Published in final edited form as:

Cell. 2021 December 09; 184(25): 6174–6192.e32. doi:10.1016/j.cell.2021.10.022.

Xist nucleates local protein gradients to propagate silencing across the X chromosome

Yolanda Markaki[1,*], **Johnny Gan Chong**[1], **Yuying Wang**[1,7], **Elsie C. Jacobson**[1,7], **Christy Luong**[1,7], **Shawn Y.X. Tan**[1,7], **Davide Maestrini**[2], **Abhik Banerjee**[3], **Bhaven A. Mistry**[2,4], **Iris Dror**[1], **Francois Dossin**[5], **Johannes Schöneberg**[6], **Edith Heard**[5], **Mitchell Guttman**[3], **Tom Chou**[2,*], **Kathrin Plath**[1,8,*]

[1.] Department of Biological Chemistry, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA

[2.] Departments of Computational Medicine and Mathematics, University of California Los Angeles, Los Angeles, CA 90095, USA

[3.] Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[4.] Claremont McKenna College, Claremont, CA 91711, USA

[5.] European Molecular Biology Laboratory, Director's Unit, Heidelberg 69117, Germany

[6.] Departments of Pharmacology & Chemistry and Biochemistry, University of California San Diego, San Diego, CA 92093, USA

[7.] Equal contribution

[8.] Lead contact

Summary

The lncRNA *Xist* forms ~50 diffraction-limited foci to transcriptionally silence one X-chromosome. How this small number of RNA foci and interacting proteins regulate a much larger number of X-linked genes is unknown. We show that *Xist* foci are locally confined, contain ~2 RNA molecules, and nucleate supramolecular complexes (SMACs) that include many copies of the critical silencing protein SPEN. Aggregation and exchange of SMAC proteins generate local protein gradients that regulate broad, proximal chromatin regions. Partitioning of numerous

[*] Correspondence: kplath@mednet.ucla.edu, tomchou@ucla.edu and gmarkaki@mednet.ucla.edu.

Author Contributions

Conceptualization, Y.M. and K.P.; Methodology, Y.M. and T.C.; Software, J.G.C., Y.M., Y.W., E.J., D.M., I.D., B.A.M. and J.S.; Validation, Y.M.; Formal Analysis, Y.M., J.G.C., Y.W., EJ., D.M., B.A.M, I.D. and C.L.; Investigation, Y.M., J.G.C., Y.W., C.L., S.T., K.P. A.B. and F.D.; Data Curation, Y.M., J.G.C., Y.W. and E.J.; Writing-Original Draft, Y.M., K.P. and T.C.; Writing–Review & Editing, Y.M., K.P. and T.C.; Visualization, Y.M., J.C.G., Y.W. and E.J.; Supervision, K.P., Y.M., T.C., M.G. and E.H.; Project Administration, Y.M. and K.P.; Funding Acquisition, K.P., Y.M. and T.C.

Declaration of Interests

K.P. is a member of Cell's advisory board. We have a patent pending related to this work.

SPEN molecules into SMACs is mediated by their intrinsically disordered regions and essential for transcriptional repression. Polycomb deposition via SMACs induces chromatin compaction and the increase in SMAC density around genes, which propagates silencing across the X chromosome. Our findings introduce a mechanism for functional nuclear compartmentalization whereby crowding of transcriptional and architectural regulators enables the silencing of many target genes by few RNA molecules.

Introduction

Mammalian genomes encode thousands of long non-coding (lnc) RNAs, many of which play key roles in regulating gene expression by localizing effector proteins to genomic targets (Engreitz et al., 2016; Rinn and Chang, 2012; Statello et al., 2021). There has been considerable debate about how lncRNAs can robustly regulate gene expression, given that they are often expressed at low levels (Cabili et al., 2015; Derrien et al., 2012). One such example is the *Xist* lncRNA, which silences transcription of a large number of genes across an entire chromosome.

Xist is transcribed from, coats, and silences one of the two X chromosomes during development of female placental mammals in a process referred to as X chromosome inactivation (XCI) (Brockdorff et al., 2020; Galupa and Heard, 2018; Jegu et al., 2017; Plath et al., 2002; Wutz, 2011). *Xist* initiates gene silencing, large-scale chromatin remodeling, and formation of a unique nuclear compartment, the inactive X chromosome (Xi), through the recruitment of chromatin-modifying proteins, transcriptional silencers, and other RNA binding proteins (Chaumeil et al., 2006; Chu et al., 2015; Giorgetti et al., 2016; McHugh et al., 2015; Minajigi et al., 2015; Wang et al., 2018). The protein SPEN is essential for initiating the silencing of virtually all X-linked genes (Chu et al., 2015; Dossin et al., 2020; McHugh et al., 2015; Moindrot et al., 2015; Monfort et al., 2015). However, a subset of X-linked genes also requires other *Xist*-interactors for silencing, including non-canonical PRC1-type Polycomb group protein complexes and the architectural protein structural-maintenance of chromosomes hinge domain containing 1 (SMCHD1) (Almeida et al., 2017; Blewitt et al., 2008; Jansz et al., 2018; Nesterova et al., 2019; Pintacuda et al., 2017; Wang et al., 2019). Moreover, some X-linked genes become repressed sooner than others (Barros de Andrade et al., 2019). Why X-linked genes differ in their silencing dynamics and require multiple repressive pathways for inactivation remain major questions.

Based on conventional fluorescence microscopy and genomic methods that measure an ensemble of millions of cells at high resolution, *Xist*, its effector proteins, and Xi chromatin modifications appear to accumulate along the entire chromosome (Clemson et al., 1996; Engreitz et al., 2013; Plath et al., 2003; Silva et al., 2003; Simon et al., 2013; Zylicz et al., 2019). These observations have led to a model in which *Xist* and its interacting proteins form ribonucleoprotein complexes that are distributed across all X-linked genes to control gene expression. However, super-resolution microscopy has shown that *Xist* distributes as only 50 to 150 diffraction-limited foci on the Xi in differentiated cells (Cerase et al., 2014; Markaki et al., 2012; Smeets et al., 2014; Sunwoo et al., 2015). Therefore, these foci cannot regulate gene expression through simultaneous accumulation at each target gene.

Thus, even though we know the effector proteins of *Xist*, the mechanism by which RNA foci exploit these proteins to induce silencing of ~1000 genes distributed over 167 million base pairs remains unknown. Here, we addressed this fundamental problem by developing quantitative super-resolution microscopy approaches to interrogate the stoichiometry and spatial relationship of *Xist* to its effector proteins and target genes during the initiation of XCI. We also performed kinetic measurements and single-particle tracking to explore the dynamics and mobility of *Xist* and associated proteins.

We discovered that *Xist* foci are locally confined and that they induce the *de novo* formation of local protein compartments that encompass *Xist*-interactors at concentrations exceeding those of the RNA. We refer to these compartments as supramolecular complexes (SMACs). SMACs are dynamic structures formed by transient protein interactions around a slowly exchanging *Xist* core. The rapid binding and dissociation of most *Xist*-interacting proteins in SMACs creates local protein concentration gradients that mediate gene silencing on the entire chromosome. We show that the intrinsically disordered regions of SPEN are essential for its integration into SMACs and for gene silencing, and that Polycomb-mediated chromatin reconfiguration propagates silencing across the X chromosome. In summary, our work reveals that protein crowding enables a limited number of locally confined seeding RNA molecules to silence a much larger number of target genes.

Results

Progressive gene silencing during XCI is associated with chromatin compaction

To explore how *Xist* orchestrates the formation of the Xi, we differentiated female mouse embryonic stem cells (ESCs) to epiblast-like cells (EpiLCs), which leads to the induction of *Xist* expression and initiation of XCI (Figures 1A and S1A). Gene silencing occurs predominantly during the transition from day 2 (D2) to D4, as shown by nascent transcript detection of rapidly (*Rlim*) and slowly (*AtrX* and *Mecp2*) silencing genes (Figures 1B, 1C, and S1B). Single cell (sc) RNA-seq analyses extended this result to all X-linked genes (Figure 1D and S1C). These data confirm that *Xist* coating and gene silencing are stepwise processes (Chaumeil et al., 2006) and establish the D2 to D4 transition as a critical window for dissecting the relationship between *Xist*, its interacting proteins and gradual gene silencing. Henceforth, we refer to the D2 X chromosome as the “pre-Xi” and the D4 X chromosome as the “Xi”.

Although architectural differences between the active X chromosome (Xa) and Xi are well known (Darrow et al., 2016; Giorgetti et al., 2016; Teller et al., 2011; Wang et al., 2019), it remains unclear when they arise during XCI. Volume and sphericity measurements upon X-painting showed that the pre-Xi is similar to the Xa and that the Xi at D4 is as compact and spherical as in somatic cells (Figures 1E and 1F) (Teller et al., 2011). Accordingly, assessing the conformation of seven loci on the X through DNA FISH revealed a moderate change of the higher-order configuration in the pre-Xi compared to the Xa and a dramatic difference between the Xa and Xi (Figures 1G–I and S1D–F). Thus, gene silencing is associated with major changes in higher-order chromatin structure and both processes need to be considered to understand the mechanism through which *Xist* foci form the Xi.

~50 *Xist* foci of 2 transcripts induce XCI

As the number of Xist foci during the initiation of XCI is unknown, we quantified them during the D2 to D4 transition by super-resolution three-dimensional Structured Illumination Microscopy (3D-SIM). To this end, we generated a female mouse ESC line that allows for Xist detection in fixed as well as living cells. Specifically, exploiting the MS2 RNA-MS2 Coat Protein (MCP) interaction (Bertrand et al., 1998), we tagged the Xist gene on one of the two X chromosomes with 24 MS2-repeats. We then expressed MCP-GFP to label *Xist* with GFP and confirmed the functionality of the Xist$^{MS2\text{-}GFP}$ allele (Figures 1J, 1K, S1G and S1H). Quantitative 3D-SIM analysis of *Xist*$^{MS2\text{-}GFP}$ signals showed that the *Xist* territory consists of, on average, 74 diffraction-limited foci on the pre-Xi and 60 on the Xi (Figure 1L), which we corroborated by RNA FISH (Figures S1H and S1I). We also found that the doubling of the X chromosome with DNA replication is accompanied by the doubling of *Xist* foci from ~50 in G1 to ~100 in G2 and that the number of foci correlates with chromosome length (Figures 1M and S2A–D). Thus, the variability in the number of *Xist* foci is largely due to differences in cell cycle across the cell population. *Xist* foci maintain their integrated density and volume, which is consistent with the constitutive transcription of the *Xist* locus, suggesting that RNA levels on the Xi are stable throughout XCI (Figures 1N, S1I, S2E and S2F). Taken together, these data reveal that XCI is induced by only ~50 *Xist* foci and that the pre-Xi to Xi transition occurs without a dramatic change in their number.

To estimate the number of *Xist* molecules per focus, we transiently expressed nanocages consisting of 60 GFP molecules (cage^{60GFP}) (Hsia et al., 2016) as internal fluorescence standards in *Xist*$^{MS2\text{-}GFP}$ cells and confirmed similar intensity profiles of cage^{60GFP} in the nucleus and cytoplasm (Figures S2G and S2H). The integrated density of one *Xist*$^{MS2\text{-}GFP}$ focus on the pre-Xi and Xi corresponds to that of one cage^{60GFP} (Figures 1O, S2I and S2J). Since ~30 MCP-GFP molecules bind to 24 MS2-repeats (Wu et al., 2012), we infer that each focus contains ~2 *Xist* molecules. This result is consistent with measured levels of *Xist* RNA in single differentiating ESCs (Pacini et al., 2021) and estimated numbers of *Xist* molecules in differentiated cells (Sunwoo et al., 2015). Thus, only ~100 *Xist* molecules orchestrate the initiation and maintenance of XCI.

Xist foci are locally confined and form at open chromatin regions

Hence, how the limited number of *Xist* foci can silence the ~1000 X-linked genes remains a puzzle. One possibility is that they regulate target genes via rapid diffusion and transient contacts. To investigate the mobility of *Xist* foci, we developed conditions for live-cell 3D-SIM of *Xist*$^{MS2\text{-}GFP}$ at D2 and D4. Single-particle tracking for at least 2min (Videos 1 and 2) showed that *Xist* foci exhibit restricted motion without fission or fusion (Figures 2A and 2B). In 90% of cases, the displacement of each *Xist* focus over time was less than 200nm and foci movement was characterized by diffusion in a local confining potential (Figures 2C–E). The confined motion of *Xist* foci is highly correlated with the motion of chromatin loci (Chen et al., 2013; Nozaki et al., 2017). We conclude that *Xist* foci are tethered to chromatin with high affinity, constraining their movement to the local Brownian motion of chromatin. Thus, XCI is initiated through ~50 sites at which *Xist* molecules are locally confined.

To investigate if the 'wiggling' of *Xist* foci around their centers occurs within a specific chromatin environment, we introduced a histone H2B-Halo transgene into *Xist*$^{MS2\text{-}GFP}$ ESCs and performed live-cell 3D-SIM (Figure 2F and Video 3). H2B signals were segmented into intensity levels that correspond to chromatin density classes, with class 1 representing DNA-free interchromatin channels (IC) and classes 2 to 7 capturing increasing chromatin densities (Markaki et al., 2012) (Figure 2G). *Xist* foci covered predominantly classes 1 to 3 (Figure S2K). Over time, the chromatin densities underlying the footprint of *Xist* foci never surpassed class 3 and the centroids remained within chromatin class 2, consistent with the linearly and incrementally increasing chromatin density (Figures 2H and S2L). Thus, *Xist* foci are spatially confined to the periphery of chromatin domains and stably maintain their positions relative to chromatin.

In agreement with these observations, RNA antisense purification (RAP) of *Xist* followed by DNA sequencing of associated chromatin (Engreitz et al., 2013) showed that *Xist* localizes to gene-rich, open chromatin regions of the A-compartment (Figure 2I). We identified 65 and 63 highly overlapping peaks of *Xist* enrichment on the pre-Xi and Xi, respectively (Figures 2J and S2M), similar to the number of foci detected by 3D-SIM. These peaks cover broad genomic regions of 1–5 megabases (Figure S2N), indicating variability in *Xist* foci locations between cells. *Xist* peaks in the Xi are broader than those in the pre-Xi despite the overall similar distributions (Pearson's correlation r=0.83) (Figures 2I, 2J and S2N), suggesting that the chromatin contacts of *Xist* foci change over time due to chromatin compaction.

Xist nucleates supramolecular complexes

To explore how *Xist* effector proteins accumulate relative to ~50 locally confined *Xist* foci, we set out to quantify, at sub-diffraction resolution, their spatial relationship to *Xist* and to each other. We initially focused on SPEN, PCGF5, CELF1 and CIZ1, four proteins that bind to distinct repeat sequences of *Xist* and have different roles in XCI (Figure 3A) (Loda and Heard, 2019). SPEN binds the A-repeat of *Xist* and activates HDAC3 on chromatin to induce gene silencing (Chu et al., 2015; McHugh et al., 2015). The non-canonical PRC1 subunit PCGF5 is recruited to the Xi via the B-repeat and supports silencing of some X-linked genes (Almeida et al., 2017; Bousard et al., 2019; Nesterova et al., 2019; Pintacuda et al., 2017). CELF1 and CIZ1 bind to the E-repeat and restrict *Xist* to the Xi (Pandya-Jones et al., 2020; Ridings-Figueroa et al., 2017; Sunwoo et al., 2017; Yue et al., 2017).

To interrogate the localization of *Xist*-interactors, we introduced Halo-tagged transgenes into *Xist*$^{MS2\text{-}GFP}$ cells, allowing imaging of an antibody-stained and a stably expressed Halo-fusion protein together with *Xist* by multispectral 3D-SIM (Figures 3B, S3A and S3B). We observed the formation of diffraction-limited protein assemblies in proximity to *Xist* foci in both the pre-Xi and Xi that are larger than nuclear accumulations (Figures 3B and S3A). To quantitatively define these distributions, we extracted nucleus-wide spatial coordinates of thousands of segmented diffraction-limited protein particles (Figures S3C–E). We paired protein to *Xist* foci and measured nearest neighbor distances between pairs of different interactors associated with the same *Xist* focus (Figure 3C), or protein pairs located in the remainder of the nucleus. For all pairs, protein foci are on average

within ~150–200nm of each other when associated with *Xist*, both on the pre-Xi and Xi, and are separated by >350nm in the rest of the nucleus (Figure 3D). Thus, *Xist* foci induce the *de novo* formation of unique protein complexes that locally concentrate effector proteins more than elsewhere in the nucleus. The density of SPEN, CELF1, PCGF5, and CIZ1 particles is significantly higher in the pre-Xi and Xi than in the nucleus, consistent with their decreased nearest and average distances in the X-territory (Figures 3E, S4A and S4B). Furthermore, their concentration increases from the pre-Xi to the Xi, along with the observed chromatin compaction. Hence, large multi-protein assemblies that are not typically found outside the Xi form around *Xist* foci. In this way, *Xist* recruitment increases the concentration of proteins within the forming Xi. We refer to the *Xist*-nucleated proteinaceous nanostructures as supramolecular complexes (SMACs). We also observed their formation when XCI is ectopically induced on an autosome (Figure S3F) consistent with SMACs being a fundamental feature of the XCI process.

SMACs contain a wide spectrum of *Xist*-interacting proteins

To explore whether integration into SMACs is the main mechanism of protein recruitment in the Xi, we probed the distribution of additional XCI effectors (Figure S4C), including the PRC1 subunit RYBP (Tavares et al., 2012); the EZH2 subunit of PRC2 (Plath et al., 2003; Silva et al., 2003); hnRNP-K, which binds the *Xist* B-repeat to recruit PCGF5 (Pintacuda et al., 2017); PTBP1 and MATR3, which regulate *Xist* localization with CELF1 (Pandya-Jones et al., 2020). For all examined proteins, we detected a particle associated with a *Xist* focus in a near 1:1 ratio and nearest neighbor measurements revealed their presence within a 200nm zone from the centroid of *Xist*, significantly more proximal than randomized protein populations (Figures 3F, S4D and Table S1). These data corroborate the *de novo* formation of a multi-protein macromolecular cloud around *Xist* foci at the onset of XCI (Figure 3G).

We next quantified the concentration of proteins in SMACs in relation to nuclear accumulations (Figures 3H and S4E). Integrated density and particle volume measurements showed that the levels of CIZ1, CELF1, SPEN, PCGF5, EZH2 and RYBP in SMACs significantly exceed those in nuclear assemblies. Protein accumulation in SMACs varies moderately across the pre-Xi to Xi transition. CIZ1 levels remain relatively stable, PCGF5, EZH2 and RYBP levels follow nuclear changes, CELF1 levels decrease, and SPEN levels dramatically increase. Thus, gene silencing is correlated with more SPEN molecules in SMACs. MATR3, PTBP1 and hnRNP-K exhibit baseline concentrations in SMACs, suggesting that their recruitment, rather than increased accumulation is essential in XCI.

To estimate numbers of protein molecules incorporated into SMACs, we focused on the critical silencing factor SPEN. We exploited our cage^{60GFP}-standard approach and a cell line in which endogenously encoded SPEN is GFP-tagged and *Xist* can be induced with doxycycline (dox) (Dossin et al., 2020) (Figure 3I). Compared to developmentally induced XCI, dox-induction results in a larger number of *Xist* foci, yet yields similar levels of SPEN in SMACs. Moreover, before plateauing, SPEN-SMAC levels increase between 6 and 18hrs of dox-addition as seen for the pre-Xi to Xi transition (Figures S4F and S4G). Comparing the integrated density of SPEN-GFP to that of cage^{60GFP}, we infer that there are up to 35 SPEN molecules per SMAC (Figure 3J). This finding suggests that each complex may

consist of 100s to 1000s of protein molecules and that many effector proteins likely are significantly enriched relative to the number of RNA molecules (Figure 3G).

Binding to *Xist* alters the kinetic behavior of interacting proteins

To explore the kinetic behavior of SMAC protein components, we introduced Halo or mCherry-tagged SPEN, PCGF5, CIZ1, CELF1 or PTBP1 fusions into *Xist*$^{MS2\text{-}GFP}$ ESCs and performed Fluorescence Recovery After Photobleaching (FRAP) over the *Xist*$^{MS2\text{-}GFP}$-demarcated Xi-territory and other same-size nuclear regions. We also examined *Xist* RNA dynamics after confirming that no recovery of the *Xist*$^{MS2\text{-}GFP}$ signals occurred in the absence of transcription (Figures S5A–C). We observed a slow exchange of photobleached *Xist*$^{MS2\text{-}GFP}$ in the Xi (Figure 4A), comparable to that of ectopically expressed *Xist* (Ng et al., 2011). A single-exponential kinetic model provided the best fit to the measured FRAP curve and inferred a slow dissociation rate (0.05/min) and an average lifetime of ~22min. This result is consistent with a single type of high-affinity interaction between *Xist* and chromatin (Figures 4B and 4C). Similar to *Xist*, CIZ1 has a ~19min recovery time in the Xi, which is much longer than that of all other interrogated proteins (Figures 4D–G, S5D and S5E). The tight kinetic and spatial relationship between *Xist* and CIZ1 (Figures 3F, 4G and Video 4) suggests that CIZ1 and *Xist* molecules form a relatively stable core within a SMAC.

Kinetic modelling of the SPEN, PCGF5, CELF1 and PTBP1 FRAP curves yielded faster exchange rates than CIZ1 and *Xist* and two effective types of binding sites (Figures 4F, 4G and S5D–F). Using two-exponential fits, we inferred parameters for the short-lived (f_1) and long-lived (f_2) bound fractions within and outside of the Xi. For each protein, the rapid binding events occurred within seconds while slow dissociation required several minutes (Figures 4G and S5G). SPEN was the most dynamic among the examined proteins. Recruitment by *Xist* extended the binding rates for these proteins compared to the nucleus, indicating that the Xi forms a unique nuclear compartment within which proteins exhibit distinct kinetic behaviors (Figure 4G). The kinetic assays indicate that *Xist* effector proteins with short residence times bind to the slowly exchanging *Xist*-CIZ1 core. Thus, SMACs are rapidly exchanging dynamic complexes that form local, high affinity concentration platforms. Accordingly, examination of SPEN and PCGF5 populations in the Xi, outside SMACs, revealed higher protein concentrations than in nuclear accumulations, demonstrating that recruitment to SMACs leads to enrichment of constituent proteins across extended local neighborhoods in the X-territory (Figures 4H and S5H–K).

Crowding of SPEN in SMACs is IDR-dependent

The formation of SMACs is consistent with a requirement for extensive protein-protein interactions (Figure 3G). SPEN contains intrinsically disordered regions (IDRs), which typically mediate weak, multivalent interactions (Banani et al., 2017; Cerase et al., 2019; Mittag and Forman-Kay, 2007; Uversky, 2015). We homozygously deleted the IDRs within the endogenous SPEN alleles in female ESCs in which SPEN is tagged with GFP (Dossin et al., 2020) and showed that ΔIDR SPEN expression did not disrupt the formation of the *Xist* cloud (Figures 5A, 5B, S6A and S6B). Deletion of the IDRs does not interfere with the binding of the protein to *Xist* but eliminates the increased SPEN levels in SMACs such

that ΔIDR SPEN levels in the pre-Xi and Xi are close to those of the wildtype (WT) protein within the nucleus (Figures 5B, 5C, S6C and S6D). We conclude that the accumulation of SPEN in SMACs is driven exclusively by their IDRs. Moreover, binding to *Xist* through its RNA binding (RRM) domains is required for the IDR-mediated concentration of SPEN into SMACs (Figures S6C–E). FRAP experiments showed that the deletion of the IDRs or RRMs abolishes the characteristic Xi-immobile fraction of SPEN and dramatically alters residence times, with ΔIDR SPEN exhibiting very long unbinding times in both the Xi and nuclear fractions, possibly due to the tight binding to the RNA through the RRMs (Figures S6F–H). Therefore, the IDRs of SPEN are also critical for creating a dynamic protein assembly (Figure 5K).

