The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3

Published in final edited form as:

Nature. 2015 May 14; 521(7551): 232–236. doi:10.1038/nature14443.

The Xist lncRNA directly interacts with SHARP to silence transcription through HDAC3

Colleen A. McHugh[1,*], **Chun-Kan Chen**[1,*], **Amy Chow**[1], **Christine F. Surka**[1], **Christina Tran**[1], **Patrick McDonel**[2], **Amy Pandya-Jones**[3], **Mario Blanco**[1], **Christina Burghard**[1], **Annie Moradian**[4], **Michael J. Sweredoski**[4], **Alexander A. Shishkin**[1], **Julia Su**[1], **Eric S. Lander**[2], **Sonja Hess**[4], **Kathrin Plath**[3], and **Mitchell Guttman**[1,†]

[1]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

[2]Broad Institute of MIT and Harvard, Cambridge, MA 02139

[3]Department of Biological Chemistry, Jonsson Comprehensive Cancer Center, Molecular Biology Institute, and Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095

[4]Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, Pasadena, CA 91125

Abstract

Many long non-coding RNAs (lncRNAs) affect gene expression[1], but the mechanisms by which they act are still largely unknown[2]. One of the best-studied lncRNAs is Xist, which is required for transcriptional silencing of one X-chromosome during development in female mammals[3,4]. Despite extensive efforts to define the mechanism of Xist-mediated transcriptional silencing, we still do not know any proteins required for this role[3]. The main challenge is that there are currently no methods to comprehensively define the proteins that directly interact with a lncRNA in the cell[5]. Here we develop a method to purify a lncRNA and identify its direct interacting proteins using quantitative mass spectrometry. We identify 10 proteins that specifically associate with Xist, three of these proteins – SHARP, SAF-A, and LBR – are required for Xist-mediated transcriptional silencing. We show that SHARP, which interacts with the SMRT co-repressor[6] that activates HDAC3[7], is not only essential for silencing, but is also required for the exclusion of RNA Polymerase II (PolII) from the inactive X. Both SMRT and HDAC3 are also required for

†Correspondence and requests for materials should be addressed to M.G. (mguttman@caltech.edu).
*These authors contributed equally to this work

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

AUTHOR CONTRIBUTIONS

CAM developed the RAP-MS method, designed, performed, and analysed RAP-MS experiments and data, CKC designed, performed, and analysed Xist functional experiments, AC designed, performed, and oversaw experiments, CFS helped develop RAP-MS and performed experiments, CT, PM, AP-J, AM, AAS, JS performed experiments, MJS, MB, CB analysed data, ESL helped develop initial ideas for adapting RAP for protein detection, SH oversaw mass spectrometry development and data analysis, KP helped design Xist RAP-MS and functional experiments and analysed data, MG conceived, designed, and oversaw the entire project and integrated the data, CAM, CKC, and MG wrote the manuscript with input from all authors.

The authors declare no competing financial interests.

silencing and PolII exclusion. In addition to silencing transcription, SHARP and HDAC3 are required for Xist-mediated recruitment of the polycomb repressive complex 2 (PRC2) across the X-chromosome. Our results suggest that Xist silences transcription by directly interacting with SHARP, recruiting SMRT, activating HDAC3, and deacetylating histones to exclude PolII across the X-chromosome.

Over the last two decades, numerous attempts have been made to define the protein complexes that interact with Xist and that are required for its various roles in XCI[3]. Most studies have used prior knowledge of the molecular events that occur on the X-chromosome to define potential Xist-interacting proteins[8,9]. While individual proteins have been identified that associate with Xist[8,10], we still do not know any of the proteins required for Xist-mediated transcriptional silencing because perturbations of these proteins, including components of the PRC2 complex, have no impact on Xist-mediated transcriptional silencing[11,12]. Current methods for identifying lncRNA-interacting proteins either require selecting specific candidate interacting proteins or fail to distinguish between direct RNA interactions that occur in the cell from those that merely associate in solution (reviewed in[5]).

