localization of various silencing proteins could be stabilized even in the absence of the Xist RNA (**Figure 4**). This model could explain why Xist is essential for initiating transcriptional repression, but is dispensable for maintaining X-linked silencing. Consistent with this idea, a recent study has shown that deletion of Ptbp1 (or its multivalent binding site on Xist) – a protein that is known to undergo phase separation *in vitro*^{62,72} – does not disrupt initiation of XCI, but impacts compartment formation and maintenance of silencing on the Xi⁷³.

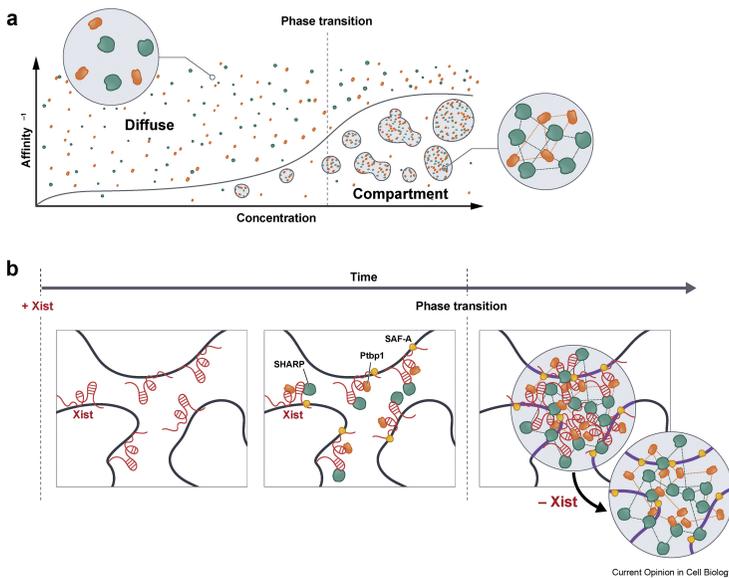


Figure 4. A liquid-liquid phase separation (LLPS) model of XCI. **a**, Molecules can transition from diffuse to compartmentalized based on the local concentration and affinity of the molecular components in 3D space. In this model, small changes in concentration can cause phase transitions that promote self-reinforcement of molecular components. **b**, Such an LLPS model

may explain many of the observed properties during the maintenance of XCI. In this model, Xist recruits a number of intrinsically disordered proteins to the Xi, including SHARP and Ptbp1, that interact and lead to a high spatial concentration of silencing complexes. After a phase transition event, these proteins can form self-reinforcing interactions with each other, thereby enabling protein enrichment within the spatial compartment on the Xi independently of the Xist RNA. In this way, transcriptional silencing may be maintained in the absence of Xist.

For this compartment to play a role in maintenance of silencing, it must be propagated across cell divisions. Although it remains unknown how this might occur, we note that various histone modifications and DNA methylation are stably associated with the Xi through cell division^{50,74,75} and that these modifications may act to seed re-establishment of the silencing compartment following mitosis. In addition, studies of transcription factors have shown that certain proteins can be retained on mitotic chromosomes to mark specific nuclear locations (referred to as ‘mitotic bookmarking’)^{76–78}. A mitotic bookmarking mechanism might similarly act to retain one or more of the proteins involved in XCI – or the phase-separated compartment itself – on the Xi across cell divisions.

1.8 | References

1. Engreitz, J. M. *et al.* The Xist lncRNA exploits three-dimensional genome architecture to spread across the X Chromosome. *Science* **341**, 1237973 (2013).
2. Simon, M. D. *et al.* High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504**, 465–469 (2013).
3. McHugh, C. A. *et al.* The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232–236 (2015).
4. Moindrot, B. *et al.* A pooled shRNA screen identifies Rbm15, Spen, and Wtap as Factors required for Xist RNA-mediated silencing. *Cell Rep.* **12**, 562–572 (2015).
5. Chu, C. *et al.* Systematic discovery of Xist RNA binding proteins. *Cell* **161**, 404–416 (2015).
6. Monfort, A. & Wutz, A. Identification of Spen as a crucial factor for Xist function through forward genetic screening in haploid embryonic stem cells. *Cell Rep.* **12**, 554–561 (2015).
7. Minajigi, A. *et al.* A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**, aab2276–aab2276 (2015).
8. Barr, M. L. & Bertram, E. G. A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature* **163**, 676–677 (1949).
9. Dyer, K. A., Canfield, T. K. & Gartler, S. M. Molecular cytological differentiation of active from inactive X domains in interphase: implications for X chromosome inactivation. *Cytogenet. Genome Res.*

- 50**, 116–120 (1989).
10. Chen, C. K. *et al.* Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* **354**, 468–472 (2016).
 11. Chaumeil, J., Le Baccon, P., Wutz, A. & Heard, E. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev.* **20**, 2223–2237 (2006).
 12. Nozawa, R.-S. *et al.* Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBiX1 pathway. *Nat. Struct. Mol. Biol.* **20**, 566–573 (2013).
 13. Giorgetti, L. *et al.* Structural organization of the inactive X chromosome in the mouse. *Nature* **535**, 575–579 (2016).
 14. Gdula, M. R. *et al.* The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome. *Nat. Commun.* **10**, 30 (2019).
 15. Wang, C.-Y., Jégu, T., Chu, H.-P., Oh, H. J. & Lee, J. T. SMCHD1 merges chromosome compartments and assists formation of super-structures on the inactive X. *Cell* **174**, 406-421.e25 (2018).
 16. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
 17. Brown, C. J. & Willard, H. F. The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. *Nature* **368**, 154–156 (1994).
 18. Wutz, A. & Jaenisch, R. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell* **5**, 695–705 (2000).
 19. Brockdorff, N. *et al.* Conservation of position and exclusive expression of mouse Xist from the inactive

- X chromosome. *Nature* **351**, 329–331 (1991).
20. Brown, C. J. *et al.* A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**, 38–44 (1991).
 21. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131–137 (1996).
 22. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. & Jaenisch, R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* **11**, 156–166 (1997).
 23. Wutz, A., Rasmussen, T. P. & Jaenisch, R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* **30**, 167–174 (2002).
 24. Csankovszki, G., Panning, B., Bates, B., Pehrson, J. R. & Jaenisch, R. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. *Nature Genetics* **22**, 323–324 (1999).
 25. Shi, Y. *et al.* Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev.* **15**, 1140–1151 (2001).
 26. You, S. H. *et al.* Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 in vivo. *Nat. Struct. Mol. Biol.* **20**, 182–187 (2013).
 27. Watson, P. J., Fairall, L., Santos, G. M. & Schwabe, J. W. R. Structure of HDAC3 bound to co-repressor and inositol tetraphosphate. *Nature* **481**, 335–340 (2012).
 28. Żylicz, J. J. *et al.* The implication of early chromatin changes in X chromosome inactivation. *Cell* **176**, 182–197 (2019).

29. Csankovszki, G., Nagy, A. & Jaenisch, R. Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. *J. Cell Biol.* **153**, 773–784 (2001).
30. Jeppesen, P. & Turner, B. M. The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* **74**, 281–289 (1993).
31. Boggs, B. A., Connors, B., Sobel, R. E., Chinault, A. C. & Allis, C. D. Reduced levels of histone H3 acetylation on the inactive X chromosome in human females. *Chromosoma* **105**, 303–309 (1996).
32. Dossin, F. *et al.* SPEN integrates transcriptional and epigenetic control of X-inactivation. *Nature* **578**, 455–460 (2020).
33. Brockdorff, N. *et al.* The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71**, 515–526 (1992).
34. Lu, Z. *et al.* RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* **165**, 1267–1279 (2016).
35. Sunwoo, H., Wu, J. Y. & Lee, J. T. The Xist RNA-PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proc. Natl. Acad. Sci. USA* **112**, E4216–E4225 (2015).
36. Clemson, C. M. *et al.* The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences. *Proc. Natl. Acad. Sci. USA* **16**, 7688–7693 (2006).
37. Hasegawa, Y. *et al.* The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev. Cell* **19**, 469–476 (2010).
38. Ridings-Figueroa, R. *et al.* The nuclear matrix

- protein CIZ1 facilitates localization of Xist RNA to the inactive X-chromosome territory. *Genes Dev.* **31**, 876–888 (2017).
39. Sunwoo, H., Colognori, D., Froberg, J. E., Jeon, Y. & Lee, J. T. Repeat E anchors Xist RNA to the inactive X chromosomal compartment through CDKN1A-interacting protein (CIZ1). *Proc. Natl. Acad. Sci. USA* **114**, 10654–10659 (2017).
 40. Olins, A. L., Rhodes, G., Welch, D. B. M., Zwerger, M. & Olins, D. E. Lamin B receptor multi-tasking at the nuclear envelope. *Nucleus* **1**, 53–70 (2010).
 41. Pintacuda, G., Young, A. N. & Cerase, A. Function by structure: spotlights on Xist long non-coding RNA. *Front. Mol. Biosci.* **4**, 90 (2017).
 42. Almeida, M. *et al.* PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. *Science* **356**, 1081–1084 (2017).
 43. Nesterova, T. B. *et al.* Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. *Nat. Commun.* **10**, 3129 (2019).
 44. Schoeftner, S. *et al.* Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J.* **25**, 3110–3122 (2006).
 45. Bousard, A. *et al.* The role of Xist-mediated Polycomb recruitment in the initiation of X-chromosome inactivation. *EMBO Rep.* **20**, e48019 (2019).
 46. Blewitt, M. E. *et al.* An N-ethyl-N-nitrosourea screen for genes involved in variegation in the mouse. *Proc. Natl. Acad. Sci. USA* **102**, 7629–7634 (2005).
 47. Blewitt, M. E. *et al.* SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat.*

- Genet.* **40**, 663–669 (2008).
48. Gendrel, A. V. *et al.* Smchd1-dependent and independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Dev. Cell* **23**, 265–279 (2012).
 49. Riggs, A. D. X inactivation, DNA methylation, and differentiation revisited. *DNA Methylation* 269–278 (1984).
 50. Hermann, A., Gowher, H. & Jeltsch, A. Biochemistry and biology of mammalian DNA methyltransferases. *Cellular and Molecular Life Sciences* **61**, 2571–2587 (2004).
 51. Lock, L. F., Takagi, N. & Martin, G. R. Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell* **48**, 39–46 (1987).
 52. Norris, D. P., Brockdorff, N. & Rastan, S. Methylation status of CpG-rich islands on active and inactive mouse X chromosomes. *Mamm. Genome* **1**, 78–83 (1991).
 53. Sado, T. *et al.* X inactivation in the mouse embryo deficient for Dnmt1: Distinct effect of hypomethylation on imprinted and random X inactivation. *Dev. Biol.* **225**, 294–303 (2000).
 54. Mohandas, T., Sparkes, R. S. & Shapiro, L. J. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* **211**, 393–396 (1981).
 55. Marshall Graves, J. A. 5-azacytidine-induced re-expression of alleles on the inactive X chromosome in a hybrid mouse cell line. *Exp. Cell Res.* **141**, 99–105 (1982).
 56. Lester, S. C., Korn, N. J. & DeMars, R. Derepression of genes on the human inactive X chromosome: evidence for differences in locus-specific rates of derepression and rates of transfer of active and inactive genes after DNA-mediated transformation.

- Somatic Cell Genet.* **8**, 265–84 (1982).
57. Wolf, S. F. & Migeon, B. R. Studies of X chromosome DNA methylation in normal human cells. *Nature* **295**, 667–671 (1982).
 58. Sado, T., Okano, M., Li, E. & Sasaki, H. De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. *Development* **131**, 975–982 (2004).
 59. Hansen, R. S. *et al.* Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant. *Hum. Mol. Genet.* **9**, 2575–2587 (2000).
 60. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
 61. Lin, Y., Protter, D. S. W., Rosen, M. K. & Parker, R. Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell* **60**, 208–219 (2015).
 62. Banani, S. F. *et al.* Compositional control of phase-separated cellular bodies. *Cell* **166**, 651–663 (2016).
 63. Jain, A. & Vale, R. D. RNA phase transitions in repeat expansion disorders. *Nature* **546**, 243–247 (2017).
 64. Cerase, A. *et al.* Phase separation drives X-chromosome inactivation: a hypothesis. *Nat. Struct. Mol. Biol.* **26**, 331–334 (2019).
 65. Fackelmayer, F. O., Dahm, K., Renz, A., Ramsperger, U. & Richter, A. Nucleic-acid-binding properties of hnRNP-U/SAF-A, a nuclear-matrix protein which binds DNA and RNA in vivo and in vitro. *Eur. J. Biochem.* **221**, 749–757 (1994).
 66. Li, P. *et al.* Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).
 67. Ying, Y. *et al.* Splicing activation by Rbfox requires

- self-aggregation through its tyrosine-rich domain. *Cell* **170**, 312–323 (2017).
68. Cirillo, D. *et al.* Quantitative predictions of protein interactions with long noncoding RNAs. *Nature Methods* **14**, 5–6 (2016).
 69. Ng, K. *et al.* A system for imaging the regulatory noncoding Xist RNA in living mouse embryonic stem cells. *Mol. Biol. Cell* **22**, 2634–2645 (2011).
 70. Clemson, C. M., McNeil, J. A., Willard, H. F. & Lawrence, J. B. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* **132**, 259–75 (1996).
 71. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
 72. Lin, Y., Currie, S. L. & Rosen, M. K. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J. Biol. Chem.* **292**, 19110–19120 (2017).
 73. Pandya-Jones, A. *et al.* A protein assembly mediates Xist localization and gene silencing. *Nature* **587**, 145–151 (2020).
 74. Casas-Delucchi, C. S. *et al.* Histone acetylation controls the inactive X chromosome replication dynamics. *Nat. Commun.* **2**, 222 (2011).
 75. Boggs, B. A. *et al.* Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes. *Nat. Genet.* **30**, 73–76 (2002).
 76. Teves, S. S. *et al.* A dynamic mode of mitotic bookmarking by transcription factors. *Elife* **5**, e22280 (2016).
 77. Deluz, C. *et al.* A role for mitotic bookmarking of SOX2 in pluripotency and differentiation. *Genes Dev.* **30**, 2538–2550 (2016).

78. Karagianni, P., Moulos, P., Schmidt, D., Odom, D. T. & Talianidis, I. Bookmarking by non-pioneer transcription factors during liver development establishes competence for future gene activation. *Cell Rep.* **30**, 1319-1328.e6 (2020).

Chapter 2

**Xist spatially amplifies SHARP/SPEN
recruitment to balance chromosome-wide
silencing and specificity to the X chromosome**

Joanna W. Jachowicz*, Mackenzie Strehle*,
Abhik K. Banerjee, Mario R. Blanco, Jasmine Thai,
and Mitchell Guttman

*Equal Contribution

A modified version of this chapter was published as:
Jachowicz, J.W.*, Strehle, M.* *et al.* Xist spatially
amplifies SHARP/SPEN recruitment to balance
chromosome-wide silencing and specificity to the X
chromosome. *Nat. Struct. Mol. Biol.* **29**, 239–249 (2022).
<https://doi.org/10.1038/s41594-022-00739-1>

2.1 | Abstract

Although thousands of long non-coding RNAs (lncRNAs) are encoded in mammalian genomes, their mechanisms of action are poorly understood, in part because they are often expressed at lower levels than their proposed targets. One such lncRNA is Xist, which mediates chromosome-wide gene silencing on one of the two X chromosomes (X) to achieve gene expression balance between males and females. How a limited number of Xist molecules can mediate robust silencing of a much larger number of target genes while maintaining specificity exclusively to genes on the X within each cell is not well understood. Here, we show that Xist drives non-stoichiometric recruitment of the essential silencing protein SHARP (also known as SPEN) to amplify its abundance across the inactive X, including at regions not directly occupied by Xist. This amplification is achieved through concentration-dependent homotypic assemblies of SHARP on the X and is required for chromosome-wide silencing. Expression of Xist at higher levels leads to increased localization at autosomal regions, demonstrating that low levels of Xist are critical for ensuring its specificity to the X. We show that Xist (through SHARP) acts to suppress production of its own RNA which may act to constrain overall RNA levels and restrict its ability to spread beyond the X. Together, our results demonstrate a spatial amplification mechanism that allows Xist to achieve two essential but countervailing regulatory objectives: chromosome-wide gene silencing and specificity to the X. This suggests a more general mechanism by which other low-abundance lncRNAs could balance specificity to, and robust control of, their regulatory targets.

2.2 | Introduction

In recent years, thousands of lncRNAs have been identified and many have been proposed to regulate gene expression¹⁻⁵. However, their precise mechanisms of action remain largely uncharacterized. One of the key issues is that lncRNAs are generally expressed at low levels such that the number of RNA molecules is less than the number of targets that they are proposed to regulate (substoichiometric)⁶⁻⁸. How an individual lncRNA molecule can control multiple distinct targets when it cannot engage with all of them simultaneously remains unknown and has led some to suggest that these lowly expressed lncRNAs may not be functionally important^{9,10}.

One example of a lncRNA that is expressed at substoichiometric levels relative to its targets is Xist. Expression of Xist is sufficient to induce transcriptional silencing of >1,000 genes across the >167 million bases of DNA on the X to achieve dosage balance of expression between males and females¹¹⁻¹⁷. Previous studies have shown that there are ~60-200 Xist molecules within an individual cell¹⁸⁻²⁰, corresponding to an average of approximately one Xist RNA for every ten genes encoded on the X.

Xist represents an ideal system in which to explore how substoichiometric levels of a lncRNA can regulate its more abundant targets because it is functionally important (developmentally essential)^{11,21} with a clear phenotype (transcriptional silencing)²²⁻²⁴ that occurs at precise and well-defined regulatory targets (X chromosome genes)¹⁵⁻¹⁷. Recent studies have begun to elucidate the mechanisms by which Xist localizes across the X and recruits silencing proteins to initiate chromosome-wide silencing. Rather than binding to specific DNA sequences, Xist diffuses from its

transcription locus to DNA sites that are in close three-dimensional (3D) proximity at both genic and intergenic regions^{16,17}. Xist binds directly to SHARP (also called SPEN)^{22,25–28}, an RNA-binding protein that associates with the SMRT and HDAC3 repressive complex to deacetylate histones^{29–31}, evict RNA polymerase II^{22,24,32}, and silence transcription on the X^{22,24,25,32–34}.

Although these discoveries have uncovered several long-sought molecular mechanisms underlying Xist-mediated silencing, they raise critical new questions about how Xist can achieve the essential quantitative features required for dosage balance. Specifically, Xist-mediated silencing needs to be both specific to ensure that only genes on the X (but not autosomes) are silenced, and robust to ensure that each of the several hundred distinct genes across the X are silenced within each individual cell.

Current models, based on ensemble measurements, cannot explain how Xist achieves these two regulatory objectives – specificity to the X and chromosome-wide silencing – within single cells. For example, Xist localization based on 3D proximity could explain its preferential localization on the X; however, because the X is not partitioned from other chromosomes, this mechanism would not preclude Xist spreading to some autosomal regions within individual cells. Because Xist can silence transcription of genes on autosomes when present in proximity^{35–37}, its specificity to the X is essential to preclude gene silencing of autosomal genes. Moreover, although Xist localizes broadly across the X when measured in a population of cells^{16,17}, it cannot localize at all of these positions simultaneously because there is only approximately one Xist RNA molecule for each megabase of genomic DNA within an individual cell (**Supplementary Note**). Accordingly, Xist must localize heterogeneously at distinct locations in single cells. Such

heterogeneous localization would be expected to lead to heterogeneous silencing where different genes are silenced in distinct cells. Yet, X silencing is not heterogeneous (**Supplementary Note**)^{24,38}. Therefore, the stoichiometric silencing model, whereby the Xist-SHARP complex localizes at each gene to silence transcription cannot explain how chromosome-wide silencing occurs at the level of single cells.

Here, we explore the mechanisms of how the Xist lncRNA can achieve chromosome-wide gene silencing while ensuring specificity to the X within each individual cell during initiation of X chromosome inactivation (XCI).

2.3 | Results

2.3.1 | SHARP accumulates on the Xi in stoichiometric excess to Xist

To explore how substoichiometric levels of Xist can silence genes across the X, we analyzed the temporal and quantitative relationships between localization of Xist and the essential silencing protein SHARP on the inactive X (Xi). SHARP binds directly to Xist and its enrichment on the Xi is dependent on Xist^{20,22,24-27}. We reasoned that if SHARP is recruited to the Xi solely through its ability to directly bind to Xist, then the concentration of SHARP would increase proportionally with the concentration of Xist across time (stoichiometric recruitment). In this case, the rate of Xist and SHARP accumulation on the X would be proportional to each other and their ratio would be constant across time (**Fig. 1a**).

To measure this, we used a female F1 hybrid (B16 x Cast) mouse embryonic stem cell (mESC) line containing a doxycycline (dox)-inducible Xist gene at its endogenous

locus and an in-frame HaloTag inserted into both copies of the endogenous SHARP protein (TX-SHARP-HALO cells)^{24,39}. This system allows for more temporally synchronized expression of Xist compared with differentiation of female mESCs (**Extended Data Fig. 1a**), achieves robust chromosome-wide silencing by 72 h of induction (**Extended Data Fig. 1b**) and utilizes the same molecular components required for initiation of XCI during development and differentiation^{23,40}. We induced Xist expression and visualized SHARP (using either a dye that conjugates directly to the HaloTag or an antibody against the HaloTag) along with Xist (using RNA fluorescence in situ hybridization (RNA FISH)) across five time-points (1-48 h) following dox induction (**Fig. 1b and Extended Data Fig. 1c,d**). Using both SHARP visualization approaches, we found that the area of the Xist-coated territory increased over time (**Fig. 1c**), whereas the total intensity of Xist over the territory increased initially, plateaued, and then remained relatively constant (**Extended Data Fig. 1e**). This means that the average Xist intensity within the territory decreased over time (**Fig. 1c**). By contrast, the average intensity of SHARP within the territory continued to increase across all time-points (**Fig. 1c and Extended Data Fig. 1f**; see **Methods** for quantification). Thus, the ratio of SHARP to Xist intensity is not constant across time, but instead increases in a nonlinear manner (**Fig. 1d**; see **Fig. 1a** for comparison). We note that similar kinetics were also recently observed using super-resolution microscopy¹⁹.