IDR-mediated crowding of SPEN in SMACs is required for XCI

The SPOC domain of SPEN is essential for silencing (Dossin et al., 2020), but whether the crowding of SPEN is necessary for the functionality of SPOC is unknown. RNA FISH for nascent transcription of five X-linked genes revealed a striking silencing defect when IDR-mediated crowding was ablated, similar to the lack of silencing caused by the deletion of SPOC (Dossin et al., 2020) (Figures 5D and 5E).To explore if the loss of gene silencing by ΔIDR SPEN extends to the entire X chromosome, we performed scRNA-seq before and 24 hours after *Xist* induction in ΔIDR, ΔSPOC and WT SPEN expressing cells. Although X-linked gene repression was observed in WT cells, both ΔIDR and ΔSPOC SPEN expressing cells displayed a dramatic X chromosome-wide loss of gene silencing, affecting both rapidly and slowly silencing genes (Barros de Andrade et al., 2019) (Figures 5F and 5G). We constitutively expressed Halo-tagged ΔIDR or full-length (FL) SPEN as rescue constructs in ESCs in which the endogenously encoded SPEN is fused to the AID degron tag and can be depleted by addition of auxin (Dossin et al., 2020) (Figures 5H and 5I). Bulk RNA-seq showed that FL but not ΔIDR SPEN can rescue X-linked gene silencing (Figure 5J). Interestingly, when only SPOC is tightly tethered to *Xist* (Dossin et al., 2020), X-linked genes are inefficiently silenced (Figures S6I–K), consistent with a dynamic SPEN protein being required for XCI. In summary, our findings demonstrate that the concentration of SPEN in SMACs and its rapid kinetic behavior, both mediated by the IDRs, are required for the protein to exert its silencing function through the SPOC domain (Figure 5K).

The B-repeat is critical for Xi compaction and late gene silencing

Non-canonical PRC1 induces the recruitment of canonical PRC1 and downstream accumulation of PRC2 (Brockdorff, 2017), and is implicated in the silencing of a subset of X-linked genes (Bousard et al., 2019; Colognori et al., 2019; Nesterova et al., 2019; Zylicz et al., 2019). Yet, PRC1 spreads into genes only after silencing has occurred (Zylicz et al., 2019), raising the question of how it contributes to XCI. Since PRC1 is critical for chromatin compaction in various developmental contexts (Boyle et al., 2020; Francis et al., 2004; Grau et al., 2011; Illingworth, 2019), we explored whether the B-repeat is required for the structural reorganization of the Xi.

We perturbed PRC1 recruitment to the X by heterozygously deleting the B-repeat (ΔB-*Xist*) on the MS2-tagged 129 Xist allele in female ESCs derived from a 129 x castaneous (Cas) cross and compared the compaction of the Xi^{Cas} formed by FL-*Xist* to the Xi^{129} induced

by ΔB-*Xist* (Figures 6A and 6B). Deletion of the B-repeat results in less compacted pre-Xi and Xi territories, larger distances between *Xist* foci and an expansion of the *Xist* cluster (Figures 6B–C and S7A–C). Accordingly, the density of SPEN-decorated SMACs is far lower in the ΔB-Xi than in the WT-Xi although the concentration of SPEN in their respective SMACs is similar (Figures S7D and S7E). Together, these results uncover a role of the B-repeat, and in turn of PRC1 and its downstream effectors, in driving the compaction of the X chromosome and densification of SMACs during XCI initiation.

To explore if X-linked gene silencing dynamics are altered in the absence of compaction, we performed scRNA-seq at D2 and D4. Upon deletion of the B-repeat, silencing is more impaired on the Xi than on the pre-Xi (Figure 6E). The silencing defect is strongest for slowly silencing genes (Figures 6F–H). These results extend to cells lacking SMCHD1 that controls the compartmentalization of the Xi and is recruited to the Xi by PRC1 (Jansz et al., 2018; Wang et al., 2019) (Figure S7E). Thus, compaction by PRC1 and SMCHD1 and the further clustering of the *Xist*-SMACs allows SPEN to act on all genes.

Xist-SMACs progressively re-configure and silence the Xi

We next explored how genes with different silencing kinetics, i.e. rapidly (early) and slowly (late) silencing genes, localize relative to *Xist* foci. To determine these spatial relationships, we applied multiple distance metrics to 20 simultaneously detected early or late genes distributed across the entire X chromosome, their nascent transcripts and *Xist* by RNA/DNA FISH and 3D-SIM (Figures 7A–D).

We first monitored the distribution of early and late gene loci on the Xa relative to the *Xist* transcription locus by exploiting the detection of *Tsix* RNA as a proxy for the X-inactivation center (Xic) where the *Xist* gene is located (Plath et al., 2002). This analysis showed that early genes are closer to the *Xist* locus than late genes (Figures 7C and 7E). Accordingly, upon differentiation, early genes are closer to individual *Xist* foci or to the center of the entire *Xist* cluster than late genes, which is more pronounced on the pre-Xi (Figures 7F and S7G). Thus, genomic regions containing early genes are spatially more proximal to the *Xist* locus at the onset of XCI and more likely to be populated by the *Xist* cluster. Intriguingly, there is no significant difference in the nearest distance of active and silenced genes to *Xist* foci, regardless of early or late silencing kinetics, yet active genes tend to be more distal to the center of the *Xist* foci cluster (Figures 7G and S7H). This finding is consistent with late genes being further away from the *Xist* cluster in the pre-Xi.

Chromosomal compaction significantly decreases the distances of both early and late genes to *Xist* foci in the pre-Xi to Xi transition, with a higher impact on late genes, which also exhibit the most dramatic repositioning (Figures 7F and 7H). The result of the conformational change is that early and late genes congregate and move closer to the center of the *Xist* cluster (Figures 7C and S7G). Consequently, the same number of *Xist* foci can progressively silence more genes. The gradual gene silencing during the XCI process can therefore be explained by the spatial organization and reconfiguration of the X chromosome that dictate the relationship of genes to the *Xist* cluster.

Finally, we explored how loss of compaction affects the organization of genes relative to *Xist* in cells expressing ΔB-*Xist*. All genes on the ΔB-Xi, regardless of their silencing kinetics, are at larger distances from the centroid of the *Xist* cluster and to each other, and the overall distances between early and late genes are enlarged compared to WT-Xi (Figures 7I–L). However, nearest neighbor measurements between *Xist* foci and early or late genes revealed no significant difference for the ΔB and WT Xi, suggesting that *Xist* foci localize similarly to target regions (Figure 7M). These findings indicate that the lack in compaction affects the reorganization of genes, which results in poor clustering of SMACs and inefficient silencing, particularly of late genes.

Discussion

SMACs are the functional units of *Xist*-mediated XCI

XCI is a powerful model for interrogating how lncRNA molecules can establish a functional nuclear compartment. Since its discovery, it has been thought that *Xist* progressively spreads on chromatin to associate with all target genes. This view was refined by the observation that *Xist* first localizes to sites in close spatial proximity to its transcription locus and then spreads chromosome-wide (Engreitz et al., 2013). Owing to the focal accumulation of *Xist* revealed by super-resolution microscopy studies, it was proposed that the RNA and interacting proteins form ribonucleoprotein complexes that sample genes along the chromosome through a “hit-and-run” model (Sunwoo et al., 2015). Our study shows that *Xist* foci are instead stably bound to chromatin and that they induce the *de novo* formation of SMACs. Each *Xist*-SMAC accumulates ~35 copies of the ~500KDa protein SPEN. A comparison with other protein levels in SMACs suggests that other *Xist* effectors likely concentrate to much higher levels. The formation of SMACs induces a phase transition in the Xi, as SMACs surrounding stably bound *Xist* molecules create a sharp increase in protein density at the boundary of the Xi. Whether SMACs exhibit features of liquid-liquid phase separation and whether the progressive coalescence of chromatin regions induces polymer-polymer phase separation (Frank and Rippe, 2020) remains to be determined.

Our results suggest that different binding environments in SMACs allow for both topological retention of proteins as well as their rapid exchange. IDRs are critical for the dynamic supramolecular aggregation of SPEN in SMACs, which is necessary for its catalytic domain SPOC to exert gene silencing. This finding is consistent with reports that catalytic rates of IDR-containing DNA modifying enzymes increase with crowding (Kuznetsova et al., 2014; Zimmerman and Pheiffer, 1983). Many *Xist*-interacting proteins contain IDRs and have the propensity to self-aggregate (Cerase et al., 2019; Pandya-Jones et al., 2020). Whether the IDRs of SPEN are involved solely in homotypic interactions and how IDRs of other *Xist*-interactors contribute to the formation and function of SMACs remain open questions. Interestingly, the binding to *Xist* is required for the IDR-dependent integration of SPEN into SMACs and may impart specificity to protein interactions within the Xi. RNA binding may induce folding of unstructured IDRs and enable ‘entry’ into SMACs, consistent with observations for other IDR-containing proteins (Uversky, 2015).

A supramolecular aggregation-based model of XCI

Our work yields a revised model of how *Xist* establishes the Xi compartment and orchestrates gradual transcriptional silencing (Figure 7N). Through expression, diffusion, sequestration, and degradation (see section "Expression-diffusion-degradation model for *Xist* confinement" in Methods S1 file), two *Xist* transcripts become localized and tightly bound to chromatin, seeding SMACs at 50 regions that are proximal to the *Xist* locus and enriched for rapidly silencing genes. By establishing high local concentrations of transiently binding effector proteins in SMACs, *Xist* induces gradients of silencing proteins, most importantly of SPEN, that can act at genomic locations on the X without their continuous association with *Xist*. This process initiates silencing on the pre-Xi. The high concentration of PRC1 and likely other architectural regulators brought about by SMACs progressively induces higher-order chromatin changes and compaction. Compaction promotes the overall densification of genes under the SMAC cluster, enabling a constant number of SMACs to gain access to an increasing number of genes. Consequently, a higher concentration of SPEN is present in the vicinity of more genes and silencing expands across the entire X. However, the presence of a SMAC *per se* is not sufficient to induce effective silencing as genomic regions that are poorly crowded by SMACs silence less efficiently. By showing that SMAC formation and chromatin reconfiguration are interdependent mechanisms to achieve robust gene silencing, our model fills the knowledge gap of how different repressive pathways cooperate in XCI.

Implications beyond *Xist*

Phase separation has recently emerged as a much-debated mechanism in the field of gene regulation (McSwiggen et al., 2019). Yet, the functional role of condensates in gene regulation remains largely undefined. Macromolecular crowding as the mode of heterochromatin formation, described here, expands transcriptional control beyond the seeding molecule. This mechanism may be particularly important for gene regulatory RNAs, typically expressed at low numbers relative to their targets (Cabili et al., 2015; Derrien et al., 2012). Intriguingly, other lncRNAs have also been found to induce spreading of Polycomb complexes (Schertzer et al., 2019), suggesting that a common mechanism in the organization of an efficient repressive nuclear compartment may involve enzymes that induce transcriptional repression together with regulators of chromatin architecture.

Limitations of the study

In this study, we applied super-resolution imaging in combination with kinetic modeling, genomic approaches and functional perturbations to investigate fundamental principles of RNA-seeded nuclear compartmentalization. While multi-spectral quantitative super-resolution imaging allowed us to explore the spatial relationships of RNA, DNA and protein in individual cells, it is subject to physical limitations. With 3D-SIM, we can resolve the distribution of XCI effectors down to few hundreds of kb along the genome that does not allow us to examine enrichment at specific genes, which would require even higher resolution (Xie and Liu, 2021). Furthermore, due to the limited number of individual fluorophores that can be employed, it is not possible to simultaneously detect several SMAC proteins *Xist*, genes and gene transcripts to directly determine what changes trigger the

switch from an active to a repressed state. Additionally, fixation, permeabilization, and heat-denaturation applied in FISH or immunodetection experiments solubilize a considerable amount of protein, thus, the protein levels in SMACs is likely underestimated. Finally, although a sharp boundary needs to be determined in segmentation-based image analysis, as employed in our study, SMACs are dynamic, formed by rapidly exchanging proteins and likely do not adopt defined, but rather graded, distributed structures. Despite these limitations, super-resolution microscopy was critical for disentangling the processes of gene silencing and chromatin reconfiguration and allowed us to distinguish *Xist* from its protein interactors as functionally distinct entities in the Xi space.

STAR Methods

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Kathrin Plath (KPlath@mednet.ucla.edu).

Materials availability—All unique materials generated in this study, such as plasmids and cell lines will be available to researchers from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

- All genomic data (bulk mRNA-seq, scRNA-seq, CLAP-seq, RAP-seq) generated in this study have been deposited in the Gene Expression Omnibus (GEO) database. The accession number is listed in the key resources table. Accession numbers of reanalyzed publicly available data are also listed in the key resources table. Super-resolution microscopy image data, segmented masks and derived features of nuclear particles will be shared by the lead contact upon request.
- This study did not generate original code. All computational approaches and software used are described in the STAR Methods and listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Cell culture—Female mouse polymorphic *129S4/SvJae / castaneus* F1 2–1 ESCs (Panning et al., 1997) and its engineered derivatives were grown on 0.5% gelatin-coated flasks seeded with irradiated DR4 feeders (obtained from day 14.5 embryos with appropriate animal protocols in place). Cultures were maintained in mouse ESC medium containing knockout medium DMEM (Life Technologies), 15% FBS (Omega), 2mM L-glutamine (Life Technologies), 1x NEAA (Life Technologies), 0.1mM β-Mercaptoethanol (Sigma), 1x Penicillin/Streptomycin (Life Technologies), and 1000U/ml mouse LIF (homemade) in 5% CO2, 37°C incubators.

Epiblast-like (EpiLC) differentiation was performed as described (Hayashi and Saitou, 2013). Briefly, prior to induction of EpiLC differentiation cells were adjusted for 3 passages to feeder-free conditions in the presence of 1000U/ml LIF and two inhibitors, CHIR99021 (3μM) and PD0325901 (0.4μM) (2i+LIF) in serum-free N2B27 medium containing 1× N2 supplement and 1× B27 supplement (Thermo Fischer), 2mM L-glutamine (Life Technologies), 1× NEAA (Life Technologies), 0.1mM β-Mercaptoethanol (Sigma), 0.5 × Penicillin/Streptomycin (Life Technologies). To induce differentiation, cells were dissociated and seeded at a density of 2×10^5 cells/ml in N2B27 media containing 20 ng/ml Activin A and 12 ng/ml bFGF on geltrex-coated flasks or coverslips. For experiments extending beyond day 4 of differentiation, we applied a protocol previously described in (Ying and Smith, 2003). Briefly, at day 4 of differentiation, EpiLCs were dissociated with accutase and seeded on geltrex-coated coverslips at a density of 5×10^5 cells/cm^2. Cells were then grown in N2B27 media supplemented with EGF and FGF (10 ng/ml each), on geltrex-coated coverslips for 4 more days (d8 of differentiation). At this developmental stage, cells have lost Tsix expression as observed in Figure 1B. Media was exchanged daily.

TX1072, female polymorphic *B6/castaneus* mouse ESCs lines, carrying a tetO-promoter driving the endogenous Xist allele on the B6 X chromosome ($B6^{tetO\text{-}Xist}$ $Cas^{WT\text{-}Xist}$), further modified by a homozygous insertion of the GFP or Halo tag in the endogenous *Spen* loci (Dossin et al., 2020) or the deletion of the SPOC domain in the endogenous locus (ΔSPOC-SPEN-GFP) (Dossin et al., 2020), as well as derivative cell lines generated in this study carrying ΔIDR-SPEN-GFP or *Rosa26* knockins, were grown on gelatin-coated flasks in feeder-free conditions (2i+LIF in mouse ESC medium) and EpiLC differentiation was performed as described for the F1 2–1 ESCs. When induction of XCI was performed from the tetO-*Xist*, 0.5μg/ml doxycycline were added to ESC media without 2i+LIF for 24hrs. Similarly, male tetO-*Xist* ESCs were grown on gelatin-coated flasks in feeder-free conditions and induction of *Xist* expression was performed with addition of 0.5μg/ml doxycycline in ESC media for 18 to 24hrs.

For auxin-mediated depletion experiments of the TX1072 ESCs expressing SPEN-AIDGFP (Dossin et al., 2020) and *Rosa26 SPEN rescue knockins,* auxin was added to 2i+LIF ESC medium at 500μM for 12hrs. Following, cells were dissociated by trypsinization and seeded in ESC medium without 2i+LIF including 500μM auxin and 0.5μg/ml doxycycline for 24hrs to induce *Xist* expression.

C127 cells were purchased from ATCC and human fibroblasts containing abnormal X-chromosomes (GM3827, GM00735, GM06960, GM07213) were obtained from Coriell Institute. These cell lines were cultured in DMEM (Life Technologies), 15% FBS (Omega), 2mM L-glutamine (Life Technologies) and 1x Penicillin/Streptomycin (Life Technologies).

METHOD DETAILS

Plasmid construction for engineered cell lines—Plasmids containing the 24×MS2 repeats (#31865) and MS2-Coat-Protein-GFP (MCP-GFP) coding sequence (#27121) were obtained from Addgene. The pBglII5k plasmid (Jonkers et al., 2008) was used for targeting the 24xMS2 repeats into Xist and contains homology arms for insertion into exon 7 of Xist, downstream of the E-repeat sequence, and a floxed neomycin resistance

cassette. The 24xMS2 repeats were excised from plasmid #31865 by restriction digest with BglII and BamHI and inserted into the pBglII5k plasmid by Infusion cloning, yielding the pBglII5k-24xMS2 plasmid (which replaces the 16xMS2 repeat array originally contained in the pBglII5k plasmid). The coding region for MCP-GFP was amplified by PCR from plasmid #27121 and introduced under control of a tetracycline-inducible promoter (tetO) into the pBS31 plasmid (pgkATGfrt) (Beard et al., 2006) by Infusion cloning, yielding pBS31-MCP-GFP. A reverse tetracycline TransActivator (rtTA3) cassette containing the PGK promoter and a BGH polyA element was amplified by PCR from the MXS_PGK::rtTA3-bGHpA plasmid (#62446, Addgene) and introduced into the unique AscI site of pBS31-MCP-GFP, downstream of the tetO-MCP-GFP-polyA insert, by Infusion cloning, resulting in the pBS31-MCP-GFP-rtTA3 plasmid. For deletion of the B-repeat of Xist the p13–5-Xist-Bdel plasmid was constructed from PCR-amplified 5’ and 3’ homology regions obtained from the PCV-XistPA plasmid and a loxP-flanked hygroTK cassette that replaces the B-repeat sequence (chrX: 103480156–103480430, mm10), inserted into a PBS-KS (+) plasmid.

Plasmid construction to generate transgenic lines—For the integration of transgenes expressing various mCherry or Halo protein fusions under the control of the endogenous *Rosa26* promoter in Xist$^{MS2\text{-}GFP}$ ESCs, we employed a parent plasmid harboring homology arms for targeting the *Rosa26* locus and a loxP-flanked puromycin cassette for antibiotic selection (R26-SA-EGFP-puro). A splice-acceptor (SA) sequence and a splice-donor (SD) coding sequence were synthesized (Genewiz) and inserted into the R26-SA-EGFP-puro plasmid after MluI/MfeI restriction digest to remove the GFP, by Infusion reaction. The resulting pYM215-R26-SA/SD-puro plasmid was used as the parent plasmid for insertion of all protein fusions in three-piece Infusion reactions. The coding sequence for CIZ1 was amplified from the donor plasmid pBS32-MCP-CIZ1(Pandya-Jones et al., 2020). Coding sequences for histone H2B and mCherry were amplified from a H2B-mCherry plasmid (Addgene, #20972) and the Halo cDNA was obtained from the Halo-EasyFusion plasmid (Addgene, #112852). The coding sequences for PTBP1, PCGF5, and CELF1 were synthesized (Genewiz).

To generate the Spen-FL-Halo, Spen-ΔRRM-Halo and Spen-ΔIDR-Halo plasmids, introduced into Xist$^{MS2\text{-}GFP}$ ESCs or TX1072 ESCs, the full-length Spen Entry Clone (Sp22) was modified using Polymerase Incomplete Primer Extension-based mutagenesis with primers designed to delete amino acids 639–3460 or 1–591, respectively. Sp22 and the Spen-ΔIDR or Spen-ΔRRM entry clones, respectively, were inserted into the PyPP-CAG-Halo-V5-IRES-Puro destination vector using Gateway LR Recombination, generating PyPP-CAG-Halo-full-length-Spen-V5, PyPP-CAG-HaloSpen-ΔIDR-V5 and PyPP-CAG-Halo-Spen-ΔRRM-V5, respectively, also containing an IRES-puromycin resistance cassette for selection. These plasmids enable constitutive expression of Spen variants with an N-terminal Halo tag and a C-terminal V5 tag and contain a polyoma episomal origin of replication for efficient propagation in mammalian cell culture.

For integration of the full length (FL-) and ΔIDR-Spen-Halo transgenes into the SPEN-AID-GFP female ESC lines the plasmid pFD46 for targeting into the *Rosa26* locus was used (previously described in (Dossin et al., 2020)). Halo was amplified from plasmid

Halo-EasyFusion (Addgene, #112852) and inserted upstream the Spen coding sequence of pFD46 by Infusion cloning, resulting in the plasmid pYM300. Similarly, the FL-Spen coding sequence was excised from pFD46 by restriction digest and Infusion cloning was performed to insert Halo-ΔIDR-Spen amplified from plasmid PyPP-CAG-Halo-Spen-ΔIDR-V5, resulting in plasmid pYM301. For the integration of FL-Spen-Halo into the Xist-FL^{Cas}Xist-ΔB^{129} female F1 2–1 ESCs and 36.11 cells, a male ESC line carrying an autosomal tetO-Xist transgene (on chromosome 11)(Wutz and Jaenisch, 2000), the same strategy was employed. All plasmids were verified by restriction digests and sequencing.

Targeting and cell line generation—For targeting, F1 2–1 female ESC lines were grown on DR4 feeders. All targetings were performed by electroporation using the GenePulserII (Biorad). Approximately 2×10^7 cells and 50μg of DNA were resuspended in 400μl PBS in 4mm diameter cuvettes and pulsed twice for 0.2msec at 800V. To target the *Rosa26* locus with plasmids pYM300 or pYM301 and pFD82, pFD83 in TX1072 ESCs expressing SPEN-AID-GFP, female F1 2–1 Xist$^{129\Delta B/CasWT}$ ESCs and male ESCs carrying an autosomal tetO-Xist transgene (36.11), the 3D Nucleofector (Lonza) was used with program CG-104 and 100μl cuvettes according to the manufacturer's instructions. Antibiotics were added to the growth media 24–36 hours after electroporation. Puromycin was used at 1.5μg/ml, hygromycin at 130μg/ml and G418 at 400μg/ml. The culture medium containing the respective antibiotics was exchanged every 2 to 3 days. Once adequate colony growth was observed (1–2 weeks), 100–200 colonies were picked under a stereoscope, dissociated by trypsinization and seeded in 96-well plate replicates. One replicate plate was used for genomic DNA extraction and subsequent genotyping PCR. All positive clones used in this study were screened to ensure that they maintain two X chromosomes in the undifferentiated state indicated by the presence of two Tsix transcripts in RNA FISH experiments. In addition, we confirmed gene silencing by *Xist* and normal *Xist* distribution across the X-territory upon induction of differentiation as applicable (Figure 1K and S1G).

