Published in final edited form as:
Nature. 2020 February ; 578(7795): 455–460. doi:10.1038/s41586-020-1974-9.

SPEN integrates transcriptional and epigenetic control of X-inactivation

François Dossin[1], **Inês Pinheiro**[#2], **Jan J. Żylicz**[#2,3], **Julia Roensch**[2], **Samuel Collombet**[1], **Agnès Le Saux**[2], **Tomasz Chelmicki**[2], **Mikaël Attia**[2], **Varun Kapoor**[2], **Ye Zhan**[4], **Florent Dingli**[5], **Damarys Loew**[5], **Thomas Mercher**[6], **Job Dekker**[4,7], **Edith Heard**[1,2,*]

[1]European Molecular Biology Laboratory, Director's Unit, Heidelberg, Germany

[2]Institut Curie, PSL Research University, CNRS UMR3215, INSERM U934, UPMC Paris-Sorbonne, 75005 Paris, France

[3]University of Cambridge, Department of Physiology, Development and Neuroscience, Cambridge CB2 3EG, UK

[4]Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, 01605, USA

[5]Institut Curie, PSL Research University, Centre de Recherche, Laboratoire de Spectrométrie de Masse Protéomique, Paris 75248 Cedex 05, France

[6]INSERM U1170, Gustave Roussy Institute, Université Paris-Sud, Villejuif 94800, France

[7]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA

[#] These authors contributed equally to this work.

Summary

Xist represents a paradigm for long non-coding RNA function in epigenetic regulation, although how it mediates X-chromosome inactivation (XCI) remains largely unexplained. Multiple *Xist*-RNA binding proteins have recently been identified, including SPEN[1–3], the loss of which has been associated with deficient XCI at multiple loci[2–6]. Here we demonstrate that SPEN is a key orchestrator of XCI *in vivo* and unravel its mechanism of action. We show that SPEN is essential

[*]Corresponding author: Edith Heard, edith.heard@embl.org.

Author Contributions

F. Dossin and E.H. conceived the experiments and F. Dossin performed them unless stated otherwise. T.M. provided the *Spen* KO mice and J.J.Z. performed the embryo experiments. E.H. and I.P. supervised the work. J. R. established the *Xist*-stem loop cell-line and acquired live-cell images. F. Dossin, J.R. and V.K. analyzed live-cell imaging experiments. V.K. wrote the image analysis software. F. Dossin and S.C. processed and analyzed the data. A.L.S. cloned the SPEN cDNA truncation constructs. T.C. helped with some experiments, and M.A. derived NPC clones. Y.Z. prepared Hi-C libraries in the lab of J.D. F.D. carried out the MS experimental work and D.L. supervised MS and data analysis. F. Dossin and E.H. wrote the manuscript with input from all authors.

Competing Interests

The authors declare no competing interests.

Data Availability

RNAseq, HiC and CUT&RUN data used in this study have been deposited in the Gene Expression Omnibus under accession number GSE131784. Source Data for Figs. 1-3 and Extended Data Figs. 1,3,4 are provided with the paper, either in the format of supplementary tables or source data files.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

for initiating gene silencing on the X chromosome in preimplantation mouse embryos and embryonic stem cells. SPEN is dispensable for maintenance of XCI in neural progenitors, although it significantly dampens expression of genes that escape XCI. We show that SPEN is immediately recruited to the X-chromosome upon *Xist* up-regulation, and is targeted to enhancers and promoters of active genes. SPEN rapidly disengages from chromatin upon gene silencing, implying a need for active transcription to tether it to chromatin. We define SPEN's SPOC domain as a major effector of SPEN's gene silencing function, and show that tethering SPOC to *Xist* RNA is sufficient to mediate gene silencing. We identify SPOC's protein partners which include NCOR/SMRT, the m6A RNA methylation machinery, the NuRD complex, RNA polymerase II and factors involved in regulation of transcription initiation and elongation. We propose that SPEN acts as a molecular integrator for initiation of XCI, bridging *Xist* RNA with the transcription machinery as well as nucleosome remodelers and histone deacetylases, at active enhancers and promoters.

To address the importance of SPEN during initiation of XCI, we used an auxin-inducible degron (AID)[7], enabling controlled and acute depletion of the endogenous SPEN protein. We used our previously described female hybrid (*Mus musculus castaneus* x C57BL/6) TX1072[8] mouse embryonic stem cells (mESCs), in which a doxycycline-inducible promoter upstream of the endogenous *Xist* locus allows conditional *Xist* RNA expression and XCI (Fig. 1a). We generated a homozygous knock-in of the AID fused to a HaloTag at the C-terminus of endogenous *Spen*, in ESCs expressing the *OsTir1* E3 ligase to ensure auxin-dependent SPEN depletion (Extended Data Fig. 1a). Efficient SPEN degradation occurred within 1 hour of auxin treatment (Fig. 1b, Extended Data Fig. 1b and Supplementary Figure 1), while removal of auxin led to rapid SPEN recovery (Fig. 1b), demonstrating potent AID-dependent modulation of SPEN levels.

To evaluate the immediate consequences of SPEN loss on initiation of XCI, we acutely depleted SPEN for 4 hours prior to inducing *Xist* expression for 24 hours and performed RNA-seq. Loss of SPEN had no effect on the formation of *Xist* RNA clouds (Extended Data Fig. 1c, e), confirming that SPEN is dispensable for *Xist* localization[2–5]. However, gene silencing was almost completely abolished in the absence of SPEN, along the entire X chromosome (Fig. 1c, d and Supplementary Table 1), while auxin had no effect on XCI in wild-type cells (Extended Data Fig. 1d). Clustering analysis highlighted three groups of genes differing by their silencing defects upon SPEN loss (Fig. 1e). Most X-linked genes (80% of 382) were found to be entirely dependent on SPEN for silencing, while only a small subset (6%) showed unaltered silencing in the absence of SPEN. This striking defect in XCI was confirmed by pyrosequencing (Fig. 1f) and nascent RNA FISH (Extended Data Fig. 1e).

We next addressed the requirement for SPEN in XCI *in vivo* during mouse early embryogenesis, using allele-specific RNA-seq in E3.5 *Spen* KO female embryos[9] harboring hybrid X chromosomes (Fig. 1g and Extended Data Fig. 1f, g). At this stage in wild-type embryos, imprinted XCI has taken place[10] and only the paternal X is inactivated (Fig. 1h and Extended Data Fig. 1h). In maternal-zygotic *Spen* knockouts, imprinted XCI is severely hindered, although paternal *Xist* is expressed. Both maternal and paternal X chromosomes are expressed equally, phenocopying *Xist* knockout E3.5 embryos[10] (Fig. 1h, Extended Data Fig. 1g, h and Supplementary Table 2). A maternal-only *Spen* KO has no effect on imprinted

XCI (Fig. 1h), suggesting that the zygotic pool of SPEN is necessary and sufficient for this process. Thus, the early gene silencing mechanism(s) involved in imprinted and random XCI are dependent on SPEN.

We next assessed precisely when SPEN is recruited during XCI. HaloTag labeling[11] of SPEN combined with *Xist* RNA FISH revealed that SPEN associates with *Xist* RNA rapidly upon *Xist* coating and throughout XCI (Fig. 2a). To capture early *Xist*/SPEN dynamics during the short time window when *Xist* becomes upregulated, we followed both *Xist* and SPEN in living cells. We thus tagged endogenous SPEN with GFP in a background where *Xist* RNA is visualized via a BglG-mCherry fusion protein binding to BglSL stem loops inserted within *Xist*[12] (Extended Data Fig. 2a, b). Live-cell imaging revealed that SPEN colocalizes with *Xist* from the very onset of *Xist* up-regulation (Extended Data Fig. 2c, d, and Supplementary Video 1). Therefore, SPEN can initiate gene silencing immediately upon *Xist* coating.

We also found SPEN robustly accumulated on the Xi following differentiation into neural progenitor cells (NPCs, Fig. 2b), where XCI is epigenetically maintained. Yet, depletion of SPEN for up to 2 days, in independent NPC clones (Fig. 2c), did not lead to reactivation of fully silenced genes (Fig. 2d and Supplementary Table 3). However, we observed moderate but significant up-regulation of genes escaping XCI (Fig. 2e, f), suggesting that SPEN buffers over-expression of X-linked escapees in female cells.

Chromosome conformation capture has revealed that, in differentiated cells, the Xi is folded into megadomains[13–15] and globally depleted of topologically associating domains (TAD) except at regions containing clusters of escapee genes[13]. *Xist* RNA has been found to play a role in Xi conformation[1,14]. To assess whether SPEN is involved, we performed allele-specific HiC in NPCs after 48 hours of SPEN depletion. No striking conformational changes were observed on the Xi (Extended Data Fig. 2e-g). Thus the structural effects mediated by *Xist* RNA in differentiated cells, occur independently of SPEN.

In summary, our data suggests that SPEN exerts its role by actively promoting gene silencing during the earliest stages of XCI. However, it has no major role in stabilizing the transcriptionally inactive state of the Xi, nor in ensuring maintenance of its conformation.

Next, we sought to identify which parts of SPEN ensure its function during XCI. SPEN is a very large protein (~400kDa), containing four RRMs (RNA recognition motifs), RID (nuclear receptor interaction domain) and SPOC (SPEN paralogue/orthologue C-terminal) domains (Fig. 3a). We overexpressed a series of SPEN cDNA truncations, stably targeted into the *ROSA26* locus in the SPEN-degron ESC line (Extended Data Fig. 3a, b, and Fig. 3a). We then induced *Xist* expression for 24 hours and assessed which SPEN fragments could rescue XCI initiation function in the context of auxin-mediated depletion of endogenous SPEN. We found that RRM1 and RID domains are dispensable for SPEN accumulation on the Xi, as well as for X-linked gene silencing (Fig. 3b, c). Conversely, a SPEN truncation lacking RRM2-4 domains failed to accumulate on the Xi and failed to rescue XCI (Fig. 3b, c). Thus, SPEN recruitment to the Xi is mediated by RRM2-4 domains and necessary for gene silencing. This is consistent with studies showing that RRM2-4

directly bind the A-repeat of *Xist* RNA *in vitro*[4,16], a region of *Xist* necessary for gene silencing[17]. On the other hand, a truncation of the SPOC domain allowed efficient SPEN accumulation on the Xi, but failed to rescue XCI (Fig. 3b, c). To validate this observation, we performed homozygous deletion of the SPOC domain at the endogenous *Spen* locus in mESCs (Extended Data Fig. 3c). Deleting SPOC had no effect neither on SPEN recruitment to the Xi nor on *Xist* RNA clouds (Extended Data Fig. 3d, e, f), but resulted in strongly deficient XCI, albeit milder than in SPEN-depleted cells (Extended Data Fig. 3g, h, i, j). Collectively, these results demonstrate that the SPOC domain is essential for XCI. However, other uncharacterised regions of SPEN contribute, albeit to a lesser extent, to ensure SPEN's full silencing potential.

To test whether the SPOC domain alone could mediate X-linked gene silencing, we used SPEN-degron ESCs to introduce an array of BglSL at the *Xist* locus (identical to the live imaging strategy). In this background, we generated several independent ESC lines expressing a BglG-GFP-SPOC protein fusion (or BglG-GFP as control) targeted into *ROSA26*. These proteins would become tethered to *Xist*-BglSL RNA via BglG (Fig. 3d). Remarkably, upon induction of *Xist* RNA in the absence of endogenous SPEN, tethering of BglG-GFP-SPOC (but not BglG-GFP alone) resulted in substantial gene silencing across the X-chromosome, with over half the genes being silenced by more than 50% (Fig. 3e, f). SPOC-specific rescuing was confirmed using pyrosequencing (Fig. 3g). Consistent with previous studies[18–20], our results reveal SPOC as a key domain of SPEN that enables gene silencing once recruited to the X-chromosome by *Xist* RNA.

The SPOC domain of SPEN was originally identified as an interactor of NCOR/SMRT in human cells[18,21,22]. Given that NCOR/SMRT interact with and activate HDAC3[23], it was proposed that SPEN triggers XCI via HDAC3[2], the activity of which is important for *Xist*-mediated silencing[2,24]. However, XCI is more dramatically affected upon loss of SPEN and SPOC than upon loss of HDAC3 (Extended Data Fig. 3j, k). These observations suggest that a model involving HDAC3 only partially explains SPEN function, and that SPOC must play its key role in gene silencing also through other, HDAC3-independent pathways. To identify such pathways, we characterised the protein interactome of the SPOC domain by performing GFP-pulldowns from mESCs stably expressing BglG-GFP-SPOC (or BglG-GFP as control, Fig. 3h and Supplementary Table 4), followed by mass spectrometry.