To ensure that this effect is not simply a product of our synthetic dox-inducible system (**Extended Data Fig. 1a**), we measured the localization of Xist and SHARP across time in female mESCs upon endogenous initiation of XCI using retinoic acid (RA)-induced differentiation. We imaged Xist and SHARP after 2 and 3 days of differentiation (**Extended Data Fig. 1g**) and observed a similar relationship

between the levels of Xist and the levels of SHARP across time: SHARP levels increased at a faster rate than Xist levels between the two time-points (**Fig. 1e** and **Extended Data Fig. 1h**). These results demonstrate that SHARP recruitment to the X occurs in a non-stoichiometric manner relative to Xist.

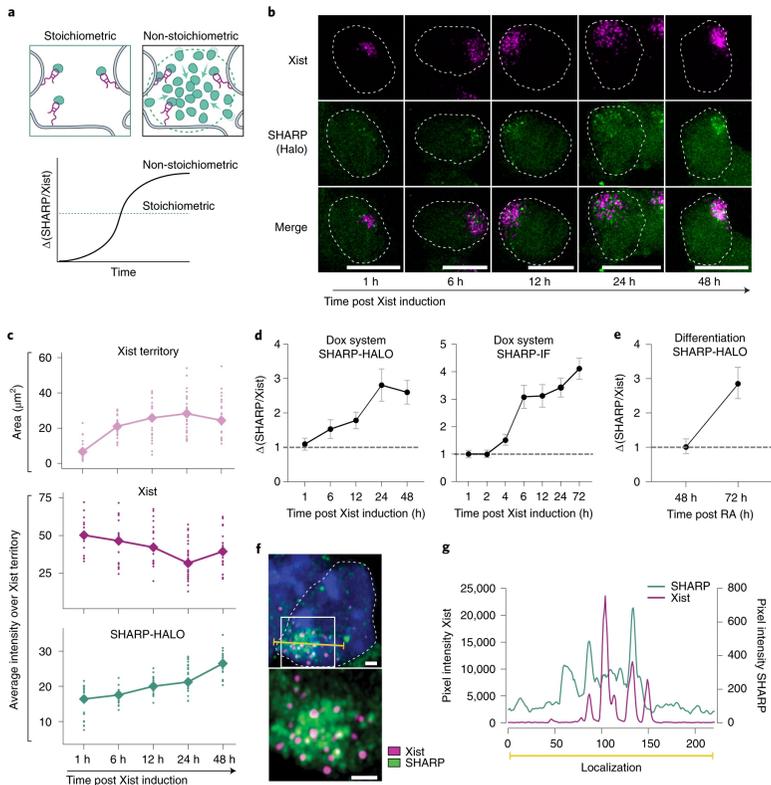


Figure 1: SHARP enrichment over the Xi increases in a non-stoichiometric manner relative to Xist. *a*, Schematic of two alternative models of SHARP recruitment to the Xi. Left: stoichiometric, SHARP (green) localizes through direct interaction with Xist (magenta). Right: non-stoichiometric, SHARP localizes even when not directly associated with Xist. In the stoichiometric model, the ratio of SHARP to Xist is directly proportional and constant across time; in the non-stoichiometric

model, the concentration of SHARP increases relative to Xist across time. **b**, Representative images of Xist and SHARP localization in TX-SHARP-HALO mESCs across 48 h of Xist expression ($n > 15$ cells per condition from two independent experiments). Xist is visualized by RNA FISH (magenta) and SHARP is visualized by direct labeling of endogenous SHARP-HALO (green). Images are shown as maximum projections. Scale bars, $10 \mu\text{m}$. **c**, Quantification of images from **b** showing area of Xist RNA territory (μm^2) (upper), average FI of Xist (RNA FISH) over a unit of Xist territory (middle), and average FI of SHARP (direct HaloTag labeling) over corresponding Xist territory (lower). Dots represent individual cell measurements; diamonds represent the average. **d**, Ratio of SHARP to Xist intensities following Xist induction normalized to 1 h. Left: SHARP visualized by direct HaloTag labeling across 48 h ($n > 15$ cells for each time-point). Right: SHARP visualized by anti-Halo IF across 72 h ($n > 15$ cells for each time-point). Data are represented as mean \pm s.d. **e**, Ratio of SHARP (direct HaloTag labeling) to Xist (RNA FISH) average intensities in RA-differentiated TX-SHARP-HALO mESCs normalized to 48 h ($n > 15$ cells for each time-point). Data are represented as mean \pm s.d. **f**, Super-resolution imaging of Xist (RNA FISH; magenta) and endogenous SHARP-HALO (direct HaloTag labeling; green) in TX-SHARP-HALO mESCs after 24 h of Xist induction ($n > 20$ cells observed with similar phenotype): (upper) single nucleus and (lower) close-up on Xist territory from the upper image demarcated by the white box. Images are shown as maximum projections. Scale bars, $1 \mu\text{m}$. The yellow line shows where the intensity profile (**g**) was measured. **g**, Line intensity profile from image in **f** showing Xist and SHARP intensities.

Based on these results, we explored whether SHARP is enriched at regions within the Xist-coated territory that are not bound by Xist. To do this, we focused on the Xist territory after 24 h of dox induction and performed super-resolution imaging of Xist and SHARP (**Fig. 1f**). We observed distinct Xist foci within the territory, whereas SHARP exhibits enrichment across the entire territory. As

such, there are clear regions of high concentration of SHARP even where Xist is not present (**Fig. 1g**).

2.3.2 | SHARP forms concentration-dependent assemblies in the nucleus

We next explored how non-stoichiometric recruitment of SHARP to the X might occur. SHARP is an ~400-kDa protein containing four RNA recognition motif (RRM) domains^{41,42} that bind to Xist^{26,32} and a SPEN paralog and ortholog C-terminal (SPOC) domain that is critical for recruiting the SMRT and HDAC3 proteins^{24,30,31}. The remainder of SHARP is predicted to consist of long intrinsically disordered regions (IDRs; **Fig. 2a**). Recently, many proteins containing long IDRs have been shown to form concentration-dependent assemblies through multivalent, high-avidity associations⁴³⁻⁴⁶. On the basis of this observation, we hypothesized that SHARP might similarly form such concentration-dependent assemblies (**Fig. 2b**). (Although some concentration-dependent assemblies have been shown to form through phase separation, this is not the only mechanism by which they form^{47,48}. In this specific context we are testing whether SHARP forms concentration-dependent assemblies, rather than the precise biophysical characteristics of their formation.)

To test this hypothesis, we explored whether SHARP exhibits known features of multivalent, high-avidity assemblies⁴³⁻⁴⁶. Specifically, we asked: (1) does SHARP form high-concentration foci in the nucleus, (2) is formation of these foci dependent on the overall concentration of SHARP, (3) are these foci dependent on multivalent associations mediated through the IDRs, and (4) are these

foci dependent on associations with other molecules of SHARP (homotypic assemblies)?

We expressed full-length SHARP tagged with monomeric enhanced green fluorescent protein (eGFP) (FL-SHARP; **Extended Data Fig. 2a**) in HEK293T cells, a cell type that allows for efficient transfection and controlled expression of the large plasmid containing SHARP, and enables characterization of its biochemical and biophysical properties independently of its functional targets. Using this system, we performed live-cell imaging and observed that FL-SHARP molecules formed discrete foci within the nucleus (**Fig. 2c and Supplementary Video 1**). These SHARP assemblies also displayed other features of multivalent, IDR-mediated assemblies in that individual molecules exchanged rapidly within a SHARP focus (**Extended Data Fig. 2b**) and SHARP foci merged into larger structures (fusion) or split into smaller structures (fission) across time^{49,50} (**Extended Data Fig. 2c,d and Supplementary Video 2**).

Next, we used the dox-inducible promoter that drives FL-SHARP expression to titrate its level across a >1,000-fold concentration range and determine whether formation of these foci depends on total protein concentration per cell. We observed that SHARP formed foci only when present at higher concentrations; at low concentrations SHARP was diffuse throughout the nucleus (**Fig. 2d,e and Extended Data Fig. 2e**; see **Methods** for quantification), similar to other proteins that do not form assemblies (**Extended Data Fig. 2f**).

To determine whether formation of SHARP assemblies is dependent on multivalent interactions driven by its IDRs, we expressed eGFP-tagged SHARP lacking its IDRs (Δ IDR-SHARP; **Extended Data Fig. 2a**) in HEK293T cells and

imaged its behavior. In contrast to the full-length protein, Δ IDR-SHARP did not form foci (Fig. 2f and Supplementary Video 3). Instead, Δ IDR-SHARP localized diffusively throughout the nucleus, even when present at concentrations where FL-SHARP formed foci (Fig. 2g).

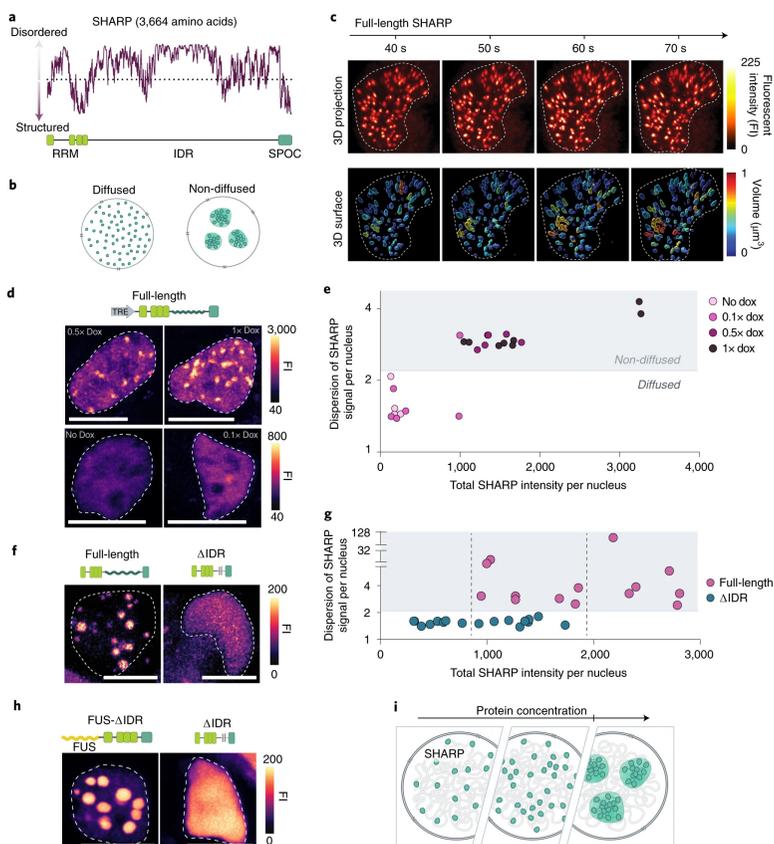


Figure 2: SHARP forms multivalent, concentration-dependent assemblies in the nucleus. *a*, Disordered scores across the SHARP protein using IUPred2 software predictions. The dotted line represents the 0.5 probability value for a given structure to be ordered. The lower visualization demarcates the position of known SHARP domains – RRM (bright green) and SPOC (dark green). *b*, Schematic representation of molecules within a nucleus

organized in a diffused or non-diffused (focal) manner. **c**, Images across four time-points from a live-cell video of eGFP-tagged FL-SHARP in transiently transfected HEK293T cells showing non-diffused, focal organization of SHARP molecules: (upper) 3D reconstructions of the FI signal; and (lower) 3D volume reconstructions color-coded based on the size of the condensate. **d**, Images representing localization patterns of eGFP-tagged FL-SHARP in transiently transfected HEK293T cells across increasing expression levels of SHARP (driven by dox-inducible promoter; 1x dox = 2 $\mu\text{g}/\text{mL}$). Images are shown as maximum projections. Scale bars, 10 μm . **e**, Quantification of images from **d** plotting the dispersion of SHARP signal across the nucleus versus average SHARP FI per nucleus. **f**, Representative images of FL-SHARP and $\Delta\text{IDR-SHARP}$ localization patterns in transiently transfected HEK293T cells. Images are shown as maximum projections. Scale bars, 10 μm . **g**, Quantification of images (Fig. 1f) plotting the dispersion of SHARP signal across the nucleus versus average SHARP FI per nucleus. The dashed line represents the range of FI that is similar for both groups. **h**, Images representing localization patterns of mCherry-tagged FUS- $\Delta\text{IDR-SHARP}$ and eGFP-tagged $\Delta\text{IDR-SHARP}$ in transiently transfected HEK293T cells. Images are shown as maximum projections. Scale bars, 10 μm . **i**, Schematic depicting formation of concentration-dependent SHARP assemblies.

Finally, we explored whether these IDR-dependent assemblies form through multivalent associations with other molecules of SHARP (homotypic assemblies) or require sequence-specific associations with other proteins (heterotypic assemblies). To do this, we fused $\Delta\text{IDR-SHARP}$ to an mCherry-tagged version of the IDR of the fused in sarcoma (FUS) protein, a domain that is known to form multivalent homotypic associations both *in vitro* and *in vivo*^{51–53} (FUS- $\Delta\text{IDR-SHARP}$; **Extended Data Fig. 2a**), and tested whether this synthetic protein rescues the ability of SHARP to form foci independently of its IDRs. We observed that FUS- $\Delta\text{IDR-SHARP}$ forms assemblies in the nucleus

that are comparable with those observed for FL-SHARP (**Fig. 2h**). Although we do not exclude the possibility that the IDRs of SHARP may form heterotypic associations with other molecules, these results indicate that homotypic associations are essential for SHARP to form the observed assemblies.

Together, these results indicate that SHARP forms concentration-dependent assemblies in the nucleus and that formation of these assemblies is dependent on homotypic multivalent interactions driven by its IDRs (**Fig. 2i**).

2.3.3 | SHARP recruitment to the Xi depends on IDR-mediated assembly

To determine whether IDR-dependent assemblies of SHARP are critical for its enrichment on the Xi, we tested whether deletion of the IDRs impacts localization over the X. To do this, we first generated a mESC line containing a deletion of both copies of the endogenous SHARP gene (SHARP-KO; **Extended Data Fig. 3a**). In parallel, we utilized mESCs containing an auxin-degradable SHARP (SHARP-AID)²⁴. Within each of these lines (SHARP-KO and SHARP-AID), we stably expressed a Halo-tagged or eGFP-tagged version of either full-length SHARP (FL-SHARP), SHARP containing a deletion of its RRM domains (Δ RRM-SHARP), or SHARP containing a deletion of its IDRs (Δ IDR-SHARP) (**Extended Data Fig. 3b,c**). We visualized each of these tagged SHARP proteins along with Ezh2 (to demarcate the Xi) after Xist expression for >72 h (**Fig. 3a and Extended Data Fig. 3d**). As expected, FL-SHARP was enriched over the Xi compartment. By contrast, Δ RRM-SHARP failed to localize on the Xi, consistent with previous observations²⁴. Interestingly, we also observed a strong decrease in the enrichment of Δ IDR-SHARP over the

Xi, comparable with that observed upon deletion of the RRM domains (**Fig. 3b**; see **Extended Data Fig. 3e** for quantification schematics). The level of Ezh2 was similar in all conditions (**Fig. 3b**).

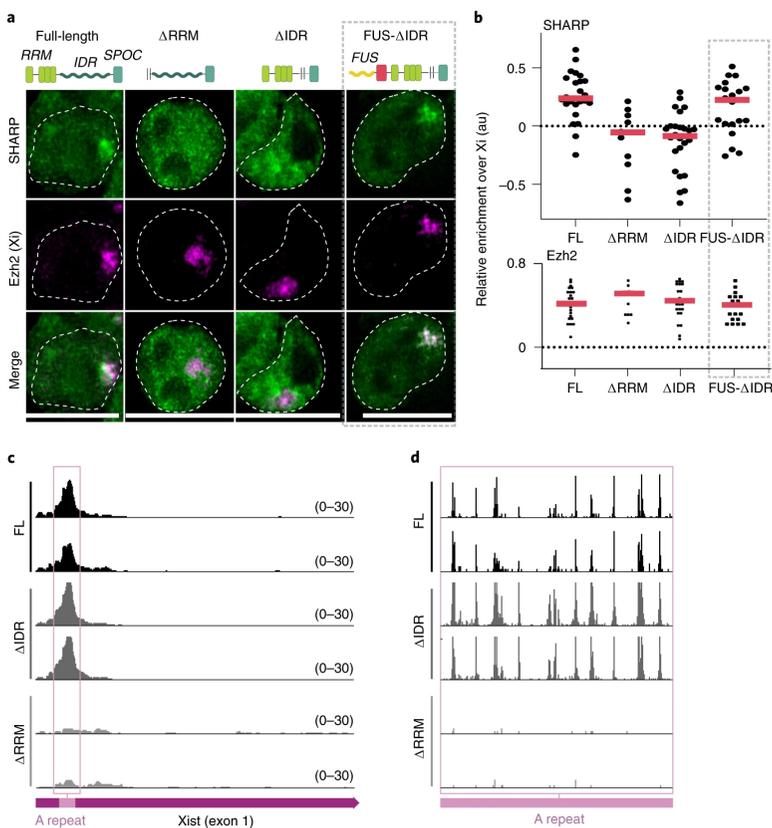


Figure 3: Formation of SHARP assemblies is required for enrichment on the Xi but not Xist binding. *a*, Representative images of SHARP enrichment (eGFP, green) over the Xi (anti-Ezh2 IF, magenta) in TX SHARP-KO mESCs containing dox-inducible Xist, genetic deletion of SHARP and stable integrations of eGFP-FL-SHARP, eGFP- Δ RRM-SHARP, eGFP- Δ IDR-SHARP, or FUS-mCherry- Δ IDR-SHARP constructs (see Extended Data Fig. 3a-c for cell line details). Xist induction for 72 h; images are shown as z-sections. Scale bars, 10 μ m. *b*,

Quantification of images from a plotting (upper) SHARP FI over the Xi (denoted by Ezh2) normalized to the FI of a random nuclear region of the same size or (lower) Ezh2 FI over the same area normalized to a random nuclear region (see Extended Data Fig. 3d for quantification details). Values for individual nuclei ($n > 11$ for each condition) are shown; red lines represent median values; 0 represents enrichment not higher than that measured over a random nuclear region. c, SHARP enrichment across the first exon of Xist after UV-crosslinking and purification using the HaloTag in female TX SHARP-AID mESCs treated with auxin. HaloTags were fused to FL-SHARP (upper), Δ IDR-SHARP (middle) or Δ RRM-SHARP (lower; see Extended Data Fig. 3b,c for cell line details). Two replicates shown for each cell line; magenta arrow represents beginning of the first Xist exon; pink square demarcates Xist A-repeat (SHARP binding site). d, Close-up of crosslink-induced truncation sites for the three rescue conditions on the A-repeat region from c (demarcated by pink square).

Because SHARP binds directly to Xist^{22,25-27}, we tested whether the Δ RRM- and Δ IDR-SHARP mutants fail to localize on the Xi simply because they cannot bind Xist. To do this, we UV-crosslinked intact cells to form a covalent crosslink between directly interacting proteins and RNA, purified the Halo-tagged SHARP proteins using fully denaturing conditions, and sequenced the associated RNAs (**Methods**). We observed that FL-SHARP forms a highly specific interaction with the A-repeat region of Xist. By contrast, expression of Δ RRM-SHARP ablated this interaction across Xist. Interestingly, Δ IDR-SHARP is still able to bind the A-repeat of Xist at levels and positions comparable with those observed for FL-SHARP (**Fig. 3c,d**). These observations are consistent with previous studies showing that the RRM domains of SHARP are sufficient to bind to Xist^{26,32}. Together, these results demonstrate that the IDRs of SHARP are essential for its enrichment on the Xi (**Fig. 3a,b**) even though they are not required for direct

binding to Xist (**Fig. 3c,d**). A recent study similarly found that the IDRs of SHARP are required for its accumulation on the Xi¹⁹.

To exclude the possibility that Δ IDR-SHARP impacts localization on the Xi because it disrupts a cryptic localization domain contained within the protein, we tested whether we could rescue the Xi localization deficits simply by promoting multivalent homotypic associations. To do this, we used our FUS- Δ IDR-SHARP system that forms foci independently of its IDRs (**Fig. 2h**) and explored whether this could rescue SHARP localization on the X in mESCs. Indeed, we observed that FUS- Δ IDR-SHARP showed levels of localization over the Xi that were comparable with FL-SHARP after 72 h of Xist induction (**Fig. 3a,b**). These results demonstrate that the ability of SHARP to form homotypic assemblies (via its IDRs) is essential for its accumulation on the Xi.

2.3.4 | SHARP assemblies are required for gene silencing on the X

Because the Δ IDR-SHARP mutant does not accumulate on the Xi, we hypothesized that the ability to form SHARP assemblies is required for Xist-mediated transcriptional silencing during initiation of XCI.

To measure silencing, we performed RNA FISH on Xist and the introns of: (1) several genes located across the X that are known to be silenced upon XCI, and (2) genes that are known to escape XCI and therefore remain active upon Xist induction (**Fig. 4a and Extended Data Fig. 4a**). This single-cell readout allows us to restrict our analyses to cells that induce Xist expression (~50% of cells) and retain both X chromosomes (~50% of cells; **Extended Data Fig.**

4b)^{39,54,55}. Of these cells, we found that ~80% successfully silenced gene expression on one of the two X chromosomes upon Xist induction in wild-type mESCs (**Fig. 4b,c and Extended Data Fig. 4c**). Next, we measured gene silencing upon genetic deletion (SHARP-KO) or auxin-mediated degradation (SHARP-AID) of SHARP and found that both conditions led to loss of Xist-mediated transcriptional silencing (**Fig. 4b,c and Extended Data Fig. 4c**).