To develop a method for identifying the proteins that directly interact with a specific lncRNA *in vivo*, we adapted our RNA Antisense Purification (RAP) method[13] to purify a lncRNA complex and identify the interacting proteins by quantitative mass spectrometry (RAP-MS) (**Methods,** Figure 1a). Briefly, RAP-MS uses UV crosslinking to create covalent linkages between directly interacting RNA and protein and purifies lncRNAs in denaturing conditions to disrupt non-covalent interactions (**Methods**). This UV-crosslinking and denaturing approach, which is utilized by methods such as CLIP, is known to identify only direct RNA-protein interactions and to separate interactions that are crosslinked in the cell from those that merely associate in solution[5,14].

Adapting this UV-crosslinking and denaturing approach to enable purification of a specific lncRNA is challenging for several reasons: **(i)** In order to purify lncRNA complexes in denaturing conditions, we need an RNA capture method that can withstand harsh denaturing conditions. **(ii)** In order to detect the proteins associated with a given lncRNA, we need to achieve high purification yields of a lncRNA complex because, unlike nucleic acids, we cannot amplify proteins prior to detection. **(iii)** Because any individual RNA is likely to be present at a very low percentage of the total cellular RNA, we need to achieve high levels of enrichment to identify specific interacting proteins. **(iv)** Because the number of background proteins will be high, even after enrichment, we need accurate and sensitive methods for protein quantification to detect specific lncRNA interacting proteins.

The RAP-MS method addresses these challenges because: **(i)** RAP uses long biotinylated antisense probes, which form very stable RNA-DNA hybrids, and therefore can be used to purify lncRNA complexes in denaturing and reducing conditions (ie. 4M urea at 67°C, **Methods**). **(ii)** We optimized the RAP method to achieve high yields of endogenous RNA complexes. In our original protocol[13], we achieved <2% yield of the endogenous RNA complex; by optimizing hybridization, washing, and elution conditions (**Methods**), we were able to reproducibly achieve ~70% yield (Extended Data 1a, **Methods**). **(iii)** Using our

optimized conditions, we increased the enrichment levels for the target lncRNA complex (~5,000-fold, Extended Data 1b) relative to our already high levels of enrichment achieved previously (~100-fold)[13]. **(iv)** To achieve sensitive quantification and to distinguish between specific proteins and background proteins, we used Stable Isotope Labeling by Amino acids in Culture (SILAC) to label proteins (**Methods,** Extended Data 1c), which enables quantitative comparisons of purified proteins by mass spectrometry[15].

We validated the RAP-MS approach by defining the proteins that interact with two well-characterized non-coding RNAs: U1 (a core component of the spliceosome) and 18S (a component of the small ribosomal subunit). In the U1 purifications, we identified 9 enriched proteins, all of which are known to interact with U1 (Supplementary Note 1). In the 18S purification, we identified 105 enriched proteins, 98 of these (93%) were previously characterized as ribosomal proteins, ribosomal processing and assembly factors, translational regulators, or other known ribosome interactors (Extended Data 2). In particular, we identified 21 of the 31 known small ribosomal subunit proteins. The few missing proteins appear to fall predominately into two categories: (i) proteins that make few direct contacts with the RNA and (ii) small proteins that contain few peptides that could be detected by mass spectrometry.

These results demonstrate that the RAP-MS method identifies the majority of known RNA interacting proteins, and that the proteins identified by RAP-MS are highly specific for the purified ncRNA complex.

To define the proteins that interact with Xist during the initiation of XCI, we UV-crosslinked SILAC-labeled mouse embryonic stem (ES) cells after Xist induction[13] and purified Xist in nuclear extracts (**Methods**). To control for background proteins or non-specific proteins that might interact with any nuclear RNA, we separately purified the abundant U1 snRNA, which is not expected to interact with the same proteins as Xist. We identified the proteins in each sample using liquid chromatography-mass spectrometry and calculated a SILAC ratio for each protein based on the intensity of all heavy or light peptides originating from the Xist or U1 purification (Figure 1a, **Methods**).