Integration of 24xMS2 repeats into the *Xist* locus—The pBglII5k-24xMS2 plasmid was electroporated into the F1 2–1 ESC line after linearization with XhoI. The cell culture was exposed to neomycin selection 36hrs post-electroporation. Colonies were picked and expanded for screening by genotyping PCR and RNA FISH with Xist and MS2 probes was performed on EpiLCs at day 4 as shown in Figure S1G. We confirmed that the 24xMS2-repeat unit was introduced into the 129 allele (Figure S1G, right). This observation comes in line with the known skewing of XCI in 129/Cas female cells, where ~80% of cells inactivate the 129 allele. The loxP-flanked neomycin resistance cassette was removed from targeted ESC clones by transient expression of Cre-recombinase. Subsequently, a FRT-recombination site-containing a landing pad (FRT-neo plasmid) (Beard et al., 2006) was targeted into the inert *Col1A* locus (on chromosome 11) in F1 2–1$^{24xMS2\text{-}Xist}$ ESCs. The MCP-GFP-rtTA3 expression cassette was then inserted into the FRT site by electroporation of a FlpO-recombinase-encoding plasmid and the pBS31-MCP-GFP-rtTA3 plasmid. The resulting ESC line was denoted as Xist$^{MS2\text{-}GFP}$.

Engineering strategy to delete the IDRs of SPEN—CRISPR/Cas9-based genome editing was used for the deletion of the IDRs of SPEN with two guide RNAs targeting

intronic sequences flanking the IDR-encoding exons, which were synthesized by Synthego. Targeting was performed in the TX1072-SPEN-GFP cell line (Dossin et al., 2020). Cells were targeted using recombinant Cas9 protein (Synthego) and gRNAs with the 4D Nucleofector (Lonza). ~2.5×10^5 cells/ml were resuspended in 20µl Lonza solution and electroporated with 1.8µl of gRNA1, 1.8µl gRNA2 and 2.4µl Cas9 using program CG-104. gRNA sequences are given in Table S2. Nucleofected cells were serially diluted and plated onto 10cm dishes. Once adequate colony growth was observed (1–2 weeks), 100 colonies were picked under a stereoscope, dissociated by trypsinization, and seeded in 96-well plate replicates. One replicate plate was used for genomic DNA extraction and subsequent genotyping PCR. Genotyping PCR was performed as shown in Figure S6A. Positive clones were selected based on an expected 605bp band and sequenced to verify deletion. Biallelic expression was confirmed using RT-PCR with oligo-dT primers and sequencing as shown in Figure S6B, scoring for SNPs in exon 5 (rs27580268) and exon 7 (rs223335536).

Engineering strategy to delete the B-repeat of Xist—F1 2–1 ESC line previously targeted with an 16xMS2 tag in the large final Xist exon on the 129 allele (Jonkers et al., 2008) or male V6.5 ESCs expressing Xist under a tetO promoter (Engreitz et al., 2013) were electroporated with the linearized plasmid 13–5 harboring homology arms for targeting into the B-repeat region of XIST and replacing it with a loxP-flanked hygroTK cassette for antibiotic selection. The loxP-flanked hygroTK resistance cassette was removed from targeted clones by transient expression of the Cre-recombinase and ganciclovir treatment. Genotyping and confirmation of deletion of the B-repeat in both cell lines and targeting on the 129 allele in F1 2–1 ESCs were performed by Southern blotting (not shown).

Expression of cage^{60GFP} in ESCs—The gene encoding ct-60 (cage^{60GFP}) was amplified by PCR from plasmid I3–01-ct60GFP (Hsia et al., 2016). The fragment was introduced under control of the CACGS promoter into the pBS32 plasmid by Infusion reaction yielding pBS32-cage^{60GFP} and positive clones were confirmed by restriction digests and sequencing. The pBS32 plasmid was derived from the pBS31 plasmid upon replacement of the tetO promoter with a CAGGS promoter. To visualize both Xist$^{MS2\text{-}GFP}$ and cage^{60GFP}, Xist$^{MS2\text{-}GFP}$ ESCs were differentiated into EpiLCs to induce endogenous *Xist* expression and doxycycline was added at 0.5µg/ml for 2hrs to induce MCP-GFP expression. Expression of the cage^{60GFP} was achieved by transient transfection of the pBS32-cage^{60GFP} plasmid into differentiating cells by Lipofectamine3000 24 hours prior to imaging, according to the manufacturer's instructions.

Halo labelling—For FRAP experiments of Halo-fused proteins, 5µM of TMR Halo ligand was added to the culture medium for 30min following a 30min incubation in media without added ligand to wash-off unbound ligand. For fixed and live-cell 3D-SIM imaging, 1µM JF549 or JF646 Halo ligands were introduced to the media for 15min, washed-off twice with PBS and exchanged with fresh medium which was incubated for another 15 min. Live-cell imaging or fixation was done as described in the corresponding sections.

Immunofluorescence staining—Immunodetection was performed as described in (Kraus et al., 2017). In brief, cells were grown on geltrex-coated high precision coverslips

at the desired differentiation state. Coverslips were then transferred to new multi-well plates, washed three times with PBS and fixed with 2% formaldehyde dissolved in PBS for 10min, followed by two washes with PBST (1xPBS, 0.05% Tween 20). Samples were then quenched for 10min with 20mM glycine in PBS. Following samples were washed with PBS and permeabilized with 0.5% Triton X-100 dissolved in PBS for 10min and washed once with PBST. Samples were then blocked for 1hr in blocking buffer (PBST, 2% BSA, 0.5% Fish skin gelatin) and incubated with primary antibodies diluted in blocking buffer for 1hr in a humidified chamber at RT. Samples where then washed three times with PBST followed by incubation with secondary antibodies for 45min and another round of PBST washes. Samples were then washed once with PBS, post-fixed with 4% formaldehyde dissolved in PBS for 10min, followed by two washes with PBST. Chromatin counterstaining was performed using DAPI dissolved in PBST at a concentration of 2μg/ml for 5min.

Samples were then washed four times with PBS, mounted on slides with Vectashield and sealed with Covergrip. For combined Halo ligand and antibody detection, cells were labelled with the Halo ligands as described in the Halo labeling section, fixed and processed by immunostaining. For the 4-color 3D-SIM imaging where we detect combinations of proteins together with $Xist^{MS2\text{-}GFP}$ (Figures 3B and S3A) we used CIZ1H-alo and CELF1 antibody staining, SPEN-Halo and CIZ1 antibody staining, PCGF5-Halo and CIZ1/CELF1 antibody staining. Halo transgenes were detected with the Halo ligand JF549 and primary antibodies with secondary antibodies conjugated to AlexaFluor647. In Figures 3H and S4C we used endogenous Halo-tagged SPEN (Dossin et al., 2020) and Halo transgenes for detection of CIZ1, PCGF5 and PTBP1 labelled by the Halo ligand JF549 and antibody stainings with primary and secondary antibodies conjugated to CF568 dye for CELF1, RYBP, EZH2 and hnRNP-K. In Figures S6C and S7D, we used Halo transgenes FL-/ΔIDR-/ΔRRM-SPEN expressed in $Xist^{MS2\text{-}GFP}$ or F1 2–1 $Xist^{129\Delta B/CasWT}$ EpiLCs labelled by the Halo ligand JF549. In Figure S3F, RYBP and CIZ1 are detected with antibodies and secondaries conjugated to CF568, while SPEN is stably integrated into the Rosa26 locus (plasmid YM301) and detected with the JF549 Halo ligand. In Figure 5C, we detect the endogenous WT-SPEN-GFP or ΔIDR-SPEN-GFP with anti-GFP antibodies. We compared the distribution of the CIZ1-Halo fusion protein and the endogenous (antibody-stained) CIZ1 protein and show the same trend (Figure S3B) similarly to the distribution of endogenously Halo- or GFP-tagged SPEN proteins to Halo tagged transgenes (Figures 3H, 5C and S6E).

Antibodies and dilutions—Endogenous CELF1 was detected with monoclonal rabbit anti-CUG-BP1 antibody ab129115 (1/800; Abcam); hnRNP-K with polyclonal rabbit antibody A300–678A (1/800, Bethyl); MATR3 with polyclonal rabbit antibody IHC-00081 (1/200, Bethyl); RYBP (DEDAF) with polyclonal rabbit antibody AB3637 (1/1000; Sigma); Ezh2 with monoclonal rabbit antibody #5246 (1/500; Cell Signaling Technology); CIZ1 with a polyclonal rabbit antibody NB100–74624 (1/800; Novus Biologicals), and histone H3 phospho-Serine 10 with the polyclonal rabbit anti-histone H3-phospho-Serine10 #39253 (1/1000; Active Motif). GFP with a polyclonal rabbit antibody ab6556 (1/500, Abcam). Two secondary antibodies were used, including high cross-absorbed donkey anti-rabbit IgG

CF568 antibody SAB4600076 (1/400; Sigma) and high cross-absorbed goat anti-rabbit IgG Alexa Fluor 647 antibody A21245 (1/400; Life Technologies).

FISH Probe synthesis—Probes for DNA and RNA FISH experiments were labelled by Nick Translation (NT) as previously described (Cremer et al., 2008). Briefly, 1μg of DNA was labelled in a 50μl NT reaction for 8hrs using 1.3–2.5μl fluorescently-labelled dUTPs, 1μl of DNA Polymerase (Thermo Fisher Scientific) and 2μl of DNAse I (Sigma-Aldrich) from a stock which was freshly prepared by a 1:200 dilution in ice-cold H_2O. NT reactions were purified using magnetic beads, probes were then resuspended in Nuclease-free H_2O and ethanolprecipitated together with Salmon sperm and Cot1 DNA at −80°C ON. After precipitation and washes with ethanol series (70–100%), probes were resuspended in deionized formamide with shaking at 37°C ON. Probes were then adjusted in hybridization buffer (50%formamide, 2×SSC, 10% dextran sulfate) and stored at −20°C. To create mouse Xist probes, we used a full-length mouse Xist cDNA plasmid (p15A-31–17.9kb Xist). Human XIST probes were created from a full-length XIST cDNA construct. For assessing X-linked gene silencing, *Atrx* probes were synthesized using BAC RP23–265D6, *Rlim* probes using fosmid WI1–2704K12 and *Mecp2* probes using fosmid WI-894A5. For the chromosome barcoding experiment, we used BACs RP23–53H15, RP2383-J1, RP23–451D5, RP24–81K23, RP24–374B8, RP23–401G5, RP23–104K18. To create an intronic probe against the first intron of Xist, the corresponding region was amplified from the Xist-encoding BAC RP23–223G18 and was labelled by NT. To create MS2 probes the corresponding region was amplified from plasmid #31865 (Addgene) and labelled by NT. For multispectral chromosome barcoding experiments, individual BACs were labelled separately, pooled in a 1:1 ratio and used at 0.1μg/cm^2. DNA from flow sorted mouse X-chromosomes was a gift from Irina Solovei and labelled using the Bioprimer kit according to the manufacturer's instructions. NT products were labelled with Atto488-dUTP, Alexa Fluor 568-dUTP, Cy3-dUTP, Cy5-dUTP, Texas Red-dUTP and chromosome paints were labelled with Atto448-dUTPs or Cy3-dUTPs.

Probes used in this study include: Figure 1B: Xist-Atto488, Rlim-AlexaFluor568, Atrx-Cy5; Figure 1E: mmX-Atto488, Xist-Alexa568; Figure 1H: Color-coded as green-Atto488, yellow-Cy3, red-Texas Red, magenta-Cy5; Figure 5B: Xist-AlexaFluore568; Figure 5D: Xist-Atto488, Rlim-, AtrX-AlexaFluor568; Figure 6B: mmX-Atto488, MS2-AlexaFluor568, Xist-Cy5; Figure 7B: Genes-Atto550, Transcript-Alexa647N; Figure 7C, 7I: Early genes-Atto550, Late-genes-Alexa647N; Figures S1B: Xist-Atto488, Mecp2-Cy3; Figures S1G: MS2-Atto488, Xist-Cy3, Atrx-Cy5; Figures S1H: Xist-Atto488; Figures S2A: Xist-Atto488; Figures S2C: XIST-Atto488; Figures S2E: Xist-Atto488, Xist Intron1-Cy3; Figures S5A and S5B: Xist-Atto488; Figure S6B and S6D: Xist-Atto488. All BACs and fosmids used in this study were purchased from CHORI-BACPAC.

Oligo FISH Probe design and synthesis—We selected 20 early (half time [0–0.4] days) and 20 very late (half time [1–1.6] days) silencing genes from published silencing kinetics data (Barros de Andrade et al., 2019) that are longer than 10kb and have high expression levels in ESCs. To ensure the assessment of genes across the length of the

chromosome, we selected no more than five genes per early or late silencing group from each 10Mb region of the X chromosome.

The resulting early silencing genes were: ENSMUSG00000025862, ENSMUSG00000055780, ENSMUSG00000036022, ENSMUSG00000023092, ENSMUSG00000000838, ENSMUSG00000016382, ENSMUSG00000025246, ENSMUSG00000025059, ENSMUSG00000035232, ENSMUSG00000050332, ENSMUSG00000079487, ENSMUSG00000034055, ENSMUSG00000056537, ENSMUSG00000046449, ENSMUSG00000031333, ENSMUSG00000031232, ENSMUSG00000025531, ENSMUSG00000025271, ENSMUSG00000041649, ENSMUSG00000025289.

The late silencing genes were: ENSMUSG00000031161, ENSMUSG00000040363, ENSMUSG00000031012, ENSMUSG00000031060, ENSMUSG00000001173, ENSMUSG00000063785, ENSMUSG00000025630, ENSMUSG00000031351, ENSMUSG00000002015, ENSMUSG00000031328, ENSMUSG00000031197, ENSMUSG00000006678, ENSMUSG00000035150, ENSMUSG00000034480, ENSMUSG00000041229, ENSMUSG00000045180, ENSMUSG00000046873, ENSMUSG00000067194,ENSMUSG00000079316,ENSMUSG00000031352.

We note that the first 40Mb of the chromosome are poor in early-silencing genes that provide high expression levels (RPKMs) for sufficient detection by oligonucleotide FISH. Custom fluorescent oligonucleotide probe pools (MyTags) targeting either genes or their corresponding transcripts were designed and synthesized by Daicel Arbor Biosciences. To detect nascent gene transcripts, ~500 45bp oligonucleotide probes targeting gene introns downstream from the transcriptional start site of the gene were synthesized spanning 31 to 68kb. To detect genes, ~500 45bp anti-sense oligonucleotide probes targeting upstream from the transcriptional start site were synthesized, spanning between 31 to 61kb. Probe sequences are given in Table S2.

RNA/DNA and immuno-RNA FISH—RNA and DNA FISH experiments were conducted as previously described (Markaki et al., 2013). Briefly, cells were grown on geltrex-coated high precision coverslips at the desired differentiation state. Coverslips were then transferred to new multi-well plates, washed three times with PBS and fixed with 3% formaldehyde dissolved in PBS for 10min, followed by two washes with PBS. Samples were then quenched for 10min with 20mM glycine in PBS. Following, samples were washed with PBS and permeabilized with 0.5% Triton X-100 dissolved in PBS for 15min, washed twice with PBST, equilibrated in 2xSSC for 10min and incubated for 30min to 2hrs with 50%formamide/ 2xSSC. Probes were denatured at 76°C for 7min and kept on ice. To hybridize the specimens, probes were spotted on slides and coverslips were placed on the probes and sealed with rubber cement. All probes were used at 0.1μg/cm^2 and oligonucleotide probes were used at 10pmol/cm^2. For DNA FISH, a denaturation step was performed for 2min at 76°C. Samples were then hybridized in a humidified chamber at 37°C ON. After demounting coverslips, unbound probes were washed-off with three 20min washes with 2×SSCT (2×SSC, 0.5% Tween 20) under mild shaking, followed by three 5min washes with 4×SSCT. For DNA FISH samples were washed with three

additional 5min washes with 0.1×SSC. Following, samples were post-fixed and chromatin was counterstained with DAPI as described in the Immunofluorescence staining section.

For sequential RNA and DNA FISH experiments with X chromosome paints (mmX paints) or oligonucleotide probes and Xist probes, RNA FISH was performed first, samples were post-fixed and DNA FISH followed. For the detection of SPEN proteins fused to GFP together with Xist RNA, detected by FISH probes, immunodetection was performed first, samples were post-fixed and RNA FISH followed.

Confocal laser scanning microscopy—Confocal and improved confocal (Airyscan detector) laser scanning microscopy was performed on the LSM880 platform equipped with 100x/1.46NA or 63x/1.4 NA plan Apochromat oil objectives and 405/488 diode and 594 Helium-Neon lasers (Carl Zeiss Microscopy, Thornwood, NY). To optimize imaging and reduce photobleaching, the regions of interest (ROIs) in each case were marked and the appropriate magnifications were used. The pixel size and z-optical sectioning were set to meet the Nyquist sampling criterion in each case. Airyscan raw data were reconstructed using the ZEN Black (v2.3) software.

For the detection of genomic regions across the X chromosome with spectral barcoding, cells were seeded on gridded coverslips and DNA FISH was performed first. 5-color optical z-stacks of 0.35μm were acquired on a confocal Zeiss LSM880 system. Grid coordinates were recorded, and spatial coordinates of the acquired positions were registered on the ZEN Black software and saved. Following, samples were equilibrated with 50% formamide in 2×SSC pH 7.2 solution for 3 hours at 37°C followed by RNA FISH with Cy3-labelled Xist probes. Specimens were returned to the microscope stage and saved spatial coordinates were revisited to acquire the *Xist* RNA signal and discriminate between the Xi and Xa in downstream analyses. Although hybridization of RNA usually precedes DNA FISH, we have found that *Xist* RNA is remarkably stable during the sequential process. Since the sequential hybridization for this experiment was only necessary for the scoring of the Xi, without the need for harsh probe strip-off steps, RNA FISH was performed last.

Detection of cell cycle stages—To discriminate between different cell cycle stages, we used a combination of EdU pulse labelling, to detect S-phase cells, and anti-histone H3-phospho-Serine10 (Active Motif, #39253), to detect G2/M phase cells, while G1 cells remained marker-free. EdU and click-iT labeling were performed according to the manufacturer's instructions. A 10mM EdU stock solution was diluted 1:1000 in growth media and cells were pulsed for 20min prior to fixation. RNA FISH with Xist probes was performed in the 488 channel and detection of EdU by click-iT reaction with CF dye Azide 568 (Biotium, #92082) was combined with immunodetection of phospho-histone H3 Serine 10 and secondary antibodies conjugated to CF568, where RNA FISH was performed first (Markaki et al., 2013). For the assessment of *Xist* foci features and number throughout the cell cycle in EpiLCs (at day 4 of differentiation), we used the $Xist^{MS2\text{-}GFP}$ cell line and detected $Xist^{MS2\text{-}GFP}$ signals after addition of 0.5μg/ml doxycycline for 2hrs (for MCP-GFP induction).

Super-resolution microscopy—3D-Structured Illumination Microscopy (3D-SIM) was performed on a DeltaVision OMX-SR system (Cytiva, Marlborough, MA, USA) equipped with a 60x/1.42 NA Plan Apo oil immersion objective (Olympus, Tokyo, Japan), sCMOS cameras (PCO, Kelheim, Germany) and 405, 488, 642nm diode lasers and a 568nm DPSS laser. Image stacks were acquired on the OMX AcquireSR software package 4.4.9934 with a z-steps of 125nm and with 15 raw images per plane (five phases, three angles). Raw data were computationally reconstructed with the soft-WoRx 7.0.0 software package (Cytiva, Marlborough, MA, USA) using a Wiener filter set at 0.001 to 0.002 (up to 0.006 for DAPI) and optical transfer functions (OTFs) measured specifically for each channel using immersion oil with different refractive indices (RIs) as described in (Demmerle et al., 2017; Kraus et al., 2017). Images from different channels were registered using alignment parameters obtained from a calibration slide of 100nm gold grid holes and a second calibration for axial alignment using 100nm diameter Tetraspeck beads according to established procedures (Demmerle et al., 2017).

Live-cell imaging—Wide-field and confocal scanning microscopy (for FRAP experiments) or 3D-SIM live-cell imaging (4D-SIM) were performed at 37°C (for 3D-SIM in conjunction with an objective heater), with 5% CO_2, controlled humidity and 10%O_2, having equilibrated the system and immersion oils for at least five hours prior to acquisitions. This equilibration was particularly important for obtaining artifact-free 3D-SIM datasets and minimize stage drift. Cells were differentiated in geltrex-coated chambers fitted with a high precision glass (ibidi) with daily exchange of media. To induce MCP-GFP expression, doxycycline was added to the cells two hours prior to acquisitions at a concentration of 1μg/ml. Imaging was performed in media containing no phenol red and supplemented with ProlongLive Antifade reagent (Thermo Fisher). For live-cell 3D-SIM imaging, typically 1μm to 2μm stacks of 125nm z-sections were acquired in 1- or 2-color 3D-SIM imaging to obtain 240–500 raw images per frame in 5–8 second intervals depending on exposure times and z-depth. Photobleaching over time was corrected by using histogram matching on the BleachCorrection plugin in ImageJ/Fiji.

FRAP experiments—FRAP experiments with z-sectioning for $Xist^{MS2\text{-}GFP}$ and CIZ1-mCherry were performed on an LSM880 equipped with an Airyscan on a Plan-Apochromat 63×1.4NA oil immersion objective, an image size of 67.5μm x 67.5μm with a pixel size of 0.085μm. Z-optical stacks of 0.5μm were obtained through a 15μm z-depth. Bleaching was performed in ROIs demarcating the *Xist* territory or corresponding nuclear (control) regions at full laser power and 4 iterations with a pixel dwell time of 4.04μsec. The first post-bleach frame was acquired immediately after bleaching. Time series were acquired every 1.3min up to 10 frames and every 2min thereafter for a total of 30min with an Argon ion 488nm laser or a DPSS 561nm laser set to 1% laser power.

Single-plane FRAP experiments for all other proteins were performed on the OMX-SR platform in widefield mode and an image size of 512×512 pixels with a pixel size of 0.08μm. In these experiments, we employed transgenic cells lines carrying mCherry-tagged CIZ1 and CELF1 and Halo-tagged FL-/ΔIDR-/ΔRMM-SPEN, PCGF5 and PTBP1, respectively (Figures 4F, S5D and S6F). Images were acquired for $Xist^{MS2\text{-}GFP}$ in the 488nm

channel (95MHz- 6% amplitude, 20msec) and for all mCherry- or Halo-fused-TMR labelled proteins in the 568nm channel (272MHz, 6% amplitude, 50–100ms exposure). Bleaching in ROIs demarcating the *Xist* territory or nuclear (control) regions was performed by using the 568nm laser line in the Ring-TIRF/PK photokinetics module with a bleach spot of 1μm for one iteration for 0.1sec.