We measured transcription of the same X-linked genes after stable expression of FL-SHARP, Δ RRM-SHARP or Δ IDR-SHARP in both SHARP-KO and SHARP-AID backgrounds (**Fig. 4d,f and Extended Data Fig. 4d,f**). As expected, expression of FL-SHARP rescued silencing of these X-linked genes. By contrast, expression of Δ RRM-SHARP failed to silence any of these X-linked genes, consistent with the fact that it can no longer bind to Xist. Importantly, expression of Δ IDR-SHARP also failed to silence these genes (**Fig. 4e,g and Extended Data Fig. 4e,g**). Consistent with this observation, a recent study similarly found that deletion of the SHARP IDRs impacts transcriptional silencing across the entire chromosome when measured by single-cell RNA sequencing¹⁹.

To confirm that silencing depends on the ability of SHARP to form assemblies via its IDRs and not on a specific sequence within the IDRs, we performed the same assay using our SHARP-KO or SHARP-AID cells expressing the synthetic FUS- Δ IDR-SHARP construct that rescues SHARP assembly formation and Xi enrichment (**Figs. 2h and 3a,b**). We observed rescue of Xist-mediated transcriptional silencing, comparable with that observed upon expression of FL-SHARP (~70% silenced cells for FUS- Δ IDR-SHARP versus ~75% for FL-SHARP) (**Fig. 4e,g and Extended Data Fig. 4e,g**).

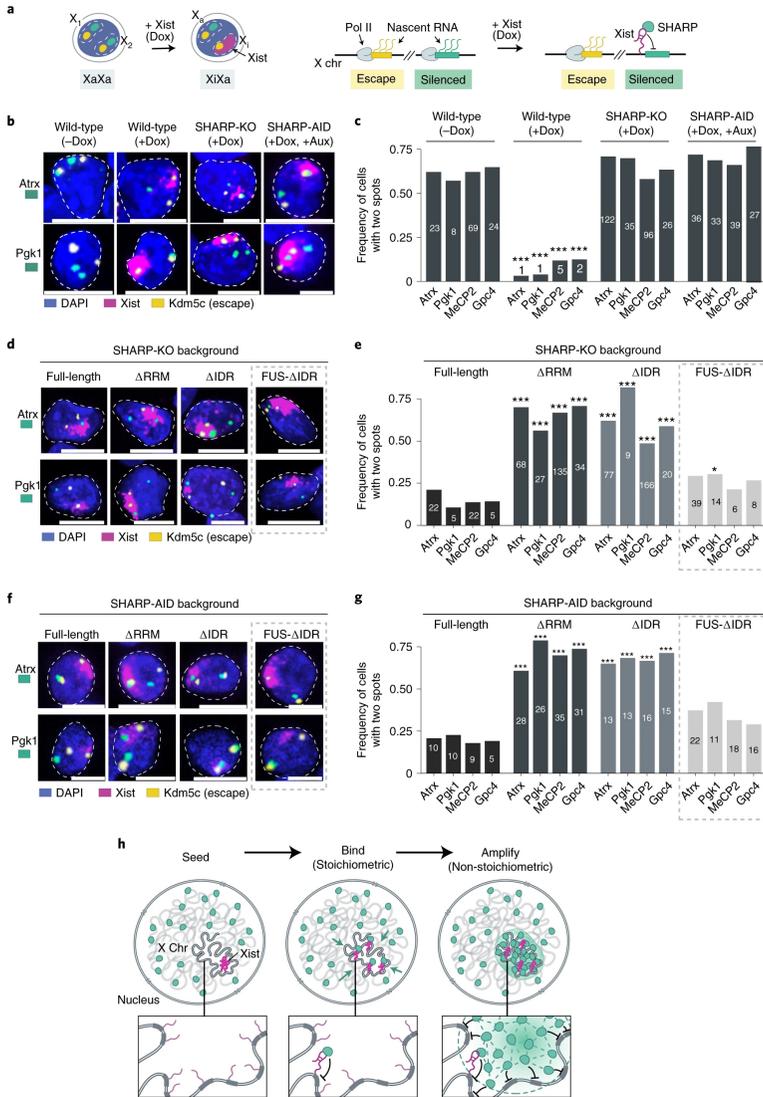


Figure 4: SHARP binding to RNA and formation of assemblies are both required for chromosome-wide gene silencing. *a*, Illustration of RNA FISH measurements in dox-inducible female mESCs. Genes that are silenced upon Xist induction are shown in green, genes that escape XCI are shown in yellow, and Xist is shown in magenta. chr, chromosome. *b*, RNA FISH images

representing (left to right): wild-type (no dox), wild-type (with dox), SHARP-KO (with dox) and auxin-treated SHARP-AID (with dox) cells. Cells were stained for DAPI (blue) and probed for *Xist* (magenta), escape gene *Kdm5c* (yellow), and silenced genes *Atrx* or *Pgk1* (green). Images are shown as maximum projections. Scale bars, 10 μ m. Total cell numbers (c) are from three independent experiments. c, Quantification of RNA FISH images from b representing the frequency of cells containing two actively transcribed alleles. Asterisks represent P values calculated for two-sided z-test; distributions compared with the FL group. *P = 0.05, ***P = 0.001. d, RNA FISH images for SHARP-KO female mESCs containing stable integrations of SHARP rescue constructs. Cells were stained as described in b. Images are shown as maximum projections. Scale bars, 10 μ m. Total cell numbers (e) are from three independent experiments. e, Quantification of RNA FISH images from d representing the frequency of cells containing two actively transcribed alleles for the various SHARP rescue constructs in SHARP-KO female mESCs. Asterisks represent P values calculated for two-sided z-test; distributions compared with the FL group. *P = 0.05, ***P = 0.001. f, RNA FISH images for SHARP-AID female mESCs containing stable integrations of SHARP rescue constructs. Cells were stained as described in b. Images are shown as maximum projections. Scale bars, 10 μ m. Total cell numbers (g) are from three independent experiments. g, Quantification of RNA FISH images from f representing the frequency of cells containing two actively transcribed alleles for the various SHARP rescue constructs in SHARP-AID female mESCs. Asterisks represent P values calculated for two-sided z-test; distributions compared with the FL group. *P = 0.05, ***P = 0.001. h, Schematic of the spatial amplification mechanism by which *Xist* RNA (magenta) can act to amplify SHARP (green) recruitment and gene silencing across the X.

Together, these results demonstrate that direct binding of SHARP to *Xist* (via its RRM domains) and its ability to form concentration-dependent homotypic assemblies (via its IDRs) are both essential and distinct components required

for chromosome-wide silencing on the Xi. Our data suggest a spatial amplification mechanism, whereby the direct interaction between Xist (which is enriched on the X) and SHARP (which is diffusible throughout the nucleus) acts to increase the local concentration of SHARP on the X. The resulting high local concentrations of SHARP on the X chromosome enable formation of IDR-mediated concentration-dependent assemblies that can occur between molecules not directly bound to Xist. In this way, these RNA-mediated assemblies can lead to the accumulation of SHARP on the Xi in stoichiometric excess of the number of Xist molecules to enable chromosome-wide silencing (**Fig. 4h**).

2.3.5 | Low Xist expression levels limit spreading to autosomes

This spatial amplification mechanism explains how Xist can achieve chromosome-wide silencing despite being expressed at substoichiometric levels relative to its target genes (**Extended Data Fig. 5a-c and Supplementary Note**). However, it does not explain why Xist expression levels are low and whether this might be critical for its functional role during XCI. Because Xist spreads to sites on the X based on 3D diffusion from its transcription locus^{16,17}, we hypothesized that its expression level might control how far it spreads in the nucleus. If true, we would expect that expressing Xist at increasing concentrations would lead to increasing localization of Xist to autosomal regions.

To test this, we used our dox-inducible Xist system, which enables induction of Xist across a range of expression levels by titrating the concentration of dox (**Fig. 5a**). We induced Xist expression across a range of dox concentrations (referred to as a 0.05x-3x Dox, where 1x = 2 $\mu\text{g/mL}$), imaged

Xist localization in individual cells (**Fig. 5b**) and observed a strong correlation between Xist expression levels and the area of the nucleus it occupies within individual cells ($r = 0.75$; **Fig. 5c**). Specifically, Xist occupies on average $\sim 6.5\%$ of the area of the nucleus when expressed upon RA-induced differentiation (endogenous control). However, cells treated with 3x dox express on average ~ 3.4 -fold higher levels of Xist (relative to average endogenous levels) and Xist occupies on average $\sim 23\%$ of the area of the entire nucleus (**Fig. 5c**).

To determine whether the larger nuclear volumes occupied by Xist correspond to increased localization on autosomes, we performed RNA antisense purification (RAP)¹⁶ on Xist and sequenced its associated DNA regions across three different induction conditions (0.25x, 1x, and 3x dox) as well as a negative (no dox) control (**Fig. 5d**). Because RAP is a bulk measurement, we first confirmed that Xist expression increases in the presence of increasing dox concentrations within a population of cells (using reverse transcription quantitative PCR (RT-qPCR)); **Extended Data Fig. 5d**). For each RAP sample, we then computed the level of Xist RNA enrichment on the X by quantifying the proportion of sequencing reads that align to the X relative to autosomes (A). We compared this with the expected X:A ratio observed in the unselected genomic DNA sample (input) (**Methods**). In all dox conditions we observed clear enrichment of Xist on the X (**Fig. 5d**). However, we observed a steady decrease in X:A enrichment as Xist concentration increased. For example, in samples treated with 3x dox (approximately fivefold above endogenous levels) we observed a more than twofold reduction in the X:A enrichment compared with samples treated with 0.25x dox (which approximates endogenous Xist levels) (**Fig. 5e**).

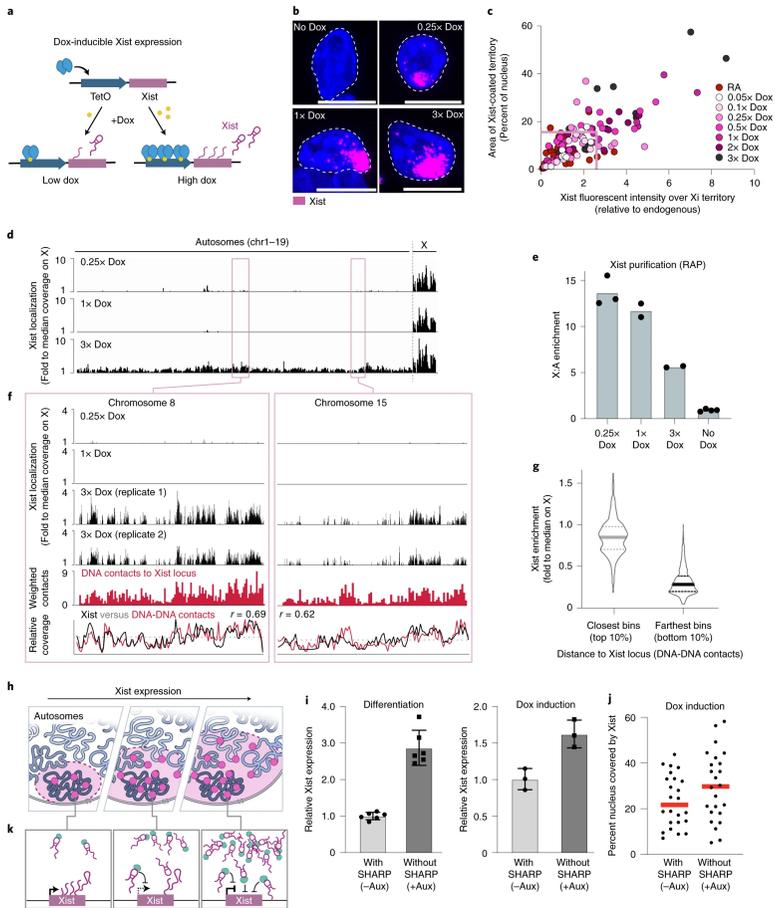


Figure 5: Low Xist expression levels limit its ability to spread to autosomes. *a*, Modulation of Xist levels using a dox-inducible system. TetO, Tet-On promoter. *b*, Images of Xist (magenta) within nucleus (DAPI) of mESCs treated with increasing dox concentrations. Scale bars, 10 μ m. Total cell numbers (*c*) from two independent experiments. *c*, Percent of nucleus covered by Xist relative to Xist levels in RA-differentiated and dox-induced mESCs; box represents 95th percentile of Xist expression and area upon RA-differentiation. *d*, Xist enrichment relative to total genomic DNA measured by RAP and normalized to the median coverage across the X in mESCs treated with 0.25x, 1x and 3x dox. *e*, Xist enrichment on the X relative to autosomes (A) in 0.25x, 1x

and 3x dox-induced and uninduced (no dox) female mESCs as measured by the proportion of sequencing reads that align to the X relative to A in RAP-DNA samples normalized to the expected X:A ratio observed in the unselected genomic DNA sample (input). Dots represent individual replicates; bars represent mean value. **f**, Xist enrichment over chr8 (left) and chr15 (right) measured by RAP; Xist enrichment relative to median coverage on the X (black bars); DNA contact frequency⁵⁶ of each region relative to the Xist locus (red bars); overlay between Xist enrichment and 3D contacts with the Xist locus. **g**, Xist enrichment in 3x dox sample over 1 Mb autosomal regions closest to Xist locus (left, top 10%) or furthest from Xist locus (right, bottom 10%) based on SPRITE data⁵⁶; bold lines represent median values; dotted lines represent 25th and 75th percentiles. **h**, Schematic depicting increased Xist spreading with increasing Xist expression levels. **i**, Relative Xist expression in RA-differentiated (left) and dox-induced (right) female SHARP-AID cells in the absence or presence of auxin. Dots represent individual replicates; $n = 6$ for differentiation conditions; $n = 3$ for dox induction conditions; data are represented as mean \pm s.d. **j**, Percent of nucleus occupied by Xist in dox-induced SHARP-AID cells in the absence or presence of auxin. Dots represent individual replicates; red bar represents the median. **k**, Model illustrating how Xist (through SHARP) may suppress production of its own RNA through negative feedback.

Because Xist spreads via 3D diffusion, we hypothesized that the autosomal regions that become occupied at increasing dox concentrations are those that are closest to the Xist locus in 3D space. To test this, we computed the 3D contact frequency between the Xist genomic locus and all 1-Mb genomic regions across autosomes⁵⁶ (**Fig. 5f and Extended Data Fig. 5e**). Interestingly, we observed a strong correlation between autosomal regions that are closest to the Xist locus and those that display increased Xist RNA occupancy in the 3x dox condition (**Fig. 5g and Extended Data Fig. 5f**). Taken together, these results indicate that

substoichiometric expression of Xist (low number of Xist molecules) is a critical mechanism by which cells limit Xist spreading to autosomal regions and ensure its specificity to the X (**Fig. 5h**).

Given that low Xist expression levels are critical for ensuring specificity to the X, we considered possible mechanisms that may act to limit its expression level *in vivo*. One long-puzzling observation is that even though Xist and SHARP accumulate in proximity to the Xist transcriptional locus^{16,17,24}, the Xist gene remains actively transcribed – an essential requirement for XCI. We hypothesized that Xist-SHARP accumulation at its own locus might act to control the level of Xist expression. To test this, we treated SHARP-AID mESCs with auxin and measured Xist expression levels upon dox induction or RA-differentiation using RT-qPCR. In both cases, we found that depletion of SHARP leads to an approximately twofold average upregulation of Xist expression (**Fig. 5i and Extended Data Fig. 5g**). Consistent with the fact that increased Xist expression leads to an increase in Xist spreading within the nucleus, we observed that degradation of SHARP led to a higher percentage of the nucleus being occupied by Xist (**Fig. 5j**). Because negative feedback loops often act as regulatory mechanisms to restrict production levels within a defined range^{57–60}, our results suggest that Xist may act to suppress its own production to ensure specificity to the X (**Fig. 5k**).

2.4 | Discussion

Our results demonstrate a critical spatial amplification mechanism by which Xist balances two essential but countervailing regulatory objectives: specificity to the X and chromosome-wide gene silencing (**Fig. 6**). We showed that low Xist RNA levels are necessary to ensure specificity to

its target sites on the X. Yet, it creates another challenge in that the RNA is expressed at substoichiometric levels compared with its regulatory targets and therefore cannot localize at each of them. We showed that Xist overcomes this challenge by driving non-stoichiometric recruitment of SHARP to amplify its abundance across the X and enable chromosome-wide gene silencing. Although a stoichiometric model (in which Xist recruits SHARP through direct binding and localizes at each of its target genes) would also enable chromosome-wide silencing, it would require Xist to be expressed at dramatically higher levels and therefore reduce Xist specificity to the X. Although the spatial amplification mechanism can achieve both specificity and robust silencing, balancing these two competing objectives requires precise quantitative control of Xist RNA levels. Our results highlight a negative feedback loop, whereby Xist (through SHARP) may act to suppress production of its own RNA to restrict its ability to spread beyond the X (Fig. 6).

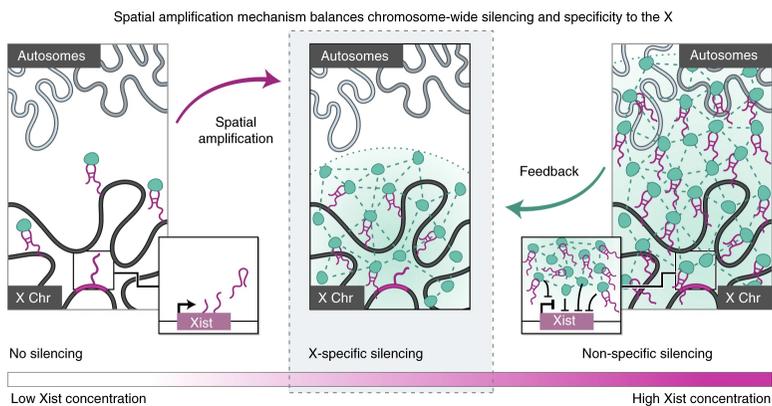


Figure 6: The spatial amplification mechanism balances chromosome-wide silencing and specificity to the X. Schematic illustrating the spatial amplification mechanism. Left: Xist is expressed, accumulates on the X through 3D diffusion from its transcription locus, binds directly to SHARP and recruits it to the

X in a stoichiometric manner. Close-up on left: at low overall expression, the *Xist* gene remains actively transcribed. Middle: once SHARP molecules achieve sufficiently high spatial concentration over the *X*, they form concentration-dependent assemblies (spatial amplification) that enable non-stoichiometric accumulation of SHARP on the *X* and chromosome-wide silencing. Right: if *Xist* expression levels get too high, the RNA would start to spread to autosomal regions. Close-up on right: at high concentrations, *Xist* recruits more SHARP molecules to its own locus which can suppress transcription of its own gene. This acts to reduce *Xist* spreading and restrain the *Xist*-SHARP complex on the *X* (feedback).

This spatial amplification mechanism is dependent on the fact that *Xist* can form a high-concentration territory on the *X* through 3D diffusion from its transcription locus (**seed**). In this way, *Xist* binding to SHARP increases its concentration on the *X* (**recruit**) and enables formation of concentration-dependent protein assemblies that amplify recruitment of repressive proteins to enable chromosome-wide gene silencing (**amplify**, **Fig. 4h**). Furthermore, because *Xist* spreads to its targets via 3D diffusion from its transcription locus, localization specificity is sensitive to its overall expression levels (**restriction**, **Fig. 5h**).

Beyond *Xist*, this spatial amplification mechanism is likely to represent a more general mechanism by which lncRNAs can balance specificity to, and robust control of, their regulatory targets because many lncRNAs share these same properties. Specifically, many hundreds of lncRNAs have been shown to form high-concentration territories in spatial proximity to their transcription sites (**seed**) and can directly bind and recruit different regulatory proteins (**recruit**), including those that contain long IDRs⁶¹ (for example, HP1 (refs. ^{62,63}) and SHARP). In this way, lncRNA-mediated recruitment may enable spatial amplification of regulatory

proteins and robust regulation of their more abundant targets (**amplification**). Because many lncRNAs localize in 3D proximity to their targets, low expression levels may similarly be important for ensuring specificity to their genomic DNA targets (**restriction**).

In this way, spatial amplification may provide a mechanistic answer to two long-standing questions in the lncRNA-field: (1) why many lncRNAs are expressed at relatively low levels, and (2) how low-abundance lncRNAs can effectively regulate their more abundant targets.

2.5 | Author Note

While we were working on this manuscript, our long-time collaborators were exploring the localization of various proteins involved in XCI using super-resolution microscopy. In parallel, they made similar observations regarding the dynamics of Xist and SHARP localization and the role of its IDRs in localization and silencing on the X. These results were reported in Markaki et al. Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **25**, 6174–6192.e32 (2021). Although this work was performed in parallel and none of the data in these papers overlap, we openly shared several reagents, including the various SHARP mutant constructs that we had generated for this manuscript and some of our early observations.

2.6 | Methods

Cell culture

Mouse embryonic stem cell (mESC) culture

Wild-type and endogenous SHARP-HALO-AID TX1072 female mESCs (gift from E. Heard laboratory) were cultured as previously described^{24,39}. Briefly, TX1072 mESCs were grown on gelatin-coated plates in serum-containing ES cell medium (high glucose DMEM (Gibco, Life Technologies), 15% FBS (Omega Scientific), 2 mM L-glutamine (Gibco, Life Technologies), 1 mM sodium pyruvate (Gibco, Life Technologies), 0.1 mM β -mercaptoethanol, 1000 U/ml leukemia inhibitory factor (Chemicon)), and 2i (3 μ M Gsk3 inhibitor CT-99021, 1 μ M MEK inhibitor PD0325901). The cell culture medium was replaced every 24 h.

Expression of Xist and/or each SHARP rescue construct (FL-SHARP, Δ RRM-SHARP, Δ IDR-SHARP, FUS- Δ IDR-SHARP; see **Supplementary Table 1** for complete plasmid list) was induced by treating cells with 2 μ g/mL dox (Sigma) for at least 72 h. Dox-containing medium was replaced every 24 h. For experiments using SHARP-AID mESCs, cells were treated with indole-3-acetic acid (IAA) for 24 h before the addition of dox to ensure complete degradation of endogenous SHARP before induction of Xist and SHARP rescue constructs. For RNA FISH and immunofluorescence (IF), cells were trypsinized into a single cell suspension, plated directly on poly-D-lysine coated coverslips and grown for at least 6 h before fixation.

