nature structural & molecular biology

ARTICLES

https://doi.org/10.1038/s41594-022-00739-1

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

Joanna W. Jachowicz[1,3], Mackenzie Strehle[1,3], Abhik K. Banerjee[1,2], Mario R. Blanco[1], Jasmine Thai[1] and Mitchell Guttman[1]✉

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

In recent years, thousands of lncRNAs have been identified and many have been proposed to regulate gene expression[1–5]. However, their precise mechanisms of action remain largely uncharacterized. One of the key issues is that lncRNAs are generally expressed at low levels such that the number of RNA molecules is less than the number of targets that they are proposed to regulate (substoichiometric)[6–8]. How an individual lncRNA molecule can control multiple distinct targets when it cannot engage with all of them simultaneously remains unknown and has led some to suggest that these lowly expressed lncRNAs may not be functionally important[9,10].

One example of a lncRNA that is expressed at substoichiometric levels relative to its targets is Xist. Expression of Xist is sufficient to induce transcriptional silencing of >1,000 genes across the >167 million bases of DNA on the X to achieve dosage balance of expression between males and females[11–17]. Previous studies have shown that there are ~60–200 Xist molecules within an individual cell[18–20], corresponding to an average of approximately one Xist RNA for every ten genes encoded on the X.

Xist represents an ideal system in which to explore how substoichiometric levels of a lncRNA can regulate its more-abundant targets because it is functionally important (developmentally essential)[11,21] with a clear phenotype (transcriptional silencing)[22–24] that occurs at precise and well-defined regulatory targets (X chromosome genes)[15–17]. Recent studies have begun to elucidate the mechanisms by which Xist localizes across the X and recruits silencing proteins to initiate chromosome-wide silencing. Rather than binding to specific DNA sequences, Xist diffuses from its transcription locus to DNA sites that are in close three-dimensional (3D) proximity at both genic and intergenic regions[16,17]. Xist binds directly to SHARP (also called SPEN)[22,25–28], an RNA-binding protein that associates with the SMRT and HDAC3 repressive complex to deacetylate histones[29–31], evict RNA polymerase II[22,24,32], and silence transcription on the X[22,24,25,32–34].

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

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

[1]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [2]Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [3]These authors contributed equally: J. W. Jachowicz, M. Strehle. ✉e-mail: mguttman@caltech.edu

cell (Supplementary Note). Accordingly, Xist must localize heterogeneously at distinct locations in single cells. Such heterogeneous localization would be expected to lead to heterogenous silencing where different genes are silenced in distinct cells. Yet, X silencing is not heterogenous (Supplementary Note)[24,38]. Therefore, the stoichiometric silencing model, whereby the Xist–SHARP complex localizes at each gene to silence transcription cannot explain how chromosome-wide silencing occurs at the level of single cells.

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

Results

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

To measure this, we used a female F1 hybrid (Bl6×Cast) mouse embryonic stem cell (mESC) line containing a doxycycline (dox)-inducible Xist gene at its endogenous locus and an in-frame HaloTag inserted into both copies of the endogenous SHARP protein (TX-SHARP-HALO cells)[24,39]. This system allows for more temporally synchronized expression of Xist compared with differentiation of female mESCs (Extended Data Fig. 1a), achieves robust chromosome-wide silencing by 72 h of induction (Extended Data Fig. 1b) and utilizes the same molecular components required for initiation of XCI during development and differentiation[23,40]. We induced Xist expression and visualized SHARP (using either a dye that conjugates directly to the HaloTag or an antibody against the HaloTag) along with Xist (using RNA fluorescence in situ hybridization (RNA FISH)) across five time-points (1–48 h) following dox induction (Fig. 1b and Extended Data Fig. 1c,d). Using both SHARP visualization approaches, we found that the area of the Xist-coated territory increased over time (Fig. 1c), whereas the total intensity of Xist over the territory increased initially, plateaued, and then remained relatively constant (Extended Data Fig. 1e). This means that the average Xist intensity within the territory decreased over time (Fig. 1c). By contrast, the average intensity of SHARP within the territory continued to increase across all time-points (Fig. 1c and Extended Data Fig. 1f; see Methods for quantification). Thus, the ratio of SHARP to Xist intensity is not constant across time, but instead increases in a nonlinear manner (Fig. 1d; see Fig. 1a for comparison). We note that similar kinetics were also recently observed using super-resolution microscopy[19].

To ensure that this effect is not simply a product of our synthetic dox-inducible system (Extended Data Fig. 1a), we measured the localization of Xist and SHARP across time in female mESCs upon endogenous initiation of XCI using retinoic acid (RA)-induced differentiation. We imaged Xist and SHARP after 2 and 3 days of differentiation (Extended Data Fig. 1g) and observed a similar relationship between the levels of Xist and the levels of SHARP across time: SHARP levels increased at a faster rate than Xist levels between the two time-points (Fig. 1e and Extended Data Fig. 1h). These results demonstrate that SHARP recruitment to the X occurs in a non-stoichiometric manner relative to Xist.

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

SHARP forms concentration-dependent assemblies in the nucleus. We next explored how non-stoichiometric recruitment of SHARP to the X might occur. SHARP is an ~400-kDa protein containing four RNA recognition motif (RRM) domains[41,42] that bind to Xist[26,32] and a SPEN paralog and ortholog C-terminal (SPOC) domain that is critical for recruiting the SMRT and HDAC3 proteins[24,30,31]. The remainder of SHARP is predicted to consist of long intrinsically disordered regions (IDRs; Fig. 2a). Recently, many proteins containing long IDRs have been shown to form concentration-dependent assemblies through multivalent, high-avidity associations[43–46]. On the basis of this observation, we hypothesized that SHARP might similarly form such concentration-dependent assemblies (Fig. 2b). (Although some concentration-dependent assemblies have been shown to form through phase separation, this is not the only mechanism by which they form[47,48]. In this specific context we are testing whether SHARP forms concentration-dependent assemblies, rather than the precise biophysical characteristics of their formation.)

To test this hypothesis, we explored whether SHARP exhibits known features of multivalent, high-avidity assemblies[43–46]. Specifically, we asked: (1) does SHARP form high-concentration foci in the nucleus, (2) is formation of these foci dependent on the overall concentration of SHARP, (3) are these foci dependent on multivalent associations mediated through the IDRs, and (4) are these foci dependent on associations with other molecules of SHARP (homotypic assemblies)?

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

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

To determine whether formation of SHARP assemblies is dependent on multivalent interactions driven by its IDRs, we expressed eGFP-tagged SHARP lacking its IDRs (ΔIDR-SHARP; Extended Data Fig. 2a) in HEK293T cells and imaged its behavior. In contrast to the full-length protein, ΔIDR-SHARP did not form foci (Fig. 2f and Supplementary Video 3). Instead, ΔIDR-SHARP localized diffusively throughout the nucleus, even when present at concentrations where FL-SHARP formed foci (Fig. 2g).

Finally, we explored whether these IDR-dependent assemblies form through multivalent associations with other molecules of SHARP (homotypic assemblies) or require sequence-specific

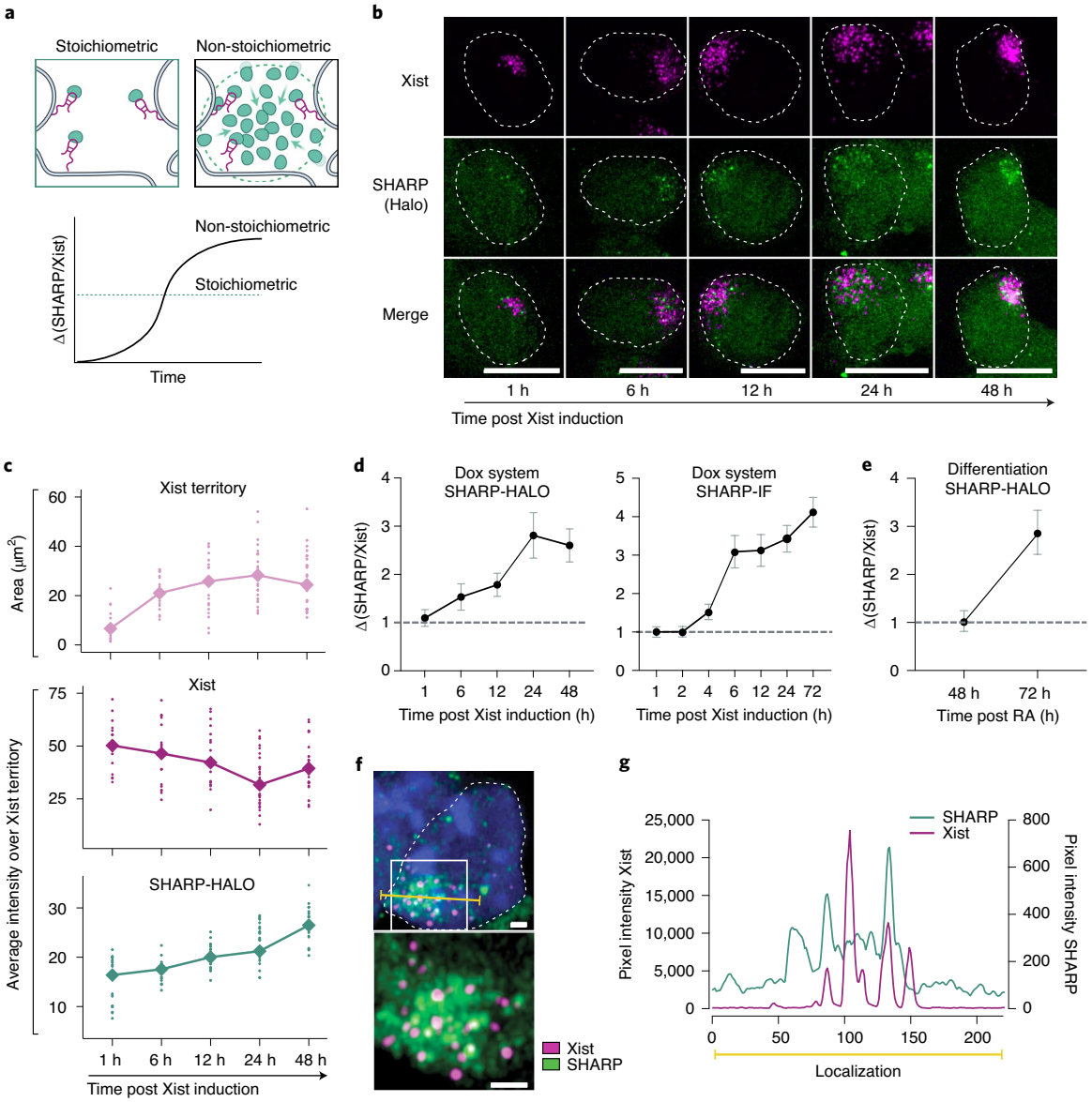


Fig. 1 | SHARP enrichment over the Xi increases in a non-stoichiometric manner relative to Xist. a, Schematic of two alternative models of SHARP recruitment to the Xi. Left: stoichiometric, SHARP (green) localizes through direct interaction with Xist (magenta). Right: non-stoichiometric, SHARP localizes even when not directly associated with Xist. In the stoichiometric model, the ratio of SHARP to Xist is directly proportional and constant across time; in the non-stoichiometric model, the concentration of SHARP increases relative to Xist across time. **b**, Representative images of Xist and SHARP localization in TX-SHARP-HALO mESCs across 48 h of Xist expression ($n > 15$ cells per condition from two independent experiments). Xist is visualized by RNA FISH (magenta) and SHARP is visualized by direct labeling of endogenous SHARP-HALO (green). Images are shown as maximum projections. Scale bars, 10 μm. **c**, Quantification of images from **b** showing area of Xist RNA territory (μm²) (upper), average FI of Xist (RNA FISH) over a unit of Xist territory (middle), and average FI of SHARP (direct HaloTag labeling) over corresponding Xist territory (lower). Dots represent individual cell measurements; diamonds represent the average. **d**, Ratio of SHARP to Xist intensities following Xist induction normalized to 1 h. Left: SHARP visualized by direct HaloTag labeling across 48 h ($n > 15$ cells for each time-point). Right: SHARP visualized by anti-Halo IF across 72 h ($n > 15$ cells for each time-point). Data are represented as mean ± s.d. **e**, Ratio of SHARP (direct HaloTag labeling) to Xist (RNA FISH) average intensities in RA-differentiated TX-SHARP-HALO mESCs normalized to 48 h ($n > 15$ cells for each time-point). Data are represented as mean ± s.d. **f**, Super-resolution imaging of Xist (RNA FISH; magenta) and endogenous SHARP-HALO (direct HaloTag labeling; green) in TX-SHARP-HALO mESCs after 24 h of Xist induction ($n > 20$ cells observed with similar phenotype): (upper) single nucleus and (lower) close-up on Xist territory from the upper image demarcated by the white box. Images are shown as maximum projections. Scale bars, 1 μm. The yellow line shows where the intensity profile (**g**) was measured. **g**, Line intensity profile from image in **f** showing Xist and SHARP intensities.

associations with other proteins (heterotypic assemblies). To do this, we fused ΔIDR-SHARP to an mCherry-tagged version of the IDR of the fused in sarcoma (FUS) protein, a domain that is known to form multivalent homotypic associations both in vitro and in vivo[51–53] (FUS-ΔIDR-SHARP; Extended Data Fig. 2a), and tested whether this synthetic protein rescues the ability of SHARP to form foci independently of its IDRs. We observed that FUS-ΔIDR-SHARP forms assemblies in the nucleus that are comparable with those observed for FL-SHARP (Fig. 2h). Although we do not exclude the possibility that the IDRs of SHARP may form heterotypic associations with other molecules, these results indicate that homotypic associations are essential for SHARP to form the observed assemblies.