As expected, we identified NCOR/SMRT, but also HDAC3 (Fig. 3h and Extended Data Fig. 3l), further supporting the proposed model for SPEN function in XCI[2]. Strikingly, we identified the m6A methyltransferase complex and the m6A reader YTHDC1 (Fig. 3h and Extended Data Fig. 3l) which have been proposed to play a role in XCI[5,20,25]. One of these factors, WTAP, co-purified with *Xist* RNA in an A-repeat dependent manner[3], although contrary to SPEN, a direct interaction between WTAP and *Xist* A-repeat has not been reported. Thus, our results suggest that SPOC may participate in m6A machinery recruitment to *Xist* RNA. We also identified the NuRD complex - a potent repressor which displaces RNA polymerase II (RNAPII) from TSS through chromatin remodeling[26] - ; and RNAPII, together with factors involved in regulation of transcription initiation and elongation (Fig. 3h, and Extended Data Fig. 3l). Together, this reveals that through its SPOC domain, SPEN bridges *Xist* with multiple factors involved in transcription and chromatin

regulation, that together mediate efficient gene silencing. Given that SPOC immunoprecipitation was performed in the absence of *Xist* induction, the identified interactions are not mediated by *Xist* RNA.

We also investigated where SPEN binds to the X chromosome during XCI and whether it has distinct binding sites or rather associates with chromatin diffusely across the entire chromosome as anticipated from our imaging results. To this end we performed allele-specific cross-linked CUT&RUN[27] for SPEN during a time course of *Xist* induction (0h, 4h, 8h, 24h dox, or 8h dox+auxin as a negative control).

We found that SPEN shows few binding sites across the genome of uninduced ES cells (Extended Data Fig. 4a). Conversely, hundreds of SPEN sites appeared specifically on the X-chromosome as early as 4 hours after *Xist* induction (Fig. 4a and Extended Data Fig. 4a). This is consistent with imaging data (Extended Data Fig. 2). Interestingly, SPEN accumulation is seen across the gene body of *Xist* (Fig. 4b), suggesting that SPEN binds *Xist* RNA while it is transcribed. In sharp contrast to the *Xist* locus, SPEN shows focal binding on the rest of the genome, with peaks falling almost exclusively on promoters and enhancers (Fig. 4c, d and Extended Data Figs. 4b, 5a-g).

Following *Xist* induction, SPEN recruitment to the Xi peaks at 4 hours (Fig. 4a and Extended Data Fig. 4c), showing highest enrichment within regions coated earliest by *Xist*[28] (entry sites, Fig. 4e). Thus, SPEN accumulation follows the spatial dynamics of *Xist* spreading. Among promoter targets on the X chromosome, SPEN preferentially binds those of actively expressed genes (Fig. 4f and Extended Data Fig. 4d), suggesting that the ability of SPEN to target chromatin depends on transcriptional. Consistently, genes classified as fully dependent on SPEN for silencing (Fig. 1e), which show more SPEN binding at their promoters within 4 hours of *Xist* coating than less dependent genes (Extended Data Fig. 4e), also show initially higher transcription levels (Fig. 4g).

Furthermore, within 4 hours of *Xist* induction, SPEN-binding is greater at promoters of efficiently silenced genes than it is at promoters of less efficiently silenced genes (Fig. 4h). Similarly, upon *Xist* coating, efficiently deacetylated enhancers[24] show higher SPEN enrichment than less efficiently deacetylated enhancers (Fig. 4i). Finally, genes subject to very little silencing, if not complete escape from XCI in our *Xist*-inducible system show significantly lower SPEN signal at their promoters (Extended Data Fig. 4f, g and Extended Data Fig. 5h-n). This discrete pattern of SPEN recruitment to the X-chromosome undergoing XCI indicates that transcriptional silencing is caused by SPEN binding to active promoters and enhancers.

To understand how SPEN might function at enhancers and promoters, we integrated CUT&RUN profiles with publicly available ChIP-seq data for transcription and chromatin-associated factors identified in our mass spectrometry analysis. We included HDAC3[24], RNAPII[26], and two members of the NuRD complex (MBD3 and CHD4)[26]. SPEN binding strongly overlaps with HDAC3 at enhancers but not at promoters (Extended Data Fig. 4h). Our recent findings revealed that HDAC3 is pre-bound predominantly at enhancers on the X-chromosome[24]. Therefore, *Xist*-mediated recruitment of SPEN to enhancers may activate

HDAC3. On the other hand, a strong overlap with SPEN binding is observed for the NuRD complex specifically at promoters but not at enhancers (Extended Data Fig. 4h). Furthermore, SPEN peaks extensively overlap with RNAPII phosphorylated on serine 5, which is associated with transcription initiation (Extended Data Fig. 4h). This analysis suggests that SPEN may operate at enhancers and promoters through distinct pathways to promote gene silencing.

Remarkably, binding of SPEN to chromatin decreases chromosome-wide after 24 hours of *Xist* induction (Fig. 4a and Extended Data Fig. 4c). Clustering of CUT&RUN profiles at SPEN-bound promoters (Extended Data Fig. 4i and Supplementary Table 5) revealed distinct groups of promoters on the basis of how efficiently SPEN was lost within 24 hours of XCI (Fig. 4j). In the “strong SPEN loss” group, binding was maximal by 4 hours but decreased after 8 hours, and even more drastically so after 24 hours (Fig. 4j). On the contrary, the “mild SPEN loss” group showed maximal and persistent SPEN binding at 4 and 8 hours of *Xist* induction respectively, with only mild reduction of SPEN binding by 24 hours (Fig. 4j). Finally, a third group, comprising fewer promoters, showed both mild SPEN enrichment at 4 hours and low loss at 24 hours (Fig. 4j). Importantly, the group which lost SPEN most efficiently, also showed most pronounced gene silencing by 24 hours compared to the groups which significantly retained SPEN (Fig. 4k and Extended Data Fig. 4j). Altogether, this analysis suggests that once recruited to the X-chromosome by *Xist* RNA, SPEN associates with enhancers and promoters in a transcription-dependent manner. This recruitment leads to gene silencing, after which the favorable transcriptional context for SPEN binding is lost, and SPEN binding to chromatin decreases. Despite loss of the chromatin-bound SPEN fraction, persistent *Xist* RNA expression and coating ensure SPEN remains strongly accumulated around the Xi (Fig. 2a, b).

Discussion

Our study demonstrates that SPEN is a crucial factor that collaborates with *Xist* RNA to initiate gene silencing chromosome-wide both during XCI *in vitro* and imprinted XCI *in vivo*. SPEN becomes dispensable for maintaining gene silencing after XCI has been established, but partially represses escapees, suggesting *Xist* may have a silencing role even in somatic cells. Remarkably, although SPEN coats the X chromosome immediately upon *Xist* induction, it only contacts chromatin at active promoters and enhancers, which serve as substrates for SPEN-mediated gene silencing. SPEN association with chromatin is favored by active transcription, as SPEN disengages from chromatin when X-linked genes get silenced. We identify SPEN’s SPOC domain as a potent transcriptional repressor, crucial for SPEN-dependent XCI. Based on our mass spectrometry analysis, we propose that the SPOC domain is key for bridging *Xist* with other factors implicated in XCI, such as HDAC3, which we find present at most X-linked enhancers where SPEN is recruited. In particular, the interaction of the SPOC domain with the NuRD complex and the transcription machinery points to a role for SPEN in direct transcriptional repression. We also identify SPOC as a novel interactor of the m6A methyltransferase complex, which plays a role in *Xist* RNA methylation, a modification important for *Xist*-dependent silencing[25]. Methylation of *Xist* is mediated by RBM15[25], which interacts with the m6A machinery directly through ZC3H13[29]: the most highly enriched m6A machinery factor identified in our mass

spectrometry. As RBM15 also carries a SPOC domain, our study raises the possibility that the interaction with the RNA methylation machinery is not restricted only to SPEN's SPOC domain, but may instead be a feature shared across SPOC-containing proteins.

SPEN binds other non-coding RNAs, including *SRA*[18] (involved in steroid-receptor regulation). Furthermore, another *SRA*-binding protein, SLIRP, was shown to bind promoters in an *SRA*-dependent manner[30], raising the possibility that similarly to *Xist*, *SRA* could guide SPEN to target gene regulatory elements.

In conclusion, our study suggests that RNA-mediated recruitment of SPEN and other SPOC-containing proteins, which are found across fungi, plants and animals, may be a widespread means to repress transcription acutely by coordinately engaging several layers of epigenetic and transcriptional control. We propose that SPEN bridges *Xist* to the transcription machinery, histone deacetylases and chromosome remodeling factors, to ensure robust and efficient XCI (Extended Data Fig. 4k).

Methods

Data reporting and statistical analysis

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. All statistical tests, resulting p-values and observation number are indicated in the figure panels or in the figure legends.

Data visualization

All heatmaps, violin plots, boxplots, density plots, barplots and piecharts were generated using ggplot2. Unless stated otherwise, boxplots always show the median as the center line, box limits correspond to upper and lower quartiles, and whiskers cover 1.5x the interquartile range.

Plasmid construction

The plasmids to target *OsTIR1* at the *TIGRE* locus (Addgene plasmid # 92141) and the *TIGRE* specific gRNA encoding plasmid (Addgene plasmid # 92144) were kindly provided to us by Elphège Nora. The additional *TIGRE* targeting plasmids BglG-mCherry-T2A-OsTir1 (pFD51) and rtTa-VP16-T2A-OsTir1 (pFD68) were cloned using PCR amplification of corresponding gene cassettes followed by traditional cloning into the #92141 backbone.

Targeting construct (pFD19 and pFD49) to tag endogenous *Spen* at its C-terminus with AID-HaloTag and AID-GFP respectively were generated as follows: 500bp homology arms (flanking both sides of, but excluding the stop codon of *Spen*) were PCR amplified from mouse genomic DNA. 1-step Gibson cloning (NEB) was subsequently used to simultaneously surround the digested AID insert (carrying a puromycin resistance gene under the control of the PGK promoter) in frame with the homology arms and clone the insert into the pBR322 vector. Synonymous mutations in the PAM/SEED target sequence (located on the 5' homology arm) were then introduced using the QuickChange II XL site-directed mutagenesis kit (Agilent) to prevent Cas9 mediated cutting of the targeting vector

upon transfection and of the AID tagged allele(s) upon integration. Targeting construct (pFD90) to replace the endogenous Spoc domain of *Spen* by GFP was generated using the same strategy. For gRNA cloning, the pX459 plasmid (gift from Feng Zhang, addgene #62988) encoding *Sp*Cas9 was digested with *BbsI* immediately downstream of the U6 promoter, and annealed DNA duplex corresponding to the target gRNA sequences were ligated.

Cell culture

Mouse XX ES cells (TX1072) were grown on 0.1% gelatin-coated flasks in 8% CO2 37°C incubators. For all experiments, cells were cultured in 2i + LIF, and batch-tested fetal calf serum ES cell medium - DMEM (Sigma), 15% FBS (Gibco), 0.1mM β-mercaptoethanol, 1000 U/ml leukemia inhibitory factor (LIF, Chemicon), CHIR99021 (3uM), PD0325901 (1uM).

Neural progenitor cell (NPC) differentiations and subcloning were performed as previously described[31]. NPCs were grown in N2B27 medium supplemented with EGF and FGF (10ng/ml each), on 0.1% gelatin-coated flasks.

Cell transfection and clone isolation

All transgenic insertions were performed using the 4-D nucleofector system from Lonza. For each nucleofection, 5 million cells were electroporated with 2.5ug each of non-linearized targeting vectors and gRNA/Cas9 encoding plasmids (MidiPreps). Nucleofected cells were then serially diluted and plated on 10-cm dish. 48 hours later, selection was added (puromycin: 0.4ug/mL, hygromycin: 250ug/mL, blasticidin: 5ug/mL) except for transfection steps involving flippase-mediated removal of resistance cassettes, during which no selection was applied. 1 week after initial plating, 80 to 96 single colonies were picked from dishes showing ideal clonal density and seeded in 96-well plates. These cells were subsequently split into one high-confluency plate used for PCR genotyping, and one low-confluency plate from which desired clones were further expanded until T25 density was reached. At this stage, some cells were kept to reconfirm correct genotype by PCR, while the remaining cells were frozen.

Cell treatments

Xist expression in TX1072 mESCs was induced upon administration of doxycycline (1ug/mL). Auxin mediated depletion of target proteins was achieved through supplementing culture media with Auxin (Sigma) at the recommended concentration of 500uM. Auxin-containing medium was renewed every 24 hours. For auxin wash-out, auxin-containing medium was removed, cells were rinsed once with PBS, and exposed to auxin-free medium.

Protein extraction and western blotting

Cells were trypsinized, washed once in medium, once in PBS and pellets immediately frozen at -80°C. Pellets were then resuspended in RIPA buffer (50mM Tris-HCl pH8.0-8.5, 150mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS) containing protease inhibitors (Roche), incubated for 30 minutes on ice and sonicated with a Bioruptor (three 10 second pulses). Lysates were then centrifuged for 20 minutes at 4°C, and supernatants were

kept. Protein concentration was determined using the Bradford (BioRad) assay. Samples were then boiled at 95°C for 10 minutes in LDS buffer (Thermo) containing 200mM DTT. For all Western blots except ones aiming at detecting SPEN, 4-12% Bis-Tris gels were used. For detection of SPEN, a high molecular weight protein (>400 kDa), 3-8% Tris-Acetate poly-acrylamide gels were used. Transfer was performed on a 0.45um nitrocellulose membrane using a wet-transfer system, at 350-400mA for 2h at 4°C.