Human HEK293T cell culture

HEK293T cells were cultured in complete media comprising DMEM (Gibco, Life Technologies) supplemented with 10% FBS (Seradigm Premium Grade HI FBS, VWR), 1x penicillin-streptomycin (Gibco, Life Technologies), 1 mM

sodium pyruvate (Gibco, Life Technologies), and 1x MEM non-essential amino acids (Gibco, Life Technologies). Cells were maintained in 37°C incubators under 5% CO₂.

Differentiation with RA

Wild-type (F1) and TX1072 female mESCs were grown for 24 h in ES cell medium. ES cell medium was then replaced with MEF medium (high glucose DMEM (Gibco, Life Technologies), 10% FBS (Omega Scientific), 2 mM L-glutamine (Gibco, Life Technologies), 1 mM sodium pyruvate (Gibco, Life Technologies), 0.1 mM MEM non-essential amino acids (Gibco, Life Technologies), 0.1 mM β -mercaptoethanol). After 24 h in MEF medium, the medium was replaced with MEF medium supplemented with 1 μ M RA (Sigma). Cells were then grown in MEF medium containing RA for 24 h (48 h differentiation total) or 48 h (72 h differentiation total). For cells differentiated for 72 h, MEF medium containing RA was replaced after 24 h.

To ensure that replacement of the endogenous Xist promoter with a dox-inducible promoter in TX1072 cells does not impair endogenous expression of Xist upon differentiation, Xist levels were measured in both TX1072 and F1 female mESCs (using RT-qPCR; protocol and quantification described below) after 72 h of differentiation with RA. On the basis of this bulk measurement, Xist levels in TX1072 mESCs were approximately half of those in F1 mESCs (**Extended Data Fig. 5h**); however, the percent of single cell nuclei occupied by Xist in both TX1072 and F1 mESCs was roughly the same (**Extended Data Fig. 5i**; protocol and quantification described below).

Cell line generation

SHARP-KO cells

To generate a plasmid targeting SHARP for deletion (see **Supplementary Table 1** for complete plasmid list), four different guide RNA (gRNA) sequences (see **Supplementary Table 2** for sequences; **Extended Data Fig. 3a**) were multiplexed into a Cas9-nickase backbone (Addgene plasmid 48140) as previously described⁶⁴. To create a SHARP knockout (SHARP-KO) cell line, two million TX1072 mESCs were transfected with 1.25 μ g of the multiplexed Cas9n-gRNA plasmid containing green fluorescent protein (GFP) using the Neon transfection system (settings: 1400 V, 10 ms width, three pulses). Successfully transfected cells were enriched by performing fluorescence-activated cell sorting (FACS) for GFP and subsequently plated at low-confluency. After 4-5 days of growth, 96 single colonies were picked and seeded in a 96-well plate. These cells were then split into one plate for PCR genotyping and another plate for maintaining growth until positive clones were identified. PCR genotyping was performed using Q5 High-Fidelity 2x Master Mix (NEB) with the primer pairs listed in **Supplementary Table 2**. SHARP-KO clone H8 was used for subsequent experimentation and all other clones were frozen.

SHARP rescue lines in SHARP-KO or SHARP-AID parent cells

To generate SHARP rescue cell lines, SHARP rescue constructs (FL-SHARP, Δ RRM-SHARP, Δ IDR-SHARP, FUS- Δ IDR-SHARP; **Extended Data Fig. 3b**; see **Supplementary Table 1** for complete plasmid list) were first made using the Gateway LR Clonase system (ThermoFisher). Specifically, Δ RRM-SHARP and Δ IDR-SHARP entry clones were created by modifying a full-length

mouse SHARP entry clone using polymerase incomplete primer extension (PIPE) mutagenesis⁶⁵. The specific amino acids deleted in the Δ RRM-SHARP and Δ IDR-SHARP entry clones are as follows:

Δ RRM-SHARP: amino acids 2-590

Δ IDR-SHARP: amino acids 639-3460

These entry clones (FL-SHARP, Δ RRM-SHARP, Δ IDR-SHARP) were then recombined into two different modified versions of the dox-inducible PiggyBac destination vector PB-TAG-ERN (Addgene plasmid 80476) containing truncated human nerve growth factor receptor (NGFR) and HALO or eGFP. This destination vector was chosen because it enables stable integration of the rescue constructs by cotransfecting with a PiggyBac transposase⁶⁶. The Halo-tagged version of this plasmid was created by replacing eGFP with NGFR (Addgene plasmid 27489) using Gibson assembly (NEB). HALO was then introduced downstream of reverse tetracycline-controlled transactivator (rtTA) using restriction enzyme digestion and ligation to create PB-HALO-IRES-NGFR. To generate the eGFP-tagged version of this plasmid, HALO was replaced with a 6-HIS-TEV-eGFP sequence using restriction enzyme digestion and ligation. Importantly, eGFP in this construct contains an amino acid substitution (A206K) to create a monomeric variant⁶⁷.

FUS- Δ IDR-SHARP was generated by recombining the Δ IDR-SHARP entry clone into a modified version of the PB-HALO-IRES-NGFR vector containing the IDR sequence from the FUS protein tagged with mCherry (Addgene plasmid 101223) in place of HALO. Importantly, the IDR from FUS exhibits no sequence homology to endogenous SHARP IDRs (that is, they have distinct amino acid compositions and distinct proportions of amino acid charge

properties), its sequence is approximately tenfold shorter than SHARP IDRs, and the locations of these two IDRs within the SHARP protein are distinct (**Extended Data Fig. 3b**).

To generate mESC lines expressing these SHARP rescue constructs, two million SHARP-KO clone H8 or SHARP-AID mESCs were transfected with 2.4 μg of the respective SHARP rescue construct tagged with HALO or eGFP (**Supplementary Table 1**), along with 0.8 μg of PiggyBac transposase plasmid (gift from M. Elowitz laboratory) and 1.2 μg of a non-targeting GFP plasmid (TurboGFP; Addgene plasmid 69072 cloned into pcDNA backbone with CMV promoter). Cells that were successfully transfected with the plasmids of interest (SHARP rescue constructs in HALO- or eGFP-tagged PiggyBac destination vector and PiggyBac transposase) were enriched by performing FACS on the cotransfected, non-targeting GFP. Cells were then cultured for 4-5 days to enable the SHARP rescue constructs to stably integrate into the genome (without inducing expression of Xist or the SHARP rescue proteins).

Next, cells were treated with indole-3-acetic acid (IAA; for SHARP-AID mESCs) and dox (previously described) to induce expression of Xist and the SHARP rescue proteins. Importantly, these cells were cultured in dox for a minimum of 72 h to ensure that any cells with toxic SHARP expression levels did not survive and were not analyzed further. For Halo-tagged rescue constructs, cells were labeled with 1 μM HaloTag Oregon Green Ligand (Promega) according to the manufacturer's instructions and both the HALO- and eGFP-tagged cell lines were sorted again to enrich for cells expressing the Halo- or eGFP-tagged SHARP rescue constructs (**Extended Data Fig. 3c**).

During FACS, laser powers and gains were set on the basis of the lowest expressing samples (FL-SHARP) and these settings were used for all other samples to enrich for cells with comparable expression levels of each rescue construct. Following FACS, cells were kept in medium supplemented with dox and used in further experiments (covalent linkage affinity purification (CLAP), IF, RNA FISH). Cells were retained only for a maximum of 14 days of culture in dox.

Overexpression of SHARP rescue constructs in HEK293T

For those experiments that required high protein expression (live-cell imaging, concentration-dependent imaging assays, CLAP, fluorescence recovery after photobleaching (FRAP)), human HEK293T cells were used instead of mESCs because they allow for much higher expression levels and enabled investigation of the biochemical and biophysical properties of each SHARP rescue construct in an independent system that is not undergoing initiation of XCI.

HEK293T cells were transfected using BioT transfection reagent (Bioland) according to the manufacturer's recommendations. Transfected constructs include FL-SHARP, Δ RRM-SHARP, Δ IDR-SHARP, FUS- Δ IDR-SHARP, EED, PTBP1, or an empty backbone (**Supplementary Table 1**); all constructs contained eGFP attached to the N terminus of each protein of interest driven by a dox-inducible promoter.

For live-cell imaging, fixed imaging and FRAP (**Fig. 2 and Extended Data Fig. 2**) ~10 μ g of DNA was used for transfection when cells were grown on a 15-cm dish or ~1 μ g of DNA when cells were grown on 3-cm glass-bottom dishes (Matek), and DNA concentrations were adjusted to

match mole numbers across constructs. Twenty-four hours after transfection, cells were treated with dox (2 $\mu\text{g}/\text{mL}$ (Sigma)) to induce expression of the proteins of interest and further experiments were performed 48 h post-dox treatment.

For assays measuring concentration-dependent assembly formation (**Fig. 2d,e**), ~ 2.5 fmol of DNA was transfected per well of 24-well plate, adjusting DNA concentration on the basis of the construct being used. Twenty-four hours after transfection, cells were treated with increasing concentrations of dox (0x, 0.1x, 0.5x, 1x where 1x = 2 $\mu\text{g}/\text{mL}$) for 24 h.

Protein and RNA visualization

Single molecule RNA fluorescence *in situ* hybridization (RNA FISH)

RNA FISH experiments were performed using the ViewRNA ISH Cell Assay (ThermoFisher, catalog no. QVC0001) protocol with minor modifications. Specifically, cells were fixed on coverslips with 4% formaldehyde in PBS for 15 min at room temperature and then permeabilized with 4% formaldehyde and 0.5% Triton X-100 in PBS for 10 min at room temperature. Cells were then washed twice with PBS, dehydrated with 70% ethanol and incubated at -20°C for at least 20 min or stored for up to 1 week. Coverslips were washed twice with PBS and then incubated with the desired combination of RNA FISH probes (**Extended Data Fig. 4a and Supplementary Table 3**; Affymetrix) in Probe Set Diluent at 40°C for at least 3 h. Coverslips were then washed once with wash buffer, twice with PBS, and once more with wash buffer before incubating in preamplifier mix solution at 40°C for 45 min. This step was repeated for the amplifier mix solution and label probe solution. Coverslips

were incubated with 1x DAPI in PBS at room temperature for 15 min and subsequently mounted onto glass slides using ProLong Gold with DAPI (Invitrogen, P36935).

Immunofluorescence (IF)

To focus our analysis specifically on nuclear SHARP, pre-extraction was performed on cells before immunostaining as previously described⁶¹. In brief, cells on coverslips were washed once with PBS and then incubated with cold 0.1% Triton X-100 in PBS for 1-3 min on ice. Next, cells were fixed on coverslips with 4% formaldehyde in PBS for 15 min at room temperature and permeabilized with 0.5% Triton X-100 in PBS for 10 min at room temperature. After washing twice with PBS containing 0.05% Tween (PBSt) and blocking with 2% BSA in PBSt for 30 min, cells were incubated with primary antibodies overnight at 4°C in 1% BSA in PBSt. After overnight incubation at 4°C, cells were washed three times in 1x PBSt and incubated for 1 h at room temperature with secondary antibodies labeled with Alexa fluorophores (Invitrogen) diluted in 1x PBSt (1:500). Next, coverslips were washed three times in PBSt, rinsed in PBS, rinsed in double-distilled H₂O, mounted with ProLong Gold with DAPI (Invitrogen, P36935) and stored at 4°C until acquisition.

Primary antibodies and the dilutions used are as follows: anti-Halo (mouse, Promega, catalog no. G9211, 1:200); anti-Ezh2 (mouse, Cell Signaling, catalog no. AC22 3147S, 1:500); anti-SHARP (rabbit, Bethyl, catalog no. A301-119A, 1:200). Secondary antibodies and the dilutions used are as follows: Alexa Fluor antibodies from ThermoFisher, dilution 1:500 in 1x PBS; anti-mouse: 488 (A32723), 555 (A32727), 647 (A32728); anti-rabbit: 647 (A32733), 555 (A32732), 488 (A32731).

RNA FISH and immunofluorescence

For IF combined with *in situ* RNA visualization, the ViewRNA Cell Plus (ThermoFisher Scientific, catalog no. 88-19000-99) kit was used according to the manufacturer's protocol with minor modifications. First, immunostaining was performed as described above but all incubations were performed in blocking buffer containing RNase inhibitor from the kit and all wash steps were done in RNase-free PBS with RNase inhibitor. Blocking buffer, PBS and RNase inhibitors were provided with the kit. After the last wash in PBS, cells underwent post-fixation with 2% formaldehyde in PBS for 10 min at room temperature, were washed three times in PBS and then RNA FISH was performed as described above.

HaloTag staining

To visualize proteins expressing Halo tags, HaloTag TMR (G8252) or OregonGreen (G2802) was used for fixed sample imaging combined with IF (**Extended Data Fig. 3d**), and Janelia549 (GA1110) was used for combined HALO staining and RNA FISH visualization (**Fig. 1b and Extended Data Fig. 1g**). Janelia549 was used for combined HALO staining and RNA FISH visualization because other HALO ligands did not survive the RNA FISH protocol. For protein labeling, cells were incubated with HaloTag ligands according to manufacturer's instructions and then directly imaged or washed with PBS, fixed in 4% formaldehyde (Pierce, ThermoFisher Scientific) and combined with immunostaining or RNA FISH.

Image acquisition and analysis

Microscopy

Fixed samples were imaged using: Zeiss LSM 800 with the 63x oil objective (RNA FISH, IF) and collected every 0.3 μm for 16 Z-stacks, Zeiss LSM 880 with Airyscan with the 63x oil objective (IF) and collected every 0.25 μm for 20 Z-stacks, or Zeiss LSM 980 with Airyscan2 with the 63x oil objective (IF, RNA FISH-IF) where zoom, scan format and number of Z-stacks were optimized on the basis of the software recommendations for the highest resolution (super-resolution module). For all images, laser power and gain were set at the beginning of acquisition and remained constant throughout the duration of acquisition to enable comparisons of fluorescent intensities. Live samples were imaged using the Leica Stellaris microscope with 63x water objective (~ 80 nm xy, ~ 300 nm z), and 16 Z-stacks were collected every 60 s for 5 min. The microscope was equipped with a stage incubator to keep cells at 37°C and 5% CO₂.

Analysis

Image analysis was performed using Icy (v2.3) or FIJI (ImageJ v2.1.0/1.53c) software. Live-cell videos and 3D reconstructions were created using Imaris software (v8) from Bitplane (Oxford Instruments Company).

Image quantification

Enrichment over inactive X territory

Xist and SHARP enrichments over the Xist territory (**Fig. 1**) were quantified using Icy (illustration **Extended Data Fig. 1c**). First, a region-of-interest (ROI) was defined that corresponded to the Xist signal across all Z-stacks by applying an intensity threshold (signal above background)

and a binary mask was created by demarcating the Xist-coated territory (ROI). Next, several features of these ROIs were quantified, namely, the areas in μm^2 (Area), total fluorescent intensities of Xist or SHARP over the entire ROI (Total Intensity), and average fluorescent intensity (FI) of Xist or SHARP per area unit of ROI (pixel/interior) (Average Intensity).

SHARP rescue construct enrichments over the Xi demarcated by Ezh2 staining (**Fig. 3a,b**) were quantified using Icy (illustration **Extended Data Fig. 3e**). First, images were processed into maximum intensity projections and two types of ROI were specified per nucleus: (1) corresponding to the Xi (X) by creating a binary mask based on Ezh2 marker, (2) and a control region corresponding to a random region (R) of the same size across all Z-stacks. Next, the average fluorescent intensities of SHARP or Ezh2 was quantified per ROI (X or R). Finally, to normalize for intercellular differences in the expression of rescue constructs, ROI-R was subtracted from ROI-X and divided by ROI-X. As such, if FI signal over the X is not higher than fluorescent signal in a comparably sized random region in the nucleus (R), the fold change should be centered around 0, whereas when there is enrichment, the signal should be greater than 0.

Pattern of SHARP localization

To determine the pattern of SHARP localization after transfecting HEK293T cells with eGFP-SHARP constructs (**Fig. 2e,g and Extended Data Fig. 2a**), images were first processed into a maximum intensity projection using Icy software. A binary mask was then created to demarcate each nucleus of a transfected cell by setting a threshold of eGFP intensity above background levels; all masks were visually verified and, if needed, manually adjusted to fit the nuclear region of cells. On the basis of these masks, an ROI was

defined that corresponded to the entire nucleus. Values for each pixel with the ROI (nucleus) were then extracted and this extracted information was used to quantify total intensity of protein per nucleus (sum of all pixels in an ROI), which corresponds to protein expression levels, and to calculate a SHARP dispersion score describing the differences in the distribution of pixel intensities across the nucleus. Specifically, for each cell, the intensity value at the 99th percentile of the distribution was computed and divided by the mode of the intensity distribution. This score was used because diffused localization shows distributed intensity across the nucleus and non-diffused localization shows accumulation of signal in defined locations, such that the tails of the intensity distributions were much longer. These quantitative assignments were visually confirmed to ensure that these scores capture our definition of diffused and non-diffused organization across cells.

Intron RNA FISH

For intron RNA FISH analysis, each image was processed into a maximum intensity projection using FIJI software. Then, the number of spots corresponding to each intron FISH probe per nucleus was manually counted and scored for the presence of Xist signal, number of spots per escape gene (Kdm5c, Kdm6a), and number of spots per silenced gene (Atrx, Pgk1, MeCP2, Gpc4) (**Fig. 4a**). Because mESCs are known to lose one of the X chromosomes or its fragments while in culture^{39,54,55} (**Extended Data Fig. 4b**), the analysis was restricted to cells containing two X chromosomes, which were determined by the presence of exactly two spots from escape gene. In addition, cells that had more than two spots for any gene or more than one Xist territory per nucleus were excluded from the analysis.

Xist percent of nucleus and FI

To calculate the percent of each nucleus occupied by Xist, each image was first processed into a maximum intensity projection using FIJI software. A binary mask was then created to demarcate each nucleus by setting a threshold intensity based on DAPI staining; all masks were visually verified and, if needed, manually adjusted to fit the nuclear region of cells. On the basis of these masks, an ROI was defined that corresponded to the entire nucleus and the size of the nucleus was calculated in FIJI based on the image metadata. Another binary mask was then created to demarcate the Xist territory by setting a threshold intensity based on Xist RNA FISH staining; all Xist masks were also visually verified and manually adjusted if necessary. An ROI was defined based on these masks and the size of this territory was calculated in FIJI based on the image metadata. The percent of each nucleus occupied by Xist was calculated by dividing the area of the Xist territory by the area of the corresponding DAPI-demarcated nucleus. The total FI of the Xist territory was calculated in FIJI based on the same Xist-defined ROI, and Xist FI values were normalized to the median intensity of RA-differentiated cells (**Fig. 5c**).

Fluorescence recovery after photobleaching

FRAP experiments were performed in HEK293T cells overexpressing eGFP-tagged FL-SHARP, PTBP1 or EED. Forty-eight hours post-transfection, cells were subjected to FRAP as previously described⁶⁸ using the Zeiss LSM 710 with the 40x water objective and equipped with a stage incubator to keep cells at 37°C and 5% CO₂. Briefly, in each nucleus an area of ~1 μm² was bleached with the argon laser to quench eGFP and fluorescence recovery was followed while imaging in the GFP channel for 235 s. FRAP experiments were analyzed first by measuring the mean

fluorescence intensity in the bleached area over time using Icy software and then normalized and averaged over n number of cells ($n > 5$) using EasyFRAP software⁶⁹. Error bars represent standard deviation of at least five replicates.

Covalent linkage affinity purification (CLAP) followed by RNA sequencing

Purification of Halo-tagged SHARP

CLAP was performed on mESCs expressing Halo-tagged SHARP constructs (**Supplementary Table 1**) as previously described⁷⁰ (**Fig. 3c,d**). Briefly, post-transfection, media was removed from cells and then crosslinked on ice using 0.25 J cm^{-2} (UV2.5k) of UV at 254 nm in a Spectrolinker UV Crosslinker. Cells were collected by scraping in 1x PBS and pelleted by centrifugation. Cell pellets were resuspended in 1 mL of ice cold lysis buffer (50 mM HEPES, pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) supplemented with 1x protease inhibitor cocktail (Promega), 200 U of Ribolock (ThermoFisher), 20 U TURBO DNase (Ambion) and 1x manganese/calcium mix (0.5 mM CaCl_2 , 2.5 mM MnCl_2). The samples were incubated on ice for 10 min and then at 37°C for 10 min at 700 rpm shaking on a ThermoMixer (Eppendorf). Lysates were cleared by centrifugation at 15,000 g for 2 min, and the supernatant was used for capture. For Halo-protein capture 50 μL of HaloLink Resin was pre-blocked using 1x blocking buffer (50 mM HEPES, pH 7.4, 100 $\mu\text{g}/\text{mL}$ BSA) for 20 min at room temperature with continuous rotation. After incubation, the resin was washed three times with 1x PBSt. The cleared lysate was mixed with 50 μl of pre-blocked HaloLink Resin and incubated at 4 °C for 3-16 h with continuous rotation. The captured protein bound to resin was washed three times with lysis buffer at room temperature and then washed three times at 90°C for 3 min while shaking on

a ThermoMixer at 1200 rpm with each of the following buffers: 1x N-lauroylsarcosine (NLS) buffer (1x PBS, 2% NLS, 10 mM EDTA), high salt buffer (50 mM HEPES, pH 7.4, 0.1% NP-40, 1M NaCl), 8M urea buffer (50 mM HEPES, pH 7.5, 0.1% NP-40, 8 M urea), Tween buffer (50 mM HEPES, pH 7.4, 0.1% Tween 20) and TEV buffer (50 mM HEPES, pH 7.4, 1 mM EDTA, 0.1% NP-40). Between each wash, samples were centrifuged at 1,000 g for 30 s and the supernatant was removed. After the last wash, samples were centrifuged at 7,500 g for 30 s and the supernatant was discarded. For elution, the resin was resuspended in 100 μ L of NLS buffer and 10 μ L of Proteinase K (NEB) and the sample was incubated at 50°C for 30 min while shaking at 1200 rpm. Capture reactions were transferred to microspin cups (Pierce, ThermoFisher), centrifuged at 2,000 g for 30 s, and the elutions were used for RNA purification by RNA Clean and Concentrate-5 kits (Zymo, >17 nucleotides protocol).