We identified 10 proteins that were enriched for Xist relative to U1 (SILAC ratio >3-fold, Figure 1b). All 10 proteins were reproducibly enriched in multiple Xist purifications from independent biological samples (**Methods**). Consistent with the notion that these proteins are direct Xist-interacting proteins, 9 proteins contain well-characterized RNA binding domains (Figure 1c).

The identified Xist-interacting proteins are SHARP, Rbm15, Myef2, Celf1, hnRNPC, LBR, SAF-A, Raly, hnRNPM, and Ptbp1 (Figure 1b). SAF-A (Scaffold Attachment Factor-A, also called hnRNPU) was previously shown to interact directly with Xist and is required for tethering Xist to the inactive X-chromosome in differentiated cells[10]. In addition, 5 of these proteins have been previously implicated in transcriptional repression, chromatin regulation, and nuclear organization. These include SHARP (SMRT and HDAC Associated Repressor Protein, also called SPEN), a member of the SPEN family of transcriptional repressors, which directly interacts with the SMRT component (also called NCoR-2) of the nuclear co-

repressor complex[16] that is known to interact with and activate HDAC3 deacetylation activity on chromatin[7] (Figure 1c). Interestingly, we also identified RBM15, another member of the SPEN family of transcriptional repressors, which shares the same domain structure as SHARP, but appears to have a distinct functional role during development[17]. Myef2 has been shown to function as a negative regulator of transcription in multiple cell types, although its mechanism of regulation is still unknown[18]. hnRNPM is a paralog of Myef2. Finally, we identified LBR (Lamin B receptor), a protein that is anchored in the inner nuclear membrane and interacts with repressive chromatin regulatory proteins and Lamin B[19] (Figure 1c).

We confirmed the specificity of the identified Xist-interacting proteins (Supplementary Note 2, **Methods**): (i) To ensure that they are not due to non-specific RNA or protein capture, we performed RAP in uninduced cells (no Xist) and identified no enriched proteins. (ii) To ensure that these proteins are crosslinked with Xist in cells and not merely associating in solution, we performed RAP in cells that were not crosslinked (no UV) and identified no enriched proteins. (iii) To ensure that these proteins do not merely interact with any nuclear-enriched long ncRNA, we compared the Xist-purified proteins to those purified with 45S (pre-ribosomal RNA) and found that all 10 Xist-interacting proteins were still enriched. (iv) To independently validate these interactions, we obtained high-quality affinity reagents for 8 of the 10 proteins (Ptbp1, hnRNPC, Celf1, Myef2, Rbm15, LBR, Raly, and SHARP) and immunoprecipitated the identified proteins in UV-crosslinked lysates. In all cases, we observed a strong enrichment for the Xist RNA (>4-fold), but not control mRNAs or lncRNAs (Extended Data 3, Supplementary Table 1).

Together, these results identify a set of highly specific and reproducible proteins that directly interact with Xist during the initiation of XCI. Given the generality of the RAP-MS approach, we expect that it will be broadly applicable for defining the proteins that directly interact with other lncRNAs.

To determine which proteins are required for Xist-mediated transcriptional silencing, we knocked down each of the identified proteins and assayed for the failure to silence gene expression on the X-chromosome upon induction of Xist expression (Figure 2a).

Specifically, we selected two X-linked genes, Gpc4 and Atrx, that are well expressed in the absence of Xist expression, but are normally silenced by 16 hours of Xist induction in our doxycycline-inducible system in male cells (Figure 2b, **Methods**). We used siRNAs to knockdown the mRNA levels of each of the proteins identified by RAP-MS along with several negative controls (**Methods,** Supplementary Table 2). We ensured that each cell examined showed both successful depletion of the siRNA-targeted mRNA (>70% reduction) as well as induction of Xist expression using single molecule RNA FISH (**Methods**). Within each of these cells, we quantified the mRNA level of each of the two X-linked genes prior to Xist induction (−dox) and after Xist induction (+dox).