RNA extraction, RT, Pyrosequencing and RNAseq

RNA extraction was performed using the RNeasy kit and on-column DNAse digestion (Qiagen). Reverse transcription was performed on 1ug total RNA using Super Script III (Life Technologies). To quantify allelic skewing, cDNA was amplified using biotinylated primers and subsequently sequenced using Q24 Pyromark (Qiagen). Only samples showing a RNA integrity number (RIN) above 9 were used to prepare RNAseq libraries (TruSeq). Paired-end 100nt sequencing was performed on HiSeq2500 or NovaSeq6000 (Illumina, San Diego, CA).

RNA FISH

Cells were dissociated using Trypsin (Invitrogen) for ESCs or Accutase (Invitrogen) for NPCs, washed twice in medium, and allowed to attach on Poly-L-Lysine (Sigma) coated coverslips for 10 min. Cells were fixed with 3% paraformaldehyde in PBS for 10 min at room temperature, washed in PBS three times, and permeabilized with ice-cold permeabilization buffer (PBS, 0.5%Triton X-100, 2mM Vanadyl- ribonucleoside complex) for 5 minutes on ice. Coverslips were stored in 70% Ethanol at −20°C. Samples were dehydrated in 4 baths of increasing Ethanol concentration (80%, 95%, 100% twice) and air-dried quickly. Probes were prepared from minipreps of intron-spanning BACs (clone RP24-157H12 for *Huwe1*, RP23-260I15 for *Atrx*) or plasmid (p510 for *Xist*). Probes were labeled by nick translation (Abbott) using dUTP labeled with spectrum green (Abbott) for *Huwe1*, spectrum red (Abbott) for *Atrx*, and Cy5 (Merck) for *Xist*. Labeled BAC probes were co-precipitated with Cot1 DNA repeats in the presence of ethanol and salt, resuspended in formamide, denatured at 75°C for 10 minutes, and competed at 37°C for 1 hour. Probes were then co-hybridized in FISH hybridization buffer (50% Formamide, 20% Dextran sulfate, 2x SSC, 1 μg/μl BSA, 10mM Vanadyl-ribonucleoside) at 37°C overnight. The next day, hybridized coverslips were washed three times for 5 minutes with 50% formamide in 2X SSC at 42°C, and three times for 5 minutes with 2X SSC. 0.2mg/ml DAPI was added to the penultimate wash and coverslips were mounted with Vectashield (Vectorlabs).

HaloTag labeling

HaloTag labeling of SPEN-Halo fusion protein was performed in live TX1072 ESCs and NPCs. Cells were labeled with HaloTag ligand conjugated Janelia Fluor[32] (JF646-HaloTag or JF549-HaloTag, kind gift from Luke Lavis) at a final concentration of 250nM in culture medium. Labeling was performed for 1 hour at 37°C, cells were then washed 4 times with generous volumes of PBS, and incubated with unlabeled medium for 15 minutes before proceeding with downstream experiments. For NPC labeling, cells were washed with unlabeled medium and not PBS, as NPCs detach when exposed to PBS. Auxin and/or doxycycline were kept in the labeling medium when necessary.

HaloTag labeling followed by RNA FISH

For co-detection of SPEN-Halo and *Xist* RNA, cells were labeled with JF549 as indicated above, and directly processed for fixation and permeabilization as detailed in the RNA FISH section. Importantly, after permeabilization, coverslips were directly washed twice with PBS, twice with 2xSSC and immediately processed for FISH.

Mouse breeding, embryo collection and single-embryo RNAseq

Timed natural matings were used for all experiments. Noon of the day when the vaginal plugs of mated females were identified was scored as E0.5. For *Spen* matings a conditional allele was used[33]. For oocyte deletions published *Rosa26:Zp3-Cre* allele was used[34]. F1 hybrid $Spen^{+/-}$ males were obtained by crossing $Spen^{+/+}$ *CAST/EiJ* females with $Spen^{+/-}$ *C57BL/6J* males. For *Spen* maternally deleted embryos, $Spen^{flox/flox}$ *Zp3-Cre*$^{+ve}$ *C57BL/6J* females were crossed with $Spen^{+/-}$ F1 hybrid males. For *Spen* control embryos, $Spen^{flox/flox}$ *Zp3-Cre*$^{-ve}$ *C57BL/6J* females were crossed with $Spen^{+/-}$ F1 hybrid males. Animal care and use for this study were performed in accordance with the recommendations of the European community (2010/63/UE). All experimental protocols were approved by the ethics committee of Institut Curie CEEA-IC118 under the number APAFIS#8812-2017020611033784v2 given by national authority in compliance with the international guidelines. Single-embryo RNA-seq was performed as previously described[35]. Briefly, E3.5 embryos were collected and morphologically assessed to ensure only viable samples were collected. Zona pellucida was removed by treatment with acidified Tyrode's solution. Single embryos were picked into individual tubes and cDNA was prepared and amplified as previously described[36]. Illumina libraries were prepared as published in [37]. Paired-end 100nt sequencing was performed with HiSeq2500 (Illumina, San Diego, CA).

Live-cell imaging and analysis

Cells were seeded on fibronectin coated 35mm glass bottom dishes (Ibidi) 24 hours before imaging. Doxycycline was added 1 hour before image acquisition. Cells were imaged on the DeltaVision OMX microscope in widefield mode (GE Healthcare) using a 1.4 NA 100x oil immersion objective. Temperature was controlled at 37°C and CO_2 at 8% during acquisition.

Images were acquired as z-stacks of 40 slices with 400nm steps every 10 minutes for at least 4 hours. Movies were deconvolved using Huygens deconvolution with the following parameters: Iteration 4; S/N 5, 10; Quality Threshold 0.1; and Widefield mode 0.7 was used for background estimation. 2 channels were registered using TetraSpec microspheres 0.1um (Invitrogen) and unwarpJ (Fiji plugin). For segmentation, z-projected deconvolved registered images were used and pixels were classified as cloud/nuclei using Ilastik. Touching nuclei were sometimes manually separated. Cut offs on resulting probability maps were set to 0.7. We next performed connected component analysis to obtain integer labelled images where each integer label corresponds to a unique nucleus. In the tailored made Fiji plugin the inputs are the raw max z-projected time-lapse images of the two channels and the integer labelled time-lapse image of the nuclei. The probability maps of the clouds give the region of interest in the time-lapse sequence in which total intensity is calculated. In the plugin, clouds are associated with their corresponding nuclei, they are then linked via Kalman filter tracker over time. These unique links constitute track id's and contain the information about

the intensity and area measurements for each cell. For each tracked cell, the first timepoint when a cloud is detected in one channel (Xist or SPEN) is labeled as reference timepoint 1.

Hi-C

Hi-C was performed as previously described[38], except that ligated DNA size selection was omitted, and dA-tailing was performed prior to biotin pull-down. Briefly, each Hi-C experiment was performed on 10 million cells (NPCs) per sample. Cells were digested with DpnII at 37°C overnight. DNA ends were filled with biotin-14-dATP at 23°C for 4 hours. DNA was then ligated with T4 DNA ligase at 16°C overnight. Binding proteins were removed by treating ligated DNA with proteinase K at 65°C overnight. Purified proximally ligated molecules were fragmented to obtain an average fragment size of 200 bp. After DNA end repair, dA-tailing and biotin enrichment, DNA molecules were ligated to Illumina TruSeq sequencing adapters at room temperature for 2h. Final library PCR productions were carried out following the Illumina TruSeq Nano DNA Sample Prep Kit manual. Paired-end 100nt sequencing was performed on HiSeq4000 (Illumina, San Diego, CA).

Genetic engineering strategy for Xist-Bgl stem loop tagging and SPEN complementation analysis constructs

To tag *Xist* with Bgl stem loops[39], we nucleofected cells with pBS-Ptight-Xist-BglSL[40] (plasmid harboring 18 repeats of BglSL inserted between homology arms to target *Xist* exon 7, carrying a G418 selection gene, kind gift from Osamu Masui). Following G418 selection and FLP-FRT mediated removal of the selection cassette, clones were picked and genotyped. Positive clones were further tested to ensure that the stem-loop tagged *Xist* could properly be induced and trigger gene silencing upon addition of doxycycline (data not shown).

Spen cDNA truncations were generated by splicing out different regions of the *Spen* ORF (Genscript, ORF clone OMu11416C) using overlap extension PCR. Each *Spen* truncation was cloned downstream of a CAGGS promoter into a vector carrying homology arms for targeted insertion at the *ROSA26*[41] locus as well as a SV40-promoter driven hygromycin resistance gene. The BglG-GFP-Spoc targeting plasmid was designed by inserting a translational fusion between a BglG-GFP cassette and SPEN aa3244-3643 into the same *ROSA26* targeting vector. Each of these “complementation” constructs were independently targeted at *ROSA26* in Spen-degron mESCs. Independent clones were picked and protein expression of each SPEN truncation was assessed by western blot. XCI complementation analysis was then performed in 2-3 independent clones for *Spen* cDNA truncations, and 4 independent clones for BglG-GFP-SPOC expressing clones. The ability of cells to accumulate BglG-mCherry, BglG-GFP and BglG-GFP-SPOC upon addition of doxycycline was assessed using microscopy (data not shown).

Immunofluorescence

ESCs were dissociated using Trypsin (Invitrogen), washed extensively in medium, and allowed to attach on Poly-L-Lysine (Sigma) coated coverslips for 10 min. Cells were then fixed with 3% paraformaldehyde in PBS for 10 min at room temperature, washed in PBS three times, and permeabilized with 0.25%Triton X-100 in PBS for 5 minutes at room temperature. Coverslips were then washed three times in PBS and blocked for 1 hour with

blocking buffer (PBS containing 2.5% BSA, 0.1% Tween20 and 10% normal goat serum). Coverslips were then incubated with primary antibodies diluted in blocking buffer at 4°C overnight, washed three times for 5 minutes in PBST (0.1% Tween20) on the next day, incubated with fluorescently labeled secondary antibodies (1/500 in blocking buffer) for 1 hour at room temperature, and washed again three times for 5 minutes in PBST. 0.2mg/ml DAPI was added to the penultimate wash and coverslips were mounted with Vectashield (Vectorlabs).

Immunoprecipitation

Nuclear extracts were prepared by resuspending 50 million fresh cells in ice-cold 10mL buffer A (10mM HEPES pH7.9, 10mM KCl, 1.5mM MgCl2, 0.1% NP-40, c0mplete EDTA free, phosSTOP) and rotating for 10 minutes at 4°C. Nuclei were centrifuged at 2000rpm for 10 minutes at 4°C and resuspended in 1mL IP buffer C150 (20mM HEPES pH7.9, 150mM NaCl, 1.5mM MgCl2, 0.2mM EDTA, 0.25% NP-40, cOmplete EDTA free, phosSTOP). Lysates were briefly sonicated followed by Benzonase (Merck) digestion for 30 minutes at 4°C. Finally, lysates were cleared through centrifugation at 13000rpm for 20 minutes prior to being incubated with 15uL of GFP trap magnetic agarose bead slurry (ChromoTek) overnight at 4°C. Beads were washed 5 times in IP buffer. For Co-IP western blot, washed beads were directly resuspended in LDS buffer (Thermo) containing 200mM DTT, and boiled at 95°C for 10 minutes.

Proteomics and Mass Spectrometry Analysis

Proteins on magnetic beads were washed twice with 100 µL of 25 mM NH4HCO3 and we performed on-beads digestion with 0.2 µg of trypsin/LysC (Promega) for 1 hour in 100 µL of 25 mM NH4HCO3. Sample were then loaded onto a homemade C18 StageTips for desalting. Peptides were eluted using 40/60 MeCN/H2O + 0.1% formic acid and vacuum concentrated to dryness. Online chromatography was performed with an RSLCnano system (Ultimate 3000, Thermo Scientific) coupled online to a Q Exactive HF-X with a Nanospay Flex ion source (Thermo Scientific). Peptides were first trapped on a C18 column (75 µm inner diameter × 2 cm; nanoViper Acclaim PepMapTM 100, Thermo Scientific) with buffer A (2/98 MeCN/H2O in 0.1% formic acid) at a flow rate of 2.5 µL/min over 4 min. Separation was then performed on a 50 cm x 75 µm C18 column (nanoViper Acclaim PepMapTM RSLC, 2 µm, 100Å, Thermo Scientific) regulated to a temperature of 50°C with a linear gradient of 2% to 30% buffer B (100% MeCN in 0.1% formic acid) at a flow rate of 300 nL/min over 91 min. MS full scans were performed in the ultrahigh-field Orbitrap mass analyzer in ranges m/z 375–1500 with a resolution of 120 000 at m/z 200. The top 20 intense ions were subjected to Orbitrap for further fragmentation via high energy collision dissociation (HCD) activation and a resolution of 15 000 with the intensity threshold kept at 1.3 × 105. We selected ions with charge state from 2+ to 6+ for screening. Normalized collision energy (NCE) was set at 27 and the dynamic exclusion of 40s. For identification, the data were searched against the *Mus musculus* (UP000000589) Uniprot database using Sequest HF through proteome discoverer (version 2.2). Enzyme specificity was set to trypsin and a maximum of two-missed cleavage sites were allowed. Oxidized methionine and N-terminal acetylation were set as variable modifications. Maximum allowed mass deviation was set to 10 ppm for monoisotopic precursor ions and 0.02 Da for MS/MS peaks. The

resulting files were further processed using myProMS[42] v3.6 (work in progress). FDR calculation used Percolator and was set to 1% at the peptide level for the whole study. The label free quantification was performed by peptide Extracted Ion Chromatograms (XICs) computed with MassChroQ version 2.2[43]. For protein quantification, XICs from proteotypic peptides shared between compared conditions (TopN) with two-missed cleavages were used. Median and scale normalization was applied on the total signal to correct the XICs for each biological replicate. To estimate the significance of the change in protein abundance, a linear model (adjusted on peptides and biological replicates) was performed and p-values were adjusted with a Benjamini–Hochberg FDR procedure with a control threshold set to 0.05. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD015699.