RNA library preparation and sequencing

RNA-seq library preparation was carried out as previously described⁷¹. Briefly, purified RNA was dephosphorylated (Fast AP) and cyclic phosphates were removed (T4 PNK). The RNA was then cleaned using Silane beads. An RNA adapter containing a reverse transcription (RT) primer binding site was ligated to the 3' end of the RNA and the ligated RNA was reverse transcribed into complementary DNA (cDNA). The RNA was then degraded using NaOH and a second adapter was ligated to the single-stranded cDNA. The DNA was amplified, and Illumina sequencing adapters were added by performing PCR with primers that are complementary to the 3' and 5' adapters that were previously added. The molarity of each PCR amplified library was measured using an Agilent TapeStation High Sensitivity DNA screentape and the samples were then pooled at equal molarity. This library pool was then size

selected on a 2% agarose gel by cutting between 150 and 800 nucleotides and performing gel purification (Zymo). To determine the loading density of the final pooled library, the sample was measured using an Agilent Bioanalyzer and Qubit dsDNA High Sensitivity assay (ThermoFisher). The final library was paired-end sequenced on an Illumina HiSeq 2500 with read length 35 x 35 nucleotides.

CLAP analysis and visualization

For HALO purifications and RNA-binding mapping, sequencing reads were aligned to the mouse genome (RefSeq mm10) using STAR aligner. All low-quality alignments (MAPQ < 255) and PCR duplicates were excluded from the analysis using the Picard MarkDuplicates function (<https://broadinstitute.github.io/picard/>). The enrichment relative to input coverage across the Xist RNA was quantified by computing the number of reads overlapping the window in the SHARP-elution sample divided by the total number of reads within the SHARP-elution sample. This ratio was normalized by dividing the number of reads in the same window contained in the input sample by the total number of reads in the input sample. Because all windows overlapping a gene should have the same expression level in the input sample (which represents RNA expression), the number of reads in the input was estimated as the maximum of either (1) the number of reads over the window or (2) the median read count over all windows within the gene. This approach provides a conservative estimate of enrichment because it prevents windows from being scored as enriched if the input values over a given window are artificially low, while at the same time accounting for any non-random issues that lead to increases in read counts over a given window (for example fragmentation biases or alignment artifacts leading to non-random assignment or pileups). These enrichment values were visualized in IGV⁷².

Crosslink-induced truncation sites

Because UV-crosslinking forms an irreversible covalent crosslink, reverse transcriptase has a well-described tendency to stall at crosslink sites. To exploit this to identify information about putative protein-binding sites at nucleotide resolution, the second adapter is ligated to the 3' end of the cDNA. In this way, the start position of the second read in a sequencing pair corresponds to this cDNA truncation point. To quantify these positions, the frequency of reads that start at each nucleotide was counted and plotted along the Xist RNA to identify the positions of direct crosslinking between the protein of interest and the RNA.

RNA affinity purification followed by DNA sequencing (RAP-DNA)

Cell treatment and preparation

For RAP-DNA sequencing, TX1072 cells were treated with increasing dox concentrations (0.25x, 0.5x, 1x, 2x, 3x where 0x = no dox and 1x = 2 $\mu\text{g/ml}$) for 72 h, changing dox-containing medium daily. Cells were harvested and crosslinked as previously described⁵⁶. Briefly, cells were pelleted, crosslinked with 2 mM disuccinimidyl glutarate for 45 min and 3% formaldehyde for 10 min, and lysed. Chromatin was then digested to 100-500 bp fragments through a combination of sonication and treatment with TURBO DNase and cell lysates were stored at -80°C until the next step of the procedure.

Purification of DNA sites bound by Xist RNA

DNA fragments occupied by Xist RNA were purified for RAP-DNA as previously described¹⁶ with minor modifications. Briefly, the lysate was diluted to hybridization conditions containing 3M guanidine

thiocyanate, precleared by adding streptavidin-coated magnetic beads and incubating for 30 min at 37°C, mixed with biotin-labeled single-stranded DNA capture probes, and incubated at 37°C for 2 h. 90-Mer single-stranded DNA oligonucleotide probes spanning the entire length of the target Xist RNA were purchased containing a 5' biotin (Eurofins Operon)²². Next, captured chromatin complexes were eluted with RNaseH and crosslinks were reversed by adding Proteinase K to the probe-bead complexes and incubating overnight at 65°C. Standard Illumina sequencing libraries were generated from eluted DNA fragments and sequenced at a depth of 5-20 million reads per sample of 75-75 or 75-140 long paired-end reads per sample.

RAP-DNA analysis and visualization

X to A enrichments were calculated by counting the number of reads that aligned to the X and the number aligned to A. This proportion was then compared with the proportion of reads that align to the X or A in the total input sample, which represent the total genomic DNA coverage without any selection. To compute enrichments per region of the genome, the number of reads for each genomic region within 10-kb windows was counted and this count was normalized by the total number of sequencing reads within each sample. Each window was then normalized by the proportion measured in the same bin within the input samples. To explore regions on autosomes that contain high Xist coverage, each bin was divided by the median values present on the X. In this way, all genomic regions containing coverage that was at least as high as half of the regions on the X could be visualized and their enrichment levels could be directly compared with the overall X coverage.

Computing 3D contact frequencies with the Xist locus

3D contact frequency between individual genomic regions and the Xist transcription locus was calculated as previously described⁵⁶. Specifically, all SPRITE clusters containing a DNA read overlapping the Xist locus (chrX:103460373-103483233, mm10) were extracted and a genome-weight contact frequency was computed by counting the total number of SPRITE clusters for each genomic region within this set. The analysis exclusively focused on clusters containing 2-100 reads per cluster and weighted the contact frequency by the cluster size from which it was present (2/cluster size) as previously described.

RT-qPCR

Dox-induced and differentiated female mESCs were lysed in RLT buffer (Qiagen) containing β -mercaptoethanol at a 1:100 dilution. RNA was then isolated using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. Genomic DNA was removed from the purified RNA samples with TURBO DNase (ThermoFisher) as per the manufacturer's protocol. Total RNA was then purified again using the RNA Clean and Concentrate-5 kit (Zymo, >17 nucleotides protocol). cDNA was generated from purified RNA using Maxima H minus reverse transcriptase (ThermoFisher) with random 9-mers according to manufacturer's specifications.

Amplification reactions were run in a Roche LightCycler 480 instrument using LightCycler 480 SYBR Green I Master (Roche) with the primer pairs listed in **Supplementary Table 2**. Each sample had between one and six biological replicates and four technical replicates. Median Ct values were used to calculate fold change with the $2^{-\Delta\Delta C_t}$ method. For differentiation and dox induction conditions in the

presence or absence of auxin (**Fig. 5i and Extended Data Fig. 5g**), each biological replicate was normalized to the median of the corresponding ‘with SHARP (-Aux)’ condition. For dox-induced samples across increasing concentrations (**Extended Data Fig. 5d**), each sample was normalized to the corresponding differentiation (RA) sample. For differentiated wild-type (F1) and TX1072 mESCs (**Extended Data Fig. 5h**), each sample was normalized to the corresponding wild-type sample.

Other data used in this study

RAP-DNA (F1 2-1 + 48 hours of RA): Xist localization across the X relative to gene density was measured using our previously published RAP-DNA dataset generated from Xist purification in F1 female mESCs differentiated with retinoic acid for 48 h¹⁶. All normalizations and analyses were performed as previously described and plotted using the normalized bedgraphs available at GEO accession GSE46918.

SPRITE (F1 2-1 mESCs): 3D contacts were measured using our previously published RNA-DNA SPRITE dataset⁶¹ that was generated in F1 female mESCs available at GEO accession GSE151515.

Data visualization

Bar graphs and violin plots were generated using GraphPad Prism (v8.4.3) or R (v4.0.3). Sequencing data was visualized using IGV (v2.8.11).

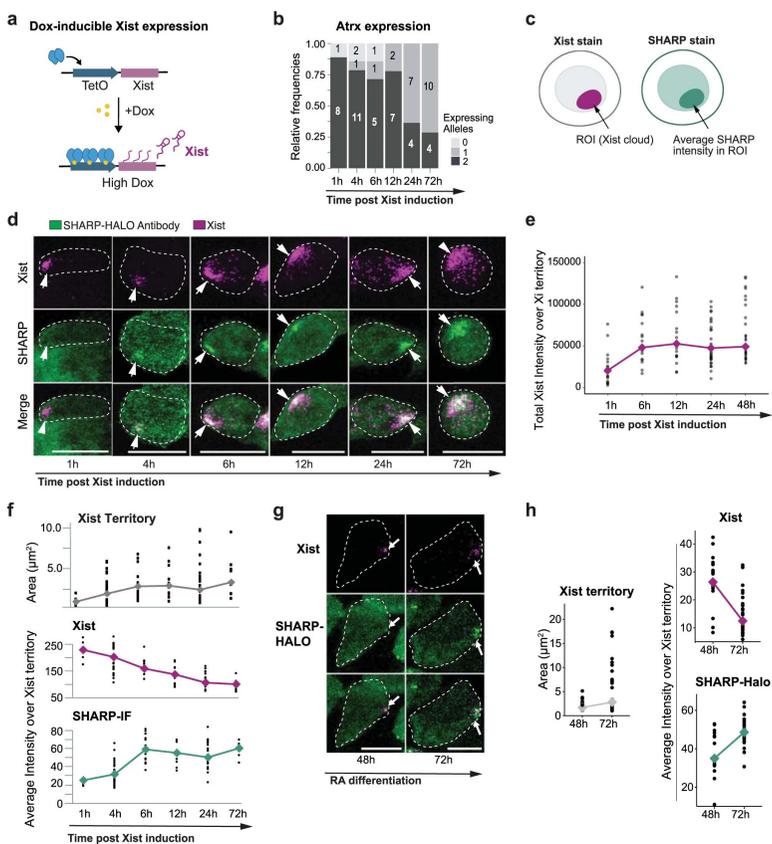
Data availability

Previously published RAP-DNA sequencing data used in **Extended Data Fig. 5b,f** is available at GEO accession GSE46918. RNA-DNA SPRITE data used in **Fig. 5f,g and Extended Data Fig. 5** is available at GEO accession GSE151515. CLAP sequencing data generated in this study and used in **Fig. 3c,d** and RAP-DNA sequencing data generated in this study and used in **Fig. 5d-f** is available at GEO accession GSE192574. Additional source data files are available for Figs. 1c,d,e,g, 2e,g, 3b, 4c,e,g and 5c,i,j. Source data are provided with this paper.

Statistics and reproducibility

Data are presented as mean/median \pm s.d., as indicated in the figure legends. Statistical analyses were performed using two-sided z-tests. Methods and details on individual statistical analyses and tests can be found in the respective figure legends. The number of times individual experiments were replicated is noted in their respective figure legends and the source data files.

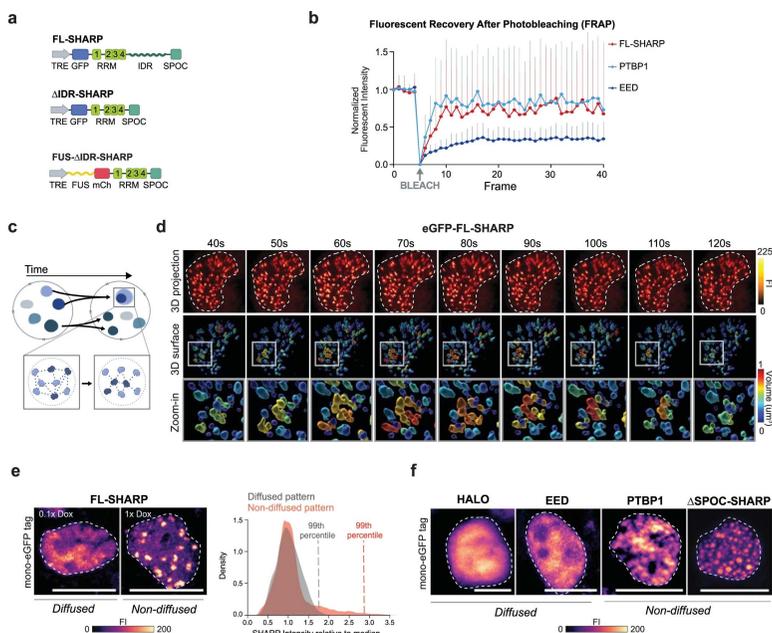
2.7 | Extended Data



Extended Data Figure 1: SHARP enrichment over the Xi increases in a non-stoichiometric manner relative to Xist.

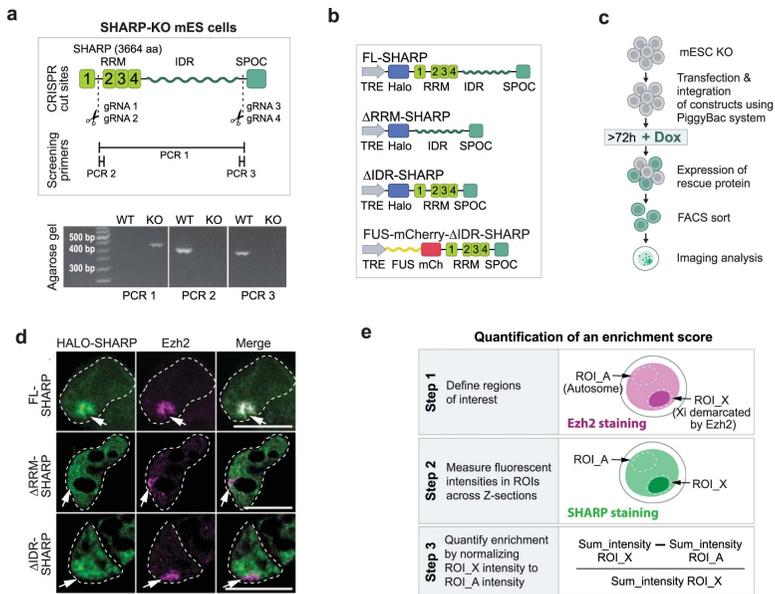
a, Schematic of dox-inducible Xist system. The endogenous Xist promoter is replaced with the tetracycline operator (TetO), which is activated upon addition of dox. **b**, Percent of cells expressing zero, one or two alleles of Atrx across 72 h of Xist induction as measured by intron RNA FISH. **c**, Illustration of SHARP enrichment over the Xi. FI values from RNA FISH, Halo or IF visualization were quantified within the defined Xi region. **d**, Representative images of Xist and SHARP localization across 72 h of Xist induction (n > 15 cells per condition from two independent experiments). Xist is visualized by RNA FISH

(magenta) and SHARP is visualized by immunofluorescence (green). Images are shown as maximum projections. Scale bars, 10 μm . **e**, Quantification of images from **Fig. 1b** plotting total FI of Xist across 48 h of dox induction. **f**, Quantification of images from **Extended Data Fig. 1d** plotting Xist and SHARP intensities across 72 h of Xist induction. Upper: area of the territory coated by Xist (μm^2). Middle: average FI of Xist per unit within the Xist territory. Lower: average FI of SHARP per unit within the Xist territory. **g**, Representative images of Xist and SHARP after 48 and 72 h of RA-induced differentiation ($n > 10$ cells for each condition). Xist is visualized by RNA FISH (magenta) and SHARP is visualized by direct Halo labeling (green). Images are shown as maximum projections. Scale bars, 10 μm . **h**, Quantification of images from **Extended Data Fig. 1g** plotting (left) area of the territory coated by Xist (μm^2), (upper right) average FI of Xist per unit within the Xist territory or (lower right) average FI of SHARP per unit within the Xist territory.



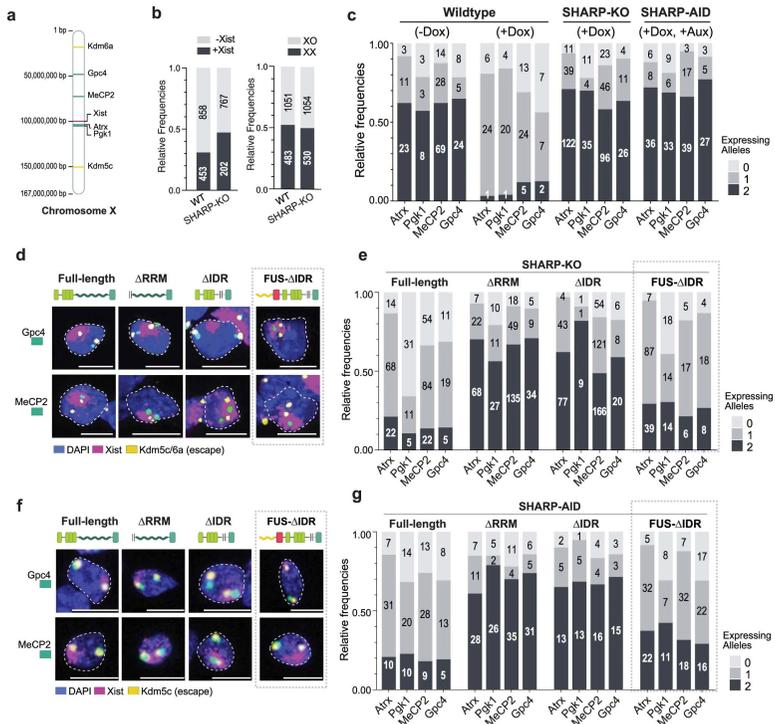
Extended Data Figure 2: SHARP forms multivalent, concentration-dependent assemblies in the nucleus.

a, Schematic of the domains included in the eGFP-tagged FL-SHARP and Δ IDR-SHARP, and the mCherry-tagged FUS- Δ IDR-SHARP rescue constructs used in **Fig. 2** and **Extended Data Fig. 2**. **b**, FRAP curves of eGFP-tagged FL-SHARP (red), positive control PTBPI (forms assemblies; light blue) and negative control EED (does not form assemblies; dark blue). Points represent mean values. Error bars represent the standard deviation of at least five replicates. **c**, Schematic depicting the physical characteristics of concentration-dependent assemblies, including (upper) foci formation, fission and fusion, and (lower) rapid diffusion of proteins within an assembly. **d**, Images across 90 s from a live-cell movie of eGFP-tagged FL-SHARP in transiently transfected HEK293T cells (**Supplementary Video 1,2**) showing non-diffused, focal organization of SHARP molecules. Upper: 3D reconstructions of the FI signal. Middle: 3D volume reconstructions color-coded based on the volume of the focus. Bottom: inset representing one region of the nucleus that changes volume across the series. **e**, Left: images representing FL-SHARP expressed with 0.1x dox (diffused) or 1x dox (non-diffused) in transiently transfected HEK293T cells. Images are shown as maximum projections. Scale bars, 10 μ m. Right: histograms representing FI for two cells showing diffused and non-diffused localization patterns of SHARP. Dashed lines correspond to the intensity at the 99th percentile of each distribution. **f**, Images representing the nuclear localization pattern of various eGFP-tagged proteins in transiently transfected HEK293T cells. Left: HALO and EED, which have not been reported to form assemblies. Right: PTBPI and Δ SPOC-SHARP, which both form assemblies. Images are shown as maximum projections. Scale bars, 10 μ m.



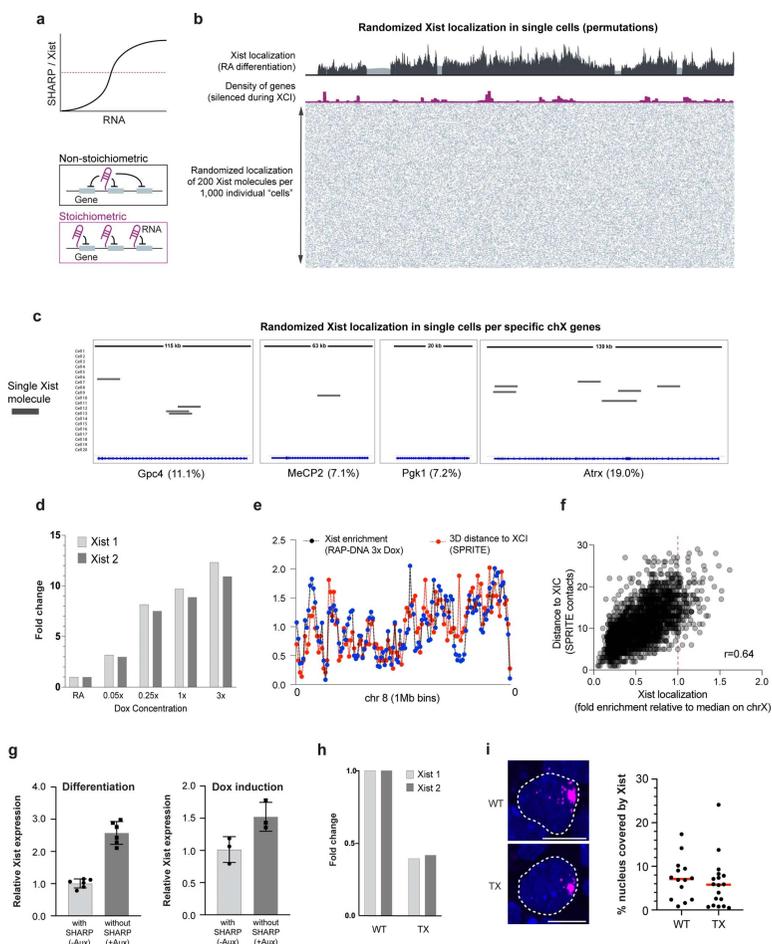
Extended Data Figure 3: Formation of SHARP assemblies is required for enrichment on the Xi but not Xist binding. **a**, Upper: schematic of CRISPR cut sites used to generate SHARP-KO mESCs and PCR primers used to screen for KO clones. Lower: agarose gel confirming homozygous deletion of SHARP in SHARP-KO clone H8 mESCs. **b**, Schematic of the constructs used to generate rescue cell lines in TX SHARP-KO or TX SHARP-HALO-AID mESCs. Grey arrow: dox-inducible promoter, blue box: HALO or eGFP tag, light green boxes: RNA recognition motifs (RRMs), wavy green line: intrinsically disordered region (IDR), dark green box: Spen paralog and ortholog C-terminal (SPOC) domain. FL-SHARP: full-length SHARP, Δ RRM-SHARP: deletion of RRM domain, Δ IDR-SHARP: deletion of IDR domain, FUS- Δ IDR-SHARP: deletion of IDR domain and insertion of alternative IDR domain from FUS protein. **c**, Schematic depicting experimental workflow for generating and enriching stable SHARP rescue mESCs (FL-SHARP, Δ RRM-SHARP, Δ IDR-SHARP, FUS- Δ IDR-SHARP) using the constructs from **Extended Data Fig. 3b**. **d**, Representative images of SHARP enrichment (Halo, green) over the Xi (anti-Ezh2 immunofluorescence, magenta) in female mESCs containing dox-inducible Xist, genetic

deletion of *SHARP*, and stable integrations of *HALO*-tagged *FL-SHARP*, Δ *RRM-SHARP* or Δ *IDR-SHARP* ($n > 10$ cells per condition). *Xist* and *SHARP* rescue constructs were induced with dox for 72 hours. Images are shown as Z-sections. Scale bars, 10 μ m. **e**, Diagram of image analysis pipeline for quantifying *SHARP* enrichment over the Xi (**Fig. 3b**).