Together, these results indicate that SHARP forms concentration-dependent assemblies in the nucleus and that formation of these

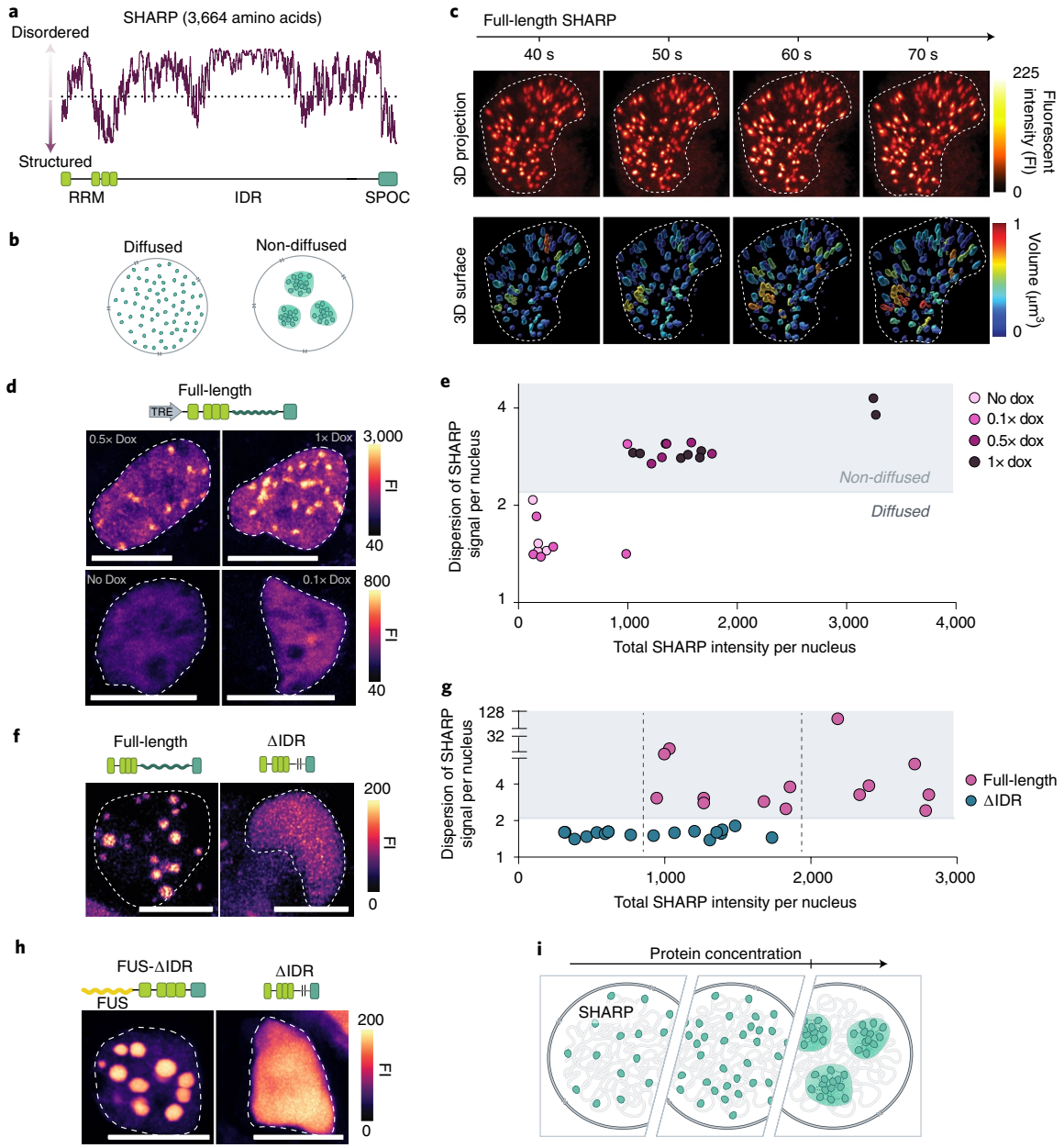


Fig. 2 | SHARP forms multivalent, concentration-dependent assemblies in the nucleus. a, Disordered scores across the SHARP protein using IUPred2 software predictions. The dotted line represents the 0.5 probability value for a given structure to be ordered. The lower visualization demarcates the position of known SHARP domains—RRM (bright green) and SPOC (dark green). **b**, Schematic representation of molecules within a nucleus organized in a diffused or non-diffused (focal) manner. **c**, Images across four time-points from a live-cell video of eGFP-tagged FL-SHARP in transiently transfected HEK293T cells showing non-diffused, focal organization of SHARP molecules: (upper) 3D reconstructions of the FI signal; and (lower) 3D volume reconstructions color-coded based on the size of the condensate. **d**, Images representing localization patterns of eGFP-tagged FL-SHARP in transiently transfected HEK293T cells across increasing expression levels of SHARP (driven by dox-inducible promoter; 1× dox = 2 μg ml^{-1}). Images are shown as maximum projections. Scale bars, 10 μm. **e**, Quantification of images (Fig. 1d) plotting the dispersion of SHARP signal across the nucleus versus average SHARP FI per nucleus. **f**, Representative images of FL-SHARP and ΔIDR-SHARP localization patterns in transiently transfected HEK293T cells. Images are shown as maximum projections. Scale bars, 10 μm. **g**, Quantification of images (Fig. 1f) plotting the dispersion of SHARP signal across the nucleus versus average SHARP FI per nucleus. The dashed line represents the range of FI that is similar for both groups. **h**, Images representing localization patterns of mCherry-tagged FUS-ΔIDR-SHARP and eGFP-tagged ΔIDR-SHARP in transiently transfected HEK293T cells. Images are shown as maximum projections. Scale bars, 10 μm. **i**, Schematic depicting formation of concentration-dependent SHARP assemblies.

assemblies is dependent on homotypic multivalent interactions driven by its IDRs (Fig. 2i).

SHARP recruitment to the Xi depends on IDR-mediated assembly. To determine whether IDR-dependent assemblies of SHARP are critical for its enrichment on the Xi, we tested whether deletion of the IDRs impacts localization over the X. To do this, we first generated a mESC line containing a deletion of both copies of the endogenous SHARP gene (SHARP-KO; Extended Data Fig. 3a). In parallel, we utilized mESCs containing an auxin-degradable SHARP (SHARP-AID)[24]. Within each of these lines (SHARP-KO and SHARP-AID), we stably expressed a Halo-tagged or eGFP-tagged

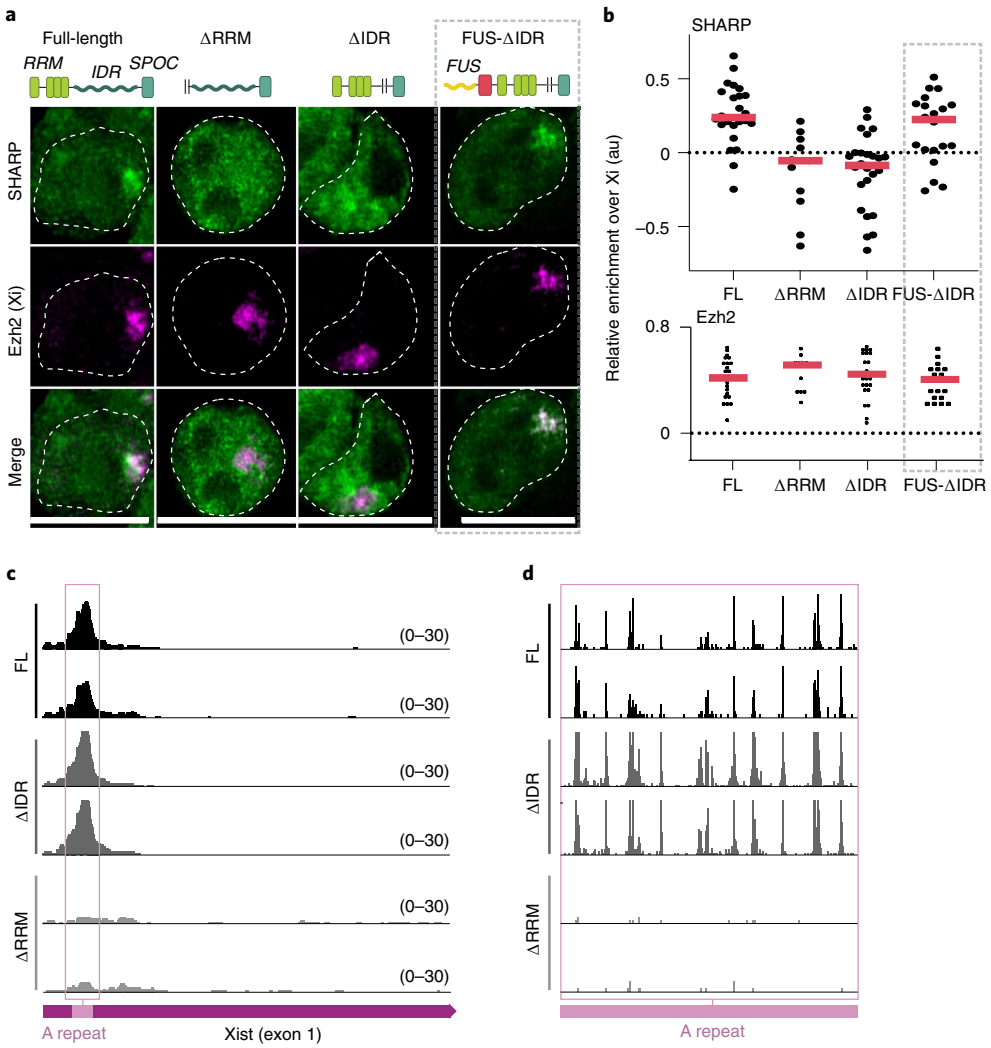
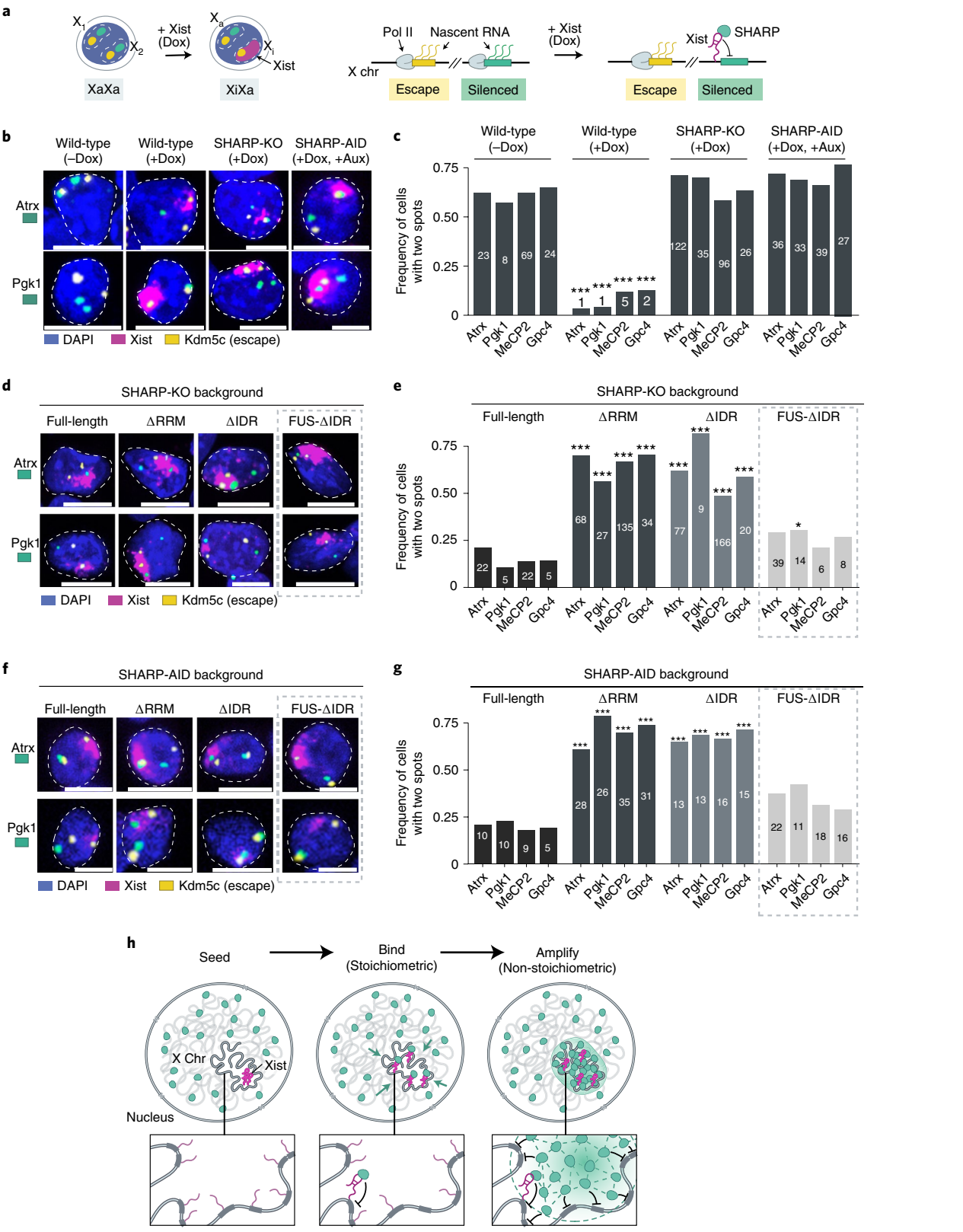


Fig. 3 | Formation of SHARP assemblies is required for enrichment on the Xi but not Xist binding. a, Representative images of SHARP enrichment (eGFP, green) over the Xi (anti-Ezh2 IF, magenta) in TX SHARP-KO mESCs containing dox-inducible Xist, genetic deletion of SHARP and stable integrations of eGFP-FL-SHARP, eGFP-ΔRRM-SHARP, eGFP-ΔIDR-SHARP, or FUS-mCherry-ΔIDR-SHARP constructs (see Extended Data Fig. 3a–c for cell line details). Xist induction for 72 h; images are shown as z-sections. Scale bars, 10 μm. **b**, Quantification of images from **a** plotting (upper) SHARP FI over the Xi (denoted by Ezh2) normalized to the FI of a random nuclear region of the same size or (lower) Ezh2 FI over the same area normalized to a random nuclear region (see Extended Data Fig. 3d for quantification details). Values for individual nuclei ($n > 11$ for each condition) are shown; red lines represent median values; 0 represents enrichment not higher than that measured over a random nuclear region. **c**, SHARP enrichment across the first exon of Xist after UV-crosslinking and purification using the HaloTag in female TX SHARP-AID mESCs treated with auxin. HaloTags were fused to FL-SHARP (upper), ΔIDR-SHARP (middle) or ΔRRM-SHARP (lower; see Extended Data Fig. 3b,c for cell line details). Two replicates shown for each cell line; magenta arrow represents beginning of the first Xist exon; pink square demarcates Xist A-repeat (SHARP binding site). **d**, Close-up of crosslink-induced truncation sites for the three rescue conditions on the A-repeat region from **c** (demarcated by pink square).

version of either full-length SHARP (FL-SHARP), SHARP containing a deletion of its RRM domains (ΔRRM-SHARP), or SHARP containing a deletion of its IDRs (ΔIDR-SHARP) (Extended Data Fig. 3b,c). We visualized each of these tagged SHARP proteins along with Ezh2 (to demarcate the Xi) after Xist expression for >72 h (Fig. 3a and Extended Data Fig. 3d). As expected, FL-SHARP was enriched over the Xi compartment. By contrast, ΔRRM-SHARP failed to localize on the Xi, consistent with previous observations[24]. Interestingly, we also observed a strong decrease in the enrichment of ΔIDR-SHARP over the Xi, comparable with that observed upon deletion of the RRM domains (Fig. 3b; see Extended Data Fig. 3e for quantification schematics). The level of Ezh2 was similar in all conditions (Fig. 3b).