Cross-linked CUT&RUN

CUT&RUN against SPEN was performed during a timecourse of *Xist* induction/SPEN degradation: 0h dox, 4h dox, 8h dox, 24h dox and 8h dox + auxin. Two biological replicates were performed. The original CUT&RUN protocol[44] was adapted for fixed cells: 10^6 cells in suspension were fixed with 2% formaldehyde diluted in PBS for 10 minutes at room temperature (2mL final volume). Fixation was quenched with 125mM glycine for 5 minutes and cells were washed twice in 1mL PBS. Fixed cells were then permeabilized with 1mL permeabilization buffer (20mM HEPES pH7.9, 150mM NaCl, 0.5mM Spermidine, 0.25% TritonX-100, c0mplete EDTA free) for 5 minutes and washed twice in 1mL PBS. Cells were then resuspended in 1mL washing buffer (20mM HEPES pH7.9, 150mM NaCl, 0.5mM Spermidine, 0.1% BSA, cOmplete EDTA free), bound to activated concanavalin beads (50uL bead slurry used per 10 million cells) for 10 minutes, blocked in 1mL blocking buffer (wash buffer + 2mM EDTA) for 5 minutes. At this stage, cells were resuspended in 500uL wash buffer containing target antibodies diluted 1/200, transferred to 0.5mL tubes, and incubated overnight at 4°C on an end-to-end rotator. Cells were washed three times in 500uL washing buffer followed by 1-hour incubation with pA-MNase (500uL of washing buffer containing 700ng/mL pA-MNase, produced by the Protein Expression and Purification Core Facility of Institut Curie) and washed again three times in 500uL washing buffer. After the last wash, cells were resuspended in 150uL washing buffer, transferred to 1.5mL tubes, and equilibrated to 0°C in a metal block for 10 minutes. To start digestion, CaCl2 was added to 1.5mM final concentration, taking care to return each sample to 0°C immediately after. Digestion was performed at 0°C for 1 hour, before being stopped by adding 150uL of 2X-STOP solution (200mM NaCl, 20mM EDTA, 5mM EGTA, 0.1% NP-40, 40ug/mL glycogen). RNAse A was added to a final concentration of 50ug/mL and samples were incubated at 37°C for 20 minutes. SDS and proteinase K were then added to final concentrations of 0.1% and 300ug/mL respectively and samples were incubated at 56°C for 2 hours followed with 68°C for 16 hours to reverse crosslinking. Total DNA was extracted using phenol-chloroform followed by two rounds of ethanol precipitation and DNA size selection (using 0.55x volume of Ampure XP beads relative to DNA sample volume) to remove the large predominating undigested DNA fragments. Each time, beads were discarded, and the supernatant (containing the selected small fragments resulting from MNase digestion) was ethanol precipitated. Following elution in 50uL TE buffer, samples were quantified and analyzed using Qubit and Tapestation assays. Cut&Run libraries were

prepared from 50ng DNA per samples, using the Accel-NGS 2S Plus DNA Library Kit (Swift) according to manufacturer's protocol. Paired-end 100nt sequencing was performed on HiSeq2500 (Illumina, San Diego, CA).

Bioinformatics analyses

All data were mapped to the mouse genome mm10, using the BL6-EiJ / CAST SNPs from the mouse genome project (v5 SNP142), and the gene annotation from ensembl (v92). Analyses were performed in R (v 3.4.2) and Bioconductor (v3.6). See (Zylicz et al.)[45] for more details.

RNA seq analysis—Reads were trimmed using Trimgalore (v 0.4.4), mapped using STAR (2.5.3a, parameters: --outFilterMultimapNmax 1 --outFilterMismatchNmax 999 --outFilterMismatchNoverLmax 0.06 --alignIntronMax 500000 --alignMatesGapMax 500000 --alignEndsType EndToEnd --outSAMattributes NH HI NM MD), and removed when mapping to the mitochondrial genome. Remaining reads were split by allele using SNPsplit (v 0.3.2). Allele specific and the unassigned bam files were sorted, duplicates removed using picard (v2.18.2, parameters: REMOVE_DUPLICATES=true ASSUME_SORTED=true) and pooled as the total reads. Quantification of expression was performed using featureCount (parameters: -p -t exon -g gene_id, -s 1 for stranded RNAseq of in vitro cell, -s 0 for non-stranded RNAseq of single embryo). Data were then analysed in R using DESeq2 (v1.18.1), calculating the sizeFactor on the count of total reads and applying it to the allele specific counts.

For all RNAseq analysis (Spen-degron mESCs, NPCs, and Spen-KO embryos), genes showing less than 10 total allelic reads in at least one sample were discarded from the analysis. Allelic ratios were then computed for genes as follows: allelic_ratio = $reads^{B6}$/($reads^{B6}$+$reads^{Cast}$). Allelic ratios were then averaged between biological replicates.

Spen-degron mESCs RNAseq analysis

Defining differential dependencies on Spen for gene silencing during XCI in mESCs: For this analysis (Figure 1D), skewed genes (i.e. genes showing allelic ratios outside of a [0.15;0.85] interval in control conditions) were removed. We then defined a silencing index, translating how much a gene is silenced after 24h of Xist induction with respect to the control condition: silencing_index = 1-(allelic_ratio$_{dox}$/allelic_ratio$_{control}$). We next filtered out genes showing less than 10% silencing (i.e. silencing_index<=0.1) in Spen non-depleted condition.

K-means with 3 clusters was then performed on the raw allelic_ratios across control, 24hdox and 24hdox+aux conditions. Clustering identified 3 groups of genes differing by their response to loss of SPEN during Xist induction. To define how dependent on Spen a gene is for silencing, we expressed the silencing defect observed upon loss of SPEN as a fraction of the total silencing that normally occurs in the presence of SPEN. Computationally, this translates in: Spen_dependence_index = 1 − (silencing_index$_{dox+aux}$/silencing_index$_{dox}$). A Spoc_dependence_index was derived identically.

Integration with HDAC3 KO RNAseq during XCI: Our Spen-degron dataset was integrated with an Hdac3 KO RNAseq dataset[45] generated from the same mESC background (TX1072) and at the same timepoint of Xist induction. The dataset was processed identically, and an Hdac3 dependence index was also computed as follows: Hdac3_dependence_index = 1 − (silencing_index$_{Hdac3KO}$/silencing_index$_{WT}$).

Spen-KO E3.5 embryo RNAseq analysis

Integration of SPEN and Xist KO embryo datasets: Our *Spen* KO E3.5 female embryo RNAseq dataset was integrated with a *Xist* KO scRNAseq (processed as pseudo-bulk for our analysis) dataset from E3.5 female embryos[46], also generated from a *Mus musculus domesticus* x *Mus musculus castaneus* mouse background.

Spen-degron NPC RNAseq analysis

Defining genes that escape from XCI: In NPCs, X-linked genes were defined as escapees if their transcript allelic ratio was superior to 0.15 in at least one condition (0h, 24h or 48h of SPEN depletion).

CUT&RUN bioinformatics analysis—Reads were trimmed using Trimgalore (v 0.4.4), mapped using STAR (2.5.3a, parameters: --outFilterMultimapNmax 1 --outFilterMismatchNmax 999 --outFilterMismatchNoverLmax 0.06 --alignIntronMax 1 --alignMatesGapMax 2000 --alignEndsType EndToEnd --outSAMattributes NH HI NM MD), and removed when mapping to the mitochondrial genome. Remaining reads were split by allele using SNPsplit (v 0.3.2). Allele specific and the unassigned bam files were sorted, duplicates removed using picard (v2.18.2, parameters: REMOVE_DUPLICATES=true ASSUME_SORTED=true) and pooled as the total reads. BigWig of coverage files were done using DeepTools bamCoverage (parameters: --extendReads --binSize 1, with --extendReads 200 for single end data). A scaling factor was calculated as 10^6 / total number of reads, and the same factor was given as the parameter --scaleFactor for both allelic signals. Peak calling was performed using macs2 (v 2-2.1.2.1, parameters for CUT&RUN : --bw 300 -f BAMPE -q 0.01 --keep-dup auto --broad for CUT&RUN and pol2S5 ChIPseq; --bw 300 -f BAMPE -q 0.01 --keep-dup auto --call_summits for other ChIPseq). For quantification of signal in peaks, reads were counted using the featureCounts function from Subread (v1.28.1, parameters: -p -s 0). Data scaling was performed in R using DESeq2 (v1.18.1), calculating the sizeFactor on the counts of total reads in 10kb windows and applying it to the allele specific counts in peaks.

Peak filtering: Spen-specific peaks were defined as having a log2FoldChange >= 1 compared to auxin treatment (negative control, Spen-degraded), and an adjusted p-value <= 0.001.

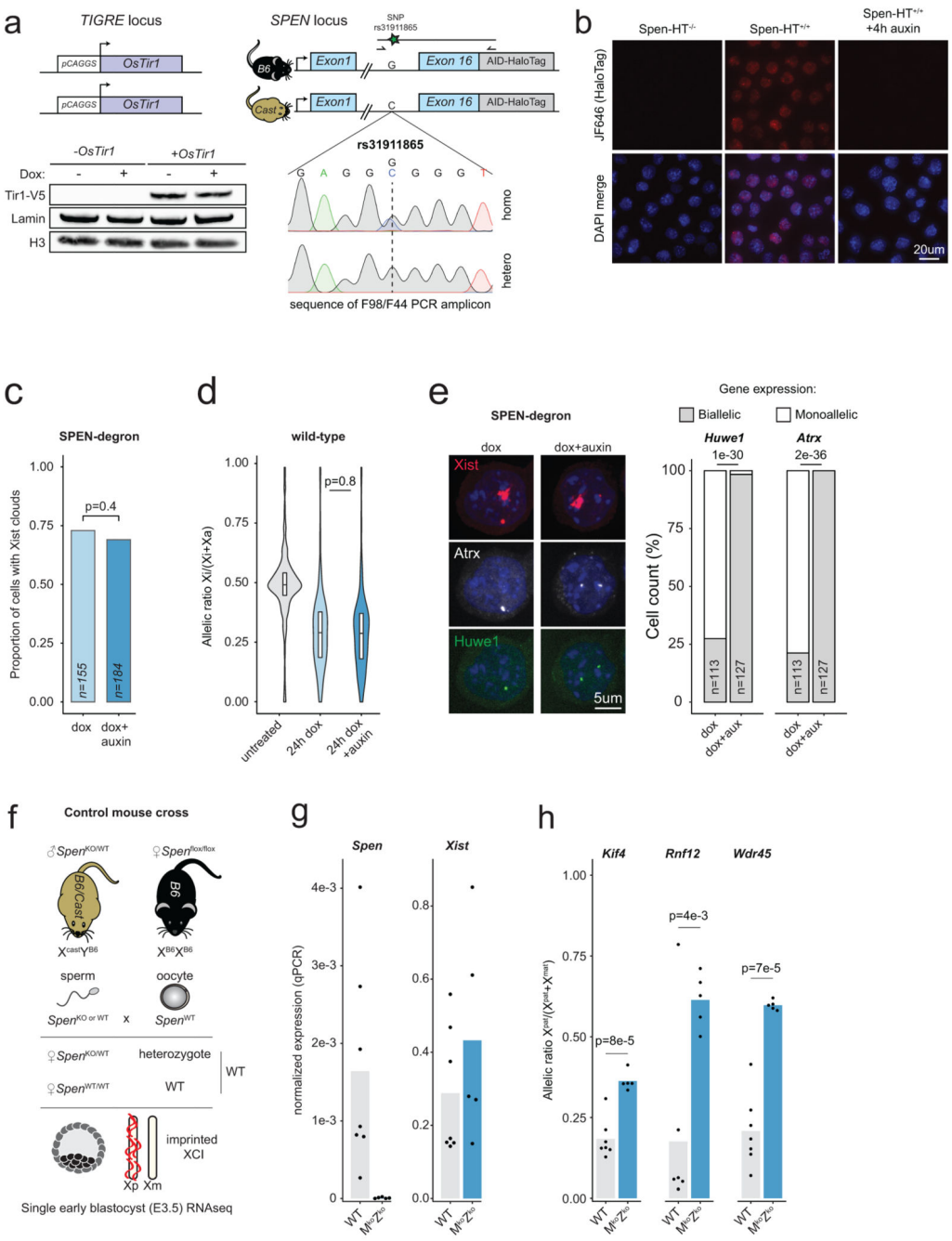
Total Spen enrichment in promoter windows: To compare SPEN accumulation among promoter of all X-linked genes unbiasedly (including genes which fail to have any peak called at their promoters), we performed DEseq analysis on counts spanning total promoter windows.