Extended Data Figure 4: SHARP binding to RNA and formation of assemblies are both required for chromosome-wide gene silencing. **a**, Schematic of mouse X chromosome showing the locations of the various genes probed in RNA FISH experiments. **b**, Frequency of *Xist* induction (left) and X chromosome ploidy (right) in wild-type and SHARP-KO mESCs based on quantification of RNA FISH images from Fig. 4. **c**, Quantification of RNA FISH images from Fig. 4b representing the frequency of

cells containing two, one or zero actively transcribed alleles. Left to right: wild-type (no dox), wild-type (with dox), SHARP-KO (with dox), auxin-treated SHARP-AID (with dox) cells. **d**, RNA FISH images representing stable integrations of (left to right): FL-SHARP, Δ RRM-SHARP, Δ IDR-SHARP or FUS- Δ IDR-SHARP in SHARP-KO cells. Cells were stained for DAPI (blue) and probed for Xist (magenta), escape gene Kdm5c (yellow), and silenced genes Gpc4 or MeCP2 (green). Images are shown as maximum projections. Scale bars, 10 μ m. Total cell numbers (**Extended Data Fig. 4e**) are from three independent experiments. **e**, Quantification of RNA FISH images from **Extended Data Fig. 4d** representing the frequency of cells containing two, one or zero actively transcribed alleles. **f**, RNA FISH images representing stable integrations of (left to right): FL-SHARP, Δ RRM-SHARP, Δ IDR-SHARP or FUS- Δ IDR-SHARP in SHARP-AID cells. Cells were stained for DAPI (blue) and probed for Xist (magenta), escape gene Kdm5c (yellow), and silenced genes Gpc4 or MeCP2 (green). Images are shown as maximum projections. Scale bars, 10 μ m. Total cell numbers (**Extended Data Fig. 4g**) are from three independent experiments. **g**, Quantification of RNA FISH images from **Extended Data Fig. 4f** representing the frequency of cells containing two, one or zero actively transcribed alleles.



Extended Data Figure 5: Low Xist expression levels limit its ability to spread to autosomes. **a**, Upper: expected ratios of SHARP to Xist based on increasing concentration of Xist RNA. Lower: diagrams illustrating non-stoichiometric and stoichiometric models. **b**, Upper: Xist localization after 48 h of RA-differentiation from bulk RAP-DNA measurements¹⁶. Middle: gene density across the X. Only genes that undergo XCI are plotted. Lower: localization of 200 Xist molecules in 1000 random permutations (“cells”). Xist is represented by grey squares. **c**, Insets representing four different genes (*Gpc4*, *MeCP2*, *Pgl1*, *Atrx*) across 20 simulated “cells”. Bottom labels correspond to

the percent of cells in which *Xist* overlaps the gene in all 1000 permutations. **d**, Expression levels of *Xist* in mESCs treated with increasing dox concentrations as measured by RT-qPCR. **e**, Comparison of *Xist* occupancy (blue lines, 3x dox, RAP-DNA) and DNA contact frequency with the *Xist* locus (red lines, SPRITE⁶¹) across 1 Mb DNA regions of chromosome 8. *Xist* enrichment and 3D distance are normalized to their median coverage to place them on the same relative scale. **f**, Scatterplot representing the 3D contact frequency between each 1 Mb autosomal bin and the *Xist* locus (y-axis, SPRITE⁶¹) and DNA sites enriched by *Xist* (x-axis, 3x dox, RAP-DNA). **g**, Relative *Xist* expression upon RA-induced differentiation (left) and dox-induction (right) of SHARP-AID mESCs in the absence or presence of auxin as measured by RT-qPCR. Dots represent individual replicates. $n=6$ for differentiation conditions. $n=3$ for dox induction conditions. Data are represented as mean \pm s.d. **h**, Relative *Xist* expression in RA-differentiated wild-type F1 cells (left) and TX cells (right) as measured by RT-qPCR. **i**, Left: images representing *Xist* clouds in wild-type F1 and TX cells. Scale bars, 10 μ m. Right: quantification of the percent of each nucleus occupied by *Xist*. Dots represent individual replicates. Red bars represent the median. Total cell numbers are from two independent experiments.

2.8 | Supplementary Information

2.8.1 | Supplementary Note

Xist localization across the X and chromosome-wide silencing

Previous studies have shown that there are between 60-200 copies of Xist within each nucleus^{18,19}. This level of expression is sufficient to drive chromosome-wide silencing across the >1500 genes encoded on the X. Based on these numbers, Xist cannot simultaneously localize to each gene within each cell because there are not enough Xist molecules present; it must instead mediate silencing over several genes at once (**Extended Data Fig. 5a**). As such, Xist localization within individual cells must be heterogenous such that in one cell it localizes at a subset of genes but in another cell, it localizes at a different subset of genes.

Based on ensemble measurements, we know that Xist does not preferentially accumulate at specific sequences (e.g., promoters) but instead localizes broadly across the chromosome (**Extended Data Fig. 5b**). This means that the Xist RNA molecules within each cell must localize randomly at distinct positions spread across the >167 megabases of the chromosome.

Using this information, we can simulate the expected occupancy of Xist across the X within single cells in a manner that would explain the ensemble pattern (**Extended Data Fig. 5b**). We find that the likelihood that Xist is present over any given gene within an individual cell is extremely low (on average <5% of genes per cell would be covered by Xist) (**Extended Data Fig. 5c**). For example, Xist would be expected to localize over any region of Pcg1 in only ~7% of cells (**Extended Data Fig. 5c**). As such, if Xist-mediated silencing was solely dependent on such localization, we

would expect that this gene would remain active in >90% of individual cells. However, using our single cell measurements we observe that this gene is silenced in >87% of single cells (**Extended Data Fig. 4c**). Therefore, these single cell measurements allow us to measure chromosome-wide silencing when focusing on a subset of X chromosome genes.

2.8.2 | Supplementary Tables

Table 1. List of plasmids generated and used in this study

Plasmid Name	Addgene number	Description
Cas9-nickase-eGFP (Cas9n-eGFP)	48140	eGFP-tagged Cas9n backbone into which four different SHARP-targeting gRNAs were inserted to generate SHARP-KO mESCs
SHARP-targeting gRNA plasmid	-	Cas9n-eGFP backbone containing four different SHARP-targeting gRNAs
FL-SHARP entry clone	-	Entry clone of full-length SHARP sequence used for Gateway cloning
Δ RRM-SHARP entry clone	-	Entry clone of SHARP sequence lacking its RRM (deletion of amino acids 2-590) used for Gateway cloning
Δ IDR-SHARP entry clone	-	Entry clone of SHARP sequence lacking its IDR (deletion of amino acids 639-3460) used for Gateway cloning
PB-TAG-ERN	80476	PiggyBac destination vector that was modified prior to Gateway cloning
PB-HALO-IRES-NGFR destination vector	-	PiggyBac destination vector used for Gateway cloning containing HALO and NGFR

PB-eGFP-IRES-NGFR destination vector	-	PiggyBac destination vector used for Gateway cloning containing eGFP and NGFR
FUS-mCherry plasmid	101223	Plasmid from which FUS-mCherry fusion was derived to create PB-FUS-mCherry-IRES-NGFR
PB-FUS-mCherry-IRES-NGFR	-	PiggyBac destination vector used for Gateway cloning containing FUS-mCherry fusion and NGFR
HALO-FL-SHARP rescue construct	-	FL-SHARP entry clone inserted into PB-HALO-IRES-NGFR destination vector
HALO- Δ RRM-SHARP rescue construct	-	Δ RRM-SHARP entry clone inserted into PB-HALO-IRES-NGFR destination vector
HALO- Δ IDR-SHARP rescue construct	-	Δ IDR-SHARP entry clone inserted into PB-HALO-IRES-NGFR destination vector
eGFP-FL-SHARP rescue construct	-	FL-SHARP entry clone inserted into PB-eGFP-IRES-NGFR destination vector
eGFP- Δ RRM-SHARP rescue construct	-	Δ RRM-SHARP entry clone inserted into PB-eGFP-IRES-NGFR destination vector
eGFP- Δ IDR-SHARP rescue construct	-	Δ IDR-SHARP entry clone inserted into PB-eGFP-IRES-NGFR destination vector
FUS-mCherry- Δ IDR-SHARP rescue construct	-	Δ IDR-SHARP entry clone inserted into PB-FUS-mCherry-IRES-NGFR destination vector
turboGFP plasmid	69072	Non-targeting plasmid co-transfected with SHARP rescue constructs to enrich for successfully transfected cells
eGFP-EED construct	-	EED entry clone inserted into PB-eGFP-IRES-NGFR used for imaging in HEK293T cells
eGFP-Ptbp1 construct	-	Ptbp1 entry clone inserted into PB-eGFP-IRES-NGFR used for imaging in HEK293T cells

Table 2. List of oligo sequences used in this study

gRNA	Target	Top Strand	Bottom Strand
gRNA 1	SHARP	TCTGGAGTCAGGTGAGACGC	GCGTCTCACCTGACTCCAGA
gRNA 2	SHARP	GTGAGTGTGGTTACACCG	CGGTGTAAGCAAACACTCAC
gRNA 3	SHARP	CTCGGTTCTTACACAGCTCC	GGAGCTGTGTAAGAACCAG
gRNA 4	SHARP	TCTTTGAGCAAGACTCCAAG	CTTGAGTCTTGTCTAAAGA
Method	Target	Forward Sequence	Reverse Sequence
gDNA PCR 1	SHARP WT	ACACACACGCAATCACAAA	ATGAGTCTCCGGCTCTTCCT
gDNA PCR 2	SHARP KO	ACACACACGCAATCACAAA	TGAAGCCCTGCATTTAGGAG
gDNA PCR 3	SHARP KO	AGGCCTATTCCGTCCTGGTA	ATGAGTCTCCGGCTCTTCCT
RT-qPCR	Xist 1	GCCTCTGATTAGCCAGCAC	GCAACCCAGCAATAGTCAT
RT-qPCR	Xist 2	AGCCAGCACTGATCTCAAGC	GCAACCCAGCAATAGTCAT
RT-qPCR	GAPDH	CATGGCCTCCGTGTTCTTA	GCCTGCTTCACCACCTTCTT

Table 3. List of RNA FISH probes used in this study and corresponding design ID

Probe Target	Probe Type	Fluorescent Label	Affymetrix Design ID
Xist	4	Alexa Fluor 488	VB4-19746
Xist	6	Alexa Fluor 647	VB6-10824
Kdm5c	1	Alexa Fluor 546	VPYMJG7-01
Kdm5c	4	Alexa Fluor 488	VB4-20659
Kdm6a	1	Alexa Fluor 546	VPZTD24-01
Pgk1	1	Alexa Fluor 546	VB1-6000089
Atrx	6	Alexa Fluor 647	VPNKRFV-06
Gpc4	4	Alexa Fluor 488	VB4-17177
MeCP2	4	Alexa Fluor 488	VPH49TK

2.9 | References

1. Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).
2. Engreitz, J. M. *et al.* Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).
3. Carninci, P. *et al.* The transcriptional landscape of the mammalian genome. *Science* **309**, 1559–1563 (2005).
4. Lee, J. T. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes Dev.* **23**, 1831–1842 (2009).
5. Nagano, T. *et al.* The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**, 1717–1720 (2008).
6. Dunagin, M., Cabili, M. N., Rinn, J. & Raj, A. Visualization of lncRNA by single-molecule fluorescence in situ hybridization. *Methods Mol. Biol.* **1262**, 3–19 (2015).
7. Wang, K. C. *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120–124 (2011).
8. Carpenter, S. *et al.* A long noncoding RNA mediates both activation and repression of immune response genes. *Science* **341**, 789–792 (2013).
9. Thomson, D. W. & Dinger, M. E. Endogenous microRNA sponges: evidence and controversy. *Nat. Rev. Genet.* **17**, 272–283 (2016).
10. Kopp, F. & Mendell, J. T. Functional classification and experimental dissection of long noncoding RNAs. *Cell* **172**, 393 (2018).
11. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131–137

- (1996).
12. Marahrens, Y., Loring, J. & Jaenisch, R. Role of the Xist gene in X chromosome choosing. *Cell* **92**, 657–664 (1998).
 13. Dyer, K. A., Canfield, T. K. & Gartler, S. M. Molecular cytological differentiation of active from inactive X domains in interphase: implications for X chromosome inactivation. *Cytogenet. Genome Res.* **50**, 116–120 (1989).
 14. Chaumeil, J., Le Baccon, P., Wutz, A. & Heard, E. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev.* **20**, 2223–2237 (2006).
 15. Clemson, C. M., McNeil, J. A., Willard, H. F. & Lawrence, J. B. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* **132**, 259–275 (1996).
 16. Engreitz, J. M. *et al.* The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**, 1237973 (2013).
 17. Simon, M. D. *et al.* High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504**, 465–469 (2013).
 18. Sunwoo, H., Wu, J. Y. & Lee, J. T. The Xist RNA-PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proc. Natl. Acad. Sci. USA* **112**, E4216–E4225 (2015).
 19. Markaki, Y. *et al.* Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **184**, 6174–6192 (2021).
 20. Rodermund, L. *et al.* Time-resolved structured

- illumination microscopy reveals key principles of Xist RNA spreading. *Science* **372**, eabe7500 (2021).
21. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. & Jaenisch, R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* **11**, 156–166 (1997).
 22. McHugh, C. A. *et al.* The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232–236 (2015).
 23. Wutz, A., Rasmussen, T. P. & Jaenisch, R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* **30**, 167–174 (2002).
 24. Dossin, F. *et al.* SPEN integrates transcriptional and epigenetic control of X-inactivation. *Nature* **578**, 455–460 (2020).
 25. Chu, C. *et al.* Systematic discovery of Xist RNA binding proteins. *Cell* **161**, 404–416 (2015).
 26. Lu, Z. *et al.* RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* **165**, 1267–1279 (2016).
 27. Minajigi, A. *et al.* A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**, aab2276 (2015).
 28. Cirillo, D. *et al.* Quantitative predictions of protein interactions with long noncoding RNAs. *Nat. Methods* **14**, 5–6 (2016).
 29. Żylicz, J. J. *et al.* The implication of early chromatin changes in X chromosome inactivation. *Cell* **176**, 182–197 (2019).
 30. You, S. H. *et al.* Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 in vivo. *Nat. Struct. Mol. Biol.* **20**, 182–187 (2013).
 31. Shi, Y. *et al.* Sharp, an inducible cofactor that

- integrates nuclear receptor repression and activation. *Genes Dev.* **15**, 1140–1151 (2001).
32. Monfort, A. & Wutz, A. Identification of Spen as a crucial factor for Xist function through forward genetic screening in haploid embryonic stem cells. *Cell Rep.* **12**, 554–561 (2015).
 33. Moindrot, B. *et al.* A pooled shRNA screen identifies Rbm15, Spen, and Wtap as factors required for Xist RNA-mediated silencing. *Cell Rep.* **12**, 562–572 (2015).
 34. Nesterova, T. B. *et al.* Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. *Nat. Commun.* **10**, 3129 (2019).
 35. Wutz, A. & Jaenisch, R. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell* **5**, 695–705 (2000).
 36. Lee, J. T., Strauss, W. M., Dausman, J. A. & Jaenisch, R. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell* **86**, 83–94 (1996).
 37. Hall, L. L. *et al.* An ectopic human XIST gene can induce chromosome inactivation in postdifferentiation human HT-1080 cells. *Proc. Natl. Acad. Sci. USA* **99**, 8677–8682 (2002).
 38. Pacini, G. *et al.* Integrated analysis of Xist upregulation and X-chromosome inactivation with single-cell and single-allele resolution. *Nat. Commun.* **12**, 3638 (2021).
 39. Schulz, E. G. *et al.* The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell* **14**, 203–216 (2014).
 40. Barros de Andrade E Sousa, L. *et al.* Kinetics of Xist-induced gene silencing can be predicted from combinations of epigenetic and genomic features.

- Genome Res.* **29**, 1087–1099 (2019).
41. Arieti, F. *et al.* The crystal structure of the Split End protein SHARP adds a new layer of complexity to proteins containing RNA recognition motifs. *Nucleic Acids Res.* **42**, 6742–6752 (2014).
 42. Newberry, E. P., Latifi, T. & Towler, D. A. The RRM domain of MINT, a novel Msx2 binding protein, recognizes and regulates the rat osteocalcin promoter. *Biochemistry* **38**, 10678–10690 (1999).
 43. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
 44. Lin, Y., Currie, S. L. & Rosen, M. K. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J. Biol. Chem.* **292**, 19110–19120 (2017).
 45. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
 46. Boija, A. *et al.* Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855 (2018).
 47. Bhat, P., Honson, D. & Guttman, M. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nat. Rev. Mol. Cell Biol.* **22**, 653–670 (2021).
 48. McSwiggen, D. T., Mir, M., Darzacq, X. & Tjian, R. Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev.* **33**, 1619–1634 (2019).
 49. Banani, S. F. *et al.* Compositional control of phase-separated cellular bodies. *Cell* **166**, 651–663 (2016).
 50. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).
 51. Kato, M. *et al.* Cell-free formation of RNA granules:

- low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
52. Bracha, D. *et al.* Mapping local and global liquid phase behavior in living cells using photo-oligomerizable seeds. *Cell* **175**, 1467–1480 (2018).
 53. Lin, Y., Protter, D. S. W., Rosen, M. K. & Parker, R. Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell* **60**, 208–219 (2015).
 54. Choi, J. *et al.* Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature* **548**, 219–223 (2017).
 55. Yagi, M. *et al.* Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. *Nature* **548**, 224–227 (2017).
 56. Quinodoz, S. A. *et al.* Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757 (2018).
 57. Wollerton, M. C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. & Smith, C. W. J. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol. Cell* **13**, 91–100 (2004).
 58. Ottens, F. & Gehring, N. H. Physiological and pathophysiological role of nonsense-mediated mRNA decay. *Eur. J. Physiol.* **468**, 1013–1028 (2016).
 59. Müller-McNicoll, M., Rossbach, O., Hui, J. & Medenbach, J. Auto-regulatory feedback by RNA-binding proteins. *J. Mol. Cell Biol.* **11**, 930–939 (2019).
 60. Rossbach, O. *et al.* Auto- and cross-regulation of the hnRNP L proteins by alternative splicing. *Mol. Cell Biol.* **29**, 1442–1451 (2009).
 61. Quinodoz, S. A. *et al.* RNA promotes the formation of spatial compartments in the nucleus. *Cell* **184**,

- 5775–5790 (2021).
62. Larson, A. G. *et al.* Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
 63. Strom, A. R. *et al.* Phase separation drives heterochromatin domain formation. *Nature* **547**, 241–245 (2017).
 64. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
 65. Klock, H. E., Koesema, E. J., Knuth, M. W. & Lesley, S. A. Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins Struct. Funct. Genet.* **71**, 982–994 (2008).
 66. Wu, S. C. Y. *et al.* piggyBac is a flexible and highly active transposon as compared to Sleeping Beauty, Tol2, and Mos1 in mammalian cells. *Proc. Natl. Acad. Sci. USA* **103**, 15008–15013 (2006).
 67. Alberti, S. *et al.* A user’s guide for phase separation assays with purified proteins. *J. Mol. Biol.* **430**, 4806–4820 (2018).
 68. Wachsmuth, M. Molecular diffusion and binding analyzed with FRAP. *Protoplasma* **251**, 373–382 (2014).
 69. Koulouras, G. *et al.* EasyFRAP-web: a web-based tool for the analysis of fluorescence recovery after photobleaching data. *Nucleic Acids Res.* **46**, W467–W472 (2018).
 70. Banerjee, A. K. *et al.* SARS-CoV-2 disrupts splicing, translation, and protein trafficking to suppress host defenses. *Cell* **183**, 1325–1339 (2020).
 71. Van Nostrand, E. L. *et al.* Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**,

- 508–514 (2016).
72. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).