Because SHARP binds directly to Xist[22,25–27], we tested whether the ΔRRM- and ΔIDR-SHARP mutants fail to localize on the Xi simply because they cannot bind Xist. To do this, we UV-crosslinked intact cells to form a covalent crosslink between directly interacting proteins and RNA, purified the Halo-tagged SHARP proteins using fully denaturing conditions, and sequenced the associated RNAs (Methods). We observed that FL-SHARP forms a highly specific interaction with the A-repeat region of Xist. By contrast, expression of ΔRRM-SHARP ablated this interaction across Xist. Interestingly, ΔIDR-SHARP is still able to bind the A-repeat of Xist at levels and positions comparable with those observed for FL-SHARP (Fig. 3c,d). These observations are consistent with previous studies showing that the RRM domains of SHARP are sufficient to bind to Xist[26,32]. Together, these results demonstrate that the IDRs of SHARP are essential for its enrichment on the Xi (Fig. 3a,b) even though they are not required for direct binding to Xist (Fig. 3c,d). A recent study similarly found that the IDRs of SHARP are required for its accumulation on the Xi[19].



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

Fig. 4 | SHARP binding to RNA and formation of assemblies are both required for chromosome-wide gene silencing. a, Illustration of RNA FISH measurements in dox-inducible female mESCs. Genes that are silenced upon Xist induction are shown in green, genes that escape XCI are shown in yellow, and Xist is shown in magenta. chr, chromosome. **b**, RNA FISH images representing (left to right): wild-type (no dox), wild-type (with dox), SHARP-KO (with dox) and auxin-treated SHARP-AID (with dox) cells. Cells were stained for DAPI (blue) and probed for Xist (magenta), escape gene Kdm5c (yellow), and silenced genes Atrx or Pgk1 (green). Images are shown as maximum projections. Scale bars, 10 μm. Total cell numbers (**c**) from three independent experiments. **c**, Quantification of RNA FISH images (Fig. 1b) representing the frequency of cells containing two actively transcribed alleles. Asterisks represent *P* values calculated for two-sided *z*-test; distributions compared with the FL group. $^{*}P = 0.05$, $^{***}P = 0.001$. **d**, RNA FISH images for SHARP-KO female mESCs containing stable integrations of SHARP rescue constructs. Cells were stained as described in **b**. Images are shown as maximum projections. Scale bars, 10 μm. Total cell numbers (**e**) from three independent experiments. **e**, Quantification of RNA FISH images (Fig. 1d) representing the frequency of cells containing two actively transcribed alleles for the various SHARP rescue constructs in SHARP-KO female mESCs. Asterisks represent *P* values calculated for two-sided *z*-test; distributions compared with the FL group. $^{*}P = 0.05$, $^{***}P = 0.001$. **f**, RNA FISH images for SHARP-AID female mESCs containing stable integrations of SHARP rescue constructs. Cells were stained as described in **b**. Images are shown as maximum projections. Scale bars, 10 μm. Total cell numbers (**g**) from three independent experiments. **g**, Quantification of RNA FISH images (Fig. 1f) representing the frequency of cells containing two actively transcribed alleles for the various SHARP rescue constructs in SHARP-AID female mESCs. Asterisks represent *P* values calculated for two-sided *z*-test; distributions compared with the FL group. $^{*}P = 0.05$, $^{***}P = 0.001$. **h**, Schematic of the spatial amplification mechanism by which Xist RNA (magenta) can act to amplify SHARP (green) recruitment and gene silencing across the X.

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

To measure silencing, we performed RNA FISH on Xist and the introns of: (1) several genes located across the X that are known to be silenced upon XCI, and (2) genes that are known to escape XCI and therefore remain active upon Xist induction (Fig. 4a and Extended Data Fig. 4a). This single-cell readout allows us to restrict our analyses to cells that induce Xist expression (~50% of cells) and retain both X chromosomes (~50% of cells; Extended Data Fig. 4b)[39,54,55]. Of these cells, we found that ~80% successfully silenced gene expression on one of the two X chromosomes upon Xist induction in wild-type mESCs (Fig. 4b,c and Extended Data Fig. 4c). Next, we measured gene silencing upon genetic deletion (SHARP-KO) or auxin-mediated degradation (SHARP-AID) of SHARP and found that both conditions led to loss of Xist-mediated transcriptional silencing (Fig. 4b,c and Extended Data Fig. 4c).

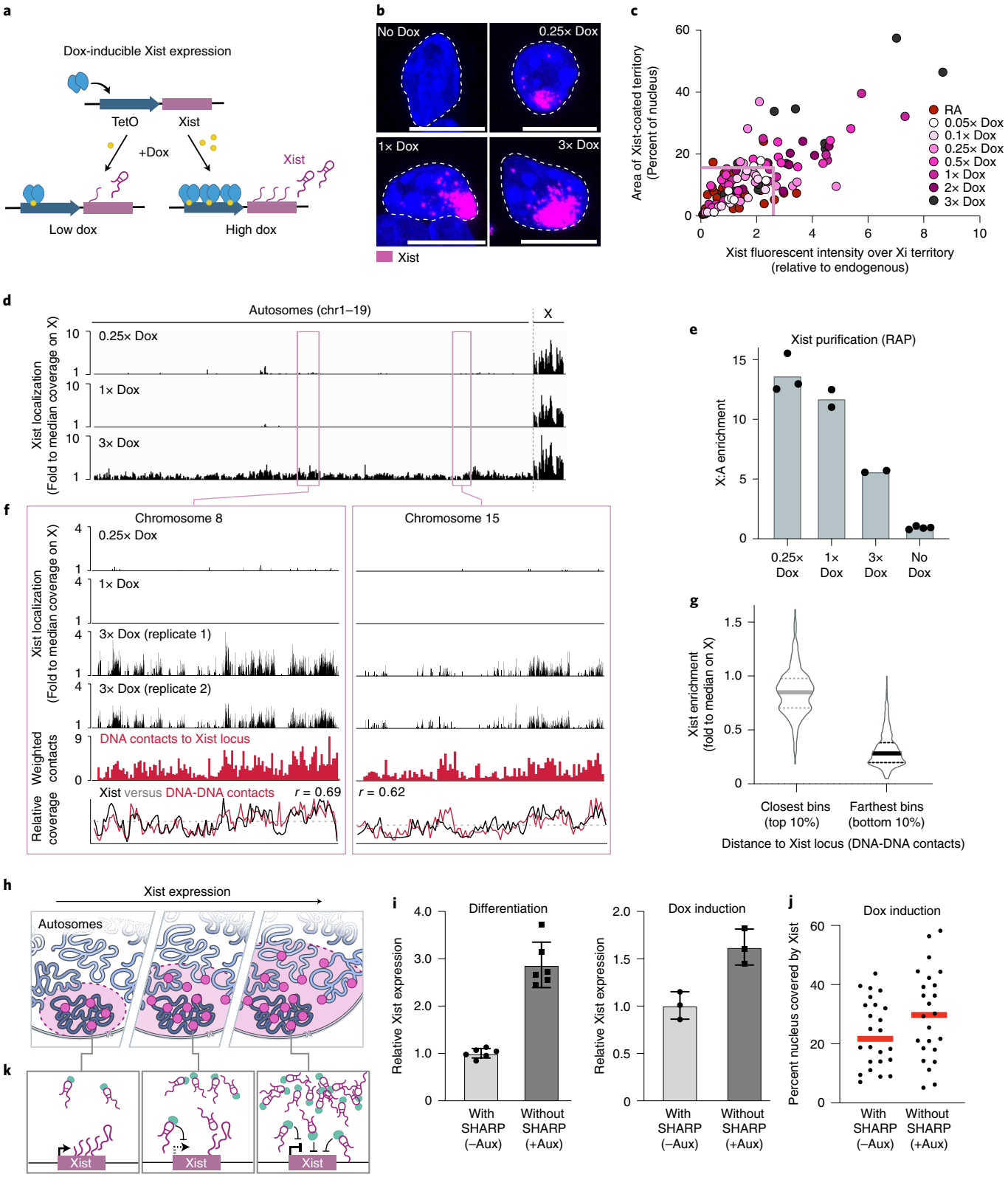
We measured transcription of the same X-linked genes after stable expression of FL-SHARP, ΔRRM-SHARP or ΔIDR-SHARP in both SHARP-KO and SHARP-AID backgrounds (Fig. 4d,f and Extended Data Fig. 4d,f). As expected, expression of FL-SHARP rescued silencing of these X-linked genes. By contrast, expression of ΔRRM-SHARP failed to silence any of these X-linked genes, consistent with the fact that it can no longer bind to Xist. Importantly, expression of ΔIDR-SHARP also failed to silence these genes (Fig. 4e,g and Extended Data Fig. 4e,g). Consistent with this observation, a recent study similarly found that deletion of the SHARP IDRs impacts transcriptional silencing across the entire chromosome when measured by single-cell RNA sequencing[19].

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

Together, these results demonstrate that direct binding of SHARP to Xist (via its RRM domains) and its ability to form concentration-dependent homotypic assemblies (via its IDRs) are both essential and distinct components required for chromosome-wide silencing on the Xi. Our data suggest a spatial amplification mechanism, whereby the direct interaction between Xist (which is enriched on the X) and SHARP (which is diffusible throughout the nucleus) acts to increase the local concentration of SHARP on the X. The resulting high local concentrations of SHARP on the X chromosome enable formation of IDR-mediated concentration-dependent assemblies that can occur between molecules not directly bound to Xist. In this way, these RNA-mediated assemblies can lead to the accumulation of SHARP on the Xi in stoichiometric excess of the number of Xist molecules to enable chromosome-wide silencing (Fig. 4h).

Low Xist expression levels limit spreading to autosomes. This spatial amplification mechanism explains how Xist can achieve chromosome-wide silencing despite being expressed at substoichiometric levels relative to its target genes (Extended Data Fig. 5a–c and Supplementary Note). However, it does not explain why Xist

Fig. 5 | Low Xist expression levels limit its ability to spread to autosomes. a, Modulation of Xist levels using a dox-inducible system. TetO, Tet-On promoter. **b**, Images of Xist (magenta) within nucleus (DAPI) of mESCs treated with increasing dox concentrations. Scale bars, 10 μm. Total cell numbers (**c**) from two independent experiments. **c**, Percent of nucleus covered by Xist relative to Xist levels in RA-differentiated and dox-induced mESCs; box represents 95th percentile of Xist expression and area upon RA-differentiation. **d**, Xist enrichment relative to total genomic DNA measured by RAP and normalized to the median coverage across the X in mESCs treated with 0.25×, 1× and 3× dox. **e**, Xist enrichment on the X relative to autosomes (A) in 0.25×, 1× and 3× dox-induced and uninduced (no dox) female mESCs as measured by the proportion of sequencing reads that align to the X relative to A in RAP-DNA samples normalized to the expected X:A ratio observed in the unselected genomic DNA sample (input). Dots represent individual replicates; bars represent mean value. **f**, Xist enrichment over chr8 (right) and chr15 (left) measured by RAP; Xist enrichment relative to median coverage on the X (black bars); DNA contact frequency[56] of each region relative to the Xist locus (red bars); overlay between Xist enrichment and 3D contacts with the Xist locus. **g**, Xist enrichment in 3× dox sample over 1 Mb autosomal regions closest to Xist locus (left, top 10%) or furthest from Xist locus (right, bottom 10%) based on SPRITE data[56]; bold lines represent median values; dotted lines represent 25th and 75th percentiles. **h**, Schematic depicting increased Xist spreading with increasing Xist expression levels. **i**, Relative Xist expression in RA-differentiated (left) and dox-induced (right) female SHARP-AID cells in the absence or presence of auxin. Dots represent individual replicates; $n = 6$ for differentiation conditions; $n = 3$ for dox induction conditions; data are represented as mean ± s.d. **j**, Percent of nucleus occupied by Xist in dox-induced SHARP-AID cells in the absence or presence of auxin. Dots represent individual replicates; red bar represents the median. **k**, Model illustrating how Xist (through SHARP) may suppress production of its own RNA through negative feedback.



expression levels are low and whether this might be critical for its functional role during XCI. Because Xist spreads to sites on the X based on 3D diffusion from its transcription locus[16,17], we hypothesized that its expression level might control how far it spreads in the nucleus. If true, we would expect that expressing Xist at increasing concentrations would lead to increasing localization of Xist to autosomal regions.