Genomic features and integration with RNAseq: Promoters were defined as ±2kb windows centered around genes TSS. Putative active enhancers and their deacetylation kinetics during XCI were obtained from Zylicz et al.[45]. Gene silencing efficiency was determined according to the silencing_index defined in the RNAseq analysis section above. We observed that our silencing_index ranges between 0 and 0.9. Hence, we split this interval in 3 to define high, medium, and low gene silencing efficiency groups with silencing_index comprised in [0.6,0.9], [0.3,0.6[, and [0,0.3[respectively.

Integration with publically available ChIPseq data: SPEN peaks were intersected with other peaks called from publically available ChIPseqs for HDAC3[45] (same cellular background, TX1072, 2i+LIF condition), RNAPII-pS5[47], CHD4[47] and MBD3[47] (all in 2i +LIF condition).

Hi-C analysis—Data were processed with HiC-Pro (v2.11.0) in allele specific mode. Only pairs with both reads having MAPQ>30 were kept. Matrices were made using cooler cload (v 0.8.5) at 1kb or 10kb resolution, using HiGlass for visualisation and snapshots. TADs were called using HicExplorer HicFindTads (parameters: --correctForMultipleTesting fdr --minDepth 250000 --maxDepth 4000000 --step 50000 --thresholdComparisons 0.1 --delta 0). Expected value for HiC signal was calculated on the non-allele specific signal using cooltool compute-expected. Average scaled matrices of observed/expected values for allele specific signal were produced with Coolpup.py (parameters: --local --rescale --rescale_size 299), using non-allele specific expected values to normalise both alleles to the same expected values. Average heatmaps were plotted using plotpup.py. For quantification, HiC signal was average over TADs upper-triangle for each allele specific matrices (10kb) using the hicExplorer hicSummarizeScorePerRegion available at https://github.com/heard-lab/HiCExplorer (parameters: --summarizeType mean --rmDiag 1).

Extended Data



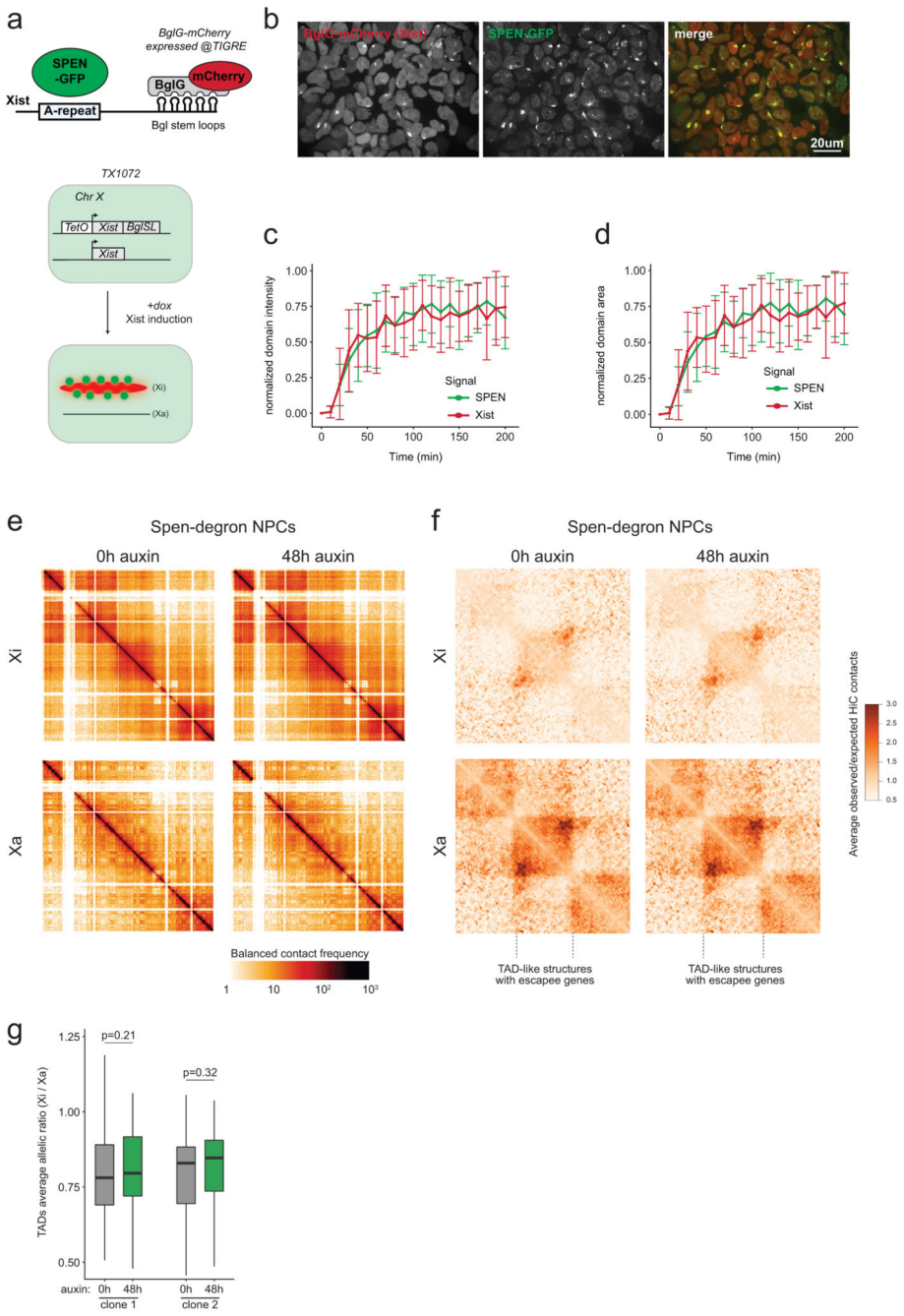
Extended Data Figure 1. SPEN mediates gene silencing across the entire X chromosome in vitro and in vivo.

a, Schematic representation of the SPEN-degron genotype with AID-HaloTag insertions in frame with the C-terminus of endogenous SPEN. Targeted homozygous insertion of V5-tagged OsTir1 at the TIGRE locus (top left) results in its constitutive protein expression as assessed by western blot (bottom left). Sanger sequencing results for a PCR amplicon specific to AID-HaloTag insertions and covering a SNP outside of the recombined left homology arm. Detection of both alleles in the amplicon confirms homozygous AID knock-

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

in. **b**, Fixed-cell imaging of HaloTag in wild-type cells (left) or in SPEN-degron mESCs (middle) exposed to auxin for 4 hours (right). Cells were labeled with Halo-JF646 before fixation. SPEN-Halo is properly localized to the nucleus, and depleted upon auxin treatment. This experiment was repeated at least twice with similar results. **c**, Barplot representing the proportion of cells displaying Xist RNA clouds (quantified using RNA FISH) before and after degradation of SPEN (n: number of cells counted, Chi-squared test). **d**, Violin plot showing X-chromosomal transcript allelic ratios distribution (RNAseq) after 0h dox, 24h dox or 24h dox+auxin treatment in Spen-degron mESCs wild-type cells (horizontal lines denote the median, box limits correspond to upper and lower quartiles, averages of two independent clones shown, n=434 genes, two-sided Student's t-test). **e**, RNA FISH for Xist (red) and two X-linked genes: Atrx (grey) and Huwe1 (green), in SPEN-degron mESCs treated with doxycycline only, or in combination with auxin for 24 hours. The proportion of Atrx/Huwe1 monoallelic and biallelic expression among *Xist*-expressing cells is shown (n: number of cells counted, Chi-squared test). **f**, Schematic of the control hybrid mouse cross set up for the experiment shown in Fig. 1g-h. **g**, qPCR quantification of Spen and Xist transcripts in wild-type (n=7) and maternal-zygotic Spen KO (n=5) E3.5 embryos. Bars show mean value. **h**, Pyrosequencing assay of 3 X-linked transcripts in maternal-zygotic Spen ko (n=5) and WT (n=7) E3.5 embryos (two-sided Student's t-test). Bars show mean value.



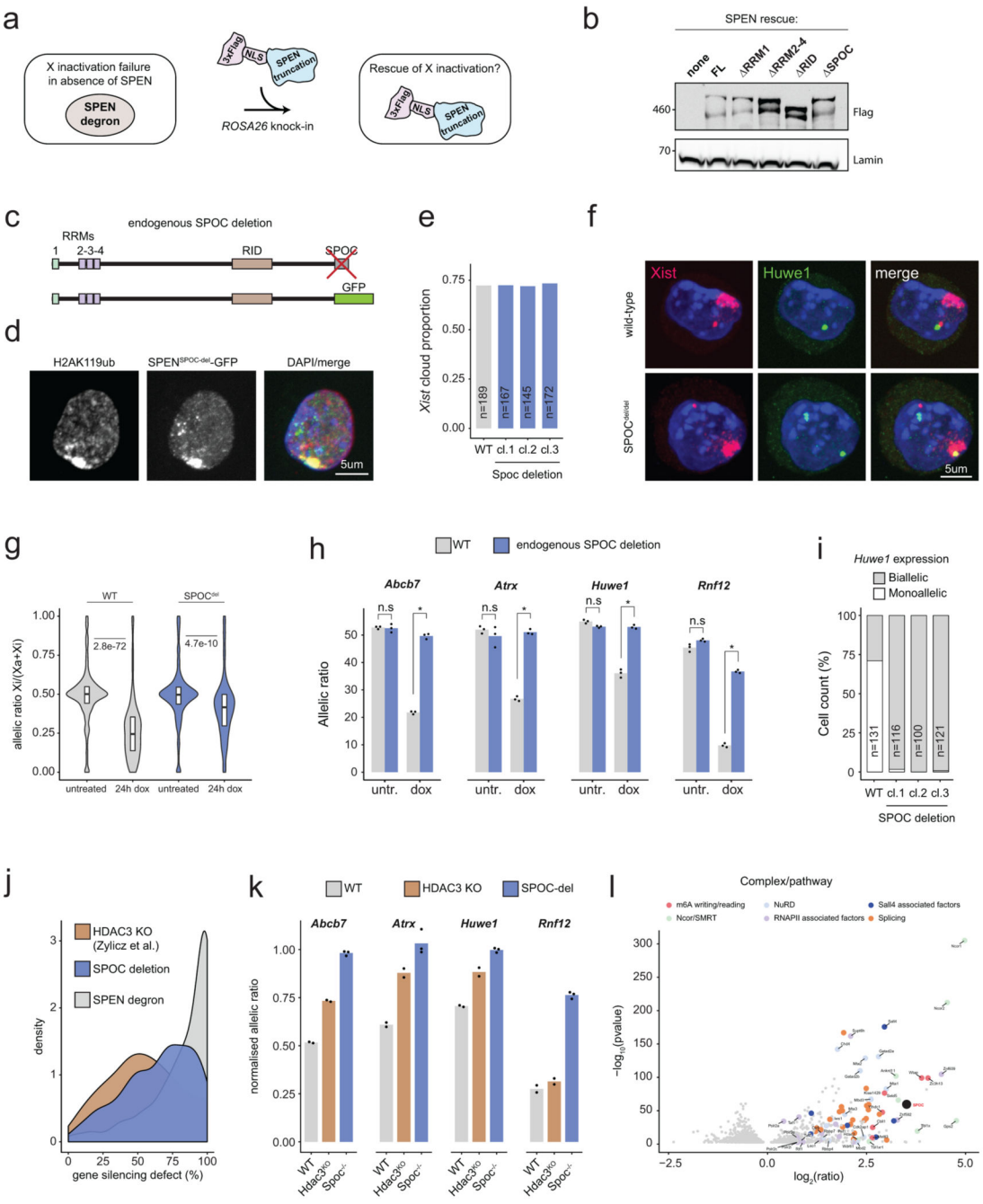
Extended Data Figure 2. SPEN localizes to the X chromosome immediately upon Xist upregulation and throughout the stages of XCI, but is dispensable for maintenance of X-linked gene silencing.

a, Scheme of the strategy for live-cell imaging of SPEN protein and Xist RNA. **b**, Live-cell snapshot after 16h of Xist induction in the cell line shown in a. This experiment was repeated at least twice with similar results. **c**, Kinetics of total intensity, and **d**, area of Xist (red) and SPEN (green) domains over time of Xist induction. **c, d**: Averages of 27 tracked cells. Error bars indicate standard deviation. Images were acquired every 10 minutes. Timepoint 1 is defined as the earliest time at which a SPEN or Xist domain is detected in

Europe PMC Funders Author Manuscripts

each cell. Intensity and area values were respectively normalized to the maximum value reached for each signal (SPEN and Xist). **e**, HiC map of the inactive (top) and active (bottom) X chromosomes (resolution: 1.024Mb) in NPCs after 0h or 48h of auxin-mediated SPEN depletion. **f**, Heatmap of the average contact enrichment on scaled TADs containing escapees in NPCs after 0h or 48h of auxin-mediated SPEN depletion. **g**, Quantification of the allelic ratio (Xi/Xa) of HiC signal within TADs (n=37) shown in f, after 0h or 48h of auxin-mediated SPEN depletion (horizontal lines denote the median, box limits correspond to upper and lower quartiles, two-sided Wilcoxon rank-sum test). **e, f**, Averages of two independent clones are shown.