Chapter 3

**Investigating the mechanisms underlying
maintenance of X chromosome inactivation**

Mackenzie Strehle

3.1 | Abstract

X chromosome inactivation (XCI) is an essential developmental process mediated by the long noncoding RNA (lncRNA) Xist. Recent discoveries of the proteins that interact with Xist have enabled elucidation of the molecular mechanisms underlying initiation of XCI; however, the precise factors required for maintenance of XCI remain poorly understood. XCI maintenance has traditionally been characterized by two observations: (1) silencing occurs independently of Xist, and (2) DNA methylation is enriched on the inactive X chromosome (Xi), which provides an epigenetic system to propagate stable gene silencing. Still, the evidence in support of these views is unclear and, in some cases, contradicted by contemporary findings. Here, we show that maintenance of XCI is indeed Xist independent but impacted by perturbation of both DNA methylation and histone deacetylation. We also develop a fluorescent-based reporter system to investigate the mechanisms underlying the transition from initiation to maintenance of XCI. Together, this work has laid the foundation for further study of the factors required to establish and maintain X-linked gene silencing.

3.2 | Introduction

X chromosome inactivation (XCI) is a developmental process during which one of the two X chromosomes (X) in female mammals is silenced. XCI is critical in early development to balance the expression of X-linked genes between males and females, a phenomenon known as dosage compensation¹. The selection of a specific X for silencing during development is random; however, once a single chromosome is silenced, it remains inactivated for the entire lifetime of the organism.

XCI is mediated by the long non-coding RNA (lncRNA) Xist, which is transcribed exclusively from the future inactive X chromosome (Xi). At the onset of XCI, Xist coats the X from which it is transcribed by directly associating with the DNA/RNA binding protein scaffold attachment factor A (SAF-A)^{2,3}. Additionally, Xist interacts with the lamin-B receptor (LBR) to tether the Xi to the nuclear lamina, leading to chromosomal compaction by restraining the chromosome in three-dimensional (3D) space^{2,4}.

Xist also recruits a variety of protein effectors to the future Xi to facilitate stable, chromosome-wide silencing. For instance, Xist directly interacts with the SMRT/HDAC-associated repressor protein (SHARP)^{2,5,6}, an RNA binding protein (RBP) that mediates histone deacetylation and subsequent transcriptional repression through its interaction with the SMRT/HDAC3 co-repressor complex^{2,7,8}. Perturbation of either SHARP or histone deacetylation profoundly disrupts gene silencing on the X^{2,8-11}, underscoring the potency of this transcriptional repressor in mediating XCI. Furthermore, Xist indirectly recruits the Polycomb repressive complexes 1 and 2 (PRC1, PRC2) to the Xi to establish repressive histone modifications across the chromosome^{12,13}.

Although the factors required for the initiation of XCI, such as LBR and SHARP, have been studied extensively, the mechanism underlying maintenance of XCI remains unclear, especially the extent to which Xist is required for stable gene silencing. Since the turn of the twentieth century, it has been proposed that Xist is only necessary for transcriptional repression at the onset of XCI – if Xist is removed during this time, silencing of the Xi is reversed. However, if Xist is removed after this developmental window, silencing is not reversed and is instead maintained even in the absence of Xist^{14–16}. Accordingly, XCI is traditionally divided into two distinct phases: initiation (Xist-dependent) and maintenance (Xist-independent).

Still, continued expression of Xist and its association with the Xi across the life of the female mammal suggests a continued role for Xist in maintaining silencing, a view that has since been supported by several lines of evidence. For example, deletion of Xist in differentiated female mouse cells¹⁷ or in the brain of female mice leads to increases in Xi gene expression¹⁸. Genetic deletion of Xist or its promoter in differentiated female human cells also leads to complete reactivation of the Xi¹⁹.

Apart from Xist, it has been proposed that DNA methylation may be the critical factor for maintenance of XCI. DNA methylation is known to be associated with transcriptional repression, but its role in mediating stable gene silencing on the X is also unclear. Previous work has shown that the Xi is enriched for DNA methylation compared to the active X chromosome (Xa)^{20,21} and disruption of DNA methylation in various XCI models leads to reactivation of expression of certain X-linked genes^{22,23}. DNA methylation has also been hypothesized to underpin XCI maintenance as it is well understood how DNA methylation can be inherited across cell division. Specifically, de novo DNA methylation is

deposited by DNA methyltransferase 3 alpha and DNA methyltransferase 3 beta (DNMT3A/B), and is re-established after cell division by DNA methyltransferase 1 (DNMT1)^{24,25}.

Nevertheless, there is evidence that contradicts the proposed role of DNA methylation as the primary mechanism underlying maintenance of XCI. First, when somatic cell lines are demethylated with the DNA methylation inhibitor 5-azacytidine, there is no reactivation of genes on the X^{16,26}. Second, it has been demonstrated that only DNMT3B is required for deposition of DNA methylation across the Xi²⁷ – neither DNMT3A nor DNMT1 are required for X-linked DNA methylation^{27,28}. This observation raises questions regarding the proposed mechanism of DNA methylation inheritance on the Xi and suggests there may be other factors that contribute to memory of silencing across cell divisions. Finally, patients with immunodeficiency, centromeric region instability or facial anomalies (ICF) syndrome – a genetic disorder caused by loss of function of the DNMT3B gene – maintain normal inactivation of the Xi, despite exhibiting X-linked hypomethylation²⁹.

Much of the previous work investigating the role of Xist and DNA methylation in XCI maintenance was performed over two decades ago and is relatively narrow in scope due to limitations in available experimental approaches. For example, prior to the advent of CRISPR-based genetic engineering, creating genetically modified cell lines relied heavily on phenotype-based screening, and X-linked silencing assays were often performed based on secondary readouts, like drug sensitivity or fluorescence intensity, as opposed to directly measuring gene expression on the X. Here, we employ modern molecular biology approaches to explore the factors required for maintenance of XCI. Specifically, we investigate the role of Xist, DNA

methylation, and histone deacetylation in mediating stable gene silencing on the X. We also establish a reporter-based system that allows for more robust interrogation of silencing during the transition from initiation to maintenance of XCI.

3.3 | Results

3.3.1 | Maintenance of XCI is Xist-independent, but sensitive to epigenetic inhibition

To determine if XCI maintenance is indeed Xist-independent, we analyzed the effects of Xist perturbation in differentiated female mouse embryonic fibroblasts (MEFs), which have already undergone XCI. Specifically, we transfected hybrid (Spretus x Bl6) Patski MEFs with a scrambled (control) or Xist locked nucleic acid (LNA) gapmer. Antisense LNA gapmers enable efficient knockdown of nuclear-retained RNAs, including lncRNAs like Xist, by recruiting ribonuclease H (RNase H) and triggering RNA degradation³⁰. To measure knockdown efficiency and X silencing 48 h and 72 h after transfection, we performed RNA fluorescence in situ hybridization (RNA FISH) on Xist and the introns of: (1) *Atrx*, a gene that is silenced upon XCI, and (2) *Kdm5c*, a gene that escapes XCI and therefore remains active after Xist induction (see chapter 2 **Fig. 4a,b** for experimental design).

We observed that cells transfected with control LNA exhibited one (Xist score = 1; ~50% cells) or more (Xist score > 1; ~50% cells) Xist clouds – concentrated regions of Xist that demarcate the Xi (**Fig. 1a,c**). Although Patski MEFs are diploid and therefore should possess a single Xist cloud (i.e. a single Xi), it has previously been reported that diploid differentiated female mouse cells can exhibit two Xist clouds^{31–33}. The mechanism underlying this phenotype

is not well understood, but it has been posited that the presence of two Xist clouds may be a normal transient state that occurs after initiation of XCI as part of the choice process to determine which X will eventually be irreversibly silenced³⁴. We also observed that a majority of cells (> 90%) transfected with Xist LNA exhibited no Xist cloud (Xist = 0) or a reduced Xist cloud (Xist = 0.5; **Fig. 1b,c**), suggesting efficient Xist knockdown. To measure silencing on the X following control or Xist LNA transfection, we restricted our analyses to cells that retain both X chromosomes (by gating on two spots for the escape gene Kd5mc) and measured the number of Atrx spots within this subpopulation. We found that for both control and Xist LNA transfected cells, > 75% of cells exhibited one or fewer Atrx spots (**Fig. 1d and Extended Data Fig. 1a**), confirming that silencing in XCI maintenance is not disrupted by perturbation of Xist.

Next, we explored the role of epigenetic inhibition on XCI maintenance in a MEF cell line that contains: (1) loxP sites flanking Xist on the Xi, and (2) a GFP transgene on the Xi¹⁷. Specifically, we treated cells with DMSO (control), 5-azacytidine (5-aza; a DNA methylation inhibitor), or RGFP-966 (RGFP; an HDAC3 inhibitor) for 48 h and then analyzed GFP expression using fluorescence activated cell sorting (FACS). As an additional negative control, we transfected cells with a Cre expressing construct to remove Xist via Cre/lox recombination. Consistent with our previous observation, we found that removal of Xist via Cre/lox knockout did not result in an increase in GFP expression compared with untreated cells (< 0.1 foldchange; ~1% cells expressing GFP; **Fig. 1e and Extended Data Fig. 1b**). DMSO treatment led to a slight increase in GFP expression compared with untreated cells (0.3 foldchange; < 1% cells expressing GFP), while 5-aza and RGFP treatment both led to modest increases in GFP expression over untreated cells (3.7 and 2.3 foldchange, respectively; 19.7% and 2.4% cells

expressing GFP, respectively; **Fig. 1e and Extended Data Fig. 1b**). These data indicate that inhibition of DNA methylation and histone deacetylation both impact maintenance of silencing on the Xi.

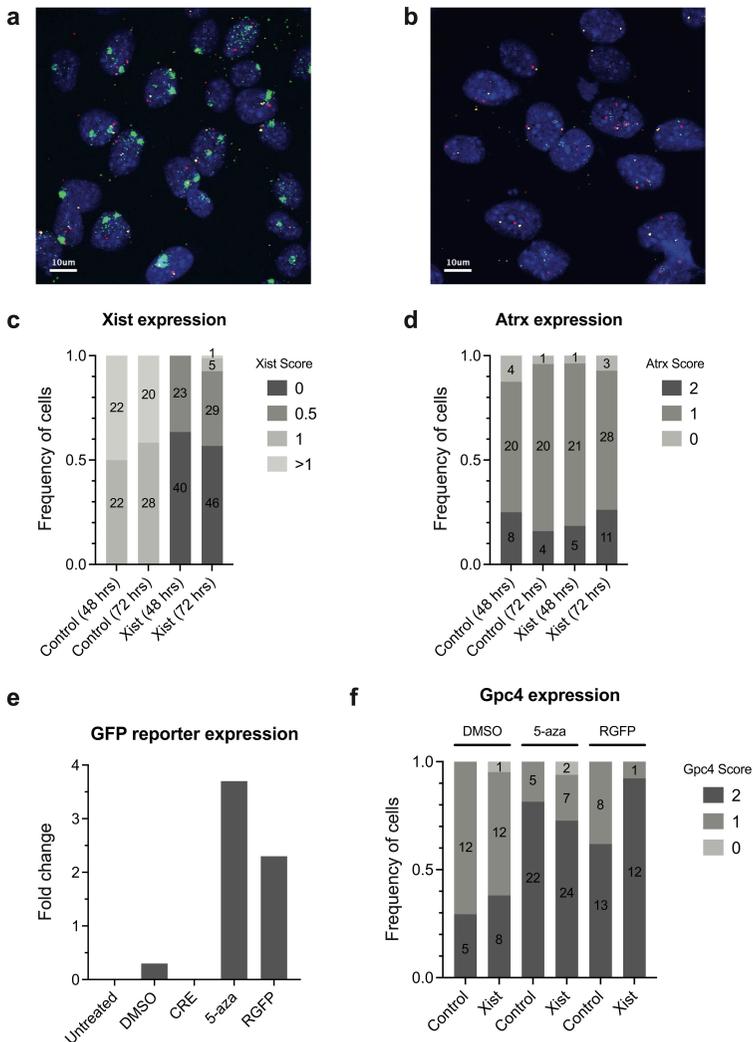


Figure 1: Effects of Xist perturbation and epigenetic inhibition on XCI maintenance. **a**, Representative RNA FISH image of Patski MEFs transfected with scrambled (control) LNA. Cells were stained for DAPI (blue) and probed for Xist (green), silenced gene *Atrx* (yellow), and escape gene *Kdm5c* (magenta). Images are shown as maximum projections. Scale bars, 10 μ m. **b**, Representative RNA FISH image of Patski MEFs transfected with Xist LNA. Cells were stained for DAPI (blue) and probed for Xist (green), silenced gene *Atrx* (yellow), and escape gene *Kdm5c* (magenta). Images are shown as maximum projections. Scale bars, 10 μ m. **c**, Quantification of RNA FISH images from **a,b** representing the frequency of cells with different Xist scores for scrambled and Xist LNA conditions. **d**, Quantification of RNA FISH images from **a,b** representing the frequency of cells containing two actively transcribed *Atrx* alleles for scrambled and Xist LNA conditions. Cells were gated on two escape gene (*Kdm5c*) spots. **e**, Fold change in GFP expression over untreated cells for DMSO (control), CRE (*Xist*), 5-azacytidine (5-mC), and RGFP-966 (HDAC3) treatments in GFP reporter MEFs¹⁷. **f**, Quantification of RNA FISH images representing the frequency of Patski MEFs containing two actively transcribed *Gpc4* alleles for scrambled and Xist LNA conditions combined with DMSO (control), 5-azacytidine (5-mC), or RGFP-966 (HDAC3) treatment. Cells were gated on two escape gene (*Kdm5c*) spots.

We next asked if Xist synergizes with these epigenetic modifications to mediate stable gene silencing during XCI maintenance. To test this hypothesis, we transfected Patski MEFs with a control or Xist LNA for 48 h then treated cells with DMSO, 5-aza, or RGFP for 24 h. To measure silencing on the X, we performed RNA FISH on Xist and the introns of *Gpc4* (silenced gene) and *Kdm6a* (escape gene), and we restricted our analyses to cells that contained two *Kdm6a* spots (i.e. two X chromosomes). We found that DMSO treatment led to a slight increase in gene expression on the X for both control and Xist LNA conditions (**Fig. 1f** and **Extended Data Fig. 1c**) compared with LNA transfection

alone (**Fig. 1d and Extended Data Fig. 1a**). We also observed that combining Xist knockdown with 5-aza treatment did not lead to an increase in X silencing compared with 5-aza treatment in the control LNA condition, whereas combining Xist knockdown with RGFP treatment led to a modest increase in gene expression on the X compared with RGFP treatment in the control LNA condition (**Fig. 1f and Extended Data Fig. 1c**).

Together, these results confirm previous reports that Xist is not essential for XCI maintenance, and indicate that both DNA methylation and histone deacetylation contribute to stable silencing on the X. Our data also suggest that XCI maintenance may depend on combinatorial effects of Xist and histone deacetylation.

3.3.2 | A fluorescent-based reporter system enables more efficient measurement of gene silencing during the transition to XCI maintenance

Previous studies investigating the role of various factors in XCI have typically relied on either (1) undifferentiated mouse embryonic stem cells (mESCs) that contain two active Xs and are competent to initiate XCI or (2) terminally differentiated cells, such as MEFs, that have already undergone gene silencing and transitioned to XCI maintenance. Limited work has focused on uncovering the temporal contributions of these factors across XCI, especially as they relate to establishment of XCI maintenance. Therefore, we wondered whether we could create a system that would allow us to better interrogate the molecular mechanisms that mediate the transition from initiation to maintenance of XCI (chapter 1 **Fig. 1a,b**). To do this, we utilized a female hybrid (B16 x Cast) mESC line that contains a doxycycline (dox)-inducible promoter at the

endogenous Xist gene on one of the X chromosomes⁸. Dox-induction of Xist in these cells is sufficient to initiate XCI, but it is unclear whether Xist expression in the absence of differentiation can also recapitulate the transition to XCI maintenance.

We reasoned that since Xist directly and indirectly recruits a variety of RBPs and chromatin modifying enzymes to the X^{2,5,6,12,13,35}, a long period of Xist expression may be sufficient to establish XCI maintenance in mESCs. To test this hypothesis, we induced Xist expression via dox-induction for 10 days and then continued to culture the cells in the absence of dox for 2, 4, or 10 days. To measure silencing on the X, we performed RNA FISH on Xist and the introns of several silenced genes (Atrx, Pgk1, Gpc4) and escape genes (Kdm5c, Kdm6a), and we restricted our analyses to cells that contained two escape gene spots (i.e. two X chromosomes). For the 10 day dox-induction condition without dox-withdrawal, we also gated on cells that expressed Xist. As expected, a low frequency of cells (< 0.2) expressed two silenced gene spots after 10 days of dox-induction compared with untreated (-dox) cells (> 0.5; **Fig. 2a**). Interestingly, for all three withdrawal conditions (-2 days, -4 days, -10 days), an intermediate frequency of cells expressed two silenced gene spots (~0.35, ~0.4, ~0.3, respectively; **Fig. 2a**). We interpret these intermediate silencing phenotypes to be the result of: (1) heterogenous dox-induction in mESCs, and (2) limitations of our RNA FISH silencing assay that do not allow us to exclude cells that did not express Xist prior to removal of dox (and therefore, did not undergo initiation of XCI).

We wondered whether a reporter-based system in the same mESCs could improve our silencing measurements by allowing us to exclude cells that do not express Xist following dox-induction. To make this reporter, we

integrated mCherry at the endogenous *Hprt* locus on the same *X* that includes a dox-inducible *Xist* promoter. In this way, we can use FACS to enrich for cells that express *Xist* after dox-induction by gating for low mCherry expression.

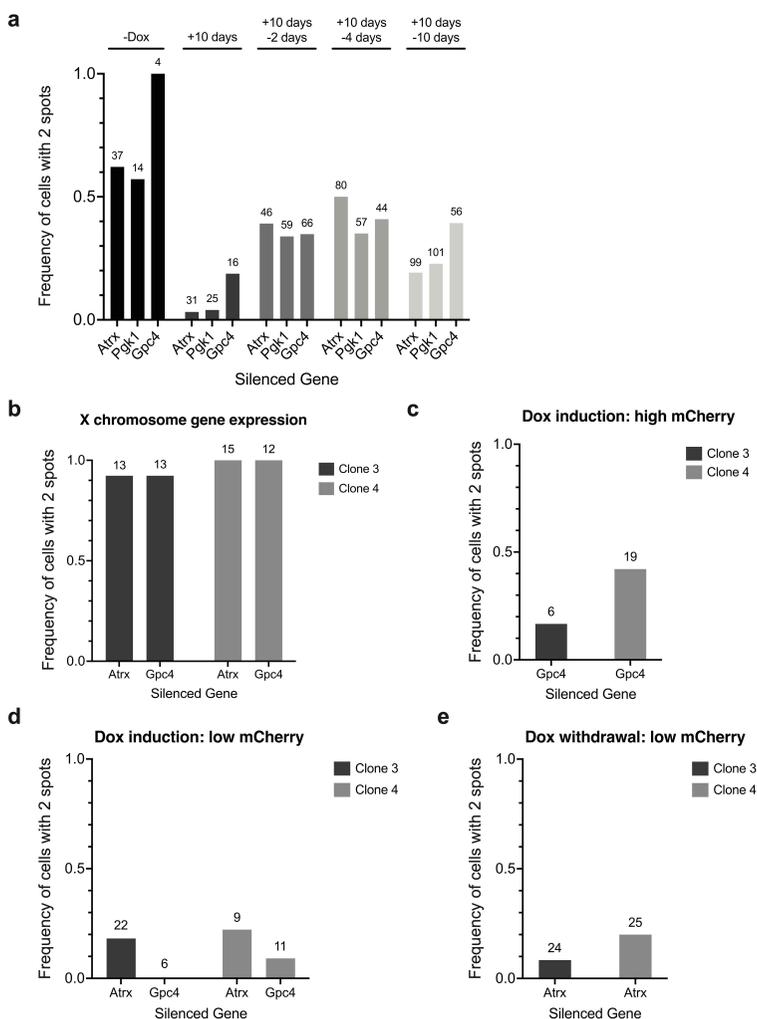


Figure 2: Establishing a reporter-based system for interrogating XCI maintenance. *a*, Quantification of RNA FISH images representing the frequency of cells containing two actively

transcribed alleles (Atrx, Pgk1, Gpc4) for various dox treatment conditions in TX1072 female mESCs. b, Quantification of RNA FISH images representing the frequency of cells containing two actively transcribed alleles (Atrx, Gpc4) in two untreated female Hpvt-mCherry reporter mESC clones. c, Quantification of RNA FISH images representing the frequency of cells containing two actively transcribed alleles (Gpc4) in two female Hpvt-mCherry reporter mESC clones treated with dox for 10 days and sorted for high mCherry expression. d, Quantification of RNA FISH images representing the frequency of cells containing two actively transcribed alleles (Atrx, Gpc4) in two female Hpvt-mCherry reporter mESC clones treated with dox for 10 days and sorted for low mCherry expression. e, Quantification of RNA FISH images representing the frequency of cells containing two actively transcribed alleles (Atrx) in two female Hpvt-mCherry reporter mESC clones treated with dox for 10 days, sorted for low mCherry expression, cultured in media without dox for 10 days, and sorted again for low mCherry expression.

After identifying two clones (clone 3, clone 4) that were responsive to dox-induction, we wanted to confirm their X chromosome ploidy and naïve state (i.e. pre-XCI) prior to performing silencing assays because introducing a clonal bottleneck can also select for undesirable phenotypes. To do this, we performed the same RNA FISH silencing assay described above, restricting our analyses to cells with two escape gene spots and counting the number of spots corresponding to Atrx or Gpc4 expression. We found that a high frequency of cells (> 0.9) in both clones expressed two silenced gene spots (**Fig. 2b**), indicating maintenance of their X ploidy and naïve state.