RNA-Antisense purification (RAP)-seq—F1 2–1 female mouse ESCs were seeded on geltrex-coated plates and differentiated for 2 or 4 days. 5×10^6 cells were collected per condition after dissociation by accutase and RAP-seq was performed (Engreitz et al., 2013). Briefly, harvested cells were incubated for 45min with 2mM DSG in PBS at RT, crosslinked with 3% formaldehyde for 10min and quenched with 500mM glycine. Following, cells were pelleted at 1,500xg for 5min and flash frozen. Pellets were resuspended in 10ml nuclear extraction buffer LB1 containing 50mM HEPES-KOH (pH 7.5), 140mM NaCl, 1mM EDTA, 10% (vol/vol) glycerol, 0.5% (vol/vol) NP-40/Igepal CA-630 and 0.25% (vol/vol) Triton X-100 and incubated for 1,500xg for 5min and flash frozen. Pellets were resuspended in 10ml nuclear extraction buffer LB1 containing 50mM HEPES-KOH (pH 7.5), 140mM NaCl, 1mM EDTA, 10% (vol/vol) glycerol, 0.5% (vol/vol) NP-40/Igepal CA-630 and 0.25% (vol/vol) Triton X-100 and incubated for 1mM EDTA, 0.5mM EGTA, 0.1% (wt/vol) sodium deoxycholate and 0.5% (vol/vol) N-lauroylsarcosine and sonicated on ice using a Misonix S-400 sonicator with microtip for 2min with 1sec pulses intermitted by 3sec pauses. Chromatin was then digested using TURBO DNase at a concentration of 0.1–0.4U/μl at 37°C for 15min. For each pulldown library, an input library was also generated. The RNA pulldown was performed using 50pmol of 90nt long biotinylated oligonucleotide probes for every 5×10^6 cells and Streptavidin C1 beads. DNA was eluted by RNase H digestion, and the crosslinks were reversed by proteinase K digestion of the eluted DNA at 60°C. The DNA libraries were prepared using NEB Next Ultra End Rpair/dA-Tailing Module (NEB) and TruSeq DNA adapters (Illumina) were ligated using Quick Ligase (NEB). Libraries were amplified by KAPA HiFi Polymerase, pooled and sequenced on the Illlumina HiSeq 6000 platform to generate 50bp pair-end reads. Probe sequences are given in Table S2.

Single cell RNA-seq—scRNA-seq was performed inF1 2–1 $Xist^{129\text{-}WT/Cas\text{-}WT}$ and $Xist^{129\text{-}\Delta B/Cas\text{-}WT}$ ESC lines at days 2 and days 4 of EpiLC differentiation. For the $B6^{tetO\text{-}Xist}Cas^{WT\text{-}Xist}$ TX1072 ESC lines expressing SPEN-GFP or ΔSPOC-SPEN-GFP (Dossin et al., 2020) and ΔIDR-SPENGFP generated in this study, scRNA-seq was performed without (0h) or with (24hr) of 0.5μg/ml doxycycline to induce tetO-Xist expression.

Cells were dissociated with accutase for 5min, washed 3 times with PBS and resuspended in PBS containing 0.04% BSA. Cell concentration was adjusted between 800 to 1200 cells/μl and cell suspension was kept on ice before loading on the 10x Genomics Chromium instrument. scRNA-seq libraries were generated using the Chromium single cell 3’ reagent kit V3.1. Individual libraries were designed to target up to 10,000 cells. Libraries were generated following manufacturer’s instructions and library fragment size distribution was determined by BioAnalyzer. Libraries were pooled and sequenced on the Illumina Novaseq 6000 platform to generate 100bp pair-end reads.

Bulk RNA-seq—Bulk mRNA-seq libraries were generated from $B6^{tetO\text{-}Xist}Cas^{WT\text{-}Xist}$ TX1072 ESC lines homozygously expressing WT-SPEN-AID-GFP (Dossin et al., 2020), with no *Rosa26* knockin rescue, full length (FL)-SPEN *Rosa26* knockin rescue, or ΔIDR-SPEN *Rosa26* knockin rescue, at 0h and 24h dox treatment and upon auxin treatment. Harvested cells were washed with PBS and collected into TRI reagent. Lysates were processed immediately or snap-frozen in liquid nitrogen and stored at −80°C for up to 3 days. All conditions were collected from three biological replicates where samples for each replicate were processed at the same time. RNA was isolated using the Zymo Research RNA miniprep isolation kit according to manufacturer's instructions. RNA-seq libraries were prepared using the TrueSeq Stranded mRNA Library Prep Kit according to manufacturer's instructions. Libraries were pooled and sequenced on the Illumina Novaseq 6000 platform to generate 100bp pair-end reads.

CLAP-seq—CLAP-seq libraries were generated from $B6^{tetO\text{-}Xist}$ $Cas^{WT\text{-}Xist}$ TX1072 derivative ESC lines expressing the FL-, ΔIDR- or ΔRRM-Halo-SPEN at day 4 of differentiation with the addition of 0.5μg/ml doxycycline for the last 24hrs to induce tetO-Xist expression. For each pulldown library, an input library was also generated. ~5×10^7 cells were collected for each replicate. Cells grown on gelatin-coated 150cm^2 culture dishes were washed three times with ice-cold PBS and UV-cross-linked on ice using 0.25 Jcm−2 (UV2.5k) at 254 nm in a Spectrolinker UV Crosslinker. Cells were then collected through scraping on ice and centrifuged at 1,500g for 5min, washed once with PBS and pelleted by centrifugation in aliquots of 5×10^6 cells. Pellets were snap frozen in liquid nitrogen for storage at −80°C. CLAP-seq using the HaloLink Resin to pulldown the Halo-tagged proteins was performed as previously described (Quinodoz et al., 2020). In brief, each pellet was lysed for 10min in 1ml Lysis buffer containing 50mM HEPES, pH 7.4, 100mM NaCl, 1% NP-40, 0.1% SDS, 0.5% Sodium Deoxycholate, 1× Protease Inhibitor Cocktail (Promega), 200U of Murine RNase Inhibitor, 20U Turbo DNase and 1X Manganese/Calcium Mix (0.5mM CaCl2, 2.5mM MnCl2) at 37°C. Lysates were pelleted at 800×g for 8min at 4°C, the supernatant was removed and pellets were resuspended in 1ml Lysis buffer and sonicated for 30sec with 0.7sec pulses intermitted by 2.3sec pauses. After removal a 10min incubation at 37°C, samples were pelleted at 15,000×g for 2min and the resulting supernatant was collected and stored on ice. The HaloLink Resin was washed three times in Wash buffer (1×PBS, 0.1% Triton) and blocked for 20min in Blocking buffer (50mM HEPES, pH 7.5, 10μg/ml Random 9-mer, 100μg/ml BSA). Following the HaloLink Resin was incubated with the supernatant for 3–16hrs under rotation at 4°C. After incubation, three washes in Wash buffer were performed at RT and three additional washes at 90C for 2min with the following buffers: NLS buffer (50mM HEPES pH 7.5, 2% NLS, 10mM EDTA, 0.1% NP-40, 10mM DTT), High Salt Buffer (50mM HEPES pH 7.5, 10mM EDTA, 0.1% NP-40, 1M NaCl), 8M Urea Buffer (50mM HEPES pH 7.5, 10mM EDTA, 0.1% NP-40, 8M Urea) and Tween buffer (50mM HEPES pH 7.5, 0.1% Tween 20, 10mM EDTA). The HaloLink Resin was then equilibrated by washing three times with Elution buffer (50mM HEPES pH 7.5, 0.5mM EDTA, 0.1% NP-40) at 30C and elution was performed by resuspension in 100μl of NLS buffer containing 10% Proteinase K and shaking at 50°C for 30min. Successful elution of the Halo-fused proteins was tested by Western blotting. After elution RNA overhangs were repaired by treatment with FastAP and T4 Polynucleotide Kinase with no ATP at 37°C for

15min and 1hr, respectively. The RNA was then reverse transcribed using Superscript III according to the manufacturer's instructions. Following, the cDNA was amplified by PCR using illumina sequencing adaptors libraries, pooled and sequenced on the Illlumina HiSeq 6000 platform to generate 100bp pair-end reads.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantitative 3D-SIM analyses—For image segmentation and image objects (particles) determination, 32-bit raw datasets were imported into ImageJ/Fiji (Schindelin et al., 2012) and converted to 16-bit tiff composite stacks. The segmentation of *Xist* and protein particles (foci) was performed as previously described (Kraus et al., 2017) using the TANGO suite (Ollion et al., 2013). Specifically, raw datasets without filtering or subtraction of signals were imported into the segmentation pipeline. Image segmentation pipelines, adjustment of thresholds and creation of seeds were performed using high-throughput batch-processing and without manual intervention. Resulting masks of segmented particles were inspected by overlays over the raw data to ensure that the majority of signals was contained in the area to be analyzed. Nuclear masks were created using the DAPI channel as the segmentation volume. For each channel, a duplicate was generated and filtered with a 3D Gaussian filter with standard deviation of 1 (σ=1) and a Tophat filter with a radius of two pixels in xy and a one-pixel radius in z. The filtered image was segmented using the 3D Suite's Watershed method. Seed threshold and image threshold for watershed were calculated by equations Mean+StdDev*2*seed multiplier and (Mean+StdDev*2*seed multiplier)/image multiplier (Signal-to-Noise Ratio, SNR), respectively, where seed multiplier and image multiplier were determined and inspected manually to ensure the inclusion of all the regions of interest (ROIs) and the removal of background noise. Object features and distance measurements were performed using the 3D ImageJ Suite's "Measure 3D", "Quantif 3D" and "Distance" option plugins for ImageJ/Fiji.

For the assessment of the cage^{60GFP} versus *Xist* signals, *Xist*$^{MS2\text{-}GFP}$ cells expressing the cage^{60GFP} plasmid were typically imaged in the same Field of View (FOV) as cells with the *Xist*$^{MS2\text{-}GFP}$ signal, allowing us to obtain data that could be directly compared. When cells expressed both entities, since the cages are located in the cytosol in the majority of cells, nuclear masks from the DAPI channel were created and *Xist*$^{MS2\text{-}GFP}$ signals were measured inside the masked regions, whereas the signal from the GFP-expressing cages was measured outside the nuclear masks. For the comparison of intensities of cages expressed in the cytosol or in the nucleus the same masking procedure was applied. The variability of the integrated density of fluorescent cages imaged under 3D-SIM conditions was compared to that obtained by a similar analysis in the original publication (Hsia et al., 2016) and found that results are within the same range.

The number of *Xist* foci is typically lower when detecting the RNA via MCP-GFP compared to detection by RNA FISH using probes that cover the entire spliced transcript. Our labelling with FISH probes captures the entire RNA molecule in contrast to MCP-GFP, which detects a region in the last large exon of *Xist* downstream to the E-repeat sequence. Therefore, RNA FISH-based detection results in slightly larger and possibly more complex *Xist* structures. Moreover, the different average number of *Xist* foci between cells at D2, D4, D8 of

differentiation or in C127 cells, defined across the cell population without discriminating cell cycle stage (Figures 1L, S1I and S2B), is likely a reflection of different populations of cells across the different cell states.

For the comparison of the integrated densities of SPEN-GFP and cage^{60GFP}, cells expressing the cage^{60GFP} plasmid and TX1072 (B6$^{tetO\text{-}Xist}$Cas$^{WT\text{-}Xist}$) derivative cells expressing SPEN-GFP after addition of 0.5μg/ml doxycycline, to induce expression of the tetO-Bgl-mCherry-*Xist* (Dossin et al., 2020), were imaged under the same settings and datasets processed with the same threshold to define seeds and objects (particles). We found that the increase in *Xist*-associated SPEN accumulation during differentiation, observed from the D2 to D4 transition, also occurs similarly upon doxycycline induction of *Xist* expression between 6 and 18hrs of *Xist* induction. Moreover, the levels of SPEN in SMACs are the same upon tetO-Xist induction at 18hrs as for WT-Xist at D4 of differentiation (not shown).

To extract global nuclear protein particle features (in and outside the Xi) such as mean intensity, integrated density (amount of fluorescence per defined particle volume) and volume, masks of the protein signals of interest were created by filtering raw data with a 3D Gaussian blur followed by automatic thresholding to include all signals and exclude nucleoli. ROIs within a 4μm radius of *Xist* centroids were selected for features extraction to limit computation time to ~1 hour per nucleus. Nearest neighbor centroid distances and all distances between ROIs within each channel and across different channels were extracted using the 3D ImageJ Suite for minimal distance and average distance analysis, respectively. Distance averaging was performed in Python. Assignment of *Xist*-associated signals was based on a proximity threshold to *Xist* centroids with a radius of 250nm. Signals 500nm away from *Xist* centroids, resulting in a ‘rim’ around the Xi due to the scattering of many *Xist* foci throughout the Xi, were defined as the nuclear fraction. To test the specificity of the tight spatial association of *Xist* and its associated protein particles in the Xi, we performed randomized controls (Figure S4D). We first determined the Xi territory by generating voxel arrays corresponding to a 350nm radius from *Xist* foci centroids. The overlapping (double-called) voxels were removed. We then generated random positions equal to *Xist*-associated protein particles within the Xi masks of each nucleus using Python’s numpy.random.choice function. We found a statistically significant difference between the nearest-neighbor *Xist*-associated protein particle to *Xist* compared to the randomized positions to *Xist*. The comparison of protein features, such as integrated density of fluorescence and volume, was performed by measurements acquired in the same laser line (568nm) for all proteins detected either with the Halo ligand JF549 or primary and secondary antibodies conjugated to CF568 dye. For each experiment, ROIs with integrated density and volume values below the 10th percentile or above the 90th percentile of the dataset were removed as outliers.

For the oligo-FISH analysis, signal centroid from 3D-SIM data were extracted as described above using the 3D ImageJ Suite. Signal coordinates were imported into Python using Pandas and separated by cell and signal type (late, early, DNA, RNA, *Xist*) into NumPy arrays. We detected on average ~10–30 foci for each gene pool per X chromosome. Numbers in the higher range can likely be explained by the detection of some genomic regions as doublets after DNA replication or the segmentation of extended structures. *Xist*

cluster centroids were calculated for each nucleus using the corresponding *Xist* array and centroid formula:

$$G_{x,y,z} = \left(\frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}, \frac{y_1 + y_2 + y_3 + \cdots + y_n}{n}, \frac{z_1 + z_2 + z_3 + \cdots + z_n}{n} \right)$$

Distance to *Xist* centroid was then calculated using the spatial coordinates of each gene centroid to *Xist* centroid for the corresponding nucleus using the distance formula. Gene transcriptional activity was determined by pairing DNA signals to RNA signals in each nucleus. Nearest neighbors (NN) of DNA and RNA were determined by calculating the distance between DNA and RNA centroids. DNA centroids with a NN RNA closer than 350 nm were paired and labeled as “active genes” while those with a NN farther away were labelled as “silent genes”.

X-territory volume and sphericity measurements—Confocal optical stacks were imported to ImageJ/Fiji and converted to 16-bit tiffs. Raw data were processed using the “Smooth” function and an automatic threshold, using either the Yen or Otsu method, was set to create 3D masks for the X chromosome territories. Assignment of the Xi was based on RNA FISH signals from the *Xist* channel. Masks were imported into 3D Suite and the volume and sphericity measurements of the X chromosomes (Xa and Xi) were extracted. Sphericity is defined as the length of the object over its width, with a maximum value of one.

Extraction of X chromosome configurations—Confocal optical stacks from sequential rounds were imported into Fiji/ImageJ and superimposed and alignment of the two sequential rounds was performed with the affine transformation of the StackReg plugin based on the DAPI channel. Data were smoothed with a 3D Gaussian blur with a standard deviation of 1 (σ=1) and background removal was performed using the “Subtract Background” plugin with a rolling ball radius of 10 pixels. The Xa and Xi (scored by the presence of *Xist* RNA) were identified and saved as separate stacks. Subsequently, each probe signal centroid was extracted using the “3D Object Counter” plugin. The 3D Object Counter generated a list of coordinates of probe signals for each channel. To assign signals to multi-spectral barcodes consisting of two labels, a nearest neighbor search between the two corresponding channels was applied based on all spatial coordinates in each channel. Once pairs of signals were assigned to the multi-spectral barcodes the coordinates obtained in the shortest wavelength were used. In cases where two adjacent signals were detected per probe, potentially due to the presence of transcripts or DNA replication, only one of the signals was used.

The coordinates of individual barcodes for the Xa and Xi at days 2 and 4 of EpiLC differentiation were reoriented in 3D space to compute spatial statistics across all cells. To obtain configurations of chromosomal backbones, for each set of probe coordinates, principal component analysis (PCA) was performed in the x and y axes using MATLAB. The z axis was unused as the segmentation resolution in that axis is significantly lower, contributing to large variations in the z coordinate (Finn et al., 2017) and confounding the reorientation method used, which is highly sensitive to anisotropic error. The principal

component is assumed to be the "backbone" of the chromosome: the expected orientation of a chromosome if initially stretched out along that component's direction before entropically relaxing into an equilibrium configuration. Each set of probes is rotated in order to align its corresponding principal components with the y-axis and translated such that the probes' centroid is aligned with the coordinate origin. Probes of the same loci were then statistically compared to locate their local spatial centroid and 95% confidence interval for Xa day 2, Xi day 2, Xa day 4, and Xi day 4 separately. Ellipsoids encompassing the 95% confidence interval were plotted around each loci centroid. In order to quantify the relative compaction between Xa and Xi from day 2 to day 4, the pairwise distances of 3D coordinates (x,y,z) between each barcode location were measured and averaged over all cells. Averages of Xa distances were subtracted from those of Xi at day 2 and the same was done for day 4 in order to measure the absolute change between chromosomes. A heatmap of this change was plotted, where large negative numbers indicate a higher compaction.

Single-Particle Tracking (SPT) of *Xist* foci—Individual *Xist* particles from live-cell 3D-SIM data were extracted using TrackMate (Tinevez et al., 2017), an ImageJ plugin. DoG Detector with a 0.2μm diameter was used to define the particles and the Simple LAP Tracker with 0.25 max linking distance, 0.3 gap-closing max distance and 2 gap-closing max frame gap were used to track the particles and generate trajectories. Trajectories that were not possible to track for over 10 consecutive frames were not used. Over 850 trajectories from 30 cells were analyzed and approximately 50% of all *Xist* foci could be tracked without manual intervention per nucleus for an average of 2min. To characterize the motion of *Xist* foci, the data extracted from the software were fed into downstream confinement analyses described in Methods S1 File (see section on "*Xist* foci position trajectories and effective confining potentials").

4D-SIM image registration—Two types of motion are captured simultaneously in 4D-SIM microscopy: a) the developmental motion of the cell and the nucleus, and b) the individual motion of *Xist* foci within the nucleus. To specifically extract (b) from live-cell 3D-SIM images we implemented the following processing pipeline using Python: 1) A set of *Xist* foci were tracked using TrackMate and spatial coordinates were extracted and imported into Python using PyTrackmate. 2) For each timestep t, the individual displacement vectors $\boldsymbol{x}_{i,t}$ of each *Xist* focus were calculated using NumPy and Pandas. 3) For each time-step, individual displacement vectors were averaged to obtain $\boldsymbol{X}_t$, an approximation of the developmental motion in that time-step. 4) This developmental motion was then subtracted from the displacement vector of each *Xist* focus to arrive at an approximation to granule i 's motion, $\boldsymbol{x}_{i,t}$ -$\boldsymbol{X}_t$.

H2B density classification and *Xist* localization—Histone H2B-Halo^{JF646} intensities were extracted on ImageJ/Fiji plugin using the "getValue" macro command, that iterates over every pixel in the image to get the intensity value of each pixel, generating a list of all the pixel intensities and their corresponding coordinates. The list of intensities was imported to Python. Then, seven intensity/density classes of equal variance were determined. The 3D Suite was used to create *Xist* masks, and *Xist* trajectories were extracted from TrackMate to obtain spatial coordinates (centroids) from each time point. The matrices were paired

within the radius of one pixel and chromatin density classes were measured under the masks. Radial distances were measured at all pixels within the respective 100, 250, 500 nm radius of the *Xist* centroid and the maximal intensity value within that range was defined. Averaged values were then plotted in a line graph as a function of time. To extract the nearest neighbors (NN) in chromatin density maps, neighboring intensities for each H2B pixel were determined as the average intensity of all adjacent pixels and stored in an array. A strip plot was used to plot the averaged intensity values where each value was assigned to one of the 7 classes based on the class of the origin pixel (Figure S2L).

FRAP data analysis—FRAP time series were imported into Fiji and converted to 16-bit tiffs. Datasets from *Xist* and CIZ1 derived from z-stacking were projected for each timepoint. To correct for drift, images were registered using the Correct3DDrift plugin and datasets that could not be registered were discarded. To measure FRAP recoveries data were normalized for fluorescence decay as described (Dundr and Misteli, 2003). In brief, ~2μm user-defined ROIs were created to define and measure the bleached region (I_t), a randomly selected unbleached nuclear region (T_t), and a randomly selected region outside the cell (B_t) for each timepoint. A relative intensity for each timepoint was calculated using the equation: $I\text{rel} = (T_0 - B_0) \times (I_{(t)} - B_{(t)})/(T_{(t)} - B_{(t)}) \times (I_0 - B_0)$ where T_0 and I_0 are derived from the average intensity of the region of interest during prebleach. To derive the FRAP recovery, I_{rel} was measured through time. For compiling figures, FRAP time series were bleach-corrected using the BleachCorrect ImageJ/Fiji plugin. To infer dissociation rates and residence times, FRAP curves for *Xist* and all proteins were fit to single or double exponential models derived from mass-action kinetics. Squared errors were minimized to obtain best-fit detachment rates, binding site densities, and freely diffusing fractions described in the Methods S1 file (“Mass-action binding and dissociation model of FRAP dynamics”).

Image data statistical analysis and visualization—Data analysis and visualization were performed using Python and executed in Google Colaboratory. All violin plots, boxplots, bar plots and point-plots were generated using Seaborn and Matplotlib. NumPy and SciPy were used for mathematical computation and Pandas for data analysis. Unless stated otherwise, all graphs show the median as the central point or the central line, and bars on point plots represent the standard deviation. Point plots of protein integrated density and volume in Figures 3H, S3B, S5J, S5K and S6D show the percentage of the maximum absolute value in each group. Statistical differences between two groups were analyzed by the two-sided Wilcoxon’s or MannWhitney rank-sum test (scipy.stats.mannwhitneyu). The Kruskal-Wallis H-test was used for statistical comparisons between multiple groups (scipy.stats.kruskal). Statistical significance was defined as a p-value less than 0.05.

RAP-seq data analysis—RAP-seq reads were trimmed using trim_galore (v0.4.1) with default parameters to remove the standard Illumina adaptor sequences. Bowtie2 (v2.2.9) was used to align reads to the mouse genome (mm9) with the default parameters. Reads with mapping quality less than 30 were removed using samtools (v1.2). Picard MarkDuplicates (v2.1.0) was used to remove PCR duplicates. Bedtools intersect (v2.26.0) was used to count reads in sliding windows (100Kb every 25Kb) along the X chromosome. *Xist* localization

across the X chromosome was defined by calculating the *Xist* enrichment scores (*Xist* pulldown/input) in the sliding windows. Unmappable regions were masked.