Europe PMC Funders Author Manuscripts

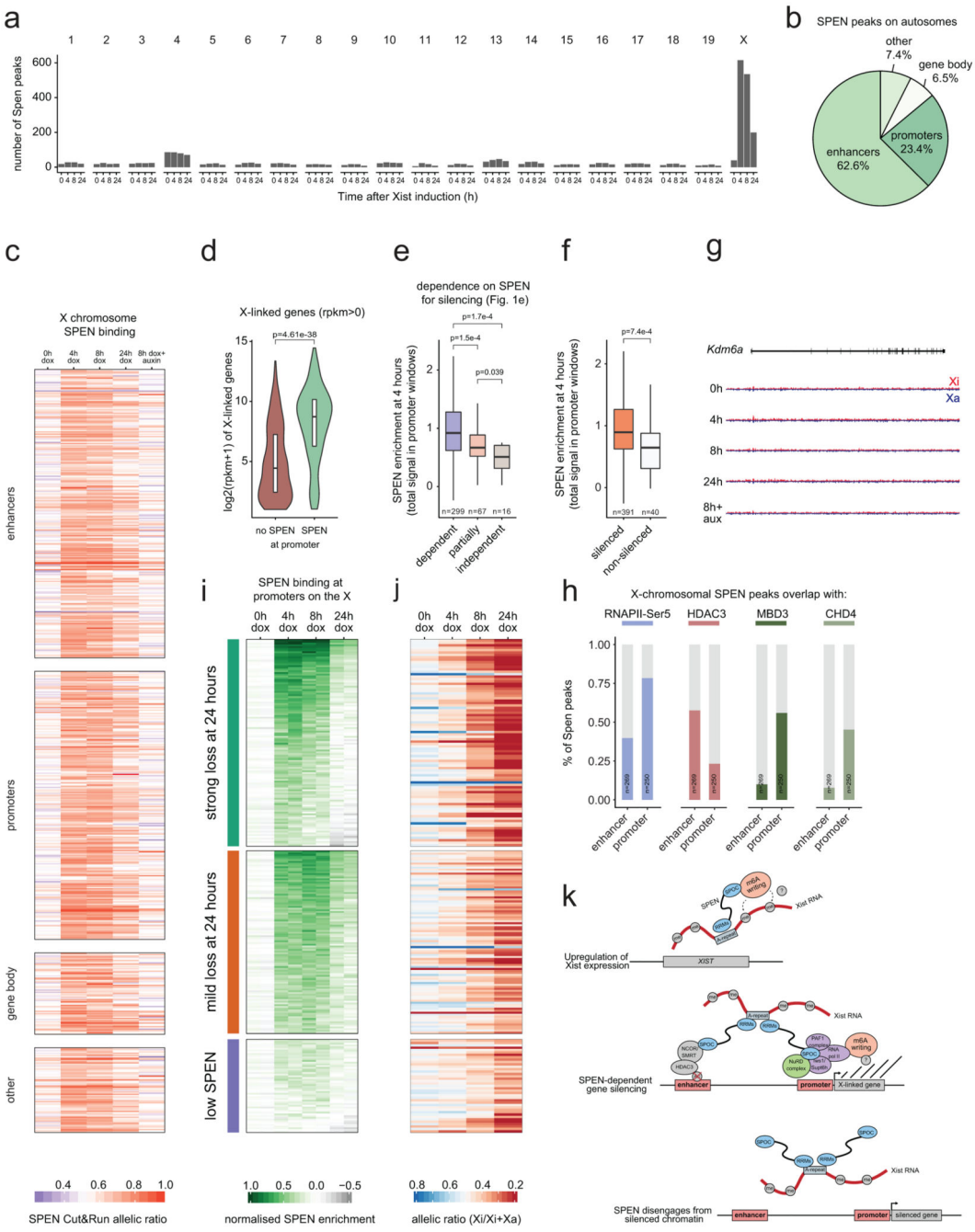


Extended Data Figure 3. The SPOC domain of SPEN mediates gene silencing and interacts with multiple molecular pathways.

a, Scheme of complementation strategy. **b**, Western blot detection of overexpressed 3x-FLAG-tagged SPEN protein rescue fragments. **c**, Scheme showing endogenous deletion of SPOC. **d**, Sub-nuclear localization of endogenous SPEN deleted of its SPOC domain upon Xist RNA induction. The Xi is identified using immunofluorescence detection of H2AK119ub1. **e**, Barplot showing the proportion of cells with Xist RNA clouds (assayed by RNA FISH) in wild-type cells and three independent Spoc deletion clones after induction of

Europe PMC Funders Author Manuscripts

Xist for 24 hours (n: number of counted cells). **f**, RNA FISH for Xist (red) and Huwe1 (green), in SPOC deletion and wild-type cells treated with doxycycline for 24 hours. **g**, Violin plot showing X-chromosomal transcript allelic ratios distribution (RNAseq) after 0h dox or 24h dox treatment in WT and SPOC deletion mESCs (horizontal lines denote the median, box limits correspond to upper and lower quartiles, averages of three independent clones shown, n=469 genes, two-sided Student's t-test). **h**, Barplot representation of transcript allelic ratio (pyrosequencing) for 4 X-linked genes in SPOC deletion (blue) or wild-type (grey) cells. Bars show mean values for 3 independent SPOC deletion clones (*P<10-4, two-sided Student's t-test). **i**, Barplot showing the proportion of cells expressing Huwe1 monoallelically (white) or biallelically (grey), assayed by RNA FISH, in wild-type cells and three independent Spoc deletion clones after induction of Xist for 24 hours (n: number of counted cells). **j**, Density plot showing the distribution of gene silencing defects (see Methods) observed across the X-chromosome in RNAseq data from HDAC3 KO (Zylicz et al.), SPEN-degron and SPOC-deletion (this study) ESCs after 24h of Xist induction. **k**, Barplot representation of normalised allelic ratio (pyrosequencing) for 4 X-linked genes in HDAC3-KO (brown), SPOC deletion (blue) and wild-type (grey) cells after 24h of Xist induction. Bars show mean values for 2 independent HDAC3 clones and 3 independent Spoc deletion clones. **l**, Volcano plot of GFP-pulldown fold-changes (BglG-GFP-SPOC versus BglG-GFP) and their adjusted p values (Benjamini-Hochberg procedure, see Methods for statistical analysis). Quantitative label-free mass spectrometry analysis was performed on 4 independent biological replicates. **b, d, f**, These experiments were repeated at least twice with similar results.

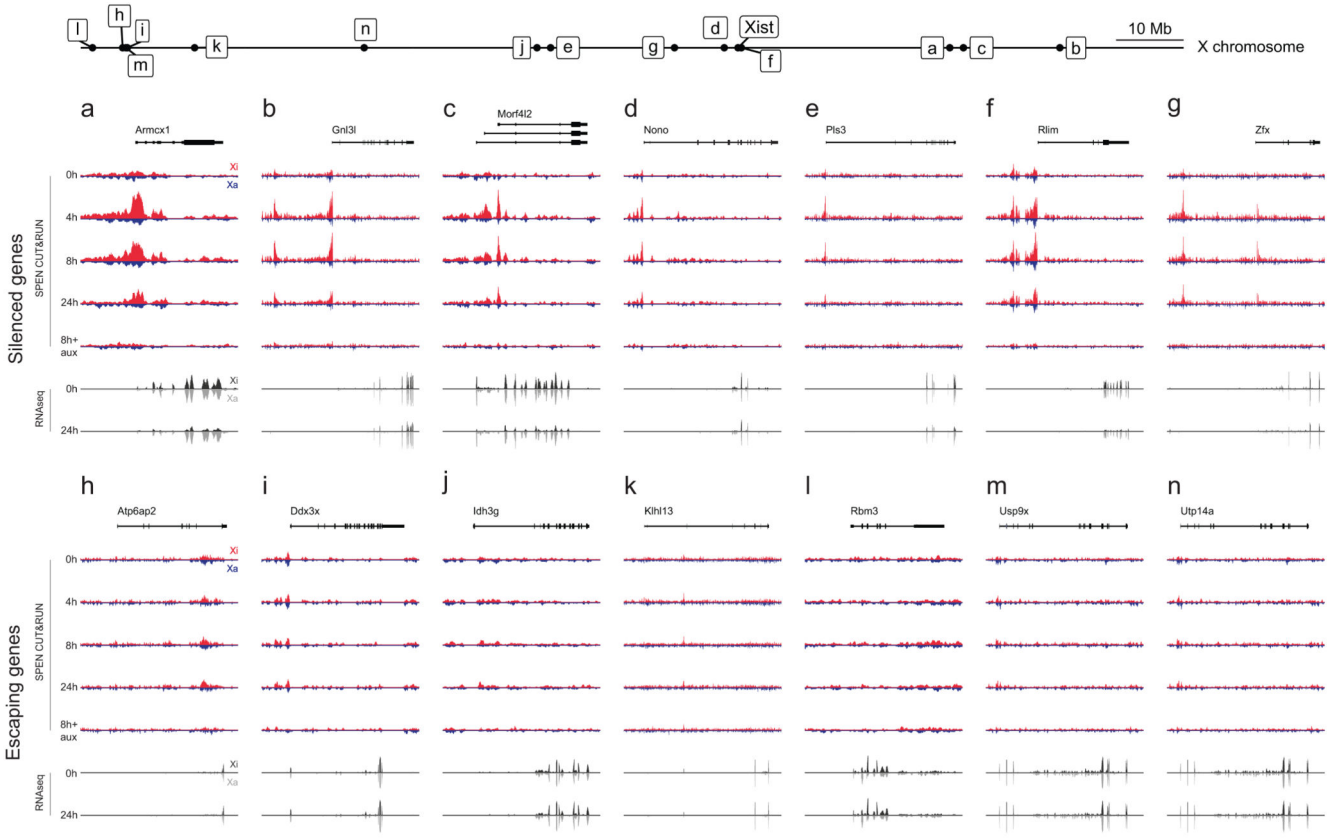


Extended Data Figure 4. SPEN is recruited by Xist to active gene promoters and enhancers where it silences transcription and subsequently disengages from chromatin.

a, Barplot showing the number of SPEN peaks on each chromosome after 0h, 4h, 8h and 24h of Xist induction in mESCs. **b**, Annotation of SPEN peaks on autosomes. **c**, Heatmap showing allelic-ratios at SPEN peaks during XCI among different X-linked genomic features. **d**, Violin plot showing gene expression (RPKM) of genes accumulating SPEN (n=259) or not (n=689) at their promoters. Genes showing 0 RPKM were excluded from this plot. **e**, Boxplots showing SPEN enrichment after 4 hours of Xist induction within promoter

Europe PMC Funders Author Manuscripts

windows of genes grouped based on their level of dependency of SPEN for gene silencing (see Figure 1e). **f**, Boxplots showing SPEN enrichment after 4 hours of Xist induction within promoter windows of genes grouped based on whether they are silenced or not at 24h of Xist induction (see Methods). **d, e, f**, two-sided Wilcoxon rank-sum test, horizontal lines denote the median, box limits correspond to upper and lower quartiles. **g**, Genome browser allele-specific track showing SPEN binding around Kdm6a, an escaping gene (Cast-Xa in blue, B6-Xi in red, all tracks are scaled identically). **h**, Barplots showing overlap between SPEN and 4 different factors' binding sites at X-linked enhancers and promoters. **i**, Heatmap showing showing normalised SPEN enrichment (log2) at promoters (both replicates are shown) and **j**, gene silencing kinetics (allelic ratio) during XCI within 3 groups of X-linked genes showing different dynamics of SPEN accumulation and loss. **k**, Model for SPEN function in XCI. **a-f, h-j**, show data from two biological replicates.



Extended Data Figure 5. Genome browser allelic tracks of SPEN binding and transcript expression at X-linked genes.

a-n, Top of each pannel: Genome brower allelic tracks of SPEN binding (CUT&RUN) at (**a-g**) silenced genes and (**h-n**), non-silenced genes, during a timecourse of Xist induction in mESCs (Cast-Xa in blue, B6-Xi in red, scaled identically within each pannel). Bottom of each pannel: Allelic tracks of transcript expression (RNAseq) at 0h and 24h of Xist induction in mESCs (Cast-Xa in light grey, B6-Xi in black, scaled identically within each pannel). Relative position of each gene along the X chromosome is shown at the top of the figure.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Katia Ancelin for her support throughout the project and significant help with *in vivo* experiments. We are grateful to Joan Barau, Daniel Holoch and Raphaël Margueron for support and help with the project, Mathieu Carrara for critical reading of the manuscript, and members of the Heard Lab for insightful discussions. We thank Elphège Nora for sharing the *OsTIR1* and AID targetting plasmids, Tasuku Honjo for sharing the *Spen*^{flox} mouse line, Osamu Masui for sharing the *Xist-Bgl*-stem loop targeting constructs and Luke Lavis for sharing Halo-JF646 and JF549 with us. We also thank the imaging—Aurélien Dauphin and PICT-IBiSA (UMR3215/U934), protein purification and sequencing platforms of Institut Curie as well as Laura Villacorta, Jan Provaznik and Vladimir Benes of GeneCore at EMBL.