To test this, we used our dox-inducible Xist system, which enables induction of Xist across a range of expression levels by titrating the concentration of dox (Fig. 5a). We induced Xist expression across a range of dox concentrations (referred to as a 0.05×–3× Dox, where 1× = 2 μg ml^{-1}), imaged Xist localization in individual cells (Fig. 5b) and observed a strong correlation between Xist expression levels and the area of the nucleus it occupies within individual cells

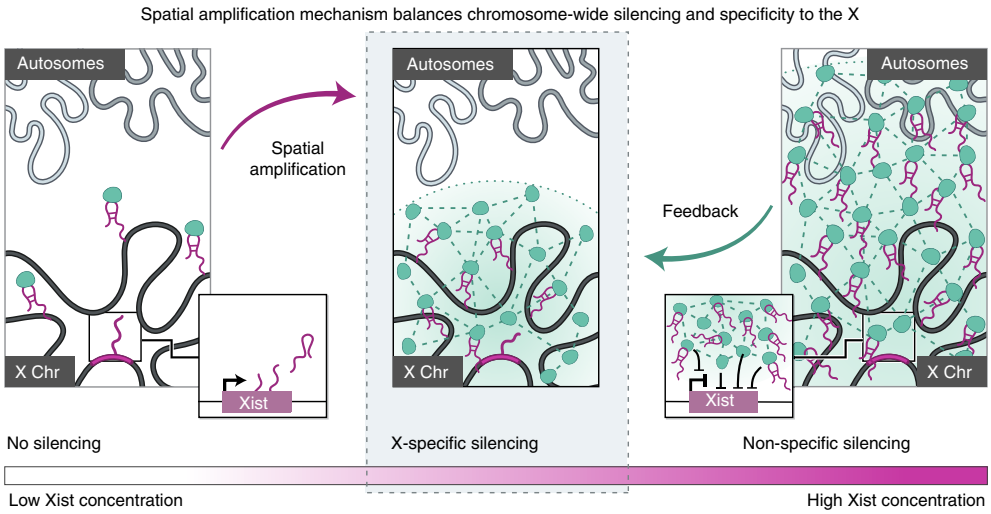


Fig. 6 | The spatial amplification mechanism balances chromosome-wide silencing and specificity to the X. Schematic illustrating the spatial amplification mechanism. Left: Xist is expressed, accumulates on the X through 3D diffusion from its transcription locus, binds directly to SHARP and recruits it to the X in a stoichiometric manner. Close-up on left: at low overall expression, the Xist gene remains actively transcribed. Middle: once SHARP molecules achieve sufficiently high spatial concentration over the X, they form concentration-dependent assemblies (spatial amplification) that enable non-stoichiometric accumulation of SHARP on the X and chromosome-wide silencing. Right: if Xist expression levels get too high, the RNA would start to spread to autosomal regions. Close-up on right: at high concentrations, Xist recruits more SHARP molecules to its own locus which can suppress transcription of its own gene. This acts to reduce Xist spreading and restrain the Xist–SHARP complex on the X (feedback).

(r=0.75; Fig. 5c). Specifically, Xist occupies on average ~6.5% of the area of the nucleus when expressed upon RA-induced differentiation (endogenous control). However, cells treated with 3× dox express on average ~3.4-fold higher levels of Xist (relative to average endogenous levels) and Xist occupies on average ~23% of the area of the entire nucleus (Fig. 5c).

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

Because Xist spreads via 3D diffusion, we hypothesized that the autosomal regions that become occupied at increasing dox concentrations are those that are closest to the Xist locus in 3D space. To test this, we computed the 3D contact frequency between the Xist genomic locus and all 1-Mb genomic regions across autosomes[56] (Fig. 5f and Extended Data Fig. 5e). Interestingly, we observed a strong correlation between autosomal regions that are closest to the Xist locus and those that display increased Xist RNA occupancy in the 3× dox condition (Fig. 5g and Extended Data Fig. 5f). Taken together, these results indicate that substoichiometric expression of Xist (low number of Xist molecules) is a critical mechanism by which cells limit Xist spreading to autosomal regions and ensure its specificity to the X (Fig. 5h).

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

Discussion

Our results demonstrate a critical spatial amplification mechanism by which Xist balances two essential but countervailing regulatory objectives: specificity to the X and chromosome-wide gene silencing (Fig. 6). We showed that low Xist RNA levels are necessary to ensure specificity to its target sites on the X. Yet, it creates another challenge in that the RNA is expressed at substoichiometric levels compared with its regulatory targets and therefore cannot localize at each of them. We showed that Xist overcomes this challenge by driving non-stoichiometric recruitment of SHARP to amplify its abundance across the X and enable chromosome-wide gene silencing. Although a stoichiometric model (in which Xist recruits SHARP through direct binding and localizes at each of its target genes) would also enable chromosome-wide silencing, it would require Xist to be expressed at dramatically higher levels and therefore reduce Xist specificity to the X. Although the spatial amplification

mechanism can achieve both specificity and robust silencing, balancing these two competing objectives requires precise quantitative control of Xist RNA levels. Our results highlight a negative feedback loop, whereby Xist (through SHARP) may act to suppress production of its own RNA to restrict its ability to spread beyond the X (Fig. 6).

This spatial amplification mechanism is dependent on the fact that Xist can form a high-concentration territory on the X through 3D diffusion from its transcription locus (seed). In this way, Xist binding to SHARP increases its concentration on the X (recruit) and enables formation of concentration-dependent protein assemblies that amplify recruitment of repressive proteins to enable chromosome-wide gene silencing (amplify, Fig. 4h). Furthermore, because Xist spreads to its targets via 3D diffusion from its transcription locus, localization specificity is sensitive to its overall expression levels (restriction, Fig. 5h). Beyond Xist, this spatial amplification mechanism is likely to represent a more general mechanism by which lncRNAs can balance specificity to, and robust control of, their regulatory targets because many lncRNAs share these same properties. Specifically, many hundreds of lncRNAs have been shown to form high-concentration territories in spatial proximity to their transcription sites (seed) and can directly bind and recruit different regulatory proteins (recruit), including those that contain long IDRs[61] (for example, HP1 (refs. [62,63]) and SHARP). In this way, lncRNA-mediated recruitment may enable spatial amplification of regulatory proteins and robust regulation of their more-abundant targets (amplification). Because many lncRNAs localize in 3D proximity to their targets, low expression levels may similarly be important for ensuring specificity to their genomic DNA targets (restriction).

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

(While we were working on this manuscript, a parallel study (Markaki et al.)[19] made similar observations about the dynamics of Xist and SHARP localization on the X and the role of the IDRs of SHARP for localization and silencing on the X. These observations further support a key aspect of our spatial amplification model, specifically that non-stoichiometric recruitment of SHARP amplifies its abundance across the X to enable chromosome-wide gene silencing.)

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-022-00739-1.

Received: 15 January 2021; Accepted: 28 January 2022;
Published online: 17 March 2022

References

1. Guttman, M. et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).
2. Engreitz, J. M. et al. Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).
3. Carninci, P. et al. The transcriptional landscape of the mammalian genome. *Science* **309**, 1559–1563 (2005).
4. Lee, J. T. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes Dev.* **23**, 1831–1842 (2009).
5. Nagano, T. et al. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**, 1717–1720 (2008).
6. Dunagin, M., Cabili, M. N., Rinn, J. & Raj, A. Visualization of lncRNA by single-molecule fluorescence in situ hybridization. *Methods Mol. Biol.* **1262**, 3–19 (2015).
7. Wang, K. C. et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120–124 (2011).
8. Carpenter, S. et al. A long noncoding RNA mediates both activation and repression of immune response genes. *Science* **341**, 789–792 (2013).
9. Thomson, D. W. & Dinger, M. E. Endogenous microRNA sponges: evidence and controversy. *Nat. Rev. Genet.* **17**, 272–283 (2016).
10. Kopp, F. & Mendell, J. T. Functional classification and experimental dissection of long noncoding RNAs. *Cell* **172**, 393 (2018).
11. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131–137 (1996).
12. Marahrens, Y., Loring, J. & Jaenisch, R. Role of the Xist gene in X chromosome choosing. *Cell* **92**, 657–664 (1998).
13. Dyer, K. A., Canfield, T. K. & Gartler, S. M. Molecular cytological differentiation of active from inactive X domains in interphase: implications for X chromosome inactivation. *Cytogenet. Genome Res.* **50**, 116–120 (1989).
14. Chaumeil, J., Le Baccon, P., Wutz, A. & Heard, E. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev.* **20**, 2223–2237 (2006).
15. Clemson, C. M., McNeil, J. A., Willard, H. F. & Lawrence, J. B. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* **132**, 259–275 (1996).
16. Engreitz, J. M. et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**, 1237973 (2013).
17. Simon, M. D. et al. High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504**, 465–469 (2013).
18. Sunwoo, H., Wu, J. Y. & Lee, J. T. The Xist RNA–PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proc. Natl Acad. Sci. USA* **112**, E4216–E4225 (2015).
19. Markaki, Y. et al. Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **184**, 6174–6192 (2021).
20. Rodermund, L. et al. Time-resolved structured illumination microscopy reveals key principles of Xist RNA spreading. *Science* **372**, eabe7500 (2021).
21. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. & Jaenisch, R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* **11**, 156–166 (1997).
22. McHugh, C. A. et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232–236 (2015).
23. Wutz, A., Rasmussen, T. P. & Jaenisch, R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* **30**, 167–174 (2002).
24. Dossin, F. et al. SPEN integrates transcriptional and epigenetic control of X-inactivation. *Nature* **578**, 455–460 (2020).
25. Chu, C. et al. Systematic discovery of Xist RNA binding proteins. *Cell* **161**, 404–416 (2015).
26. Lu, Z. et al. RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* **165**, 1267–1279 (2016).
27. Minajigi, A. et al. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**, aab2276 (2015).
28. Cirillo, D. et al. Quantitative predictions of protein interactions with long noncoding RNAs. *Nat. Methods* **14**, 5–6 (2016).
29. Żylicz, J. J. et al. The implication of early chromatin changes in X chromosome inactivation. *Cell* **176**, 182–197 (2019).
30. You, S. H. et al. Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 in vivo. *Nat. Struct. Mol. Biol.* **20**, 182–187 (2013).
31. Shi, Y. et al. Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev.* **15**, 1140–1151 (2001).
32. Monfort, A. & Wutz, A. Identification of Spen as a crucial factor for Xist function through forward genetic screening in haploid embryonic stem cells. *Cell Rep.* **12**, 554–561 (2015).
33. Moindrot, B. et al. A pooled shRNA screen identifies Rbm15, Spen, and Wtap as factors required for Xist RNA-mediated silencing. *Cell Rep.* **12**, 562–572 (2015).
34. Nesterova, T. B. et al. Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. *Nat. Commun.* **10**, 3129 (2019).
35. Wutz, A. & Jaenisch, R. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell* **5**, 695–705 (2000).
36. Lee, J. T., Strauss, W. M., Dausman, J. A. & Jaenisch, R. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell* **86**, 83–94 (1996).
37. Hall, L. L. et al. An ectopic human XIST gene can induce chromosome inactivation in postdifferentiation human HT-1080 cells. *Proc. Natl Acad. Sci. USA* **99**, 8677–8682 (2002).
38. Pacini, G. et al. Integrated analysis of Xist upregulation and X-chromosome inactivation with single-cell and single-allele resolution. *Nat. Commun.* **12**, 3638 (2021).

39. Schulz, E. G. et al. The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell* **14**, 203–216 (2014).
40. Barros de Andrade E Sousa, L. et al. Kinetics of Xist-induced gene silencing can be predicted from combinations of epigenetic and genomic features. *Genome Res.* **29**, 1087–1099 (2019).
41. Arieti, F. et al. The crystal structure of the Split End protein SHARP adds a new layer of complexity to proteins containing RNA recognition motifs. *Nucleic Acids Res.* **42**, 6742–6752 (2014).
42. Newberry, E. P., Latifi, T. & Towler, D. A. The RRM domain of MINT, a novel Msx2 binding protein, recognizes and regulates the rat osteocalcin promoter. *Biochemistry* **38**, 10678–10690 (1999).
43. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid–liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
44. Lin, Y., Currie, S. L. & Rosen, M. K. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J. Biol. Chem.* **292**, 19110–19120 (2017).
45. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
46. Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855 (2018).
47. Bhat, P., Honson, D. & Guttman, M. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nat. Rev. Mol. Cell Biol.* **22**, 653–670 (2021).
48. McSwiggen, D. T., Mir, M., Darzacq, X. & Tjian, R. Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev.* **33**, 1619–1634 (2019).
49. Banani, S. F. et al. Compositional control of phase-separated cellular bodies. *Cell* **166**, 651–663 (2016).
50. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).
51. Kato, M. et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
52. Bracha, D. et al. Mapping local and global liquid phase behavior in living cells using photo-oligomerizable seeds. *Cell* **175**, 1467–1480 (2018).
53. Lin, Y., Protter, D. S. W., Rosen, M. K. & Parker, R. Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell* **60**, 208–219 (2015).
54. Choi, J. et al. Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature* **548**, 219–223 (2017).
55. Yagi, M. et al. Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. *Nature* **548**, 224–227 (2017).
56. Quinodoz, S. A. et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757 (2018).
57. Wollerton, M. C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. & Smith, C. W. J. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol. Cell* **13**, 91–100 (2004).
58. Ottens, F. & Gehring, N. H. Physiological and pathophysiological role of nonsense-mediated mRNA decay. *Eur. J. Physiol.* **468**, 1013–1028 (2016).
59. Müller-McNicoll, M., Rossbach, O., Hui, J. & Medenbach, J. Auto-regulatory feedback by RNA-binding proteins. *J. Mol. Cell Biol.* **11**, 930–939 (2019).
60. Rossbach, O. et al. Auto- and cross-regulation of the hnRNP L proteins by alternative splicing. *Mol. Cell. Biol.* **29**, 1442–1451 (2009).
61. Quinodoz, S. A. et al. RNA promotes the formation of spatial compartments in the nucleus. *Cell* **184**, 5775–5790 (2021).
62. Larson, A. G. et al. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
63. Strom, A. R. et al. Phase separation drives heterochromatin domain formation. *Nature* **547**, 241–245 (2017).