MACS2 (Zhang et al., 2008) was used to call broadPeaks in D2 and D4 *Xist* pulldown data, using the input as control. Peaks within 50kb were merged. To identify overlapping *Xist* peaks between D2 and D4, the respective X-linked peak sets were intersected using bedtools intersect. The most significant peaks were selected with a log10(qvalue)>13 cutoff. The R package makeVennDiagram with connectedPeaks=min was used to make the Venn Diagram from the two peak ranges.

To plot the average *Xist* enrichment around the summits of *Xist* peaks, *Xist* peak summits defined for D2 were used and filtered for summit scores>100 on the X chromosome. The RAP-seq enrichment scores for 2500kb up and downstream of these D2 summits were extracted from the D2 and D4 data using normalizeToMatrix function in the R package EnrichedHeatmap.

Hi-C compartments—To compare Hi-C compartmentalization and *Xist* enrichment, Hi-C PC1 values were downloaded from GSE99991 (Wang et al., 2018). Analysis was performed in R (v3.6.0). Hi-C PC1 values from undifferentiated ESCs were correlated with *Xist* enrichment at D2 and Hi-C PC1 values at differentiation D4 with *Xist* enrichment at D4. In addition, *Xist* enrichment at D2 and D4 were correlated. Datasets were intersected with plyranges (v1.4.4) (Lee et al., 2019). Plots were made with ggplot2 (v3.3.2) and pearson's correlation coefficients and p-values (two-sided t-test, r≠0) were calculated with the function stat_cor in the R package ggpubr (v0.4.0).

SNP Calling—Single nucleotide polymorphisms (SNPs) were identified to distinguish the *mus castaneous* (Cas) genome from the 129S4/SvJae (129) or C57BL/6J (B6, reference) strains of *mus musculus*. Parental genome sequencing data were downloaded from publicly available databases (Cas genome sequence (EMBL-EBI: ERP000042); 129 genome sequence (SRA: SRX037820). In order to identify distinguishing SNPs, we first aligned whole genome sequencing reads of each strain to the mm10 genome with bowtie2 (v2.3.5.1), filtered multimapped reads (mapq<10) with samtools (v1.7), and removed duplicates with GATK (v4.1.4.1) MarkDuplicates. For allelic analysis in Cas/129 F1 hybrid cells, we jointly called SNPs with the aligned Cas data and the aligned 129 data and filtered out low quality SNPs with bcftools (v1.8). For allelic analysis in Cas/B6 F1 hybrid cells, we called SNPs with the aligned Cas data alone and filtered out low quality SNPs with bcftools. SNPs were required to have a minimum depth of 5, with at least 90% of the reads supporting the alternate allele, a minimum of 3 reads supporting the alternate allele and no more than 2 reads supporting the reference allele. Indels and complex SNPs were not included. To filter the Cas/129 SNPs, SNPs that were alternate in both strains were excluded. When a SNP is alternate in one strain, the other strain must have fewer than 10% of reads supporting the alternate allele, at least 2 reads supporting the reference allele, and fewer than 2 reads supporting the alternate allele. The resulting Cas/129 and Cas/B6 variant call files (VCFs) were used in subsequent allelic analysis.

Single cell RNA-seq data analysis—Cellranger count (v5.0.1) was used to align and process scRNA-seq reads to the mm10 genome with the option [--include-introns]. Seurat (v3.9.9.9024) in R (v3.6.1) to identify and exclude low quality cells, defined by a high percent of reads aligning to the mitochondrial genome (>7%) or a low number of unique feature counts. As the feature count per cell is partially a function of read depth per library, the threshold used was based on the number of features in the bulk of the cells in each library and ranged between 1500 and 2000. Vartrix (v1.1.14) was used to identify and count reads with informative SNPs (allelic reads) with the options [--scoring-method coverage –umi]. Either the Cas/129 VCF or the Cas/B6 VCF described above were used, as appropriate for the strain. All subsequent analyses were done in R. Allelic reads from the X chromosome were summarized for each gene to generate Xi allelic ratios (Xi/(Xi+Xa) reads) per X-linked gene. Lowly expressed genes were filtered out, by requiring a minimum of 10 cells in each library to have at least one allelic read of that gene. Xi allelic ratios for each gene were calculated per cell, requiring a minimum of 3 allelic reads to get an Xi ratio for a given gene in a cell.

For scRNA-seq data of the female F1 2–1 $Xist^{129\ WT\ /\ Cas\ WT}$ or $Xist^{129\ \Delta B\ /\ Cas\ WT}$ EpiLCs, cells with biallelic or undetectable *Xist* expression (<3 allelic reads), and XO cells, were excluded, and the Xi was determined based on monoallelic *Xist* expression. Xi allelic ratios between WT^{129} Xi and ΔB^{129} Xi were compared using a one-sided welch t test ($p<0.05$) with the R function t.test, where each cell is a replicate.

For the scRNA-seq data of TX1072 ($B6^{tetO\text{-}Xist}$ $Cas^{WT\text{-}Xist}$) ESC lines expressing WT-SPEN-GFP, ΔIDR-SPEN-GFP, and ΔSPOC-SPEN-GFP at 0h and 24h of dox treatment, where *Xist* is induced by doxycycline from the B6 allele, XO cells were excluded. All cells were assumed to inactivate the B6 X chromosome. The Xi ratios between 0 and 24 hours dox induction of *Xist* for the respective cell line were compared using a one-sided welch t-test ($p<0.05$) with the R function t.test, where each cell is a replicate.

The mean Xi allelic ratios per gene were found for each set of cells and plotted with ggplot2. Statistical tests comparing distributions were performed with the function stat_compare_means in the R package ggpubr. The Wilcoxon rank sum test was used for pairwise comparisons, while the Kruskal-Wallis test was used to compare multiple groups.

Bulk RNA-seq data analysis—Reads were aligned to the mm10 genome with STAR (v2.7.1a), filtered multimapped reads (mapq<10) with samtools (v1.7) and allelic ratios were found using GATK (v4.1.4.1) ASEReadCounter and the Cas/B6 VCF described above. Allelic reads were assigned to genes using bedtools intersect with all genes from the gencode mm10 annotation file (gencode.vM24.annotation.gtf). All subsequent analyses were done in R. Due to the tetO Xist promoter being on the B6 allele, the B6 X chromosome was inactivated in all cells. Allelic reads were summed across genes and the allelic ratios (Xi/(Xi+Xa) reads) were found. Genes with fewer than 10 reads with informative SNPs (allelic reads) were filtered out from each sample. Xi ratios between 0h and 24h dox were compared with a one-sided welch t-test ($p<0.05$) with the R function t.test, with 3 replicates per condition. The mean Xi allelic ratios per gene were found for each condition and plotted with ggplot2. Statistical tests comparing distributions were performed with the function

stat_compare_means in the R package ggpubr. The Wilcoxon rank sum test was used for pairwise comparisons, while the Kruskal-Wallis test was used to compare multiple groups.

X-linked gene silencing dynamics—The silencing half-times of genes during XCI were downloaded from (Barros de Andrade et al., 2019) and analyzed in R. X-linked genes with a half-time range of 0–0.4 days were classified as early-silencing genes, of 0.4–1 days as late-silencing genes, of 1–2 days as very late silencing genes, and >2 days as escapee genes.

CLAP-seq data analysis—CLAP-seq libraries of FL, ΔIDR and ΔRRM SPEN were aligned to the mm10 genome with STAR (v2.7.1a), filtered multimapped reads (mapq<10) with samtools (v1.7), duplicate reads were removed with GATK (v4.1.4.1) MarkDuplicates. RPKM values were calculated for each sample in 100bp bins smoothed across 300bp with deeptools, only counting the forward strand in order to avoid double counting fragments, (v3.5.0) with the options [--normalizeUsing RPKM --binSize 100 --samFlagInclude 64 --skipNAs]. Enrichment was calculated in R by dividing SPEN pulldown RPKM over input RPKM in each bin. The R package ggplot2 was used to plot the enrichment scores.

***Xist*-tethered SPOC silencing—**To identify genes that were repressed by SPOC tethered to *Xist* in the absence of SPEN, we used previously published allele-specific RNA-seq count data from GSE131784 of TX1072 $B6^{tetO\text{-}Xist}Cas^{WT\text{-}Xist}$ ESC lines expressing SPEN-AID-GFP, where depletion of SPEN was rescued through tethering of BglG-GFP-SPOC or BglG-GFP (as a control) to tetO-*Xist*–Bgl stem–loop RNA via BglG (Dossin et al., 2020).

SPOC silencing values for each gene were calculated as described (Dossin et al., 2020). Briefly, we filtered out genes that were skewed or not silenced under control conditions. We then calculated a silencing index under normal conditions ($silencing\_index_{DOX} = 1 - (allelic\_ratio_{DOX}/allelic\_ratio_{control})$) and after degrading SPEN and expressing SPOC tethered to *Xist* ($silencing\_index_{SPOC} = 1 - (allelic\_ratio_{SPOC}/allelic\_ratio_{control})$). To quantify the silencing defect in cells expressing only *Xist*-tethered SPOC, we calculated the silencing defect ($silencing\_defect = 1 - (silencing\_index_{SPOC}/silencing\_index_{DOX})$). We calculated the silencing defects for the control BglG-GFP rescue identically. The silencing defect per gene was found for each condition and plotted with ggplot2. Statistical tests comparing distributions were performed with the function stat_compare_means in the R package ggpubr. The Wilcoxon rank sum test was used for pairwise comparisons, while the Kruskal-Wallis test was used to compare multiple groups.

SMCHD1 sensitivity—To identify genes with a silencing defect as a result of the knockout of *Smchd1*, we downloaded RNA-seq allelic counts from GSE99991 and classified genes as SMCHD1-sensitive or -insensitive following the previously analysis pipeline (Wang et al). All analysis was performed in R. Briefly, the Xi allelic ratio (Xi/(Xi+Xa) reads) was determined per X-linked gene. Genes with fewer than 13 reads with informative SNPs (allelic reads), a skewed allelic ratio, or that escape X inactivation were filtered out. SMCHD1-sensitive genes were defined as having Xi allelic ratio in SMCHD1 knockout neural progenitor cells (NPCs) of 3-fold greater than Xi allelic ratio in WT NPCs.
The silencing half-time from (kinetics paper ref) was plotted for SMCHD1-sensitive and

-insensitive genes with the R package ggplot2, and distributions were compared with a Wilcoxon rank sum test using the function stat_compare_means from the R package ggpubr.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank David Baker for sharing the ct-60 gene and Yi-Yun Ho, Tsotne Chitiashvili, Amy Pandya-Jones and Mario Blanco for help with the study. We thank Lars Dreier, Douglas Black, Emilie Marcus, and Plath lab members for critical discussions. We thank the DGSOM at UCLA, David Williams and the Department of Biological Chemistry for support. The imaging was supported by the NIH (R01GM115233); YM by NIH (R03HD095086); KP by an Innovation Award from the BSCRC at UCLA, NIH (R01GM115233, 1R01MH109166, R21HD094172), the Keck Foundation and a HHMI Faculty Scholar grant; DM and TC by the NSF (DMS-1814364) and NIH (R01HL146552); and AB by NIH (F30HL136080).

References

Almeida M, Pintacuda G, Masui O, Koseki Y, Gdula M, Cerase A, Brown D, Mould A, Innocent C, Nakayama M, et al. (2017). PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. Science 356, 1081–1084. [PubMed: 28596365]

Banani SF, Lee HO, Hyman AA, and Rosen MK (2017). Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol 18, 285–298. [PubMed: 28225081]

Barros de Andrade ESL, Jonkers I, Syx L, Dunkel I, Chaumeil J, Picard C, Foret B, Chen CJ, Lis JT, Heard E, et al. (2019). Kinetics of Xist-induced gene silencing can be predicted from combinations of epigenetic and genomic features. Genome Res 29, 1087–1099. [PubMed: 31175153]

Beard C, Hochedlinger K, Plath K, Wutz A, and Jaenisch R (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis 44, 23–28. [PubMed: 16400644]

Berg BA, and Harris RC (2008). From data to probability densities without histograms. Computer Physics Communications 179, 443–448.

Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, and Long RM (1998). Localization of ASH1 mRNA particles in living yeast. Mol Cell 2, 437–445. [PubMed: 9809065]

Blelloch RH, Hochedlinger K, Yamada Y, Brennan C, Kim M, Mintz B, Chin L, and Jaenisch R (2004). Nuclear cloning of embryonal carcinoma cells. Proc Natl Acad Sci U S A 101, 13985–13990. [PubMed: 15306687]

Blewitt ME, Gendrel AV, Pang Z, Sparrow DB, Whitelaw N, Craig JM, Apedaile A, Hilton DJ, Dunwoodie SL, Brockdorff N, et al. (2008). SmcHD1, containing a structuralmaintenance-of-chromosomes hinge domain, has a critical role in X inactivation. Nat Genet 40, 663–669. [PubMed: 18425126]

Bousard A, Raposo AC, Zylicz JJ, Picard C, Pires VB, Qi Y, Gil C, Syx L, Chang HY, Heard E, et al. (2019). The role of Xist-mediated Polycomb recruitment in the initiation of X-chromosome inactivation. EMBO Rep 20, e48019. [PubMed: 31456285]

Boyle S, Flyamer IM, Williamson I, Sengupta D, Bickmore WA, and Illingworth RS (2020). A central role for canonical PRC1 in shaping the 3D nuclear landscape. Genes Dev 34, 931–949. [PubMed: 32439634]

Brackley CA, and Marenduzzo D (2020). Bridging-induced microphase separation: photobleaching experiments, chromatin domains and the need for active reactions. Brief Funct Genomics 19, 111–118. [PubMed: 31971237]

Braga J, McNally JG, and Carmo-Fonseca M (2007). A reaction-diffusion model to study RNA motion by quantitative fluorescence recovery after photobleaching. Biophys J 92, 2694–2703. [PubMed: 17259280]

Brockdorff N (2017). Polycomb complexes in X chromosome inactivation. Philos Trans R Soc Lond B Biol Sci 372.

Brockdorff N, Bowness JS, and Wei G (2020). Progress toward understanding chromosome silencing by Xist RNA. Genes & development 34, 733–744. [PubMed: 32482714]

Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, Rinn JL, and Raj A (2015). Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol 16, 20. [PubMed: 25630241]

Cerase A, Armaos A, Neumayer C, Avner P, Guttman M, and Tartaglia GG (2019). Phase separation drives X-chromosome inactivation: a hypothesis. Nat Struct Mol Biol 26, 331–334. [PubMed: 31061525]

Cerase A, Smeets D, Tang YA, Gdula M, Kraus F, Spivakov M, Moindrot B, Leleu M, Tattermusch A, Demmerle J, et al. (2014). Spatial separation of Xist RNA and polycomb proteins revealed by superresolution microscopy. Proc Natl Acad Sci U S A 111, 2235–2240. [PubMed: 24469834]

Chang JC, Fok PW, and Chou T (2015). Bayesian Uncertainty Quantification for Bond Energies and Mobilities Using Path Integral Analysis. Biophys J 109, 966–974. [PubMed: 26331254]

Chang JC, Savage VM, and Chou T (2014). A path-integral approach to bayesian inference for inverse problems using the semiclassical approximation. Journal of Statistical Physics 109, 966–974.

Chaumeil J, Le Baccon P, Wutz A, and Heard E (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev 20, 2223–2237. [PubMed: 16912274]

Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, et al. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell 155, 1479–1491. [PubMed: 24360272]

Chu C, Zhang QC, da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, Magnuson T, Heard E, and Chang HY (2015). Systematic discovery of Xist RNA binding proteins. Cell 161, 404–416. [PubMed: 25843628]

Clemson CM, McNeil JA, Willard HF, and Lawrence JB (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. J Cell Biol 132, 259–275. [PubMed: 8636206]

Colognori D, Sunwoo H, Kriz AJ, Wang CY, and Lee JT (2019). Xist Deletional Analysis Reveals an Interdependency between Xist RNA and Polycomb Complexes for Spreading along the Inactive X. Mol Cell 74, 101–117 e110. [PubMed: 30827740]

Cremer M, Grasser F, Lanctot C, Muller S, Neusser M, Zinner R, Solovei I, and Cremer T (2008). Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. Methods Mol Biol 463, 205–239. [PubMed: 18951171]

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. (2021). Twelve years of SAMtools and BCFtools. Gigascience 10.

Darrow EM, Huntley MH, Dudchenko O, Stamenova EK, Durand NC, Sun Z, Huang SC, Sanborn AL, Machol I, Shamim M, et al. (2016). Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture. Proc Natl Acad Sci U S A 113, E4504–4512. [PubMed: 27432957]

Demmerle J, Innocent C, North AJ, Ball G, Muller M, Miron E, Matsuda A, Dobbie IM, Markaki Y, and Schermelleh L (2017). Strategic and practical guidelines for successful structured illumination microscopy. Nat Protoc 12, 988–1010. [PubMed: 28406496]

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 22, 1775–1789. [PubMed: 22955988]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Dossin F, Pinheiro I, Zylicz JJ, Roensch J, Collombet S, Le Saux A, Chelmicki T, Attia M, Kapoor V, Zhan Y, et al. (2020). SPEN integrates transcriptional and epigenetic control of X-inactivation. Nature 578, 455–460. [PubMed: 32025035]

Dundr M, and Misteli T (2003). Measuring dynamics of nuclear proteins by photobleaching. Curr Protoc Cell Biol Chapter 13, Unit 13 15.

Eliscovich C, Buxbaum AR, Katz ZB, and Singer RH (2013). mRNA on the move: the road to its biological destiny. J Biol Chem 288, 20361–20368. [PubMed: 23720759]

Engreitz JM, Ollikainen N, and Guttman M (2016). Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. Nat Rev Mol Cell Biol 17, 756–770. [PubMed: 27780979]

Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES, et al. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science 341, 1237973. [PubMed: 23828888]

Finn EH, Pegoraro G, Shachar S, and Misteli T (2017). Comparative analysis of 2D and 3D distance measurements to study spatial genome organization. Methods 123, 47–55. [PubMed: 28179124]

Francis NJ, Kingston RE, and Woodcock CL (2004). Chromatin compaction by a polycomb group protein complex. Science 306, 1574–1577. [PubMed: 15567868]

Frank L, and Rippe K (2020). Repetitive RNAs as Regulators of Chromatin-Associated Subcompartment Formation by Phase Separation. J Mol Biol 432, 4270–4286. [PubMed: 32320688]

Fusco D, Accornero N, Lavoie B, Shenoy SM, Blanchard JM, Singer RH, and Bertrand E (2003). Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. Curr Biol 13, 161–167. [PubMed: 12546792]

Galupa R, and Heard E (2018). X-Chromosome Inactivation: A Crossroads Between Chromosome Architecture and Gene Regulation. Annu Rev Genet 52, 535–566. [PubMed: 30256677]

Giorgetti L, Lajoie BR, Carter AC, Attia M, Zhan Y, Xu J, Chen CJ, Kaplan N, Chang HY, Heard E, et al. (2016). Structural organization of the inactive X chromosome in the mouse. Nature 535, 575–579. [PubMed: 27437574]

Grau DJ, Chapman BA, Garlick JD, Borowsky M, Francis NJ, and Kingston RE (2011). Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. Genes Dev 25, 2210–2221. [PubMed: 22012622]

Gu B, Posfai E, and Rossant J (2018a). Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. Nat Biotechnol 36, 632–637. [PubMed: 29889212]

Gu Z, Eils R, Schlesner M, and Ishaque N (2018b). EnrichedHeatmap: an R/Bioconductor package for comprehensive visualization of genomic signal associations. BMC Genomics.

Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, et al. (2020). Array programming with NumPy. Nature 585, 357–362. [PubMed: 32939066]

Hayashi K, and Saitou M (2013). Generation of eggs from mouse embryonic stem cells and induced pluripotent stem cells. Nat Protoc 8, 1513–1524. [PubMed: 23845963]

Hendrich BD, Plenge RM, and Willard HF (1997). Identification and characterization of the human XIST gene promoter: implications for models of X chromosome inactivation. Nucleic Acids Res 25, 2661–2671. [PubMed: 9185579]

Hsia Y, Bale JB, Gonen S, Shi D, Sheffler W, Fong KK, Nattermann U, Xu C, Huang PS, Ravichandran R, et al. (2016). Design of a hyperstable 60-subunit protein dodecahedron. [corrected]. Nature 535, 136–139. [PubMed: 27309817]

Hunter JD (2007). Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering 9, 90–95.

Illingworth RS (2019). Chromatin folding and nuclear architecture: PRC1 function in 3D. Curr Opin Genet Dev 55, 82–90. [PubMed: 31323466]

Jansz N, Nesterova T, Keniry A, Iminitoff M, Hickey PF, Pintacuda G, Masui O, Kobelke S, Geoghegan N, Breslin KA, et al. (2018). Smchd1 Targeting to the Inactive X Is Dependent on the Xist-HnrnpK-PRC1 Pathway. Cell Rep 25, 1912–1923 e1919. [PubMed: 30428357]

Jegu T, Aeby E, and Lee JT (2017). The X chromosome in space. Nat Rev Genet 18, 377389.

Jonkers I, Monkhorst K, Rentmeester E, Grootegoed JA, Grosveld F, and Gribnau J (2008). Xist RNA is confined to the nuclear territory of the silenced X chromosome throughout the cell cycle. Mol Cell Biol 28, 5583–5594. [PubMed: 18625719]

Kang M, Day CA, DiBenedetto E, and Kenworthy AK (2010). A quantitative approach to analyze binding diffusion kinetics by confocal FRAP. Biophys J 99, 2737–2747. [PubMed: 21044570]

Kranz A, Fu J, Duerschke K, Weidlich S, Naumann R, Stewart AF, and Anastassiadis K (2010). An improved Flp deleter mouse in C57Bl/6 based on Flpo recombinase. Genesis 48, 512520.

Kraus F, Miron E, Demmerle J, Chitiashvili T, Budco A, Alle Q, Matsuda A, Leonhardt H, Schermelleh L, and Markaki Y (2017). Quantitative 3D structured illumination microscopy of nuclear structures. Nat Protoc 12, 1011–1028. [PubMed: 28406495]

Kuznetsova IM, Turoverov KK, and Uversky VN (2014). What macromolecular crowding can do to a protein. Int J Mol Sci 15, 23090–23140. [PubMed: 25514413]

Langmead B, Trapnell C, Pop M, and Salzberg SL (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10, R25. [PubMed: 19261174]

Lee S, Cook D, and Lawrence M (2019). plyranges: a grammar of genomic data transformation. Genome Biol 20, 4. [PubMed: 30609939]

Loda A, and Heard E (2019). Xist RNA in action: Past, present, and future. PLoS Genet 15, e1008333. [PubMed: 31537017]

Markaki Y, Smeets D, Cremer M, and Schermelleh L (2013). Fluorescence in situ hybridization applications for super-resolution 3D structured illumination microscopy. Methods in molecular biology (Clifton, NJ 950, 43–64.