Europe PMC Funders Author Manuscripts

This work was funded by a Boehringer Ingelheim doctoral Fellowship (20017-2019 to F.Dossin), an ERC advanced investigator award (ERC-ADG-2014 671027 to E.H.), Labellisation La Ligue (to E.H.), ANR (DoseX 2017: ANR-17-CE12-0029, Labex DEEP: ANR-11- LBX-0044, ABS4NGS: ANR-11-BINF-0001, and part of the IDEX PSL: ANR-10-IDEX-0001-02 PSL to E.H.), a Sir Henry Wellcome Postdoctoral Fellowship (201369/Z/16/Z to J.J.Z.), "Région Ile-de-France" and Fondation pour la Recherche Médicale grants (to D.L.), EMBO long-term fellowship (ALTF 549-2014 to I.P., ALTF 301-2015 to T.C.), and Fondation pour la Recherche Medicale (SPF 20140129387 to I.P.) and NIH (grant HG003143 to J.D.). J.D. is an investigator of the Howard Hughes Medical Institute.

References

1. Minajigi A, et al. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. Science (80-.). 2015; 349:aab2276–aab2276.
2. McHugh CA, et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature. 2015; 521:232–236. [PubMed: 25915022]
3. Chu C, et al. Systematic Discovery of Xist RNA Binding Proteins. Cell. 2015; 161:404–416. [PubMed: 25843628]
4. Monfort A, et al. Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells. Cell Rep. 2015; 12:554–561. [PubMed: 26190100]
5. Moindrot B, et al. A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. Cell Rep. 2015; 12:562–72. [PubMed: 26190105]
6. Nesterova TB, et al. Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. Nat Commun. 2019; 10
7. Nishimura K, Fukagawa T, Takisawa H, Kakimoto T, Kanemaki M. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat Methods. 2009; 6:917–922. [PubMed: 19915560]
8. Schulz EG, 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. 2014; 14:203–216. [PubMed: 24506884]
9. Yabe D, et al. Generation of a conditional knockout allele for mammalian Spen protein Mint/SHARP. genesis. 2007; 45:300–306. [PubMed: 17457934]
10. Borensztein M, et al. Xist-dependent imprinted X inactivation and the early developmental consequences of its failure. Nat Struct Mol Biol. 2017; doi: 10.1038/nsmb.3365
11. Grimm JB, et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. Nat Methods. 2015; 12:244–250. [PubMed: 25599551]
12. Masui O, Heard E, Koseki H. Live Imaging of Xist RNA. Methods in molecular biology (Clifton, N.J.). 2018; 1861:67–72.
13. Giorgetti L, et al. Structural organization of the inactive X chromosome in the mouse. Nature. 2016; 535:575–579. [PubMed: 27437574]
14. Deng X, et al. Bipartite structure of the inactive mouse X chromosome. Genome Biol. 2015; 16:152. [PubMed: 26248554]
15. Rao SSP, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014; 159:1665–80. [PubMed: 25497547]
16. Lu Z, et al. RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. Cell. 2016; 165:1267–1279. [PubMed: 27180905]
17. Wutz A, Rasmussen TP, Jaenisch R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat Genet. 2002; 30:167–174. [PubMed: 11780141]
18. Shi Y, et al. Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. Genes Dev. 2001; 15:1140–51. [PubMed: 11331609]
19. Oswald F, et al. RBP-Jkappa/SHARP recruits CtIP/CtBP corepressors to silence Notch target genes. Mol Cell Biol. 2005; 25:10379–90. [PubMed: 16287852]
20. Ha N, et al. Live-Cell Imaging and Functional Dissection of Xist RNA Reveal Mechanisms of X Chromosome Inactivation and Reactivation. iScience. 2018; 8:1–14. [PubMed: 30266032]

21. Ariyoshi M, Schwabe JWR. A conserved structural motif reveals the essential transcriptional repression function of Spen proteins and their role in developmental signaling. Genes Dev. 2003; 17:1909–20. [PubMed: 12897056]
22. Oswald F, et al. A phospho-dependent mechanism involving NCoR and KMT2D controls a permissive chromatin state at Notch target genes. Nucleic Acids Res. 2016; 44:4703–4720. [PubMed: 26912830]
23. Guenther MG, Barak O, Lazar MA. The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. Mol Cell Biol. 2001; 21:6091–101. [PubMed: 11509652]
24. Żylicz JJ, et al. The Implication of Early Chromatin Changes in X Chromosome Inactivation. Cell. 2019; 176:182–197.e23. [PubMed: 30595450]
25. Patil DP, et al. m6A RNA methylation promotes XIST-mediated transcriptional repression. Nature. 2016; 537:369–373. [PubMed: 27602518]
26. Bornelöv S, et al. The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. Mol Cell. 2018; 71:56–72.e4. [PubMed: 30008319]
27. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife. 2017; 6
28. Engreitz JM, et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science. 2013; 341
29. Knuckles P, et al. Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m ^{6} A machinery component Wtap/Fl(2)d. Genes Dev. 2018; 32:415–429. [PubMed: 29535189]
30. Hatchell EC, et al. SLIRP, a Small SRA Binding Protein, Is a Nuclear Receptor Corepressor. Mol Cell. 2006; 22:657–668. [PubMed: 16762838]
31. Giorgetti L, et al. Structural organization of the inactive X chromosome in the mouse. Nature. 2016; 535:575–579. [PubMed: 27437574]
32. Grimm JB, et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. Nat Methods. 2015; 12:244–250. [PubMed: 25599551]
33. Yabe D, et al. Generation of a conditional knockout allele for mammalian Spen protein Mint/SHARP. genesis. 2007; 45:300–306. [PubMed: 17457934]
34. de Vries WN, et al. Expression of Cre recombinase in mouse oocytes: a means to study maternal effect genes. Genesis. 2000; 26:110–2. [PubMed: 10686600]
35. Zylicz JJ, et al. G9a regulates temporal preimplantation developmental program and lineage segregation in blastocyst. Elife. 2018; 7
36. Tang F, et al. RNA-Seq analysis to capture the transcriptome landscape of a single cell. Nat Protoc. 2010; 5:516–35. [PubMed: 20203668]
37. Huang Y, et al. Stella modulates transcriptional and endogenous retrovirus programs during maternal-to-zygotic transition. Elife. 2017; 6
38. Belaghzal H, Dekker J, Gibcus JH. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. Methods. 2017; 123:56–65. [PubMed: 28435001]
39. Chen J, et al. High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. Proc Natl Acad Sci U S A. 2009; 106:13535–40. [PubMed: 19628694]
40. Masui O, Heard E, Koseki H. Live Imaging of Xist RNA. Methods in molecular biology (Clifton, N.J.). 2018; 1861:67–72.
41. Barau J, et al. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science (80-.). 2016; 354:909–912.
42. Poullet P, Carpentier S, Barillot E. myProMS, a web server for management and validation of mass spectrometry-based proteomic data. Proteomics. 2007; 7:2553–6. [PubMed: 17610305]
43. Valot B, Langella O, Nano E, Zivy M. MassChroQ: A versatile tool for mass spectrometry quantification. Proteomics. 2011; 11:3572–3577. [PubMed: 21751374]

44. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife. 2017; 6

45. Żylicz JJ, et al. The Implication of Early Chromatin Changes in X Chromosome Inactivation. Cell. 2019; 176:182–197.e23. [PubMed: 30595450]

46. Borensztein M, et al. Xist-dependent imprinted X inactivation and the early developmental consequences of its failure. Nat Struct Mol Biol. 2017; doi: 10.1038/nsmb.3365

47. Bornelöv S, et al. The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. Mol Cell. 2018; 71:56–72.e4. [PubMed: 30008319]

Europe PMC Funders Author Manuscripts

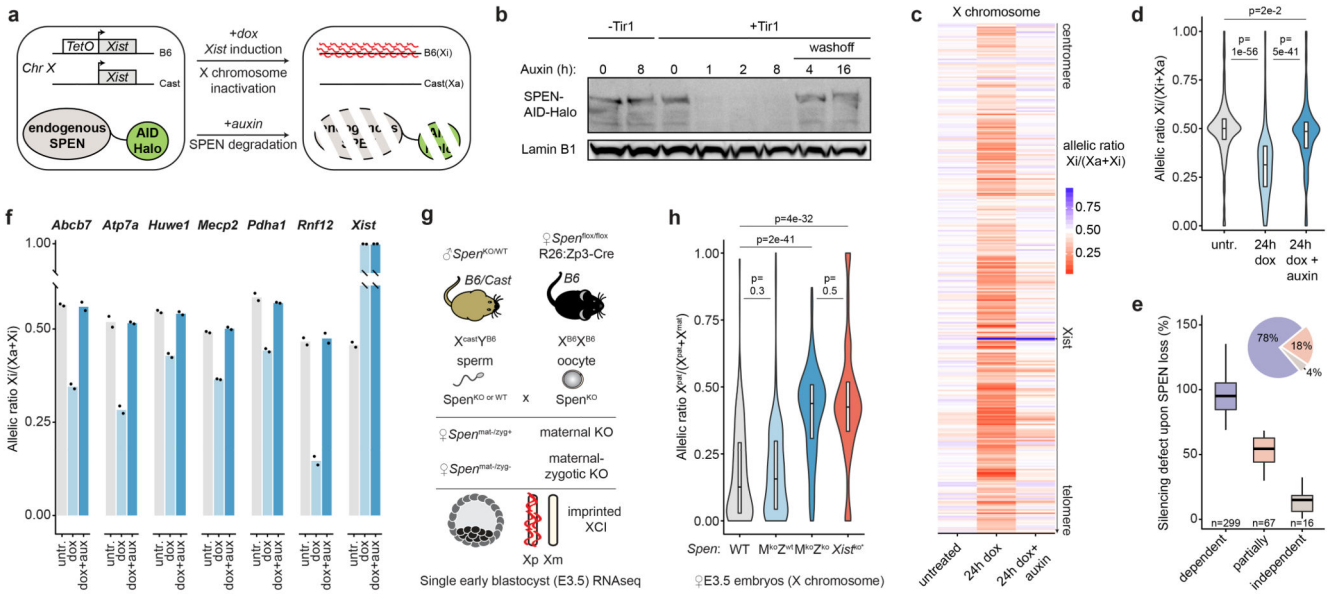


Figure 1. SPEN mediates gene silencing across the entire X chromosome in vitro and in vivo.

a, Schematic of SPEN-degron Xist-inducible mESCs. **b**, Western blot showing auxin-induced degradation of endogenous Halo-tagged SPEN. This experiment was repeated at least twice with similar results. **c**, Heatmap and **d**, violin plots showing X-chromosomal transcript allelic ratios after 0h, 24h dox or 24h dox+auxin treatment in SPEN-degron mESCs (n=434 genes, two-sided Student's t-test). **e**, Boxplot representation of gene silencing defect upon SPEN loss in three groups of genes differing by their SPEN-dependence level for *Xist*-mediated silencing. Piechart shows relative number of genes. **f**, Pyrosequencing assay of 7 X-linked transcripts in mESCs after 0h, 24h dox or 24h dox +auxin. **c, d, e**, and **f**, show averages of 2 independent clones. **g**, Mouse cross schematic for *Spen* KO experiment. **h**, X-chromosomal transcript allelic ratio distribution (n=256 genes) in WT (N=2), maternal-only *Spen* ko (N=3), maternal-zygotic *Spen* ko (N=5), and *Xist* ko E3.5 embryos (N=30 single-cells, *see Borensztein et al., two-sided Wilcoxon rank-sum test). **d, e, h**, horizontal lines denote the median, box limits correspond to upper and lower quartiles.