Methods

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

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

Human HEK293T cell culture. HEK293T cells were cultured in complete media comprising DMEM (Gibco, Life Technologies) supplemented with 10% FBS (Seradigm Premium Grade HI FBS, VWR), 1× penicillin–streptomycin (Gibco, Life Technologies), 1 mM sodium pyruvate (Gibco, Life Technologies) and 1× MEM non-essential amino acids (Gibco, Life Technologies). Cells were maintained in 37 °C incubators under 5% CO_2.

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

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

Cell line generation. *SHARP-KO cells.* To generate a plasmid targeting SHARP for deletion (see Supplementary Table 1 for complete plasmid list), four different guide RNA sequences (see Supplementary Table 2 for sequences; Extended Data Fig. 3a) were multiplexed into a Cas9–nickase backbone (Addgene plasmid 48140) as previously described[64]. To create a SHARP knockout (SHARP-KO) cell line, two million TX1072 mESCs were transfected with 1.25 μg of the multiplexed Cas9n–gRNA plasmid containing green fluorescent protein (GFP) using the Neon transfection system (settings: 1,400 V, 10 ms width, three pulses). Successfully transfected cells were enriched by performing fluorescence-activated cell sorting (FACS) for GFP and subsequently plated at low-confluency. After 4–5 days of growth, 96 single colonies were picked and seeded in a 96-well plate. These cells were then split into one plate for PCR genotyping and another plate for maintaining growth until positive clones were identified. PCR genotyping was performed using Q5 High-Fidelity 2× Master Mix (NEB) with the primer pairs listed in Supplementary Table 2. SHARP-KO clone H8 was used for subsequent experimentation and all other clones were frozen.

SHARP rescue lines in SHARP-KO or SHARP-AID parent cells. To generate SHARP rescue cell lines, SHARP rescue constructs (FL-SHARP, ΔRRM-SHARP, ΔIDR-SHARP, FUS-ΔIDR-SHARP; Extended Data Fig. 3b; see Supplementary Table 1 for complete plasmid list) were first made using the Gateway LR Clonase system (ThermoFisher). Specifically, ΔRRM-SHARP and ΔIDR-SHARP entry clones were created by modifying a full-length mouse SHARP entry clone using polymerase incomplete primer extension mutagenesis[65]. The specific amino acids deleted in the ΔRRM-SHARP and ΔIDR-SHARP entry clones are as follows:

ΔRRM-SHARP: amino acids 2–590;

ΔIDR-SHARP: amino acids 639–3460.

These entry clones (FL-SHARP, ΔRRM-SHARP, ΔIDR-SHARP) were then recombined into two different modified versions of the dox-inducible PiggyBac destination vector PB-TAG-ERN (Addgene plasmid 80476) containing truncated human nerve growth factor receptor (NGFR) and HALO or eGFP. This destination vector was chosen because it enables stable integration of the rescue constructs by cotransfecting with a PiggyBac transposase[66]. The Halo-tagged version of this plasmid was created by replacing eGFP with NGFR (Addgene plasmid 27489) using Gibson assembly (NEB). HALO was then introduced downstream of reverse tetracycline-controlled transactivator using restriction enzyme digestion and ligation to create PB-HALO-IRES-NGFR. To generate the eGFP-tagged version of this plasmid, HALO was replaced with a 6-HIS-TEV-eGFP sequence using restriction enzyme digestion and ligation. Importantly, eGFP in this construct contains an amino acid substitution (A206K) to create a monomeric variant[67].

FUS-ΔIDR-SHARP was generated by recombining the ΔIDR-SHARP entry clone into a modified version of the PB-HALO-IRES-NGFR vector containing the IDR sequence from the FUS protein tagged with mCherry (Addgene plasmid 101223) in place of HALO. Importantly, the IDR from FUS exhibits no sequence homology to endogenous SHARP IDRs (that is, they have distinct amino acid compositions and distinct proportions of amino acid charge properties), its sequence is approximately tenfold shorter that SHARP IDRs, and the locations of these two IDRs within the SHARP protein are distinct (Extended Data Fig. 3b).

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

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

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

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

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

For live-cell imaging, fixed imaging and FRAP (Fig. 2 and Extended Data Fig. 2) ~10 μg of DNA was used for transfection when cells were grown on a 15-cm dish or ~1 ug of DNA when cells were grown on 3-cm glass-bottom dishes (Matek), and DNA concentrations were adjusted to match mole numbers across constructs. Twenty-four hours after transfection, cells were treated with dox (2 μg ml^{-1} (Sigma)) to induce expression of the proteins of interest and further experiments were performed 48 h postdox treatment.

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

Protein and RNA visualization. *Single molecule RNA fluorescence in situ.* RNA FISH experiments were performed using the ViewRNA ISH Cell Assay (ThermoFisher, catalog no. QVC0001) protocol with minor modifications. Specifically, cells were fixed on coverslips with 4% formaldehyde in PBS for 15 min at room temperature and then permeabilized with 4% formaldehyde and 0.5% Triton X-100 in PBS for 10 min at room temperature. Cells were then washed twice with PBS, dehydrated with 70% ethanol and incubated at −20 °C for at least 20 min or stored for up to 1 week. Coverslips were washed twice with PBS and then

incubated with the desired combination of RNA FISH probes (Extended Data Fig. 4a and Supplementary Table 3; Affymetrix) in Probe Set Diluent at 40 °C for at least 3 h. Coverslips were then washed once with wash buffer, twice with PBS and once more with wash buffer before incubating in preamplifier mix solution at 40 °C for 45 min. This step was repeated for the amplifier mix solution and label probe solution. Coverslips were incubated with 1× DAPI in PBS at room temperature for 15 min and subsequently mounted onto glass slides using ProLong Gold with DAPI (Invitrogen, P36935).

Immunofluorescence. To focus our analysis specifically on nuclear SHARP, pre-extraction was performed on cells before immunostaining as previously described[61]. In brief, cells on coverslips were washed once with PBS and then incubated with cold 0.1% Triton X-100 in PBS for 1–3 min on ice. Next, cells were fixed on coverslips with 4% formaldehyde in PBS for 15 min at room temperature and permeabilized with 0.5% Triton X-100 in PBS for 10 min at room temperature. After washing twice with PBS containing 0.05% Tween (PBSt) and blocking with 2% BSA in PBSt for 30 min, cells were incubated with primary antibodies overnight at 4 °C in 1% BSA in PBSt. After overnight incubation at 4 °C, cells were washed three times in 1× PBSt and incubated for 1 h at room temperature with secondary antibodies labeled with Alexa fluorophores (Invitrogen) diluted in 1× PBSt (1:500). Next, coverslips were washed three times in PBSt, rinsed in PBS, rinsed in double-distilled H_2O, mounted with ProLong Gold with DAPI (Invitrogen, P36935) and stored at 4 °C until acquisition.

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

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

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

Image acquisition and analysis. *Microscopy.* Fixed samples were imaged using: Zeiss LSM 800 with the ×63 oil objective (RNA FISH, IF) and collected every 0.3 µm for 16 Z-stacks, Zeiss LSM 880 with Airyscan with the ×63 oil objective (IF) and collected every 0.25 µm for 20 Z-stacks, or Zeiss LSM 980 with Airyscan2 with the ×63 oil objective (IF, RNA FISH-IF) where zoom, scan format and number of Z-stacks were optimized on the basis of the software recommendations for the highest resolution (super-resolution module). For all images, laser power and gain were set at the beginning of acquisition and remained constant throughout the duration of acquisition to enable comparisons of fluorescent intensities. Live samples were imaged using the Leica Stellaris microscope with ×63 water objective (~80 nm *xy*, ~300 nm *z*), and 16 Z-stacks were collected every 60 s for 5 min. The microscope was equipped with a stage incubator to keep cells at 37 °C and 5% CO_2.

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

Enrichment over inactive X territory. Xist and SHARP enrichments over the Xist territory (Fig. 1) were quantified using Icy (illustration Extended Data Fig. 1c). First, a region-of-interest (ROI) was defined that corresponded to the Xist signal across all Z-stacks by applying an intensity threshold (signal above background) and a binary mask was created by demarcating the Xist-coated territory (ROI). Next, several features of these ROIs were quantified, namely, the areas in µm^2 (Area), total fluorescent intensities of Xist or SHARP over the entire ROI (Total Intensity), and average fluorescent intensity (FI) of Xist or SHARP per area unit of ROI (pixel/interior) (Average Intensity).

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

Pattern of SHARP localization. To determine the pattern of SHARP localization after transfecting HEK293T cells with eGFP-SHARP constructs (Fig. 2e,g and Extended Data Fig. 2a), images were first processed into a maximum intensity projection using Icy software. A binary mask was then created to demarcate each nucleus of a transfected cell by setting a threshold of eGFP intensity above background levels; all masks were visually verified and, if needed, manually adjusted to fit the nuclear region of cells. On the basis of these masks, an ROI was defined that corresponded to the entire nucleus. Values for each pixel with the ROI (nucleus) were then extracted and this extracted information was used to quantify total intensity of protein per nucleus (sum of all pixels in an ROI), which corresponds to protein expression levels, and to calculate a SHARP dispersion score describing the differences in the distribution of pixel intensities across the nucleus. Specifically, for each cell, the intensity value at the 99th percentile of the distribution was computed and divided by the mode of the intensity distribution. This score was used because diffused localization shows distributed intensity across the nucleus and non-diffused localization shows accumulation of signal in defined locations, such that the tails of the intensity distributions were much longer. These quantitative assignments were visually confirmed to ensure that these scores capture our definition of diffused and non-diffused organization across cells.

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

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

Fluorescence recovery after photobleaching. FRAP experiments were performed in HEK293T cells overexpressing eGFP-tagged FL-SHARP, PTBP1 or embryonic ectoderm development. Forty-eight hours post-transfection, cells were subjected to FRAP as previously described[68] using the Zeiss LSM 710 with the ×40 water objective and equipped with a stage incubator to keep cells at 37 °C and 5% CO_2. Briefly, in each nucleus an area of ~1 µm^2 was bleached with the argon laser to quench eGFP and fluorescence recovery was followed while imaging in the GFP channel for 235 s. FRAP experiments were analyzed first by measuring the mean fluorescence intensity in the bleached area over time using Icy software and then normalized and averaged over n number of cells ($n > 5$) using EasyFRAP software[69]. Error bars represent standard deviation of at least five replicates.

Covalent linkage affinity purification followed by RNA sequencing. *Purification of Halo-tagged SHARP.* CLAP was performed on mESCs expressing Halo-tagged SHARP constructs (Supplementary Table 1) as previously described[70] (Fig. 3c,d). Briefly, post-transfection, media was removed from cells and then crosslinked on ice using 0.25 J cm^{-2} (UV2.5k) of UV at 254 nm in a Spectrolinker UV Crosslinker.

Cells were collected by scraping in 1× PBS and pelleted by centrifugation. Cell pellets were resuspended in 1 ml of ice cold lysis buffer (50 mM HEPES, pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) supplemented with 1× protease inhibitor cocktail (Promega), 200 U of Ribolock (ThermoFisher), 20 U TURBO DNase (Ambion) and 1× manganese/calcium mix (0.5 mM $CaCl_2$, 2.5 mM $MnCl_2$). The samples were incubated on ice for 10 min and then at 37 °C for 10 min at 700 r.p.m. shaking on a ThermoMixer (Eppendorf). Lysates were cleared by centrifugation at 15,000*g* for 2 min, and the supernatant was used for capture. For Halo-protein capture 50 µl of HaloLink Resin was preblocked using 1× blocking buffer (50 mM HEPES, pH 7.4, 100 µg ml^{-1} BSA) for 20 min at room temperature with continuous rotation. After incubation, the resin was washed three times with 1× PBSt. The cleared lysate was mixed with 50 µl of preblocked HaloLink Resin and incubated at 4 °C for 3–16 h with continuous rotation. The captured protein bound to resin was washed three times with lysis buffer at room temperature and then washed three times at 90 °C for 3 min while shaking on a ThermoMixer at 1,200 r.p.m. with each of the following buffers: 1× N-lauroylsarcosine (NLS) buffer (1× PBS, 2% NLS, 10 mM EDTA), high salt buffer (50 mM HEPES, pH 7.4, 0.1% NP-40, 1 M NaCl), 8 M urea buffer (50 mM HEPES, pH 7.5, 0.1% NP-40, 8 M urea), Tween buffer (50 mM HEPES, pH 7.4, 0.1% Tween 20) and TEV buffer (50 mM HEPES, pH 7.4, 1 mM EDTA, 0.1% NP-40). Between each wash, samples were centrifuged at 1,000 g for 30 s and the supernatant was removed. After the last wash, samples were centrifuged at 7,500*g* for 30 s and the supernatant was discarded. For elution, the resin was resuspended in 100 µl of NLS buffer and 10 µl of Proteinase K (NEB) and the sample was incubated at 50 °C for 30 min while shaking at 1,200 r.p.m. Capture reactions were transferred to micropsin cups (Pierce, ThermoFisher), centrifuged at 2,000*g* for 30 s, and the elutions were used for RNA purification by RNA Clean and Concentrate-5 kits (Zymo, >17 nucleotides protocol).