Markaki Y, Smeets D, Fiedler S, Schmid VJ, Schermelleh L, Cremer T, and Cremer M (2012). The potential of 3D-FISH and super-resolution structured illumination microscopy for studies of 3D nuclear architecture: 3D structured illumination microscopy of defined chromosomal structures visualized by 3D (immuno)-FISH opens new perspectives for studies of nuclear architecture. Bioessays 34, 412–426. [PubMed: 22508100]

McHugh CA, Chen CK, Chow A, Surka CF, Tran C, McDonel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A, et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature 521, 232–236. [PubMed: 25915022]

McNally JG (2008). Quantitative FRAP in analysis of molecular binding dynamics in vivo. Methods Cell Biol 85, 329–351. [PubMed: 18155469]

McSwiggen DT, Mir M, Darzacq X, and Tjian R (2019). Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. Genes Dev 33, 1619–1634. [PubMed: 31594803]

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

Minkovsky A, Sahakyan A, Rankin-Gee E, Bonora G, Patel S, and Plath K (2014). The Mbd1-Atf7ip-Setdb1 pathway contributes to the maintenance of X chromosome inactivation. Epigenetics Chromatin 7, 12. [PubMed: 25028596]

Mittag T, and Forman-Kay JD (2007). Atomic-level characterization of disordered protein ensembles. Curr Opin Struct Biol 17, 3–14. [PubMed: 17250999]

Mohammed H, Taylor C, Brown GD, Papachristou EK, Carroll JS, and D'Santos CS (2016). Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. Nat Protoc 11, 316–326. [PubMed: 26797456]

Moindrot B, Cerase A, Coker H, Masui O, Grijzenhout A, Pintacuda G, Schermelleh L, Nesterova TB, and Brockdorff N (2015). A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. Cell Rep 12, 562–572. [PubMed: 26190105]

Monfort A, Di Minin G, Postlmayr A, Freimann R, Arieti F, Thore S, and Wutz A (2015). Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells. Cell Rep 12, 554–561. [PubMed: 26190100]

Mor A, Suliman S, Ben-Yishay R, Yunger S, Brody Y, and Shav-Tal Y (2010). Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. Nat Cell Biol 12, 543–552. [PubMed: 20453848]

Nam HS, and Benezra R (2009). High levels of Id1 expression define B1 type adult neural stem cells. Cell Stem Cell 5, 515–526. [PubMed: 19896442]

Nesterova TB, Wei G, Coker H, Pintacuda G, Bowness JS, Zhang T, Almeida M, Bloechl B, Moindrot B, Carter EJ, et al. (2019). Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. Nat Commun 10, 3129. [PubMed: 31311937]

Ng K, Daigle N, Bancaud A, Ohhata T, Humphreys P, Walker R, Ellenberg J, and Wutz A (2011). A system for imaging the regulatory noncoding Xist RNA in living mouse embryonic stem cells. Mol Biol Cell 22, 2634–2645. [PubMed: 21613549]

Nozaki T, Imai R, Tanbo M, Nagashima R, Tamura S, Tani T, Joti Y, Tomita M, Hibino K, Kanemaki MT, et al. (2017). Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. Mol Cell 67, 282–293 e287. [PubMed: 28712725]

Ollion J, Cochennec J, Loll F, Escude C, and Boudier T (2013). TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. Bioinformatics 29, 1840–1841. [PubMed: 23681123]

Pacini G, Dunkel I, Mages N, Mutzel V, Timmermann B, Marsico A, and Schulz EG (2021). Integrated analysis of Xist upregulation and X-chromosome inactivation with single-cell and single-allele resolution. Nat Commun 12, 3638. [PubMed: 34131144]

Pandya-Jones A, Markaki Y, Serizay J, Chitiashvili T, Mancia Leon WR, Damianov A, Chronis C, Papp B, Chen CK, McKee R, et al. (2020). A protein assembly mediates Xist localization and gene silencing. Nature.

Panning B, Dausman J, and Jaenisch R (1997). X chromosome inactivation is mediated by Xist RNA stabilization. Cell 90, 907–916. [PubMed: 9298902]

Pintacuda G, Wei G, Roustan C, Kirmizitas BA, Solcan N, Cerase A, Castello A, Mohammed S, Moindrot B, Nesterova TB, et al. (2017). hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing. Mol Cell 68, 955–969 e910. [PubMed: 29220657]

Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, de la Cruz CC, Otte AP, Panning B, and Zhang Y (2003). Role of histone H3 lysine 27 methylation in X inactivation. Science 300, 131–135. [PubMed: 12649488]

Plath K, Mlynarczyk-Evans S, Nusinow DA, and Panning B (2002). Xist RNA and the mechanism of X chromosome inactivation. Annu Rev Genet 36, 233–278. [PubMed: 12429693]

Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. [PubMed: 20110278]

Quinodoz SA, Bhat P, Ollikainen N, Jachowicz JW, Banerjee AK, Chovanec P, Blanco MR, Chow A, Markaki Y, Plath K, et al. (2020). RNA promotes the formation of spatial compartments in the nucleus. bioRxiv, 2020.2008.2025.267435.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, and Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44, W160–165. [PubMed: 27079975]

Ridings-Figueroa R, Stewart ER, Nesterova TB, Coker H, Pintacuda G, Godwin J, Wilson R, Haslam A, Lilley F, Ruigrok R, et al. (2017). The nuclear matrix protein CIZ1 facilitates localization of Xist RNA to the inactive X-chromosome territory. Genes Dev 31, 876–888. [PubMed: 28546514]

Rinn JL, and Chang HY (2012). Genome regulation by long noncoding RNAs. Annu Rev Biochem 81, 145–166. [PubMed: 22663078]

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA.

Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, and Eliceiri KW (2017). ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinformatics 18, 529. [PubMed: 29187165]

Schertzer MD, Braceros KCA, Starmer J, Cherney RE, Lee DM, Salazar G, Justice M, Bischoff SR, Cowley DO, Ariel P, et al. (2019). lncRNA-Induced Spread of Polycomb Controlled by Genome Architecture, RNA Abundance, and CpG Island DNA. Mol Cell 75, 523–537 e510. [PubMed: 31256989]

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676–682. [PubMed: 22743772]

Silva J, Mak W, Zvetkova I, Appanah R, Nesterova TB, Webster Z, Peters AH, Jenuwein T, Otte AP, and Brockdorff N (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. Dev Cell 4, 481–495. [PubMed: 12689588]

Simon MD, Pinter SF, Fang R, Sarma K, Rutenberg-Schoenberg M, Bowman SK, Kesner BA, Maier VK, Kingston RE, and Lee JT (2013). High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. Nature 504, 465–469. [PubMed: 24162848]

Sladitschek HL, and Neveu PA (2015). MXS-Chaining: A Highly Efficient Cloning Platform for Imaging and Flow Cytometry Approaches in Mammalian Systems. PLoS One 10, e0124958. [PubMed: 25909630]

Smeets D, Markaki Y, Schmid VJ, Kraus F, Tattermusch A, Cerase A, Sterr M, Fiedler S, Demmerle J, Popken J, et al. (2014). Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. Epigenetics & chromatin 7, 8. [PubMed: 25057298]

Statello L, Guo CJ, Chen LL, and Huarte M (2021). Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol 22, 96–118. [PubMed: 33353982]

Sunwoo H, Colognori D, Froberg JE, Jeon Y, and Lee JT (2017). Repeat E anchors Xist RNA to the inactive X chromosomal compartment through CDKN1A-interacting protein (CIZ1). Proc Natl Acad Sci U S A 114, 10654–10659. [PubMed: 28923964]

Sunwoo H, Wu JY, and Lee JT (2015). The Xist RNA-PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. Proc Natl Acad Sci U S A 112, E4216–4225. [PubMed: 26195790]

Tavares L, Dimitrova E, Oxley D, Webster J, Poot R, Demmers J, Bezstarosti K, Taylor S, Ura H, Koide H, et al. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. Cell 148, 664–678. [PubMed: 22325148]

Teller K, Illner D, Thamm S, Casas-Delucchi CS, Versteeg R, Indemans M, Cremer T, and Cremer M (2011). A top-down analysis of Xa- and Xi-territories reveals differences of higher order structure at >/= 20 Mb genomic length scales. Nucleus 2, 465–477. [PubMed: 21970989]

The pandas development team (2020). pandas-dev/pandas: Pandas. Zenodo, 10.5281/zenodo.3509134

Tinevez JY, Perry N, Schindelin J, Hoopes GM, Reynolds GD, Laplantine E, Bednarek SY, Shorte SL, and Eliceiri KW (2017). TrackMate: An open and extensible platform for single-particle tracking. Methods 115, 80–90. [PubMed: 27713081]

Uversky VN (2015). The multifaceted roles of intrinsic disorder in protein complexes. FEBS Lett 589, 2498–2506. [PubMed: 26073257]

Van der Auwera GA, and O'Connor BD (2020). Genomics in the Cloud: Using Docker, GATK, and WDL in Terra (1st Edition). O'Reilly Media.

van Rossum G, and Drake FL (2009). Python 3 Reference Manual. Scotts Valley, CA.

van Zon R, and Schofield J (2010). Constructing smooth potentials of mean force, radial distribution functions and probability densities from sampled data. Journal of Chemical Physics 132, 154110.

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 17, 261–272. [PubMed: 32015543]

Wang CY, Colognori D, Sunwoo H, Wang D, and Lee JT (2019). PRC1 collaborates with SMCHD1 to fold the X-chromosome and spread Xist RNA between chromosome compartments. Nat Commun 10, 2950. [PubMed: 31270318]

Wang CY, Jegu T, Chu HP, Oh HJ, and Lee JT (2018). SMCHD1 Merges Chromosome Compartments and Assists Formation of Super-Structures on the Inactive X. Cell 174, 406–421 e425. [PubMed: 29887375]

Waskom ML (2021). seaborn: statistical data visualization. Journal of Open Source Software 6, 3021.

Wickham et al. (2019). Welcome to the tidyverse. Journal of Open Source Software 4 (43).

Wu B, Chao JA, and Singer RH (2012). Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. Biophys J 102, 2936–2944. [PubMed: 22735544]

Wutz A (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat Rev Genet 12, 542–553. [PubMed: 21765457]

Wutz A, and Jaenisch R (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol Cell 5, 695–705. [PubMed: 10882105]

Xie L, and Liu Z (2021). Single-cell imaging of genome organization and dynamics. Mol Syst Biol 17, e9653. [PubMed: 34232558]

Ying QL, and Smith AG (2003). Defined conditions for neural commitment and differentiation. Methods Enzymol 365, 327–341. [PubMed: 14696356]

Yue M, Ogawa A, Yamada N, Charles Richard JL, Barski A, and Ogawa Y (2017). Xist RNA repeat E is essential for ASH2L recruitment to the inactive X and regulates histone modifications and escape gene expression. PLoS Genet 13, e1006890. [PubMed: 28686623]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol 9, R137. [PubMed: 18798982]

Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS, and Green MR (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics 11, 237. [PubMed: 20459804]

Zimmerman SB, and Pheiffer BH (1983). Macromolecular crowding allows blunt-end ligation by DNA ligases from rat liver or Escherichia coli. Proc Natl Acad Sci U S A 80, 5852–5856. [PubMed: 6351067]

Zylicz JJ, Bousard A, Zumer K, Dossin F, Mohammad E, da Rocha ST, Schwalb B, Syx L, Dingli F, Loew D, et al. (2019). The Implication of Early Chromatin Changes in X Chromosome Inactivation. Cell 176, 182–197 e123. [PubMed: 30595450]

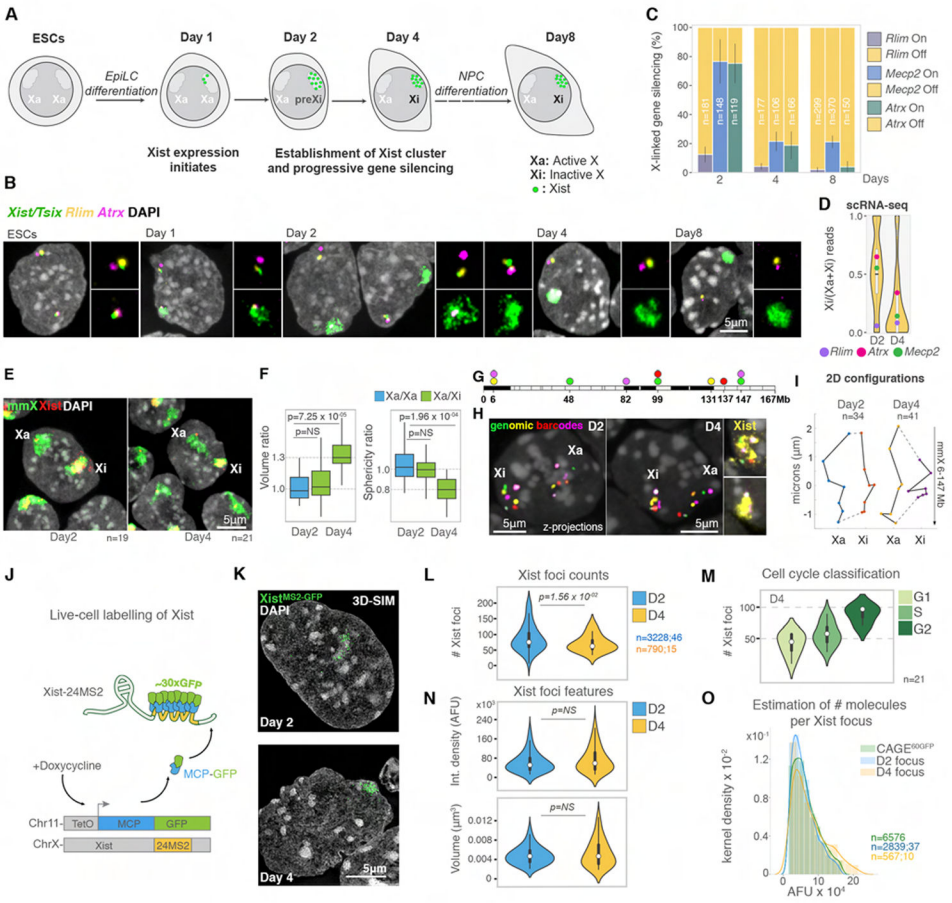

Figure 1: ~50 *Xist* foci initiate XCI

(A) Schematic of the ESC differentiation protocol.

(B) RNA FISH for *Xist/Tsix* (green), *Rlim* (yellow) and *Atrx* (magenta) transcripts during differentiation. DAPI staining is shown in grey. Small images show magnifications of the *Tsix* signals on the Xa or the *Xist* signals on the Xi and nascent gene transcripts.

(C) Percentage of cells with given nascent transcripts of *Rlim, Atrx* and *Mecp2* under *Xist* clouds during differentiation from two replicates. Error bars denote standard deviation; *n* is the number of cells analyzed.

(D) Violin plots of the Xi expression ratios of X-linked genes averaged across single cells expressing *Xist* from the 129 allele at D2 and D4. The ratios of *Rlim, Atrx* and *Mecp2* are highlighted.

(E) DNA/RNA FISH with X-chromosome paints (mmX, green) and Xist probes (red) of cells at D2 and D4. DAPI is shown in grey. *n* is the number of cells analyzed in **F**.

(F) Boxplots showing the volume and sphericity ratio of Xa/Xa (blue) in cells not expressing *Xist* at D2 and Xa/Xi (green) in cells with an *Xist* cloud at D2 and D4. MannWhitney-Wilcoxon (MWW) *p*-values are given.

(G) Schematic of the spectral barcoding strategy applied to map X chromosome configuration.

(H) DNA FISH of the spectrally barcoded genomic regions described in **G**. Overlay with *Xist* RNA FISH signals (far right, yellow) was used to score for the Xi at D2 (top) and D4 (bottom).

(I) 2D configuration plots of average coordinates of genomic barcodes from **H**, extracted with 95% confidence. *n* is the number of cells analyzed from three experiments.

(J) Illustration of live-cell *Xist* labelling strategy.

(K) 3D-SIM projections showing $Xist^{MS2\text{-}GFP}$ signals (green) and DAPI staining (grey) at the indicated differentiation day.

(L) Violin plots of the 3D-SIM quantification of $Xist^{MS2\text{-}GFP}$ foci number at D2 and D4. *n* denotes the number of Xist foci measured, followed by the number of cells analyzed from two replicates. MWW *p*-value is given.

(M) As in **L** at D4 after scoring for cell cycle stages.

(N) As in **L**, except for showing integrated density (AFU) and volume (μm^3) of $Xist^{MS2\text{-}GFP}$ foci at D2 and D4 for the same sets of foci.

(O) Histograms depicting integrated density (AFU) of $Xist^{MS2\text{-}GFP}$ foci (~30 GFP molecules per *Xist*) and GFP nanocages ($cage^{60GFP}$) per pixel kernel density, detected by 3D-SIM. *n* denotes the number of foci measured followed by the number of cells analyzed from two replicates.

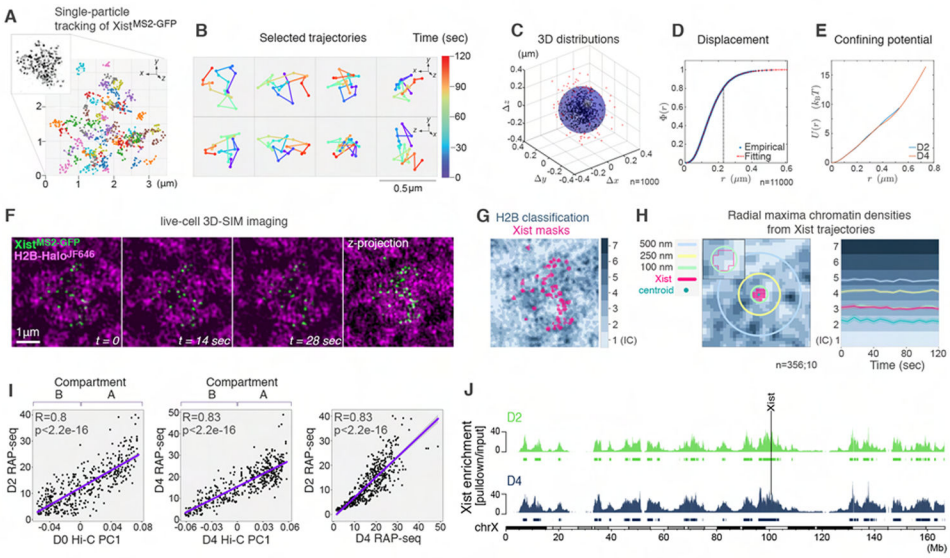

Figure 2. *Xist* foci are locally confined at open chromatin regions

(A) Trajectories of *Xist*$^{MS2\text{-}GFP}$ foci from live-cell 3D-SIM imaging for 2 minutes (5sec/frame) at D4. Inset: Projection of one frame showing an *Xist*$^{MS2\text{-}GFP}$ cluster.

(B) Selected trajectories from **A** showing the displacement of *Xist* foci over time (color-gradient) in top (xyz, top) and side (zyx, bottom) views.

(C) D4 *Xist* foci displacements derived from ~100 trajectories, each centered about their centers of mass. *n* denotes the number of foci analyzed from four experiments.

(D) Cumulative distribution function $\Phi(r)$ of the number of displacement positions at D4 with distance from origin $<r$. The distance marked with the dashed line at r=0.22μm corresponds to the radius of the shaded sphere in **C** where ~80% of all distances lie (Methods S1 file). *n* denotes the number of foci analyzed from ~800 trajectories from four experiments.

(E) Effective spherically symmetric confining potential inferred from the spatial distribution of displacement distances of *Xist* foci at D2 and D4. We assume an equilibrium Boltzmann distribution over an effective potential energy well that is a function of *r*.

(F) Image sequence from t=0 to t=28sec and z-projection (from t=0) of *Xist*$^{MS2\text{-}GFP}$ (green) and H2B-Halo^{JF646} (magenta) based on live-cell 3D-SIM.

(G) Segmentation of H2B-Halo^{JF646} from live-cell 3D-SIM data into seven density classes with overlay of *Xist* masks.

(H) Left: Schematic for the assessment of the chromatin landscape around one *Xist* focus. Mask (bright pink) of one *Xist* focus showing radial distances denoted by circles from its centroid (dark green). Inset: magnification showing the outline of the *Xist* mask. Right: Plot of *Xist* foci trajectories showing the average radial maxima of chromatin density reached at indicated timepoints. Light shaded areas show 95% confidence interval. *n* denotes the number of foci and cells analyzed from three experiments.

(I) Correlation of *Xist* enrichment determined by RAP-seq at D2 and D4 to the first principal component of ESC (D0) and D4 Hi-C data (A-compartment = positive values, B-compartment = negative values). Far right panel shows the correlation between D2 and D4 *Xist* RAP-seq data. Pearson correlation r-coefficients and associated *p*-values are given.

(J) *Xist* enrichment along the X chromosome, defined based on RAP-seq data for *Xist* over the input, at D2 and D4, with peak calls below. The *Xist* locus is indicated.

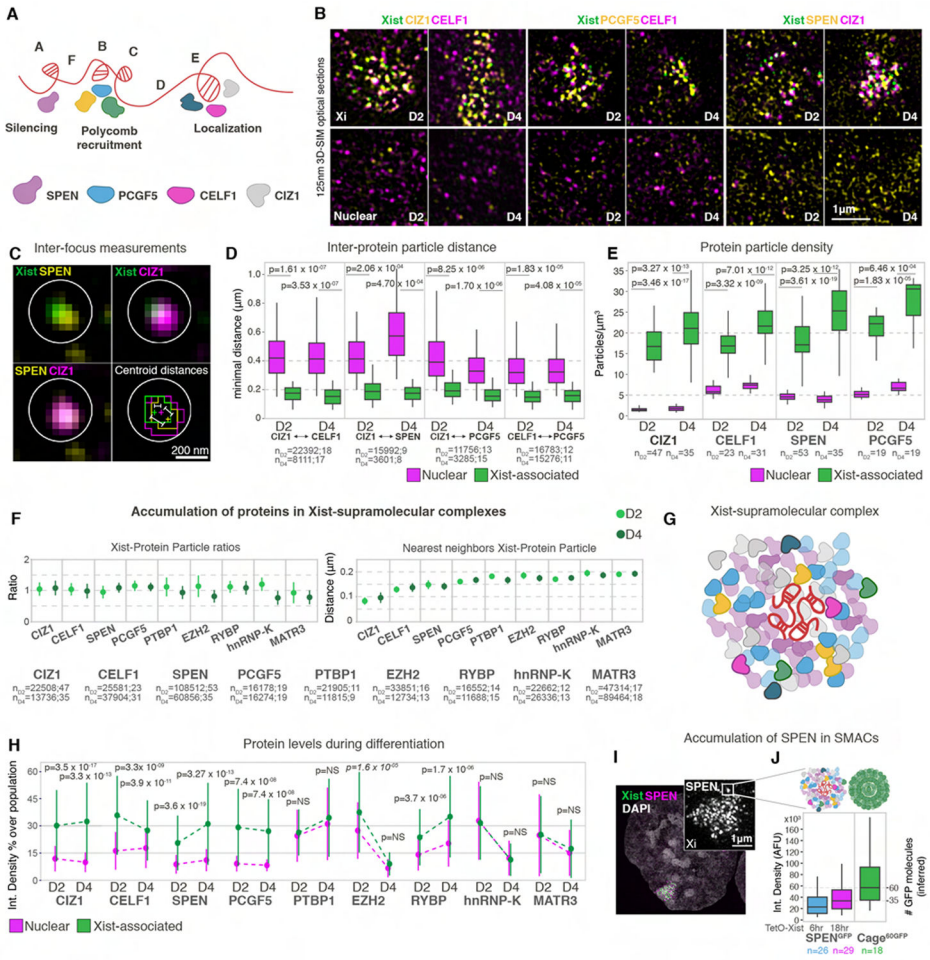

Figure 3. *Xist* nucleates supramolecular complexes

(A) Schematic of *Xist* RNA with its repeat sequences A-F, different repeat-binding proteins, and repeat functions. Proteins examined in **B** are indicated.