Europe PMC Funders Author Manuscripts

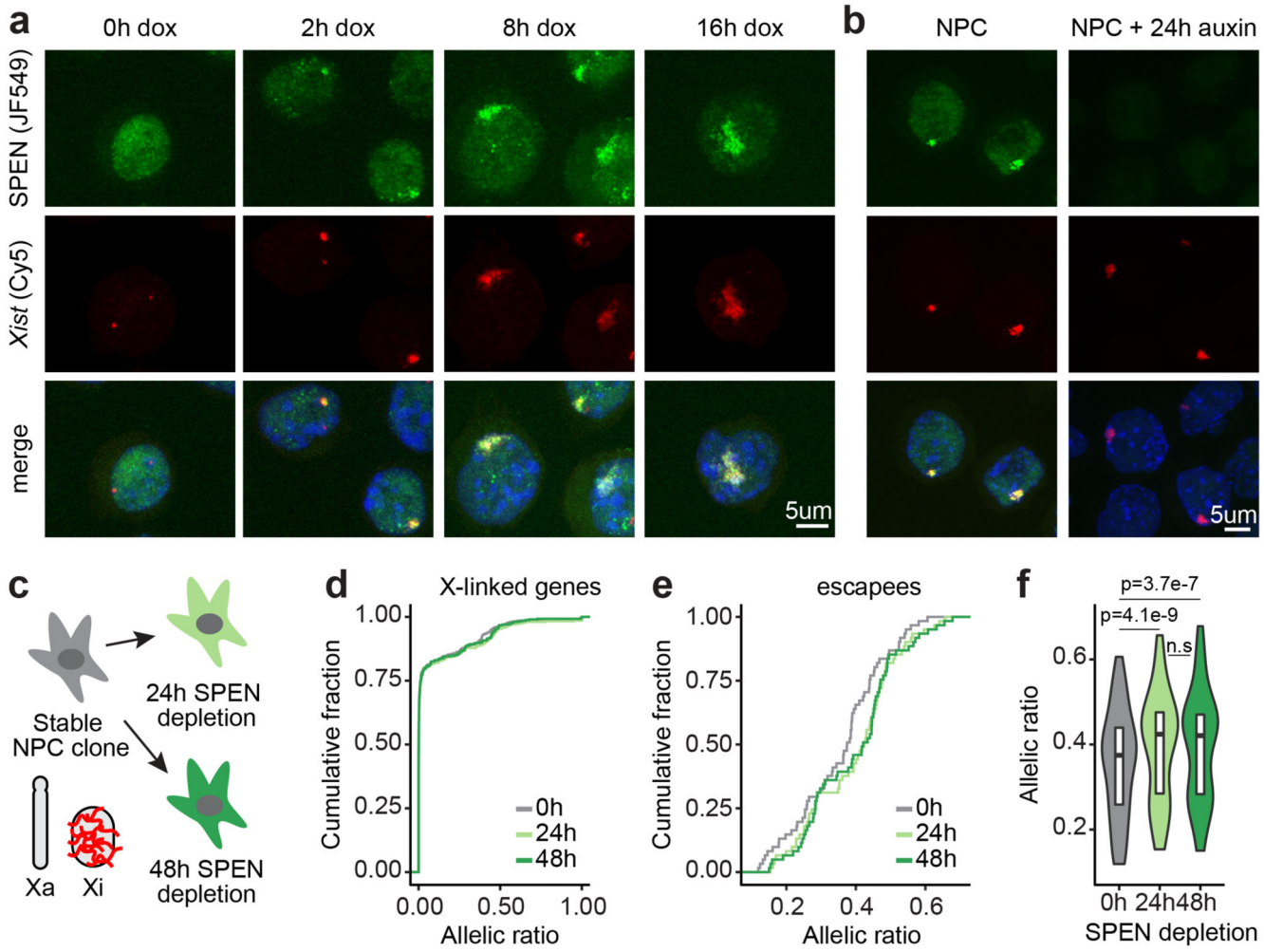


Figure 2. SPEN localizes to the X chromosome immediately upon Xist upregulation and throughout the stages of XCI, but is dispensable for maintenance of X-linked gene silencing.

a, Images from combined HaloTag labeling of SPEN (green) and FISH for Xist RNA (red) in mESCs during a time course of *Xist* induction, and **b**, in NPCs. **c**, Schematic of the SPEN-degron experiment in NPCs. **d**, Cumulative distribution of transcript allelic ratios across the X chromosome (n=387 genes) after SPEN depletion in NPCs. **e**, Cumulative distribution and **f**, violin plot representation of transcript allelic ratio of escapees after SPEN depletion in NPCs (two-sided Wilcoxon signed-rank test, n=65, horizontal lines denote the median, box limits correspond to upper and lower quartiles). **d, e** and **f**, show averages of two independent NPC clones. **a, b**, These experiments were repeated at least twice with similar results.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

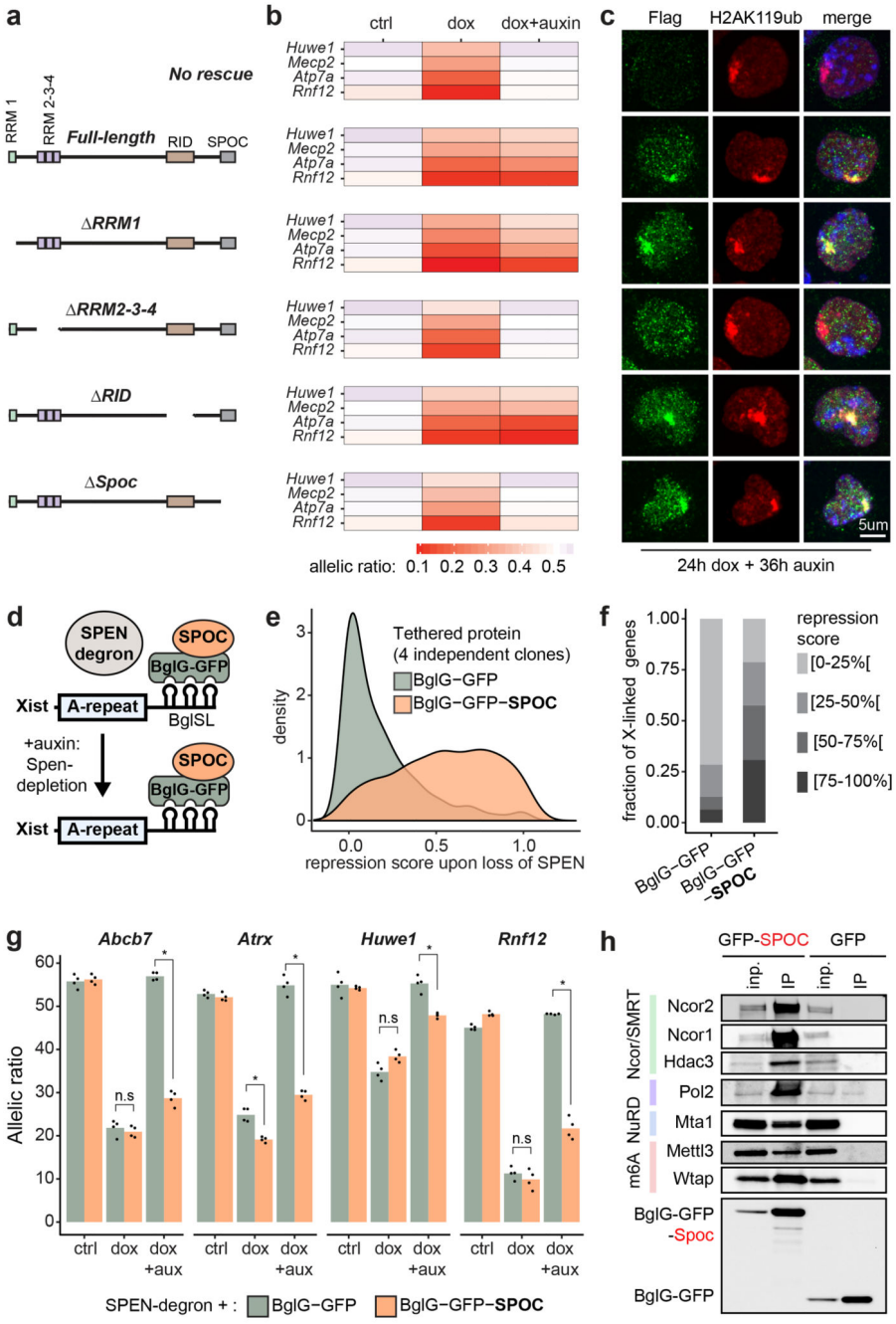


Figure 3. The SPOC domain of SPEN mediates gene silencing and interacts with multiple molecular pathways.

a, SPEN cDNA fragments used for the rescue experiment. **b**, Heatmap representation of 4 X-linked transcript allelic ratios (pyrosequencing) in control, 24h dox and 24hdox+36h auxin SPEN-degron mESCs overexpressing each cDNA construct. Data represents averages of 2-3 independent clones. **c**, Immunofluorescence detection of Flag-tagged SPEN truncations (green) and H2AK119ub1 (red), a marker of the Xi, in SPEN-degron mESCs treated with dox and auxin.

d, Scheme of BglG-SPOC tethering to *Xist*. **e**, Distribution of gene repression scores observed across the X-chromosome upon depletion of endogenous SPEN and tethering of BglG-GFP (green) or BglG-GFP-SPOC (orange) to *Xist*. **f**, Barplots showing the fraction of X-linked genes comprised within four windows of repression score and **g**, transcript allelic ratio (pyrosequencing) for 4 X-linked genes upon depletion of endogenous SPEN and tethering of BglG-GFP or BglG-GFP-SPOC to *Xist* (*P<.01, two-sided Student's t-test). **h**, Western blot detection of co-immunoprecipitated proteins in BglG-GFP and BglG-GFP-SPOC immunoprecipitation experiments. 1% of the input were loaded (0.1% for RNAPII), and 10% of the pulldown. **c, h**, These experiments were repeated at least twice with similar results. **e, f, g** show average data for 4 independent clones.

Europe PMC Funders Author Manuscripts

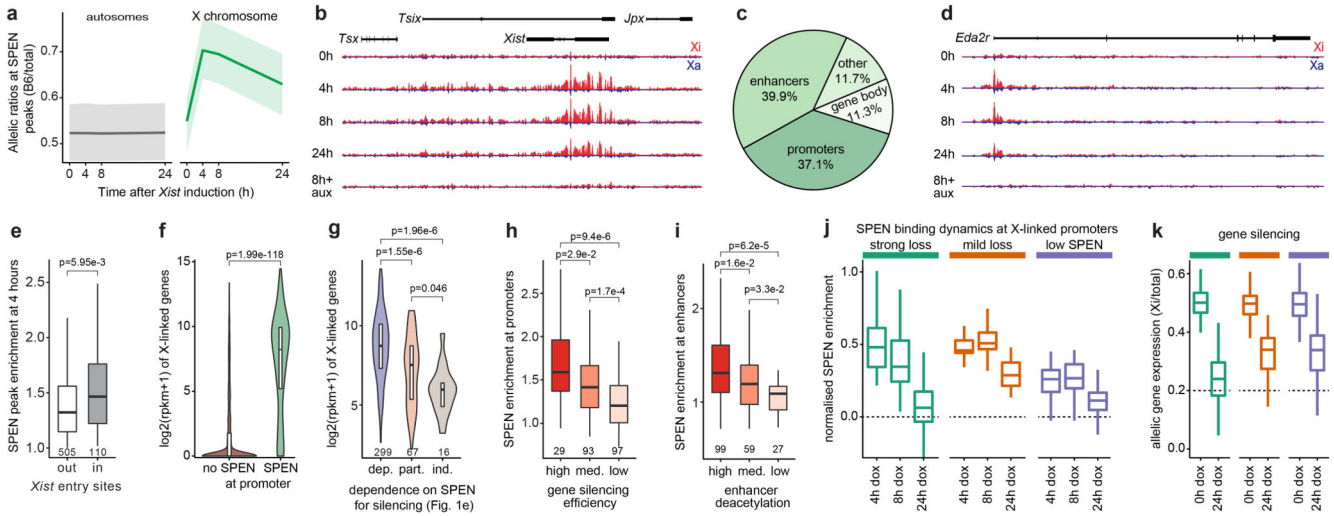


Figure 4. SPEN is recruited by Xist to active gene promoters and enhancers where it silences transcription and subsequently disengages from chromatin.

a, SPEN allele-specific accumulation (CUT&RUN) on peaks at autosomes (grey, n=948) and on the X chromosome (green, n=635) after 0h, 4h, 8h and 24h of Xist induction in mESCs. Shown are average allelic-ratios (shading is the interquartile range) of all peaks. **b**, Genome browser allele-specific track showing SPEN binding around *Xist*. **c**, Annotation of SPEN peaks on the X-chromosome. **d**, Genome browser allele-specific track showing SPEN binding around Eda2r, an X-linked gene. **b, d**, blue: Cast-Xa, red: B6-Xi, tracks are scaled identically. **e**, Boxplot showing SPEN enrichment at 4 hours in peaks outside or within *Xist* entry sites. **f**, Violin plot showing gene expression (RPKM) of genes accumulating SPEN (n=289) or not (n=2325) at their promoters. **g**, Violin plots showing gene expression levels (RPKM in control conditions) of genes grouped based on their SPEN-dependence level for gene silencing (see Fig. 1e). **h**, Boxplots showing SPEN enrichment after 4 hours of Xist induction within peaks at promoters grouped based on how efficiently their respective genes are silenced or **i**, at enhancers grouped based on how efficiently they are deacetylated during XCI. **j**, boxplots showing normalised SPEN enrichment at promoters and **k**, gene silencing (transcript allelic ratio) during XCI within 3 groups of X-linked genes showing different dynamics of SPEN accumulation and loss (n=86 strong loss, n=92 mild loss, n=39 low SPEN). **e-i**, two-sided Wilcoxon rank-sum test. **e-k**, Horizontal lines denote the median, box limits correspond to upper and lower quartiles.

Europe PMC Funders Author Manuscripts