RNA library preparation and sequencing. RNA-seq library preparation was carried out as previously described[71]. Briefly, purified RNA was dephosphorylated (Fast AP) and cyclic phosphates were removed (T4 PNK). The RNA was then cleaned using Silane beads. An RNA adapter containing a reverse transcription (RT) primer binding site was ligated to the 3′ end of the RNA and the ligated RNA was reverse transcribed into cDNA. The RNA was then degraded using NaOH and a second adapter was ligated to the single-stranded complementary DNA. The DNA was amplified, and Illumina sequencing adapters were added by performing PCR with primers that are complementary to the 3′ and 5′ adapters that were previously added. The molarity of each PCR amplified library was measured using an Agilent Tapestation High Sensitivity DNA screentape and the samples were then pooled at equal molarity. This library pool was then size selected on a 2% agarose gel by cutting between 150 and 800 nucleotides and performing gel purification (Zymo). To determine the loading density of the final pooled library, the sample was measured using an Agilent Bioanalyzer and Qubit dsDNA High Sensitivity assay (ThermoFisher). The final library was paired-end sequenced on an Illumina HiSeq 2500 with read length 35×35 nucleotides.

Covalent linkage affinity purification analysis and visualization. For HALO purifications and RNA-binding mapping, sequencing reads were aligned to the mouse genome (RefSeq mm10) using STAR aligner. All low-quality alignments (MAPQ < 255) and PCR duplicates were excluded from the analysis using the Picard MarkDuplicates function (https://broadinstitute.github.io/picard/). The enrichment relative to input coverage across the Xist RNA was quantified by computing the number of reads overlapping the window in the SHARP-elution sample divided by the total number of reads within the SHARP-elution sample. This ratio was normalized by dividing the number of reads in the same window contained in the input sample by the total number of reads in the input sample. Because all windows overlapping a gene should have the same expression level in the input sample (which represents RNA expression), the number of reads in the input was estimated as the maximum of either (1) the number of reads over the window or (2) the median read count over all windows within the gene. This approach provides a conservative estimate of enrichment because it prevents windows from being scored as enriched if the input values over a given window are artificially low, while at the same time accounting for any non-random issues that lead to increases in read counts over a given window (for example fragmentation biases or alignment artifacts leading to non-random assignment or pileups). These enrichment values were visualized in IGV[72].

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

RNA affinity purification followed by DNA sequencing. *Cell treatment and preparation.* For RAP-DNA sequencing, TX1072 cells were treated with increasing dox concentrations (0.25×, 0.5×, 1×, 2×, 3× where 0× = no dox and 1× = 2 µg ml^{-1}) for 72 h, changing dox-containing medium daily. Cells were harvested and crosslinked as previously described[56]. Briefly, cells were pelleted, crosslinked with 2 mM disuccinimidyl glutarate for 45 min and 3% formaldehyde for 10 min, and lysed. Chromatin was then digested to 100–500-bp fragments through a combination of sonication and treatment with TURBO DNase and cell lysates were stored at −80 °C until the next step of the procedure.

Purification of DNA sites bound by Xist RNA. DNA fragments occupied by Xist RNA were purified for RAP-DNA as previously described[16] with minor modifications. Briefly, the lysate was diluted to hybridization conditions containing 3 M guanidine thiocyanate, precleared by adding streptavidin-coated magnetic beads and incubating for 30 min at 37 °C, mixed with biotin-labeled single-stranded DNA capture probes, and incubated at 37 °C for 2 h. 90-Mer single-stranded DNA oligonucleotide probes spanning the entire length of the target Xist RNA were purchased containing a 5′ biotin (Eurofins Operon)[22]. Next, captured chromatin complexes were eluted with RNaseH and crosslinks were reversed by adding Proteinase K to the probe–bead complexes and incubating overnight at 65 °C. Standard Illumina sequencing libraries were generated from eluted DNA fragments and sequenced at a depth of 5–20 million reads per sample of 75–75- or 75–140-long paired-end reads per sample.

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

Computing 3D contact frequencies with the Xist locus. 3D contact frequency between individual genomic regions and the Xist transcription locus was calculated as previously described[56]. Specifically, all SPRITE clusters containing a DNA read overlapping the Xist locus (chrX:103460373-103483233, mm10) were extracted and a genome-weight contact frequency was computed by counting the total number of SPRITE clusters for each genomic region within this set. The analysis exclusively focused on clusters containing 2–100 reads per cluster and weighted the contact frequency by the cluster size from which it was present (2/cluster size) as previously described.

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

Amplification reactions were run in a Roche LightCycler 480 instrument using LightCycler 480 SYBR Green I Master (Roche) with the primer pairs listed in Supplementary Table 2. Each sample had between one and six biological replicates and four technical replicates. Median Ct values were used to calculate fold change with the $2^{-\Delta\Delta Ct}$ method. For differentiation and dox induction conditions in the presence or absence of auxin (Fig. 5i and Extended Data Fig. 5g), each biological replicate was normalized to the median of the corresponding 'with SHARP (-Aux)' condition. For dox-induced samples across increasing concentrations (Extended Data Fig. 5d), each sample was normalized to the corresponding differentiation (RA) sample. For differentiated wild-type (F1) and TX1072 mESCs (Extended Data Fig. 5h), each sample was normalized to the corresponding wild-type sample.

Other data used in this study. RAP-DNA (F1 2-1 + 48 h of RA):. Xist localization across the X relative to gene density was measured using our previously published RAP-DNA dataset generated from Xist purification in F1 female mESCs differentiated with retinoic acid for 48 h[16]. All normalizations and analyses were performed as previously described and plotted using the normalized bedgraphs available at GEO accession GSE46918. **SPRITE (F1 2-1 mESCs):** 3D contacts were measured using our previously published RNA-DNA SPRITE dataset[61] that was generated in F1 female mESCs available at GEO accession GSE151515.

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

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

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

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

References

64. Ran, F. A. et al. Genome engineering using the CRISPR−Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
65. Klock, H. E., Koesema, E. J., Knuth, M. W. & Lesley, S. A. Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins Struct. Funct. Genet.* **71**, 982–994 (2008).
66. Wu, S. C. Y. et al. piggyBac is a flexible and highly active transposon as compared to Sleeping Beauty, Tol2, and Mos1 in mammalian cells. *Proc. Natl Acad. Sci. USA* **103**, 15008–15013 (2006).
67. Alberti, S. et al. A user's guide for phase separation assays with purified proteins. *J. Mol. Biol.* **430**, 4806–4820 (2018).
68. Wachsmuth, M. Molecular diffusion and binding analyzed with FRAP. *Protoplasma* **251**, 373–382 (2014).
69. Koulouras, G. et al. EasyFRAP-web: a web-based tool for the analysis of fluorescence recovery after photobleaching data. *Nucleic Acids Res.* **46**, W467–W472 (2018).
70. Banerjee, A. K. et al. SARS-CoV-2 disrupts splicing, translation, and protein trafficking to suppress host defenses. *Cell* **183**, 1325–1339 (2020).
71. Van Nostrand, E. L. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016).
72. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).

Acknowledgements

We thank M. Elowitz, S. Chong, I. Goronzy and D. Honson (Caltech) for their critical comments on the manuscript, A. Pandya-Jones, Y. Markaki and K. Plath (UCLA) for initial discussions and guidance on SHARP visualization, and F. Dossin, A. Loda and E. Heard (EMBL) for sharing their cell lines and cell culture protocols. We would like to thank A. Chow (Caltech) for helpful comments and support with the cell culture work done at the Guttman laboratory, the Biological Imaging Facility at Caltech for their help with microscopy, the Flow Cytometry Facility at Caltech for their help with cell sorting, and the Millard and Muriel Jacobs Genetics and Genomics Laboratory at Caltech for their help with sequencing. We also thank S. Hiley for contributions to writing and editing this manuscript, G. Riddihough and M. Bao (Life Science Editors) for editorial assistance, and I.M. Strazhnik (Caltech) for helping with illustrations. This work was supported by the National Institutes of Health (NIH) 4DN program (U01 DA040612 and U01 HL130007), NIH Directors' Transformative Research Award (R01 DA053178), the New York Stem Cell Foundation and funds from the California Institute of Technology. M.G. is a NYSCF-Robertson Investigator. J.W.J. was supported by a Biology and Biological Engineering postdoctoral fellowship from Caltech. A.K.B. was funded by National Heart, Lung, and Blood Institute F30-HL136080 and the University of Southern California MD/PhD Program.

Author contributions

J.W.J. conceived of this project with M.G. J.W.J. and M.S. performed experiments, analyzed and interpreted data, generated figures and wrote the paper. A.K.B. performed all CLAP sequencing experiments and provided comments and edits for the manuscript. J.T. created the SHARP rescue constructs with A.K.B. and assisted with cell culture. M.R.B. worked with A.K.B. on CLAP sequencing experiments, worked with J.W.J. on RAP sequencing experiments, analyzed sequencing data and provided comments and edits for the manuscript. M.G. oversaw all experiments and analysis, performed analyses and generated figures, and wrote the paper with J.W.J. and M.S.

Competing interests

The authors declare no competing interests.

Additional information

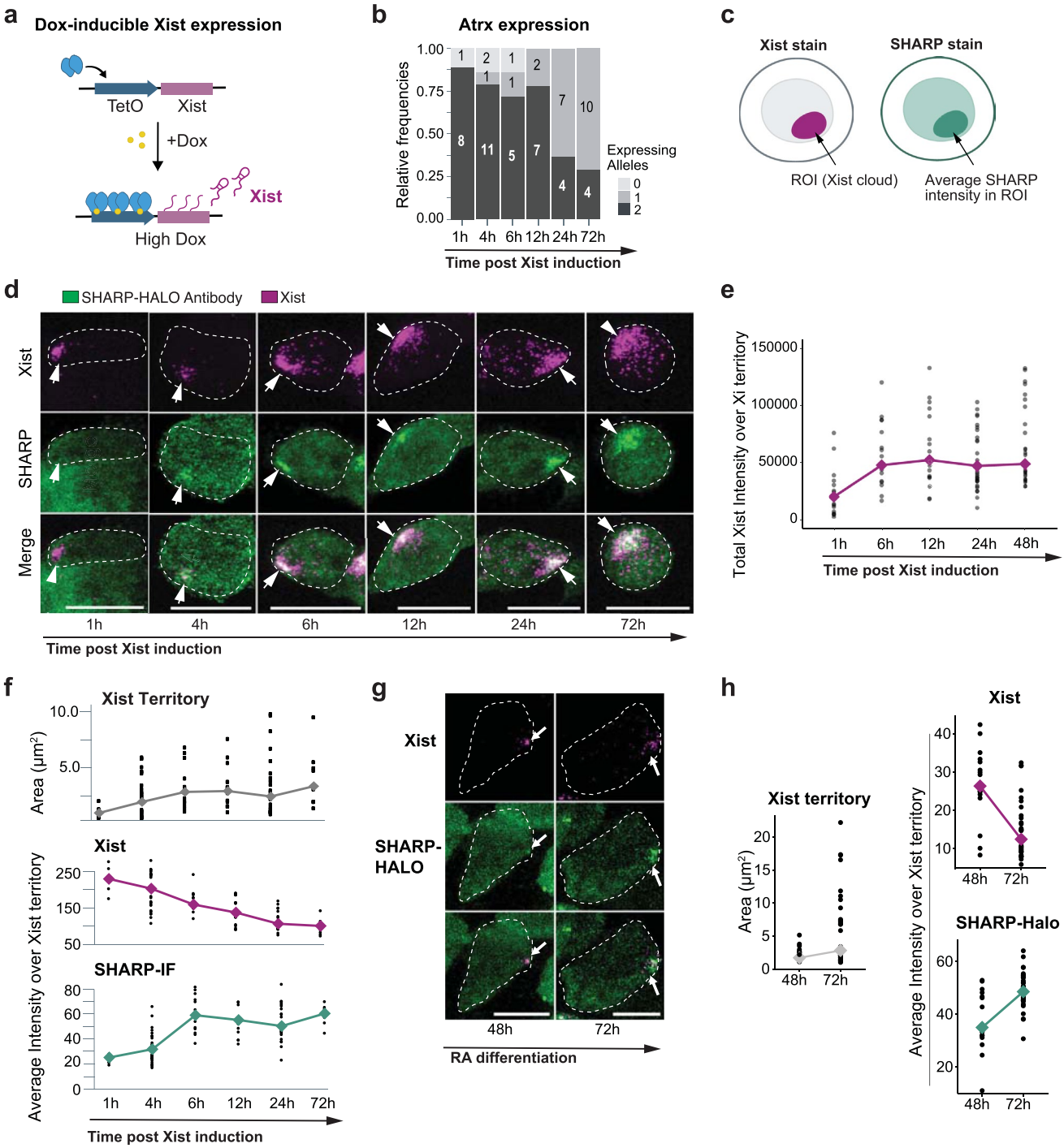
Extended data is available for this paper at https://doi.org/10.1038/s41594-022-00739-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-022-00739-1.

Correspondence and requests for materials should be addressed to Mitchell Guttman.