(B) 125nm 3D-SIM optical sections showing detection of indicated Halo protein-fusions labelled with JF549 (yellow) and immunodected proteins (magenta) in *Xist*$^{MS2\text{-}GFP}$ cells at D2 and D4. Top panels show the *Xist*-demarcated X-territory (pre-Xi/Xi); bottom panels a nuclear region (Nuclear). Note the distinctive enrichment of all pairs of interactors around *Xist* foci.

(C) Overview of inter-protein particle distance measurements for protein foci associated with one *Xist* focus, based on data in **B**. Overlay: *Xist* (green), SPEN (yellow) and CIZ1 (magenta). Bottom right panel shows mask outlines after image segmentation, depiction of protein and *Xist* foci centroids (crosses) and measurement of inter-particle distances performed in **D**. Circles denote a 200nm radius.

(D) Boxplots from data in **B** showing the nearest-neighbor distances between the indicated pairs of protein particles in nuclear and *Xist*-associated fractions obtained as shown in C for D2 and D4. *n* denotes the number of protein particles followed by the number of cells analyzed. MWW *p*-values are given.

(E) Boxplots of the distribution of the density of indicated protein particles (number of particles per μm^3) in the Xi and in nuclear regions on D2 and D4. *n* denotes the number of cells analyzed. MWW *p*-values are given.

(F) Point-plots showing the average ratio of the number of indicated protein particles per *Xist* particle within 250nm radial search (left) and their nearest distance (right) on D2 and D4. The bars denote the standard deviation and *n* the number of particles followed by number of cells analyzed.

(G) Schematic of a *Xist*-supramolecular complex.

(H) Point-plots from data in F showing the integrated density of fluorescence of indicated protein particles in *Xist*-associated (green) and nuclear (magenta) fractions, on D2 and D4 from two experiments. Dots denote the median, bars the standard deviation. Dotted lines are included to visualize changes. Data are normalized to the highest signal observed across the entire population of each protein. Absolute values are shown in Figure S4E. MWW *p*-values are given.

(I) Projection of a nucleus imaged with 3D-SIM expressing SPEN-GFP from the endogenous locus (magenta) and *Xist*-Bgl-mCherry (green) at 18hrs post tetO-*Xist* induction. Inset shows SPEN signals in the Xi.

(J) Boxplots showing integrated densities of cages^{60GFP}, *Xist*-associated SPEN at 6 or 18hrs after tetO-*Xist* induction. *n* denotes the number of cells analyzed from two replicates.

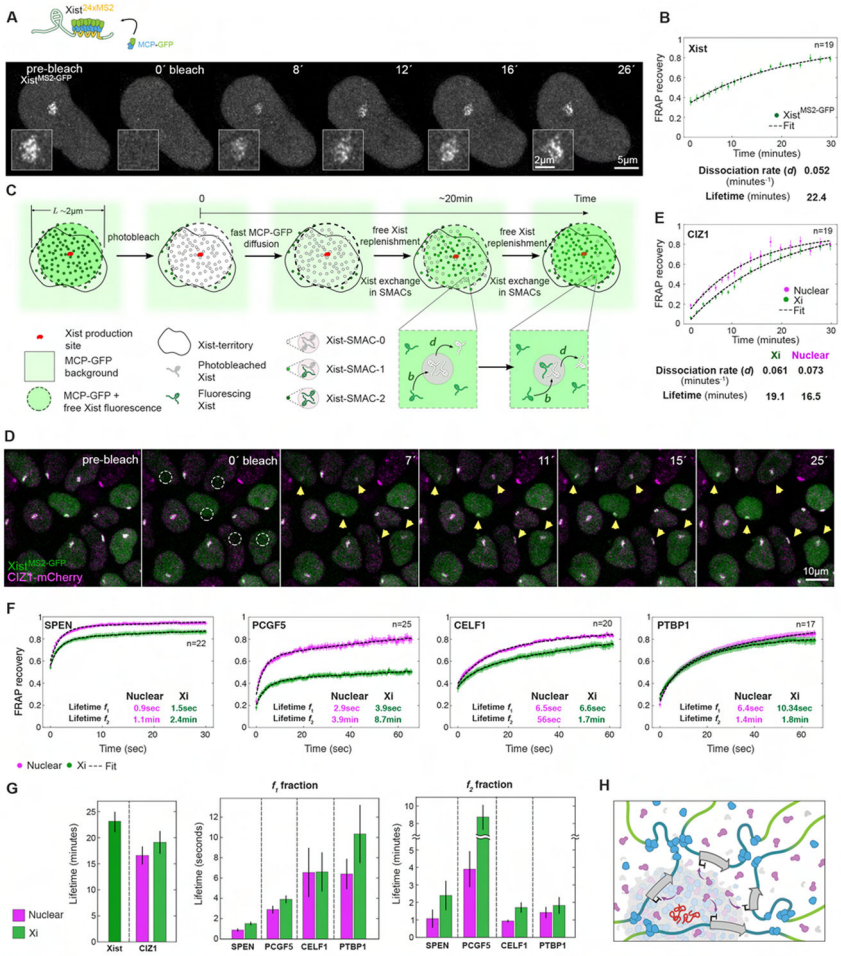

Figure 4. FRAP of *Xist* interactors identifies diverse protein behaviors in the Xi

(A) Top: Schematic of *Xist* live-cell labeling. Bottom: Image sequence from an Airyscan FRAP experiment of $Xist^{MS2\text{-}GFP}$ at D4. Insets show the *Xist* territory.

(B) $Xist^{MS2\text{-}GFP}$ FRAP recovery at D4 and fitting. Error bars indicate the standard error. Dissociation rate and lifetime inferred from fitting are given. *n* denotes the number of cells analyzed from four experiments.

(C) Model for the *Xist* FRAP process. The expression and replenishment of *Xist* from its expression site is assumed to be fast and free MCP-GFP replenishment is assumed almost instantaneous after $t=0$. The exchange of photobleached with fluorescing *Xist* is assumed fast in the Xi-territory outside *Xist*-SMACs (free pool) and slow within *Xist*SMACs. *Xist*-SMACs with zero, one, and two fluorescing *Xist* molecules are denoted *Xist*-SMAC-0, −1, and −2. Binding of *Xist* to sites in SMACs occurs at rate *b* and dissociation at rate *d,* which sets the timescale for FRAP recovery. The FRAP curves for *Xist* were fit with a single exponential.

(D) Image sequence showing a FRAP experiment of $Xist^{MS2\text{-}GFP}$ (green) and CIZ1-mCherry (magenta) at D4. Dashed circles indicate bleached *Xist*-territories and yellow arrows monitor recovery.

(E) FRAP recovery and fitting (dashed black lines) of the nuclear and *Xist*-associated populations of CIZ1-mCherry. Error bars denote the standard error. Parameters from fitting with a single exponential are given. *n* denotes the number of cells analyzed from four experiments.

(F) FRAP recovery and fitting (dashed black lines) of the nuclear and *Xist*-associated populations of SPEN-Halo^{TMR}, PCGF5-Halo^{TMR}, CELF1-mCherry and PTBP1-Halo^{TMR} at D4. Error bars denote the standard error. Every fifth timepoint is shown. Lifetimes for the slow (f_1) and fast (f_2) detaching fractions inferred for each protein from bi-exponential fitting are indicated. *n* denotes the number of cells analyzed from two experiments.

(G) Bargraphs showing the lifetimes for *Xist* and CIZ1 (left) and for the two subpopulations (f_1, f_2) of bi-exponentially fitted proteins (right). Error bars denote the standard error.

(H) Schematic showing an *Xist*-SMAC and its dynamic regulatory compartment. The increased accumulation of proteins surrounding *Xist* (red) and their rapid cycling results in gradients over broad chromosomal regions in the vicinity to the SMAC. SPEN is depicted in purple, accumulation of PCGF5 in blue, and purple arrows indicate the gene silencing function of SPEN.

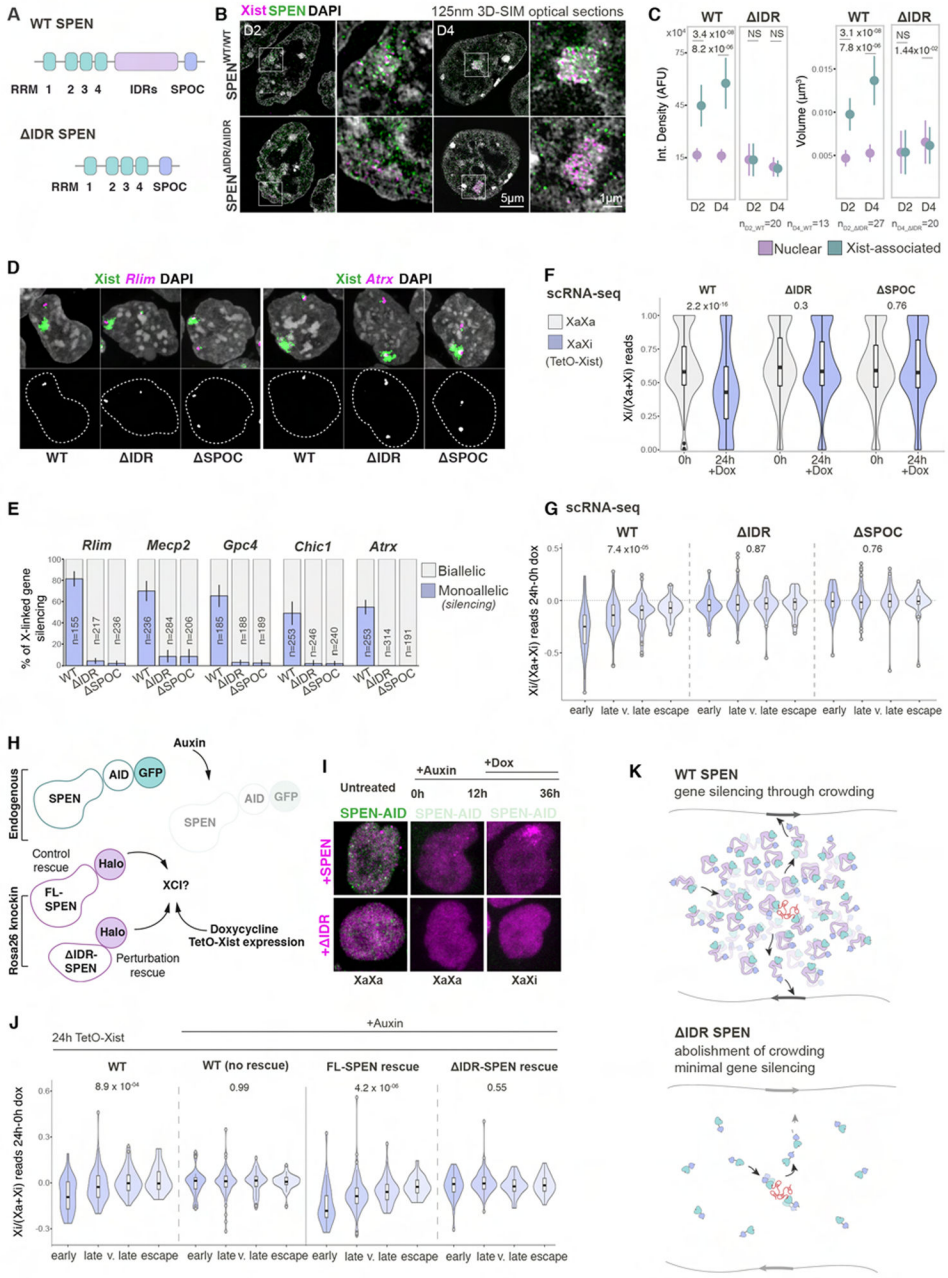

Figure 5. IDR-mediated crowding of SPEN in SMACs is required for gene silencing

(A) Schematic of the domains of WT and ΔIDR SPEN.

(B) 3D-SIM optical sections of immuno-RNA-FISH with Xist probes (magenta) and GFP antibodies (green) at D2 and D4 in cells homozygously expressing GFP-tagged WT- or ΔIDR-SPEN. Second columns show the Xi. Staining by DAPI is shown in grey.

(C) Point-plots of integrated density and volume measurements for WT- and ΔIDR-SPEN particles that are *Xist*-associated (green) or in the nuclear (purple) fraction, on D2 and D4 from **B**. *n* denotes the number of cells analyzed from two experiments. MWW *p*-values are given.

(D) RNA FISH of *Xist* (green) and nascent transcripts of *Rlim* or *Atrx* (magenta) at 24hrs after doxycycline induction of tetO-*Xist* expression in female Xist$^{wt/tetO}$ cells homozygously expressing WT-, ΔIDR- and ΔSPOC-SPEN. Chromatin is stained with DAPI (grey). Second row images show *Rlim* or *Atrx* signals (grey) and nuclei masks (dashed lines).
(E) Quantification of experiment in **D** showing percentage of *Xist* clouds with a nascent transcript spot (purple, monoallelic) versus no transcripts (grey, biallelic). *n* denotes the number of cells analyzed from two replicates.
(F) Violin plots of Xi expression ratios of X-linked genes averaged across single Xist$^{wt/tetO}$ cells homozygously expressing WT-, ΔIDR- or ΔSPOC-SPEN without and with 24 hours of doxycycline addition. MWW *p*-values are given.
(G) Violin plots of the change in Xi expression ratio for data in **F**, grouped according to gene silencing dynamics in normal cells. Kruskal-Wallis *p*-values are given.
(H) Schematic of rescue assay used in **I** and **J**. SPEN-AID-GFP encoded from the endogenous locus is depleted with addition of auxin for 12hrs and FL- or ΔIDR-SPEN-Halo are constitutively expressed from the R26 locus. *Xist* expression and XCI are induced by addition of doxycycline for 24hrs in the presence of auxin.
(I) Images of SPEN-AID-GFP (green) with transgenic FL-SPEN (top) and ΔIDR-SPEN (bottom) rescue proteins (magenta). Columns from left to right: untreated; 12hrs auxin treated; and 24hrs doxycycline and 36hrs auxin treated cells. This strategy was used in **J** to explore rescue of XCI by transgenically encoded SPEN proteins after depletion of endogenously encoded SPEN and induction of tetO-*Xist*.
(J) Violin plots of the change in Xi expression ratios of X-linked genes over 24hrs of doxycycline-induced *Xist* expression, grouped according to gene silencing dynamics in normal cells, for conditions described in **H**. The Xi ratio was averaged across 3 replicates. Kruskal-Wallis *p*-values are given.
(K) Model of the augmented and dynamic distribution of WT-SPEN (top) in a *Xist*-SMAC compared to ΔIDR-SPEN (bottom). *Xist* is shown in red and SPEN domains are annotated as in **A**. Silent and active X-linked genes are indicated with black and grey arrows.

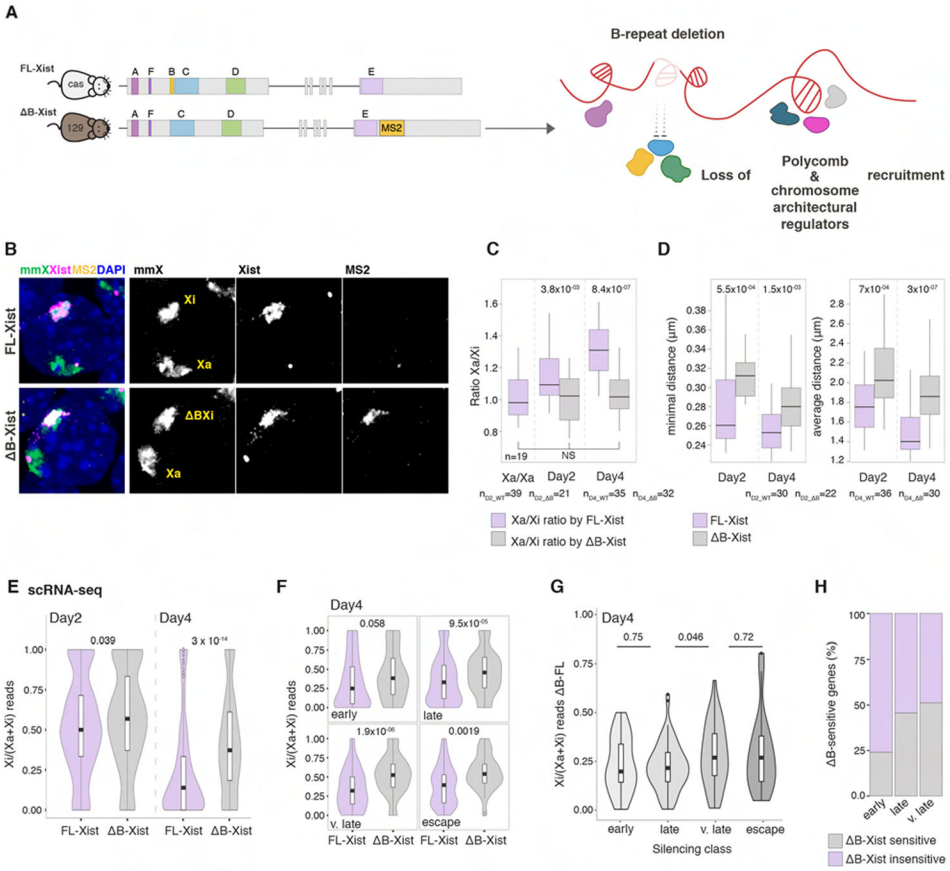

Figure 6. The B-repeat is critical for Xi compaction and late gene silencing

(A) Schematic of the heterozygous deletion of the B-repeat of Xist and insertion of a MS2 tag on the 129 X chromosome in female mouse 129/cas ESCs ($Xist^{\Delta B+MS2/WT}$ ESCs). The effect of the B-repeat deletion is indicated.

(B) RNA/DNA FISH of $Xist^{\Delta B+MS2/WT}$ cells at D4 using X-chromosome paints (mmX, green), Xist (magenta) and MS2 probes (yellow). DAPI staining is shown in blue. Greyscale images show individual channels as indicated.

(C) Boxplots showing the ratio of Xa/Xa at D2 in cells not expressing *Xist* and of Xa/Xi in FL-*Xist* (purple) or ΔB-*Xist* (grey) expressing $Xist^{\Delta B+MS2/WT}$ cells at D2 and D4. *n* is the number of cells analyzed from two replicates.

(D) Boxplots showing the minimal (left) and average (right) distance between *Xist* foci in FL- or ΔB-*Xist* expressing cells at D2 and D4. *n* is the number of cells analyzed from three experiments.

(E) Violin plots of Xi ratios of X-linked gene expression at D2 and D4 in $Xist^{\Delta B+MS2/WT}$ cells or parent WT ($Xist^{WT-MS2/WT}$) cells silencing the WT- or ΔB-*Xist* 129 X chromosome. Mean Xi ratio per gene was averaged across single cells based on scRNA-seq data. MWW *p*-values are given.

(F) As in **E**, except that genes are grouped by gene silencing dynamics. MWW *p*-values are given.

(G) Violin plots of the difference in Xi ratio between the ΔB- and WT-*Xist* expressing Xi shown in **F**. MWW *p*-values are given.

(H) Bargraphs showing the proportion of ΔB-sensitive or insensitive genes from **F**. Genes were considered ΔB-sensitive based on a one-sided Welch t-test comparing ΔB and WT Xi^{129} ratios, *p*-value<0.05.

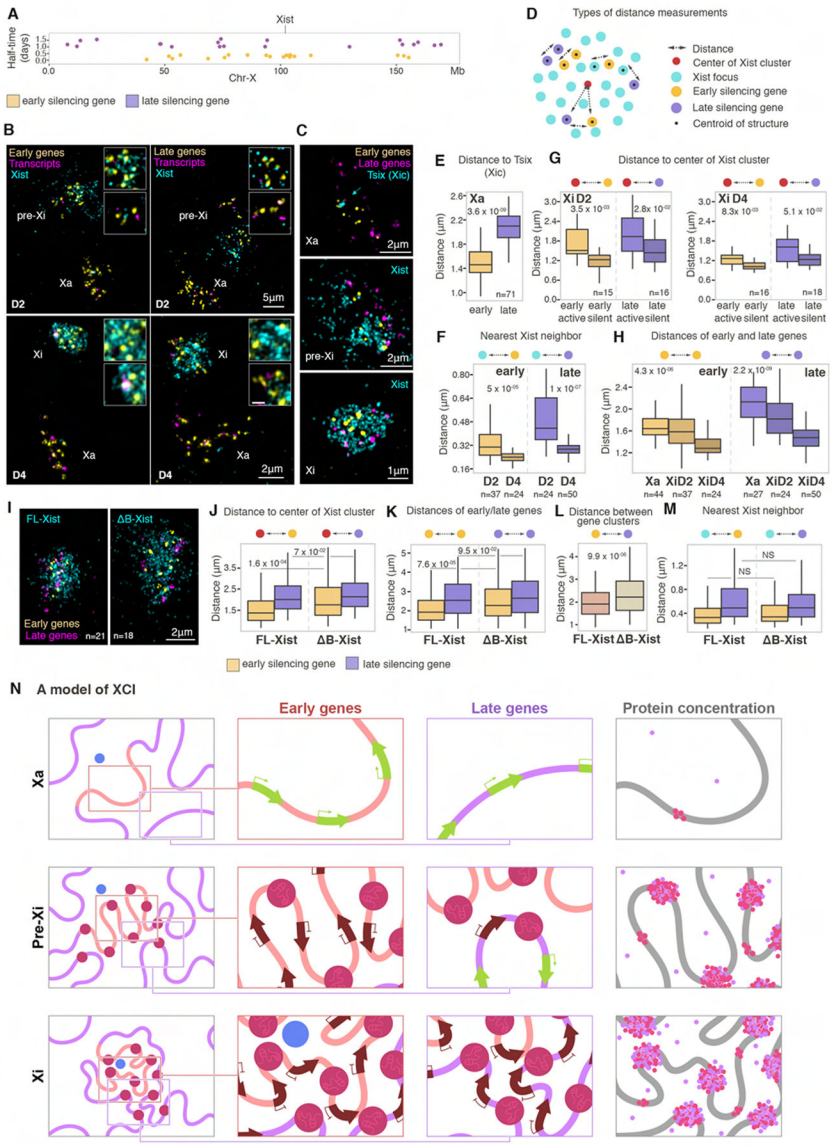

Figure 7. *Xist*-SMACs progressively re-configure and silence the Xi

(A) Annotation of the position of early (yellow) and late (purple) genes on the X chromosome simultaneously detected with oligo probes in RNA/DNA FISH experiments and their silencing half-time.

(B) 3D-SIM projections of nuclei after RNA/DNA FISH, showing indicated gene sets (yellow), their corresponding transcripts (magenta), and *Xist* signals (cyan) at D2 and D4. Xa, pre-Xi and Xi are indicated. Insets show magnifications of pre-Xi or Xi areas with high (top) and low (bottom) *Xist* density. Images are smoothed with a 3×3 px for clarity.