Next, we dox-induced both clones for 10 days and then performed FACS to sort for cells with high or low mCherry expression. We measured silencing in these two populations via RNA FISH and found that the frequency of cells expressing two silenced gene spots in the high mCherry

condition (0.17 – 0.42; **Fig. 2c**) was higher than the frequency of cells expressing two silenced gene spots in the low mCherry condition (0 – 0.22; **Fig. 2d**). We continued to culture the low mCherry cell population for an additional 10 days in the absence of dox and sorted these cells again for low mCherry expression. We observed that the frequency of cells expressing two silenced gene spots in this population (0.08 – 0.2; **Fig. 2e**) remained similar to the frequency observed prior to dox withdrawal (**Fig. 2d**). This result indicates that our mCherry reporter enables successful enrichment of cells that have undergone XCI and suggests that Xist expression by dox-induction in mESCs may be sufficient to establish maintenance of silencing on the X.

3.4 | Discussion

Our results demonstrate that maintenance of XCI is Xist-independent, but dependent on both DNA methylation and histone deacetylation. We showed that DNA methylation is important for silencing independent of Xist, while histone deacetylation acts synergistically with Xist to mediate maintenance of silencing. We also showed that dox-induction of Xist in mESCs may be sufficient to establish XCI maintenance, providing a novel system by which to interrogate the transition from initiation to maintenance of XCI.

Factors required for XCI maintenance

Our work confirms previous findings that maintenance of XCI is both Xist- and DNA-methylation dependent. However, our results are limited in scope by the low-throughput methods used to measure gene silencing. Specifically, our data rely on single gene measurements to

quantify X-linked silencing, which may not be an accurate reflection of chromosome-wide transcriptional states. Many of the studies we set out to replicate here were performed using similar low-throughput silencing assays and therefore possess the same limitations. As such, revisiting these experiments while utilizing new molecular biology approaches – namely high-throughput sequencing methods such as RNA sequencing (RNAseq) – will further clarify the role of Xist and DNA methylation in maintenance of XCI.

Furthermore, we investigated the effects of perturbation of three different factors – Xist, DNA methylation, and histone deacetylation – on maintenance of XCI, but numerous other proteins and epigenetic modifications may contribute to stable gene silencing on the X. For instance, it has been observed that deletion of the noncanonical Polycomb group RING finger 3/5 (PCGF3/5) protein, which forms a complex with PRC1 to deposit H2AK119ub1 across the chromosome, abrogates transcriptional repression on the X¹², suggesting a potential role in maintenance of XCI. Additionally, while PRC2 is not essential for gene silencing on the X, it has long been proposed that H3K27me3 deposited by PRC2 may synergize with other silencing factors and epigenetic marks on the X to mediate stable gene silencing^{36,37}. Continued investigation of the role of PRC1 and PRC2, as well as the myriad other factors that are known to be recruited to the Xi, in gene silencing on the X is necessary to fully elucidate the molecular mechanisms underlying transcriptional repression during XCI maintenance.

Transition from initiation to maintenance of XCI

In addition to studying the effects of various transcriptional repressors and chromatin modifiers on maintenance of silencing, further work is needed to describe their temporal

requirements and relationships during the transition from initiation to maintenance of XCI. Specifically, combinatorial perturbation of these factors after the initiation of silencing, either in the presence or absence of Xist, may provide a better understanding of the molecular synergies required to mediate lifelong transcriptional repression on the X.

To interrogate the transition to XCI maintenance after induction of Xist expression, we developed a novel fluorescence-based mESC model that enables enrichment of cells that have successfully initiated XCI. While additional work is required to confirm our results, including chromosome-wide RNAseq measurements of gene silencing in the two clones generated, preliminary data presented here suggests that mESCs may be competent to recapitulate silencing maintenance in the absence of differentiation. Using such a system, perturbation of silencing factors can be performed at well-defined points during XCI, and these perturbations can also be combined with Xist abrogation via dox-withdrawal. Collectively, this highly tunable model may help uncover the molecular mechanisms underlying maintenance of XCI.

3.5 | Methods

Cell culture

Mouse embryonic fibroblast (MEF) cell culture

Patski MEFs and GFP reporter MEFs¹⁷ were cultured on uncoated plates in serum-containing MEF cell media (high glucose DMEM (Gibco, Life Technologies), 10% FBS (Seradigm), 2 mM L-glutamine (Gibco, Life Technologies), 1 mM sodium pyruvate (Gibco, Life Technologies), 1x MEM non-essential amino acids (Gibco, Life Technologies), 0.1 mM β -mercaptoethanol).

Mouse embryonic stem cell (mESC) culture

TX1072 female mESCs (gift from E. Heard laboratory) were cultured as previously described⁸. Briefly, TX1072 mESCs were grown on gelatin-coated plates in serum-containing ES cell media (high glucose DMEM (Gibco, Life Technologies), 15% FBS (Omega Scientific), 2 mM L-glutamine (Gibco, Life Technologies), 1 mM sodium pyruvate (Gibco, Life Technologies), 0.1 mM β -mercaptoethanol, 1000 U/ml leukemia inhibitory factor (Chemicon)), and 2i (3 μ M Gsk3 inhibitor CT-99021, 1 μ M MEK inhibitor PD0325901). The cell culture media was replaced every 24 h.

Cell treatments and transfections

Drug treatments

To induce expression of Xist in female mESCs, dox (Sigma) was added to cell media at a final concentration of 2 μ g/mL. Dox-containing media was replaced every 24 h.

For epigenetic drug inhibition treatments, Patski or GFP reporter MEFs were plated in a six-well plate and grown to 60-70% confluency. 5-azacytidine and RGFP-966 were resuspended in DMSO and added to cells at a final concentration of 10 μ M. An identical volume of DMSO was added to cells for the control condition. Cells were grown for 24 or 48 h before being processed for downstream experiments (RNA FISH, FACS).

Locked nucleic acid (LNA) transfection

Patski MEFs plated in a six-well plate were grown to 60% confluency and transfected with scrambled (control) or Xist LNA using RNAiMAX (Invitrogen) according to the manufacturer's protocol. Briefly, 9 μ L of RNAiMAX reagent was diluted in 150 μ L of Opti-MEM media (Thermo Fisher Scientific) and 30 pmol of LNA was diluted in 150 μ L of Opti-MEM. The diluted LNA sample was added to the diluted RNAiMAX sample, and this mixture was incubated at room temperature for five min. After incubation, 250 μ L of the LNA/RNAiMAX mixture was added to one well of the cell culture plate. Cells were then cultured for 48 or 72 h before performing downstream experiments (drug treatment, RNA FISH).

Plasmid transfection

To test the effects of Xist perturbation on maintenance of silencing, GFP reporter MEFs plated in a six-well plate were grown to 70% confluency and transfected with a EF1a-driven Cre plasmid using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. Briefly, 10 μ L of Lipofectamine 2000 reagent was diluted in 100 μ L of Opti-MEM media and 1.5 μ g of Cre plasmid was diluted in 100 μ L of Opti-MEM media. The diluted plasmid sample was added to the diluted Lipofectamine 2000 sample, and this mixture was incubated at room temperature for five min.

After incubation, the entire volume of the plasmid/Lipofectamine 2000 mixture was added to one well of the cell culture plate. Cells were then cultured for 48 h before being analyzed by FACS.

To create Hprt-mCherry reporter cell lines, we utilized the CRISPR/Cas9-mediated Homology-independent PCR-product integration (CHoP-In) system³⁸. A guide RNA (gRNA; tacaactaatcattatgccg) targeting the N-terminus of Hprt was inserted into SapI-digested pZB-Sg3 plasmid³⁹. The mCherry insertion DNA was generated by performing polymerase chain reaction (PCR) on an mCherry plasmid with primers containing the Cas9 recognition sequence of the target locus flanking the insert. To generate enough PCR product for transfection, 10 identical reactions were set up, treated with DpnI to remove any remaining plasmid template, and then purified using the DNA Clean & Concentrator kit (Zymo). Two million TX1072 mESCs were then transfected with 1.2 picomoles each of the mCherry PCR product, the PX459 Cas9 plasmid containing a puromycin resistance selection marker and the Hprt-gRNA plasmid using the Neon transfection system (settings: 1400 V, 10 ms width, three pulses).

48 h after transfection, puromycin (puro) was added to the culture media at a final concentration of 1 $\mu\text{g}/\text{mL}$ to enrich for successfully transfected cells. The surviving cells were sorted for mCherry expression after 72 h of puro selection and subsequently plated at low-confluency. After 4-5 days of growth, 24 mCherry-positive colonies were picked and seeded in a 96-well plate. These cells were then split into one plate for FACS-based genotyping and another plate for maintaining growth until positive clones were identified. FACS-based genotyping was performed by identifying clones that silenced mCherry after 48 h of dox induction. We assumed that clones that were positive for mCherry in the

absence of dox but negative for mCherry in the presence of dox contained a single mCherry integration at the Hprt locus on the same X that includes the dox-inducible Xist promoter. Clones 3 and 4 were used for subsequent experimentation and all other clones were frozen.

Single molecule RNA fluorescence *in situ* hybridization (RNA FISH)

RNA FISH experiments were performed using the ViewRNA ISH Cell Assay (ThermoFisher, catalog no. QVC0001) protocol with minor modifications. Specifically, cells were fixed on coverslips with 4% formaldehyde in PBS for 15 min at room temperature and then permeabilized with 4% formaldehyde and 0.5% Triton X-100 in PBS for 10 min at room temperature. Cells were then washed twice with PBS, dehydrated with 70% ethanol and incubated at -20°C for at least 20 min or stored for up to 1 week. Coverslips were washed twice with PBS and then incubated with the desired combination of RNA FISH probes (see chapter 2 **Extended Data Fig. 4a and Supplementary Table 3**; Affymetrix) in Probe Set Diluent at 40°C for at least 3 h. Coverslips were then washed once with wash buffer, twice with PBS, and once more with wash buffer before incubating in preamplifier mix solution at 40°C for 45 min. This step was repeated for the amplifier mix solution and label probe solution. Coverslips were incubated with 1x DAPI in PBS at room temperature for 15 min and subsequently mounted onto glass slides using ProLong Gold with DAPI (Invitrogen, P36935).

Image acquisition and quantification

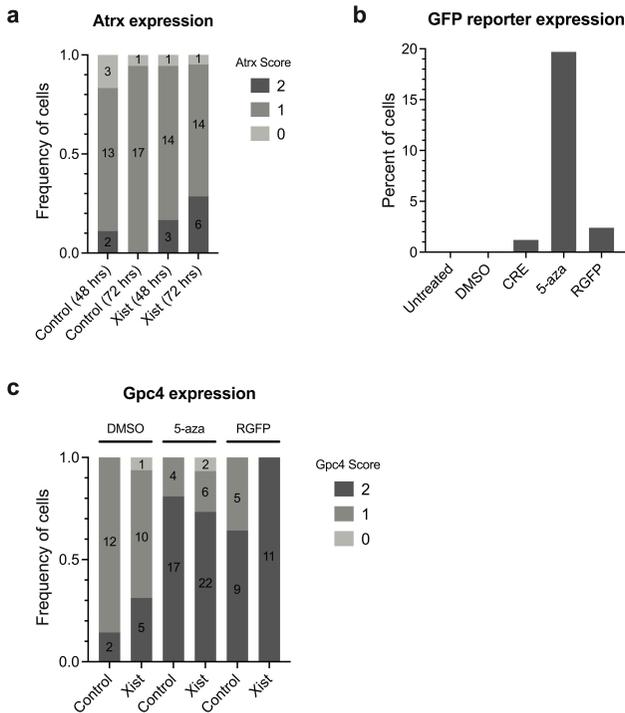
Microscopy

Fixed samples were imaged using the Zeiss LSM 800 with the 63x oil objective and collected every 0.3 μm for 16 Z-stacks. For all images, laser power and gain were set at the beginning of acquisition and remained constant throughout the duration of acquisition.

RNA FISH analysis

Image analysis was performed using FIJI (ImageJ v2.1.0/1.53c) software. Each image was processed into a maximum intensity projection using FIJI software. Then, the number of spots corresponding to each intron FISH probe per nucleus was manually counted and scored for the presence of Xist signal, number of spots per escape gene (Kdm5c, Kdm6a), and number of spots per silenced gene (Atrx, Pgc1, Gpc4) (see chapter 2 **Fig. 4a**). Because mESCs are known to lose one of the X chromosomes or its fragments while in culture⁴⁰⁻⁴² (see chapter 2 **Extended Data Fig. 4b**), the analysis was restricted to cells containing two X chromosomes, which were determined by the presence of exactly two spots from the probed escape gene. In addition, cells that had more than two spots per nucleus for any gene were excluded from the analysis.

3.6 | Extended Data



Extended Data Figure 1: Effects of Xist perturbation and epigenetic inhibition on XCI maintenance. **a**, Quantification of RNA FISH images from **Fig. 1a,b** representing the frequency of cells containing two actively transcribed *Atrx* alleles for scrambled and *Xist* LNA conditions. Cells were gated on two escape gene (*Kdm5c*) spots and one (control LNA) or zero (*Xist* LNA) *Xist* clouds. **b**, Percent of GFP reporter MEFs¹⁷ expressing GFP for DMSO (control), CRE (*Xist*), 5-azacytidine (5-mC), and RGFP-966 (HDAC3) treatments. **c**, Quantification of RNA FISH images representing the frequency of Patski MEFs containing two actively transcribed *Gpc4* alleles for scrambled and *Xist* LNA conditions combined with DMSO (control), 5-azacytidine (5-mC), or RGFP-966 (HDAC3) treatment. Cells were gated on two escape gene (*Kdm5c*) spots and one (control LNA) or zero (*Xist* LNA) *Xist* clouds.

3.7 | References

1. Brockdorff, N. & Turner, B. M. Dosage compensation in mammals. *Cold Spring Harb. Perspect. Biol.* **7**, (2015).
2. McHugh, C. A. *et al.* The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232–236 (2015).
3. Hasegawa, Y. *et al.* The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev. Cell* **19**, 469–476 (2010).
4. Chen, C. K. *et al.* Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* **354**, 468–472 (2016).
5. Chu, C. *et al.* Systematic Discovery of Xist RNA Binding Proteins. *Cell* **161**, 404–416 (2015).
6. Lu, Z. *et al.* RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. *Cell* **165**, 1267–1279 (2016).
7. Wutz, A., Rasmussen, T. P. & Jaenisch, R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* **30**, 167–174 (2002).
8. Dossin, F. *et al.* SPEN integrates transcriptional and epigenetic control of X-inactivation. *Nature* **578**, 455–460 (2020).
9. Żylicz, J. J. *et al.* The Implication of Early Chromatin Changes in X Chromosome Inactivation. *Cell* **176**, 182–197 (2019).
10. Markaki, Y. *et al.* Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **184**, 6174–6192.e32 (2021).
11. Jachowicz, J. W. *et al.* Xist spatially amplifies SHARP/SPEN recruitment to balance chromosome-wide silencing and specificity to the X chromosome. *Nat. Struct. Mol. Biol.* **2022** 293 **29**, 239–249 (2022).

12. Almeida, M. *et al.* PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. *Science* **356**, 1081–1084 (2017).
13. Pintacuda, G. *et al.* hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing. *Mol. Cell* **68**, 955-969.e10 (2017).
14. Wutz, A. & Jaenisch, R. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell* **5**, 695–705 (2000).
15. Csankovszki, G., Panning, B., Bates, B., Pehrson, J. R. & Jaenisch, R. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. *Nature Genetics* **22**, 323–324 (1999).
16. Brown, C. J. & Willard, H. F. The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. *Nature* **368**, 154–156 (1994).
17. Csankovszki, G., Nagy, A. & Jaenisch, R. Synergism of Xist Rna, DNA Methylation, and Histone Hypoacetylation in Maintaining X Chromosome Inactivation. *J. Cell Biol.* **153**, 773–784 (2001).
18. Adrianse, R. L. *et al.* Perturbed maintenance of transcriptional repression on the inactive X-chromosome in the mouse brain after Xist deletion. (2018). doi:10.1186/s13072-018-0219-8
19. Lee, H. J. *et al.* En bloc and segmental deletions of human XIST reveal X chromosome inactivation-involving RNA elements. *Nucleic Acids Res.* **47**, 3875–3887 (2019).
20. Norris, D. P., Brockdorff, N. & Rastan, S. Methylation status of CpG-rich islands on active and inactive mouse X chromosomes. *Mamm. Genome* **1**, 78–83 (1991).
21. Hellman, A. & Chess, A. Gene Body-Specific

- Methylation on the Active X Chromosome. (2007). doi:10.1126/science.1136352
22. Sado, T. *et al.* X inactivation in the mouse embryo deficient for Dnmt1: Distinct effect of hypomethylation on imprinted and random X inactivation. *Dev. Biol.* **225**, 294–303 (2000).
 23. Marshall Graves, J. A. 5-Azacytidine-induced re-expression of alleles on the inactive X chromosome in a hybrid mouse cell line. *Exp. Cell Res.* **141**, 99–105 (1982).
 24. Hermann, A., Gowher, H. & Jeltsch, A. Biochemistry and biology of mammalian DNA methyltransferases. *Cellular and Molecular Life Sciences* **61**, 2571–2587 (2004).
 25. Li, E. & Zhang, Y. DNA methylation in mammals. *Cold Spring Harb. Perspect. Biol.* **6**, a019133 (2014).
 26. Wolf, S. F. & Migeon, B. R. Studies of X chromosome DNA methylation in normal human cells. *Nature* **295**, 667–671 (1982).
 27. Gendrel, A. V. *et al.* Smchd1-Dependent and-Independent Pathways Determine Developmental Dynamics of CpG Island Methylation on the Inactive X Chromosome. *Dev. Cell* **23**, 265–279 (2012).
 28. Wiseman, A. K. *et al.* Chromosome-specific retention of cancer-associated DNA hypermethylation following pharmacological inhibition of DNMT1. *Commun. Biol.* **2022 51** **5**, 1–10 (2022).
 29. Hansen, R. S. *et al.* Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant. *Hum. Mol. Genet.* **9**, 2575–2587 (2000).
 30. Nishikawa, M. & Yanagawa, N. Knockdown of nuclear Incrnas by locked nucleic acid (LNA) gapmers in nephron progenitor cells. *Methods Mol.*

- Biol.* **2161**, 29–36 (2020).
31. Monkhorst, K., Jonkers, I., Rentmeester, E., Grosveld, F. & Gribnau, J. X Inactivation Counting and Choice Is a Stochastic Process: Evidence for Involvement of an X-Linked Activator. *Cell* **132**, 410–421 (2008).
 32. Shiura, H. & Abe, K. Xist/Tsix expression dynamics during mouse peri-implantation development revealed by whole-mount 3D RNA-FISH. *Sci. Reports 2019 91* **9**, 1–11 (2019).
 33. Pacini, G. *et al.* Integrated analysis of Xist upregulation and X-chromosome inactivation with single-cell and single-allele resolution. *Nat. Commun.* **12**, 1–17 (2021).
 34. Sousa, E. J. *et al.* Exit from Naive Pluripotency Induces a Transient X Chromosome Inactivation-like State in Males. *Cell Stem Cell* **22**, 919-928.e6 (2018).
 35. Jansz, N. *et al.* Smchd1 Targeting to the Inactive X Is Dependent on the Xist-HnrnpK-PRC1 Pathway. *Cell Rep.* **25**, 1912-1923.e9 (2018).
 36. Nesterova, T. B. *et al.* Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. *Nat. Commun.* **10**, 3129 (2019).
 37. Dixon-Mcdougall, T. & Brown, C. The making of a Barr body: the mosaic of factors that eXIST on the mammalian inactive X chromosome 1. (2016). doi:10.1139/bcb-2015-0016
 38. Manna, P. T., Davis, L. J. & Robinson, M. S. Fast and cloning-free CRISPR/Cas9-mediated genomic editing in mammalian cells. *Traffic* **20**, 974–982 (2019).
 39. Engreitz, J. M. *et al.* Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).

40. Schulz, E. G. *et al.* The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell* **14**, 203–216 (2014).
41. Choi, J. *et al.* Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature* **548**, 219–223 (2017).
42. Yagi, M. *et al.* Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. *Nature* **548**, 224–227 (2017).

Chapter 4

Conclusion and future directions

Mackenzie Strehle and Mitchell Guttman

A modified portion of this chapter was published in: Strehle, M. & Guttman, M. Xist drives spatial compartmentalization of DNA and protein to orchestrate initiation and maintenance of X inactivation. *Curr. Opin. Cell. Biol.* **64**, 139–147 (2020). <https://doi.org/10.1016/j.ceb.2020.04.009>

4.1 | Conclusion and future directions

The recent identification of proteins that interact with Xist has provided a framework for dissecting many of the molecular mechanisms of Xist function and XCI in greater detail. Here, we uncovered how the transcriptional repressor SHARP is recruited to the X in super-stoichiometric excess relative to Xist to mediate gene silencing during the initiation of XCI. We also revealed why low Xist copy number in individual cells is important to mitigate off-target targeting of Xist and how SHARP recruitment to the X contributes to maintaining low levels of Xist expression. Together, these results provide a unified model that describes how Xist silences the X, the whole X, and nothing but the X.

Additionally, we began to explore the role of various factors in mediating maintenance of XCI. Specifically, we showed that XCI maintenance is Xist independent, but dependent on two different epigenetic modifications – DNA methylation and histone deacetylation. Further, we developed a novel mESC-based model by which to better interrogate the factors required during the transition from initiation to maintenance of XCI. Collectively, these experiments have laid the groundwork for further investigation into the mechanism of stable gene silencing on the X.

Still, there are many open questions related to XCI that need to be addressed. For example, though IDR-dependent recruitment of SHARP to the X points to a potential LLPS mechanism underlying the initiation of gene silencing, this hypothesis remains to be functionally tested. Specifically, the biophysical and biochemical properties of the Xi should be studied in further detail to better understand the nature of this silent compartment. Furthermore, it remains unclear how this compartment is maintained after initiation of XCI and what, if any, other molecular components are required

for heritable transcriptional repression. We expect that new molecular and cell biology tools will enable exploration of these hypotheses and provide other critical insights into this essential developmental process. Answering these outstanding questions will not only fill a fundamental gap in our understanding of XCI, but it may also provide new insights into broader aspects of epigenetic silencing beyond the X chromosome.