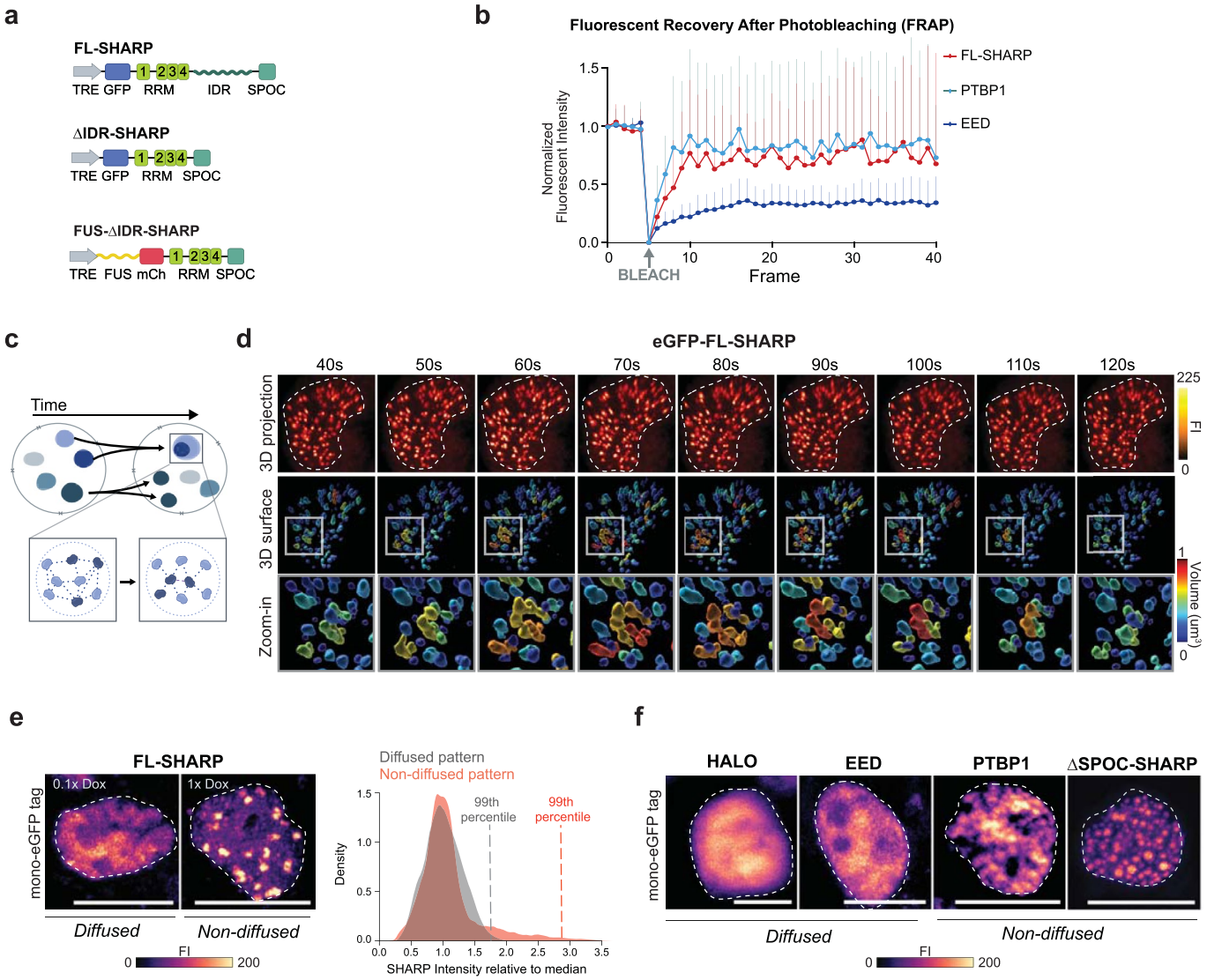
Peer review information *Nature Structural & Molecular Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Anke Sparmann, Beth Moorefield and Carolina Perdigoto were the primary editors on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

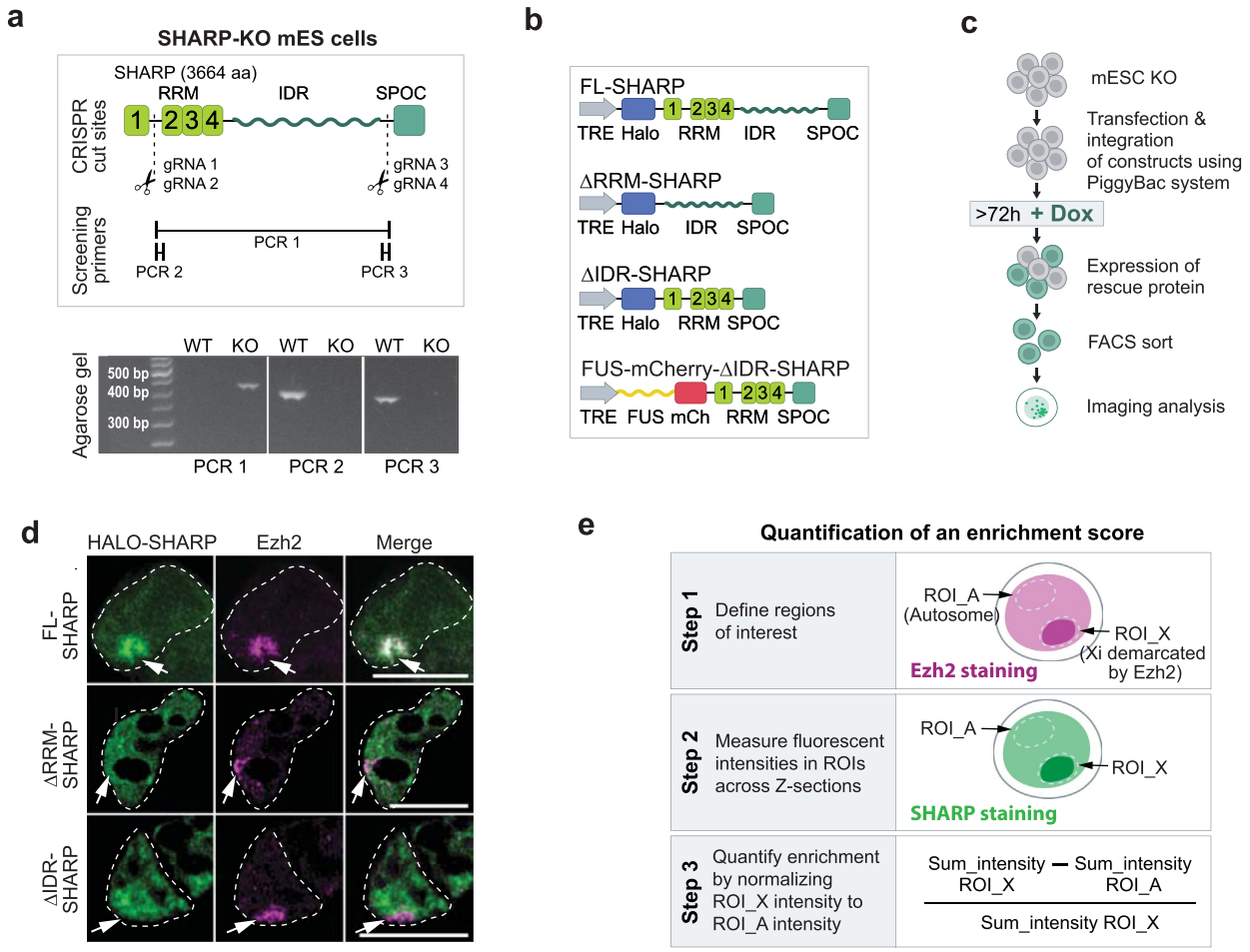


Extended Data Fig. 1 | See next page for caption.

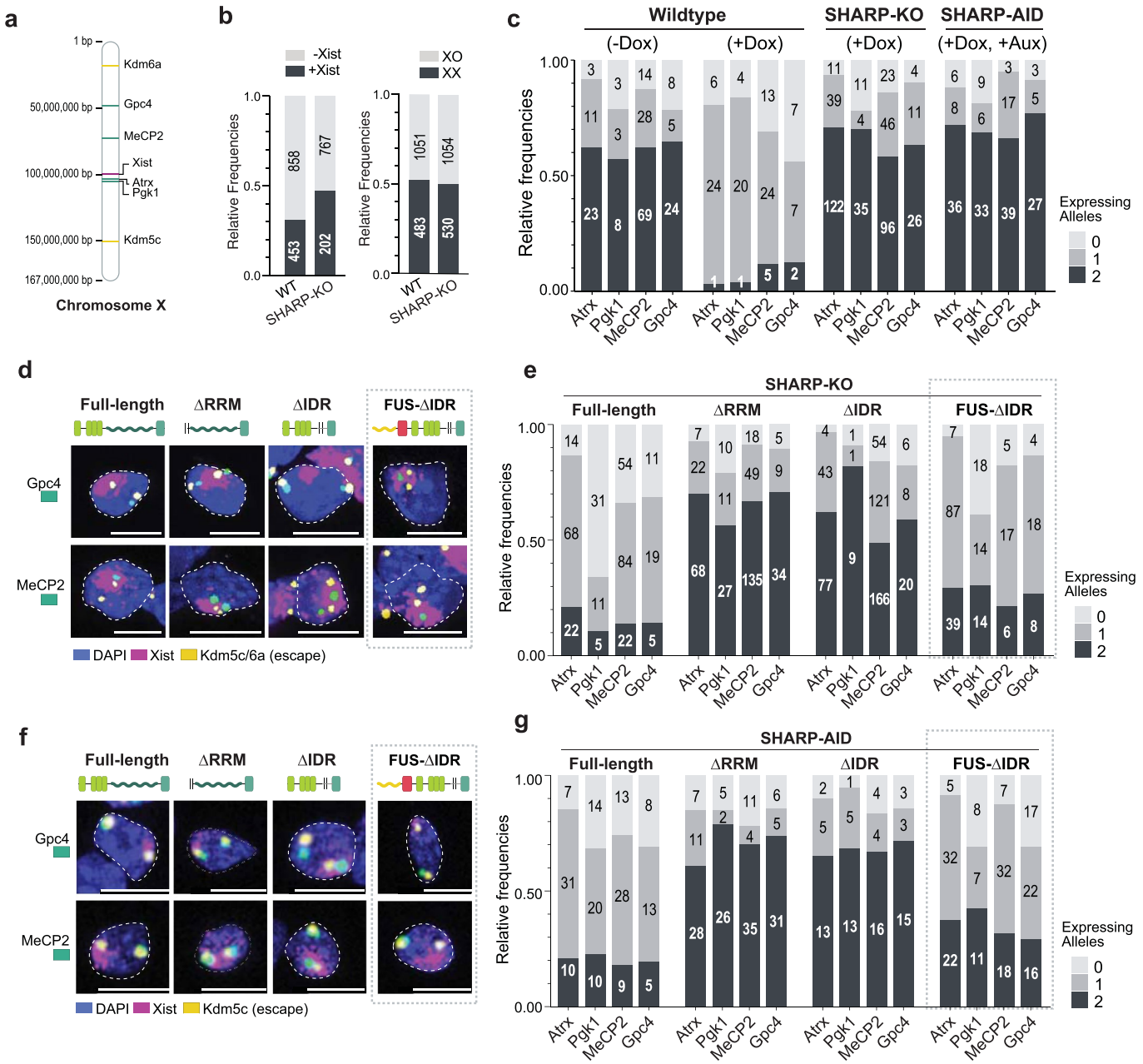
Extended Data Fig. 1 | SHARP enrichment over the Xi increases in a non-stoichiometric manner relative to Xist. a) Schematic of dox-inducible Xist expression system. The endogenous Xist promoter is replaced with a TetO element that can be activated upon the addition of doxycycline. b) Percent of cells expressing zero, one, or two alleles of the silenced Atrx gene as measured by RNA-FISH at various timepoints after Xist induction. c) Illustration of SHARP enrichment analysis over the Xi in TX-SHARP-HALO female mESCs. The Xi region was demarcated based on Xist RNA-FISH; SHARP was demarcated by either direct HALO labelling or immunofluorescence (anti-HALO). Fluorescent intensities of RNA-FISH probes, HALO tag, or anti-HALO immunofluorescence were then quantified within the defined Xi region and plotted. d) Representative images of Xist and SHARP localization in TX-SHARP-HALO female mESCs across 72 hours of Xist expression ($n > 15$ cells per timepoint from two independent experiments); Xist visualized by RNA-FISH (magenta); SHARP visualized by immunofluorescence labelling with anti-HALO antibody (green). Images shown as max. projections; scale bars 10 μm. e) Quantification of total fluorescence intensity of Xist (RNA-FISH) in multiple individual cells over 48 hours of Xist expression (Fig. 1b). f) Quantification of Xist and SHARP intensities in multiple individual cells over 72 hours of Xist expression (Extended Data Fig. 1d). Top panel: area of the territory coated by Xist RNA (μm2); middle panel: average fluorescent intensity of Xist (RNA-FISH) per unit within the Xist territory; bottom panel: average fluorescent intensity of SHARP (anti-Halo antibody) per unit within the Xist territory. g) Representative images of Xist and SHARP localization in TX-SHARP-HALO female mESCs after 48 and 72 hours of RA-induced differentiation ($n > 10$ cells for each timepoint); Xist visualized by RNA-FISH (magenta), SHARP visualized by direct Halo labelling (green). Images shown as max. projections; scale bars 10 μm. h) Quantification of Xist and SHARP in individual differentiated cells (Extended Data Fig. 1g). Left panel: area of the territory coated by Xist RNA (μm2); top right panel: average fluorescent intensity of Xist (RNA-FISH) per unit within the Xist territory; bottom right panel: average fluorescent intensity of SHARP (direct Halo labelling) per unit within the Xist territory.