(C) 3D-SIM projections of Xa, pre-Xi or Xi regions after RNA/DNA FISH for *Xist* (cyan), early (yellow) and late (magenta) genes at D2 and D4.

(D) Schematic of different types of distance metrics performed in this figure.

(E) Boxplots of distances of early (yellow) or late (purple) genes relative to Tsix signals on the Xa. *n* denotes the number of cells analyzed from three experiments. MWW *p*-value is given.

(F) Boxplots of the nearest distance of *Xist* foci to early (yellow) or late (purple) genes at D2 and D4. *n* denotes the number of cells analyzed from three experiments. MWW *p*-values are given.

(G) Boxplots of distances of early (yellow) or late (purple) genes to the center of the *Xist* cluster at D2 and D4, divided into silent or active based on nascent transcripts detection. *n* is the number of cells analyzed from three experiments. MWW *p*-values are given.

(H) Intra-genic distances of early (yellow) or late (purple) genes on the Xa, pre-Xi, or Xi at D2 and D4. *n* denotes the number of cells analyzed from three experiments. Kruskal-Wallis *p*-values are given.

(I) 3D-SIM projections of the Xi after RNA/DNA FISH for *Xist* (cyan), early (yellow) and late (magenta) genes in male FL- or ΔB-*Xist* expressing ESCs after 18hrs doxycycline induction of tetO-*Xist*. *n* denotes the number of cells analyzed in **J** to **M** from two experiments. MWW *p*-values are given.

(J) Boxplots of distances of early (yellow) or late (purple) genes to the center of the *Xist* cluster in FL- or ΔB-*Xist* expressing cells described in I.

(K) As in **J**, except for intra-genic distances of early (yellow) or late (purple) genes.

(L) As in **J**, except for the distances of early to late silencing genes.

(M) As in **J**, except for nearest neighbor distance of early (yellow) or late (purple) genes to *Xist* foci.

(N) SMAC-based model of XCI. Left column shows the changes in the higher-order chromatin organization between the Xa (top), pre-Xi (middle) and Xi (bottom). The *Xist* production site is shown in blue, SMACs in dark red, and genomic regions harboring early and late silencing genes with orange and purple lines, respectively. Insets indicate regions of early (top) or late (bottom) silencing genes magnified in the second and third columns to show the progression of silencing. Arrows indicate active (green) and silent (brown) genes. Fourth column shows the increase in protein concentration upon establishment of *Xist*-SMACs. Pink dots indicate Polycomb-group proteins and purple dots SPEN. Free protein dots indicate increased concentrations in the Xi due to the presence of SMACs. Architectural protein-mediated chromosomal compaction is depicted by pink islets on the DNA fiber.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti-CUG-BP1	Abcam	RRID:AB_11141441; Cat #: ab129115
Rabbit polyclonal anti-hnRNP-K	Bethyl	RRID:AB_530281; Cat #: A300-678A
Rabbit polyclonal anti-MATR3	Bethyl	RRID:AB_2141651; Cat #: IHC-00081
Rabbit polyclonal anti-RYBP (DEDAF)	Millipore Sigma	RRID:AB_2285466; Cat #: AB3637
Rabbit monoclonal anti-EZH2	Cell signaling Technology	RRID:AB_10694683; Cat #: 5246
Rabbit polyclonal anti-CIZ1	Novus Biologicals	RRID:AB_1048573; Cat #: NB100-74624
Rabbit polyclonal anti-histone H3 phospho-Serine 10	Active Motif	RRID:AB_2793206; Cat #: 39253
Rabbit polyclonal anti-GFP	Abcam	Cat #: ab6556
Donkey anti-rabbit IgG Alexa Fluor 488	Invitrogen	RRID: AB_2535792; Cat #: A-21206
Donkey anti-rabbit IgG CF568	Sigma	Cat #: SAB4600076
Goat anti-rabbit IgG Alexa Fluor 647	Life Technologies	RRID: AB_2535813; Cat #: A21245
Bacterial and virus strains		
Stellar Competent Cells	Clontech	Cat #: 636766
One Shot TOP10 Chemically Competent	Thermo Fisher Scientific	Cat #: C404003
10-beta Competent E.coli (High Efficiency)	NEB	Cat #: C3019H
Biological samples		
N/A		
Chemicals, peptides, and recombinant proteins		
HaloTag Ligands for Super Resolution Microscopy JF 549	Promega	Cat #: GA1111
HaloTag Ligands for Super Resolution Microscopy JF 646	Promega	Cat #: GA1121
HaloTag TMR Ligand	Promega	Cat #: PRG8252
HaloLink Resin	Promega	Cat #: G1912
Aminoallyl dUTP	Sigma-Aldrich	Cat #: A 0410
ATTO 488-NHS Ester	Sigma-Aldrich	Cat #: 41698-1MG-F
Alexa Fluor 568 NHS Ester	Thermo Fisher Scientific	Cat #: A20003
Cy3 Mono-NHS Ester	VWR	Cat #:PA13101
Cy5 NHS-Ester	VWR	Cat #: 95017-506
CF Dye Azide 568	Biotium	Cat #: 92082
Texas Red-X, Succinimidyl Ester, mixed isomers	Thermo Fisher Scientific	Cat #: T6134
ProLong Live Antifade Reagent	Thermo Fisher Scientific	Cat #: P36975
DNase I recombinant, RNase-free	Sigma-Aldrich	Cat #: 4716728001
DNA Polymerase I (10 U/μL)	Thermo Fisher Scientific	Cat #: EP0042
Phusion High-Fidelity DNA Polymerase (2 U/μL)	Thermo Fisher Scientific	Cat #: F530L
Superscript III Reverse Transcriptase	Life Technologies	Cat #: 18080-044
RNase H	NEB	Cat #: M0297L
TURBO DNase (2 U/μL)	Thermo Fisher Scientific	Cat #: AM2238
KAPA HiFi hot start Taq	Kapa Biosystems	Cat #: kk2502

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Dynabeads MyOne Streptavidin C1	Life Technologies	Cat #: 65002
Cytiva Sera-Mag SpeedBeads Carboxyl Magnetic Beads, hydrophobic	Fisher Scientific	Cat #: 09-981-123
DSG (disuccinimidyl glutarate)	Thermo Fisher Scientific	Cat #: 20593
Proteinase K, Molecular Biology Grade	NEB	Cat #: P8107S
RNase Inhibitor, Murine	NEB	Cat #: M0314L
T4 Polynucleotide Kinase	NEB	Cat #: M0201L
FastAP Thermosensitive Alkaline Phosphatase	Thermo Fisher Scientific	Cat #: EF0654
Protease Inhibitor Cocktail	Promega	Cat #: G6521
ProTEV Plus	Promega	Cat #: V6101
Vectashield	Vector Labs	Cat #: H-1000
DAPI	Thermo Fisher Scientific	Cat #: D1306
SpCas9 2NLS Nuclease (1000 pmol)	Synthego	N/A
DMSO	Sigma-Aldrich	Cat #: D2650
DMEM/F-12, HEPES, no phenol red	Life Technologies	Cat #: 11039021
Geltrex LDEV-Free Reduced Growth Factor Basement Membrane Matrix	Life Technologies	Cat #: A1413202
N2 Supplement (100X)	Thermo Fisher Scientific	Cat #: 17502048
B27 Supplement (50X), minus Vitamin A	Thermo Fisher Scientific	Cat #:12587010
Animal Free Human Activin-A(e.coli)	PeproTech	Cat #:AF-120-14E
Recombinant Human FGF-basic (FGF)	PeproTech	Cat #: 100-18B
Recombinant Human EGF Protein, CF	RnD (Perseus Proteomics)	Cat #: 2028-EG-200
Mouse LIF	Homemade	N/A
PD 0325901	Fisher Scientific	Cat #: 4192
CHIR99021	Stemgent	Cat #: 04-0004
Neurobasal Medium	Life Technologies	Cat #: 21103-049
DMEM: F12	Life Technologies	Cat #: 11320-082
Knockout DMEM	Life Technologies	Cat #: 10829018
DMEM	Life Technologies	Cat #: 11995073
FBS	Thermo Fisher Scientific	Cat #: 10437028
Glutamax I	Life Technologies	Cat #: 35050061
MEM NEAA	Life Technologies	Cat #: 11140-050
20x Penicillin/Streptomycin	Life Technologies	Cat #: 15140-163
Gelatin from porcine skin, Type A	Sigma-Aldrich	Cat #: G2500
Accutase (cell dissociation)	Life Technologies	Cat #: A11105-01
Trypsin	Life Technologies	Cat #: 25200114
DPBS	Fisher Scientific	Cat #: SH3002802
UltraPure BSA (50 mg/mL)	Thermo Fisher Scientific	Cat #: AM2616
Bovine Serum Albumin (BSA)	Sigma-Aldrich	Cat #: A7906
Fish Skin Gelatin	Sigma-Aldrich	Cat #: G7765

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Triton X-100	Sigma-Aldrich	Cat #: T8787
Tween 20	Sigma-Aldrich	Cat #: P9416
20X SSC	Life Technologies	Cat #:AM9765
PBS (10X), pH 7.4	Life Technologies	Cat #: 70011069
Dextran sulphate sodium salt	Sigma-Aldrich	Cat #: D8906
Formamide	Fisher Scientific	Cat #: F84-1
Omnipur deionized formamide	VWR	Cat #: EM-4610
Lipofectamine 3000	Life Technologies	Cat #: L3000015
RNAseOUT	Life Technologies	Cat #: 10777019
Opti-MEM I Reduced Serum Medium	Thermo Fisher Scientific	Cat #: 31985070
Tri Reagent	Zymo Research	Cat #: R2050-1-200
Thermo Scientific Pierce Methanol free Formaldehyde Ampules	Thermo Fisher Scientific	Cat #: 28908
Ribonucleoside Vanadyl Complex	NEB	Cat #: S1402S
Actinomycin D	Sigma-Aldrich	Cat #: A9415
Indole-3-Acetic Acid	Cayman Chemical Company	Cat #: 16954
Glycine,bioultra, for molecular biology, ≥99.0% (NT)	Sigma-Aldrich	Cat #: 50046
Critical commercial assays		
TrueSeq Stranded mRNA Library Prep Kit	Illumina	Cat #: 20020594
Chromium single cell 3’ reagent kit V3.1	10xGenomics	Cat #: PN-1000121
Click-IT EdU Cell Proliferation Kit for Imaging	Thermo Fisher Scientific	Cat #: C10337
In-Fusion HD Cloning	Clontech	Cat #: 639649
BioPrime Array CGH Labeling System	Life Technologies	Cat #: 18095011
Quick Ligation Kit	NEB	Cat #: M2200L
NEBNext Ultra End Repair/dA-tailing	NEB	Cat #: E7442L
4D-NucleofectorTM X Kit	Lonza	Cat #: V4XP-3024
P3 Primary Cell 4D-Nucleofector Kit S	Lonza	Cat #: V4XP-3032
GeneJET Plasmid Miniprep Kit	Thermo Fisher Scientific	Cat #: FERK0503
Direct-zol RNA MiniPrep Kit with TRI-Reagent, Zymo-Spin IIC Columns	Zymo Research	Cat #: R2051
MinElute Gel Extraction Kit	Qiagen	Cat #: 28606
NucleoBond Xtra BAC	Clontech	Cat #: 740436.25
NucleoBond Xtra Maxi	Clontech	Cat #: 740414.50
Deposited data		
Bulk RNA-seq, scRNA-seq, CLAP-seq, RAP-seq	This study	GSE181236
mus musculus Cas genome sequence	EMBL-EBI	ERP000042
mus musculus 129 genome sequence	EMBL-EBI	SRA: SRX037820
Hi-C and RNA-seq allelic counts in Smchd1−/− female NPCs	(Wang et al., 2018)	GSE99991
RNA-seq from SPOC-Bgl tethering to Xist	(Dossin et al., 2020)	GSE131784
Experimental models: Cell lines		
Mouse ESCs 129S4/SvJae/castaneus F1 2-1	(Panning et al., 1997)	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
XistMS2-GFP ESCs (F1 2-1-XIST24MS2/MCP-GFP)	This study	N/A
XistMS2-GFP/R26CIZ1mCherry	This study	N/A
XistMS2-GFP/R26CIZ1Halo	This study	N/A
XistMS2-GFP/R26CELF1mCherry	This study	N/A
XistMS2-GFP/R26PCGF5Halo	This study	N/A
XistMS2-GFP/R26PTBP1 Halo	This study	N/A
XistMS2-GFP/PyP-CAG-HaloSPEN	This study	N/A
XistMS2-GFP/PyP-CAG-HaloΔIDRSPEN	This study	N/A
XistMS2-GFP/PyP-CAG-HaloΔRRMSPEN	This study	N/A
Mouse ESCs 129S4/SvJae/castaneus F1 2-1-XIST12MS2	(Jonkers et al., 2008)	N/A
129S4ΔB/SvJae/castaneus F1 2-1 ESCs-XIST12MS2	This study	N/A
129S4ΔB/SvJae/castaneus F1 2-1 ESCs-XIST12MS2-R26SPENHalo	This study	N/A
pSM33 tetO-Xist V6.5 male mouse ESCs	(Engreitz et al., 2013)	N/A
pSM9 tetO-XistΔB V6.5 male mouse ESCs	This study	N/A
36.11 tetO-Xist cDNA transgene chr 11, male mouse ESCs	(Wutz and Jaenisch, 2000)	N/A
36.11 tetO-Xist cDNA transgene chr 11, male mouse ESCs-R26SPENHalo	This study	N/A
Mouse ESCs TX1072-Spen-GFP/Spen-GFP-BglXist-mCherry	(Dossin et al., 2020)	N/A
TX1072 ESCs -Spen-Halo/Spen-Halo	(Dossin et al., 2020)	N/A
TX1072 ESCs-Spen-GFP/Spen-GFP	(Dossin et al., 2020)	N/A
TX1072 ESCs-Spen-AID-GFP/Spen-AID-GFP	(Dossin et al., 2020)	N/A
TX1072 ESCs -ΔIDRSpen-GFP/ΔIDRSpen-GFP	This study	N/A
TX1072 ESCs-Spen-AID-GFP/Spen-AID-GFP-R26Spen-Halo	This study	N/A
TX1072 ESCs-Spen-AID-GFP/Spen-AID-GFP-R26ΔIDRSpen-Halo	This study	N/A
C127I	ATCC	Cat #: CRL-1616
Human fibroblasts - Abnormal Xi-Chromosome deletion	Coriell	Cat #: GM3827
Human fibroblasts - Abnormal Xi-Turner Syndrome	Coriell	Cat #: GM00735
Human fibroblasts - Abnormal Xi-Dicentric chromosome	Coriell	Cat #: GM06960
Human fibroblasts - Abnormal Xi-Dicentric chromosome	Coriell	Cat #: GM07213
Experimental models: Organisms/strains		
DR4 mice (for feeders)	The Jackson Laboratory	Cat #: 003208
Oligonucleotides		
Fluorescently labelled oligonucleotides used in oligoFISH, probes for RAP-seq and gRNAs for the IDRs deletion of SPEN	This study	See Table S2
Recombinant DNA		
pMS2-GFP	(Fusco et al., 2003)	Addgene plasmid cat #: #27121
pCR4-24XMS2SL-stable	(Bertrand et al., 1998)	Addgene plasmid cat #: 31865
pBglII5k plasmid	(Jonkers et al., 2008)	N/A
pBglII5k-24xMS2 plasmid	This study	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pBS31 (pgkATGfrt) plasmid	(Beard et al., 2006)	N/A
pBS32 plasmid	(Minkovsky et al., 2014)	N/A
pBS32-MCP-CIZ1 plasmid	(Pandya-Jones et al., 2020)	N/A
FRT-neo plasmid	(Beard et al., 2006)	N/A
I3-01-ct60GFP plasmid	(Hsia et al., 2016)	N/A
pBS32-cage-60GFP plasmid	This study	N/A
MXS_PGK::rtTA3-bGHpA	(Sladitschek and Neveu, 2015)	Addgene plasmid cat #: 62446
pBS31-MCP-GFP-rtTA3 plasmid	This study	N/A
H2B-mCherry	(Nam and Benezra, 2009)	Addgene plasmid cat #: 20972
EasyFusion Halo-mAID	(Gu et al., 2018a)	Addgene plasmid cat #: 112852
R26-SA-EGFP-puro	(Blelloch et al., 2004)	Addgene plasmid cat #: 26890
pYM215-R26-SA/SD-puro	This study	N/A
R26-CELF1-mCherry-puro plasmid	This study	N/A
R26-PCGF5-Halo-puro plasmid	This study	N/A
R26-PTBP1-Halo-puro plasmid	This study	N/A
R26-CIZ1-Halo-puro plasmid	This study	N/A
R26-CIZ1-mCherry-puro plasmid	This study	N/A
R26-H2B-Halo-puro plasmid	This study	N/A
pYM300-R26-Halo-SPEN-hygro plasmid	This study	N/A
pYM301-R26-Halo-ΔIDR-SPEN-hygro plasmid	This study	N/A
PyPP-CAG-Halo-full-length-Spen-V5 plasmid	This study	N/A
PyPP-CAG-Halo-Spen-ΔIDR-V5 plasmid	This study	N/A
PyPP-CAG-Halo-Spen-ΔRRM-V5 plasmid	This study	N/A
full-length mSpen Entry Clone (Sp22)	Alexander Shiskin	N/A
pFD46-R26-SPEN-hygro plasmid	(Dossin et al., 2020)	N/A
pFD82-Cas9D10A-gRNA1 plasmid	(Dossin et al., 2020)	N/A
pFD83-Cas9D10A-gRNA2 plasmid	(Dossin et al., 2020)	N/A
p15A-31-17.9kb Xist plasmid	(Pandya-Jones et al., 2020)	N/A
pCMV-Xist-PA	(Wutz and Jaenisch, 2000)	Addgene plasmid cat #: 26760
p13-5-Xist-Bdel plasmid	This study	N/A
pPGK-Cre-bpA	Klaus Rajewsky	Addgene plasmid cat #: 11543
FlpO plasmid	(Kranz et al., 2010)	N/A
BAC plasmid used to generate AtrX probe	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: RP23-265D6
Fosmid plasmid used to generate Mecp2 probe	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: WI-894A5
Fosmid plasmid used to generate Rlim probe	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: WI1-2704K12

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BAC plasmid used to generate multispectral X chromosome barcoding	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: RP23-53H15
BAC plasmid used to generate multispectral X chromosome barcoding	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: RP23-83J1
BAC plasmid used to generate multispectral X chromosome barcoding	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: RP23-451D5
BAC plasmid used to generate multispectral X chromosome barcoding	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: RP24-81K23
BAC plasmid used to generate multispectral X chromosome barcoding	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: RP24-374B8
BAC plasmid used to generate multispectral X chromosome barcoding	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: RP23-401G5
BAC plasmid used to generate multispectral X chromosome barcoding	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: RP23-104K18
BAC plasmid to generate the Xist intron 1 probe	BacPac Consortium at Children's Hospital Oakland Research Institute	Cat #: RP23-223G18
Software and algorithms		
Fiji	(Schindelin et al., 2012)	https://fiji.sc/
ImageJ	(Rueden et al., 2017)	https://imagej.nih.gov/ij/
TrackMate	(Tinevez et al., 2017)	https://imagej.net/plugins/trackmate/
3D ImageJ Suite	(Ollion et al., 2013)	https://imagej.net/plugins/3d-imagej-suite/
Python	(van Rossum and Drake, 2009)	https://www.python.org/
Google Colaboratory	Google Research	https://research.google.com/colaboratory
PyTrackmate	Hadrien Mary	https://github.com/hadim/pytrackmate
pandas	(The pandas development team, 2020)	https://pandas.pydata.org/
NumPy	(Harris et al., 2020)	https://numpy.org/
SciPy	(Virtanen et al., 2020)	https://www.scipy.org/index.html
Matplotlib	(Hunter, 2007)	https://matplotlib.org/
Seaborn	(Waskom, 2021)	https://seaborn.pydata.org/index.html
Benchling	The Benchling Life Sciences R&D	https://benchling.com.
MACS2	(Zhang et al., 2008)	https://github.com/taoliu/MACS/
Bowtie2	(Langmead et al., 2009)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
TrimGalore (v0.4.1)	Babraham Bioinformatics	https://github.com/FelixKrueger/TrimGalore
bedtools (2.26.0)	(Quinlan and Hall, 2010)	http://bedtools.readthedocs.io/en/latest
samtools (v1.7)	(Danecek et al., 2021)	http://www.htslib.org/
bcftools	Wellcome Sanger Institute	http://www.htslib.org/

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Picard (v2.1.0)	The Broad Institute	https://broadinstitute.github.io/picard/
Plyranges (v1.4.4)	(Lee et al., 2019)	https://bioconductor.org/packages/release/bioc/html/plyranges.html
Cellranger (v5.0.1)	10xGenomics	https://support.10xgenomics.com/single-cell-vdj/software/pipelines/latest/installation
Vartrix (v1.1.14)	10xGenomics	https://github.com/10XGenomics/vartrix/releases
STAR (v2.7.1a)	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
GATK (v4.1.4.1)	(Van der Auwera and O'Connor, 2020)	https://gatk.broadinstitute.org/hc/en-us
R Software Package (v3.6)	(R Core Team, 2021)	https://www.r-project.org/
RStudio	(RStudio Team, 2020)	https://www.rstudio.com/
Tidyverse	(Wickham et al., 2019)	https://www.tidyverse.org/
ggpubr	CRAN	https://cran.r-project.org/web/packages/ggpubr/index.html
Deeptools	(Ramirez et al., 2016)	https://deeptools.readthedocs.io/en/3.4.3/index.html
ChIPpeakAnno	(Zhu et al., 2010)	http://bioconductor.org/packages/release/bioc/html/ChIPpeakAnno.html
EnrichedHeatmap	(Gu et al., 2018b)	https://bioconductor.org/packages/release/bioc/html/EnrichedHeatmap.html
MATLAB	Mathworks	https://www.mathworks.com/products/matlab.html
Mathematica (v10.1)	Wolfram Research, Inc.	https://www.wolfram.com/mathematica
Other		
Salmon Sperm DNA	Sigma-Aldrich	Cat #: D9156
XMP X Green (mmX probe)	MetaSystems Probes	Cat #: D-1420-050-FI
Mouse Cot1 DNA	Life Technologies	Cat #: 18440016
Mouse flow sorted chromosome X	Gift form I. Solovei	N/A
high precision coverslips 12 mm round	Azer Scientific	Cat #: ES0117520
Correlative microscopy coverslips	Ted Pella	Cat #: 260511
Fixogum rubber cement	Fisher Scientific	Cat #: 11FIXO0125
μ-Slide 8 Well Glass Bottom	ibidi	Cat #: 80827
μ-Slide 4-well Glass Bottom	ibidi	Cat #: 80427
CoverGrip Coverslip Sealant	VWR	Cat #: 89411-108
Gene Pulser/MicroPulser Electroporation Cuvettes, 0.4 cm gap	Bio-rad	Cat #: 1652088
TetraSpeck Microspheres, 0.1 μm, fluorescent blue/green/orange/dark red	Thermo Fisher Scientific	Cat #: T7279