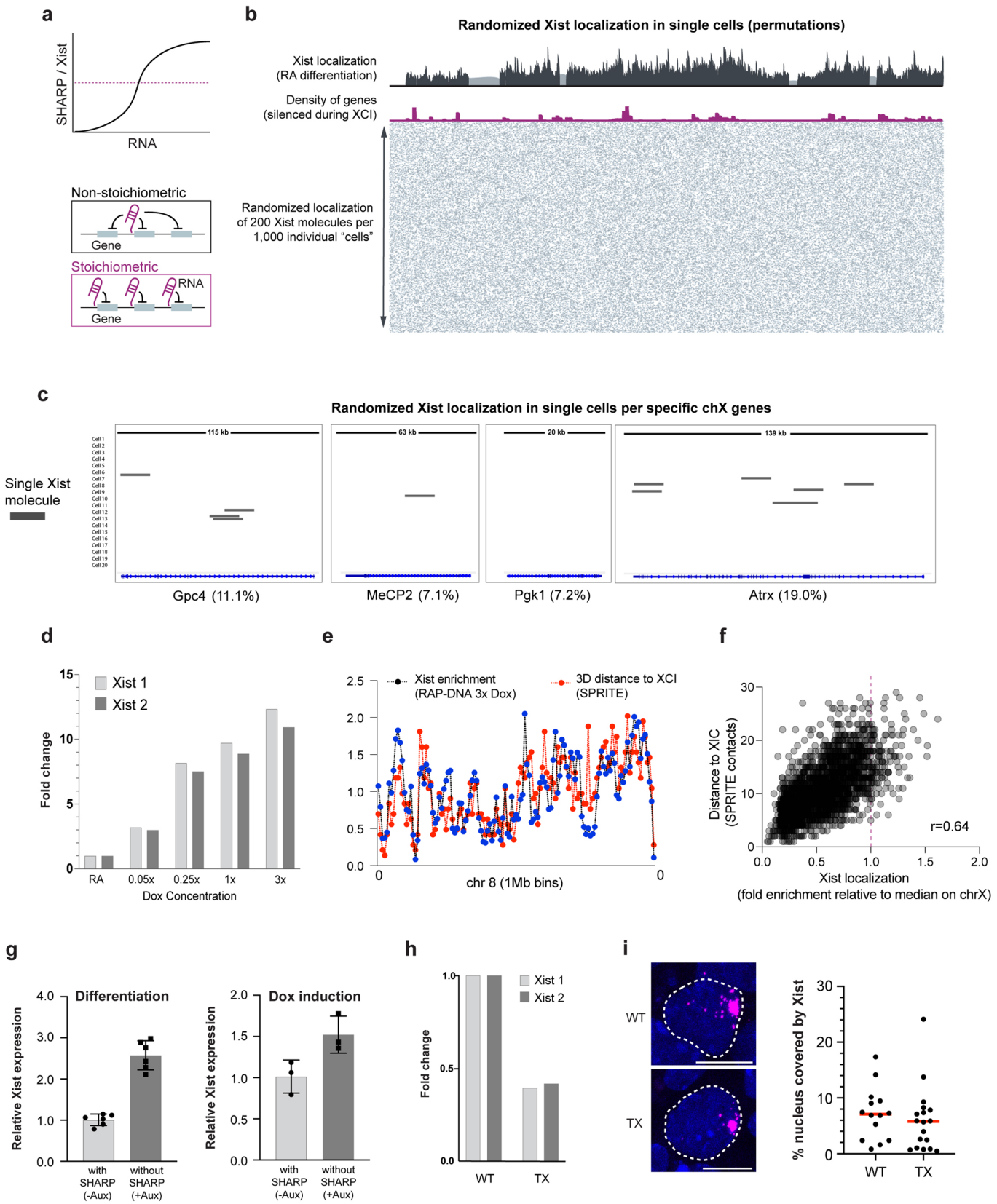
Extended Data Fig. 2 | SHARP forms multivalent, concentration-dependent assemblies in the nucleus. a) Schematic of the domains included in the eGFP-tagged FL-SHARP and ΔIDR-SHARP, and the mCherry-tagged FUS-ΔIDR-SHARP rescue constructs used in Fig. 2 and Extended Data Fig. 2. b) FRAP recovery curve of eGFP-tagged FL-SHARP (red), positive control PTBP1 (forms assemblies; light blue), and negative control EED (does not form assemblies; dark blue). Points represent mean value; error bars represent standard deviation of at least five replicates. c) Schematic depicting physical characteristics of concentration-dependent assemblies, including foci formation, fission and fusion, and rapid diffusion of proteins within an assembly (inset). d) Images across nine time-points from a live-cell movie of eGFP-tagged FL-SHARP in transiently transfected HEK293T cells (Supplementary Video 1,2) showing non-diffused, focal organization of SHARP molecules. Top panel: 3D reconstructions of the fluorescent intensity signal; middle panel: 3D volume reconstructions color-coded based on the volume of the focus; bottom panel: zoom-in representing one region of the nucleus that changes volume; Fluorescent Intensity (FI) e) Comparison of diffused or non-diffused localization patterns of FL-SHARP at different dox concentrations. Left: images representing FL-SHARP expressed with either 0.1x dox (diffused) or 1x dox (non-diffused) in transiently transfected HEK293T cells; images shown as max projections; scale bars 10 μm. Right: histograms representing fluorescent intensities for two cells showing diffused and non-diffused localization patterns. The intensity at the 99th percentile of each distribution is shown with the dashed lines. f) Images representing nuclear localization pattern of eGFP-tagged proteins in transiently transfected HEK293T cells. Left: proteins that have not been reported to form assemblies (HALO and EED); Right: an eGFP tagged protein that has been reported to form assemblies (Ptbp1) and ΔSPOC-SHARP that also forms assemblies. Images shown as max projections; scale bars 10 μm.



Extended Data Fig. 3 | Formation of SHARP assemblies is required for SHARP enrichment on the Xi and is dispensable for Xist binding. a) Generation of SHARP-KO cell line in TX mESCs. Top: schematic of CRISPR cut sites used to generate SHARP-KO mESCs and PCR primers used to screen for KO clones; bottom: agarose gel confirming homozygous deletion of SHARP in SHARP-KO clone H8 mESCs. b) Schematics of constructs used to generate rescue cell lines in TX SHARP-KO or TX SHARP-HALO-AID backgrounds. Grey arrow represents dox-inducible promoter; blue box represents HALO (or eGFP) tags used; light green boxes represent RNA Recognition Motifs (RRM); wavy green line represents the Intrinsically Disordered Regions (IDRs); dark green box represents the Spen Paralog and Ortholog C-terminal (SPOC) domain. Full-length SHARP (FL-SHARP), deletion of RRM domain (ΔRRM-SHARP), deletion of IDR domain (ΔIDR-SHARP), deletion of IDR domain and insertion of alternative IDR domain from FUS protein (FUS-ΔIDR-SHARP). c) Schematic showing experimental workflow for generating and enriching stable SHARP rescue mESCs (FL-SHARP, ΔRRM-SHARP, ΔIDR-SHARP, FUS-ΔIDR-SHARP) using constructs from Extended Data Fig. 3b. d) Representative images of SHARP enrichment (HALO, green) over the Xi (anti-Ezh2 immunofluorescence, magenta) in female mESCs containing dox-inducible Xist, genetic deletion of SHARP, and stable integrations of HALO-tagged FL-SHARP, ΔRRM-SHARP, or ΔIDR-SHARP ($n > 10$ cells per condition). Xist and SHARP rescue constructs induced with doxycycline for 72 hours; images shown as Z-sections; scale bars 10 μm e) Diagram of image analysis workflow for quantifying SHARP enrichment over the Xi (Fig. 3b).



Extended Data Fig. 4 | SHARP binding to RNA (via RRMs) and formation of assemblies (via IDRs) are both required for chromosome-wide gene silencing. a) Schematic of mouse X chromosome showing the locations of the various genes probed in RNA-FISH experiments. b) Frequency of Xist induction (left) and X chromosome ploidy (right) in wildtype and SHARP-KO mESCs based on quantification of RNA-FISH images. c) Quantification of RNA-FISH images (Fig. 4b) representing the frequency of cells containing two, one, or zero actively transcribed alleles. Left to right: wildtype (-dox); wildtype (+dox); dox-induced SHARP-KO; dox-induced, auxin-treated SHARP-AID female mESCs. d) RNA-FISH images from SHARP-KO female mESCs containing stable integrations of (left to right): FL-SHARP, ΔRRM-SHARP, ΔIDR-SHARP, or FUS-ΔIDR-SHARP after >72 hours of dox induction. Cells were stained for DAPI (blue) and probed for Xist (magenta), escape gene Kdm5c (yellow), and silenced genes Gpc4 or MeCP2 (green). Images shown as max projections; scale bars 10 μm; total cell numbers (Extended Data Fig. 4e) from three independent experiments. e) Quantification of RNA-FISH images (Extended Data Fig. 4d) representing the frequency of cells containing two, one, or zero actively transcribed alleles for the various SHARP rescue constructs in SHARP-KO female mESCs. f) RNA-FISH images from SHARP-AID female mESCs containing stable integrations of (left to right): FL-SHARP, ΔRRM-SHARP, ΔIDR-SHARP, or FUS-ΔIDR-SHARP after >72 hours of dox induction. Cells were stained for DAPI (blue) and probed for Xist (magenta), escape gene Kdm5c (yellow), and silenced genes Gpc4 or MeCP2 (green). Images shown as max projections; scale bars 10 μm; total cell numbers (Extended Data Fig. 4g) from three independent experiments. g) Quantification of RNA-FISH images (Extended Data Fig. 4f) representing the frequency of cells containing two, one, or zero actively transcribed alleles for the various SHARP rescue constructs in SHARP-KO female mESCs.



Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Low Xist expression levels limit its ability to spread to autosomes. a) Top: schematic depicting expected ratios of SHARP to Xist based on increasing concentration of Xist RNA; bottom: diagrams illustrating non-stoichiometric and stoichiometric gene silencing by Xist. b) Randomized Xist localization in simulated single cells (permutations) compared to experimental data. Top: Xist localization after 48 hours of RA-differentiation in female mESCs from bulk RAP-DNA experiments16; middle: gene density across X, only genes that undergo XCI are plotted; bottom: randomized localization of 200 Xist molecules in 1000 random "cells"; Xist represented by grey squares. c) Simulation visualizing localization of Xist molecules over genes that undergo transcriptional silencing during XCI across 20 cells (zoom-in from Extended Data Fig. 5b). Percent reflects proportion of Xist molecules overlapping the gene in all 1000 permutations. d) Expression levels of Xist in female mESCs treated with increasing dox concentrations as measured by RT-qPCR. e) Comparison of Xist occupancy (black lines; 3x dox RAP data) and DNA contact frequency with Xist locus (red lines; SPRITE data61) across 1 Mb DNA regions of chr8. Xist enrichment and 3D distance are normalized to their median coverage across chromosome 8 to place them on the same relative scale. f) Scatterplot representing the frequency of 3D contacts between each 1 Mb autosomal bin with the Xist locus (SPRITE data61; y axis) and DNA sites enriched by Xist when expressed at high concentration (3x dox; RAP-DNA data; X-axis). g) Relative Xist expression upon RA-induced differentiation (left) and dox-induction (right) of SHARP-AID mESCs in the absence or presence of auxin as measured by RT-qPCR (primer pair Xist 2). Dots represent individual replicates; n = 6 for differentiation conditions; n = 3 for dox induction conditions; data represented as mean ± s.d. h) Relative Xist expression in RA differentiated wildtype cells that do not contain dox-inducible Xist (left) and TX cells with dox-inducible Xist (right) as measured by RT-qPCR (primer pairs Xist 1, 2). i) Left: images representing Xist territories in wildtype cells without dox-inducible Xist and TX cells with dox-inducible Xist; total cell numbers (right) from two independent experiments. Right: quantification of percent of nucleus occupied by Xist in the same cell lines as measured by RNA-FISH (Xist) and DAPI staining (nucleus); scale bars 10 μm. Dots represent individual replicates; red bars represent median.

nature portfolio

Corresponding author(s): Mitchell Guttman

Last updated by author(s): Dec 28, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Images were collected and analyzed using standard packages and plug-ins available in Icy (v2.3), ImageJ (v2.1.0/1.53c), Imaris (v8), and EasyFRAP software systems.
Data analysis	Images were quantified using standard packages and plug-ins available in Icy (v2.3), ImageJ (v2.1.0/1.53c), Imaris (v8), and EasyFRAP software systems.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data generated for this study is available at GEO accession GSE192574.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were not pre-determined. For imaging experiments, sample sizes were chosen to ensure that a representative population was captured. For sequencing experiments, number of cells included was based on technical requirements.
Data exclusions	For RNA-FISH spot quantification, we excluded cells that did not have signal from Xist and did not have two spots from either Kdm5c or Kdm6a (described in methods).
Replication	RNA-FISH experiments were performed in at least three replicates and RNA-FISH/IF in at least two replicates: biological replicates represented by cells from different passages, and technical replicates represented by multiple cells from the same passage or cover-slip. Additionally, alternative validation using different antibodies or labeling methods was performed. Sequencing data was performed in replicates as described in the figure legends and methods, where separate culture plates were considered biological replicates. All attempts at replication were successful.
Randomization	Cells used in experiments were randomized insomuch as all cells within a biological or technical replicate were included (provided they met exclusionary criteria as described above).
Blinding	Experiments were not blinded due to technical and analytical requirements. However, quantification for RNA-FISH spot counting was performed by the co-authors independently and the results were cross-validated.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Primary: anti-Halo (mouse, Promega G9211); anti-Ezh2 (mouse, Cell Signaling AC22 3147S); anti-SHARP (rabbit, Bethyl A301-119A); Secondary: Alexa Fluor antibodies from Thermo Fisher, anti-mouse: 488 (# A32723), 555 (# A32727), 647 (# A32728); anti-rabbit: 647 (# A32733), 555 (# A32732), 488 (# A32731)
Validation	All antibodies used in this study are commercially available and have been validated by the provider using WB or/and IF. Additionally, secondary antibodies were incubated without primary to test for their specificity.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	All mouse embryonic stem cells (TX lines) used in this study were gifts from the Edith Heard lab (EMBL, Heidelberg); for human cell line HEK293T refer to 293T/17 [HEK 293T/17] (ATCC CRL-11268).
Authentication	RNAseq analyses were performed in the lab (for other projects) to verify the identify of the cell lines (e.g. using corresponding gene expression patterns).

Mycoplasma contamination	Cells negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	There are no commonly misidentified cell lines used in this study.