As a control, we transfected several non-targeting siRNAs (**Methods**). In these negative controls, we observed the expected silencing of the X-linked genes studied (Gpc4 expression decreased from an average of 20 copies (−dox) to 2 copies (+dox) per cell and Atrx

expression decreased from 22 to 3 copies per cell; Figure 2b,c). Consistent with previous observations, we found no effect on X-chromosome gene silencing upon knockdown of EED[42,43], a required component of PRC2[20] (Figure 2b), or other proteins previously associated with Xist that do not appear to be required for transcriptional silencing (Extended Data 4, **Methods**). Similarly, knockdown of Rbm15, Myef2, Ptbp1, Celf1, hnRNPC, Raly, or hnRNPM did not alter gene silencing on the X-chromosome (Figure 2b, Extended Data 5).

In contrast, knockdown of SHARP, LBR, or SAF-A largely abolished the silencing of X-chromosome genes following Xist induction (Figure 2b,c, Supplementary Note 3, Extended Data 5, Extended Data 6). Indeed, the expression levels of the X-chromosome genes studied did not significantly change following Xist expression (Figure 2c, Extended Data 5). These same silencing defects were observed with several independent siRNAs (Extended Data 7). Importantly, we observed the same X-chromosome silencing defects upon knockdown of SHARP, LBR, or SAF-A in differentiating female ES cells (Extended Data 8, **Methods**).

These results demonstrate that SHARP, LBR, and SAF-A are required for Xist-mediated transcriptional silencing of the X-chromosome. Although the remaining seven Xist-interacting proteins showed no effect on X-chromosome gene silencing, they may still be important for Xist function: (i) some may have redundant functions (e.g. Myef2 and hnRNPM, which are known paralogs), (ii) in some of these cases, the small amount of protein remaining after knockdown may still be sufficient for Xist function, or (iii) some of these proteins may be important for alternative Xist-mediated roles, such as the maintenance of XCI, which would not be captured by this silencing assay.

Xist initiates XCI by spreading across the future inactive X-chromosome, excluding RNA Polymerase II (PolII), and repositioning active genes into a transcriptionally silenced nuclear compartment[3,10,13,21]. All of these roles – localization, RNA PolII exclusion, and repositioning – are required for proper silencing of transcription during the initiation of XCI[3].

Consistent with previous observations that SAF-A is required for Xist localization to chromatin in differentiated cells[10], we observed a diffuse Xist localization pattern in the nucleus upon knock down of SAF-A (Extended Data 5). This suggests that SAF-A is required for transcriptional silencing by localizing Xist, and its silencing proteins, to the X-chromosome during the initiation of XCI.

To determine the proteins responsible for establishing the initial silenced compartment on the X-chromosome, we explored whether SHARP or LBR are required for the exclusion of PolII from the Xist-coated region. Specifically, we measured the co-localization of Xist and PolII in single cells (**Methods**). In wild-type cells after 16 hours of Xist induction, we observed a depletion of PolII over the Xist-coated territory (Figure 3a, **Methods**). We observed a similar exclusion of PolII from the Xist-coated region in the negative controls and upon knockdown of EED or LBR (Figure 3a,b). In contrast, upon knockdown of SHARP, we observed higher levels of PolII over the Xist-coated territory relative to the control samples (Figure 3b). We confirmed that SHARP, but not LBR or EED, is similarly

required for PolII exclusion in differentiating female ES cells (Figure 3c, Extended Data 9, Supplementary Note 4).

These results demonstrate that SHARP is required to exclude PolII on the inactive X-chromosome and may be required for creating the initial silenced compartment upon Xist localization[3]. While LBR is not required for PolII exclusion, it is likely to play an alternative role during the initiation of Xist-mediated transcriptional silencing, such as repositioning genes into this PolII-excluded compartment[13,21].

Having identified SHARP as the direct Xist-interacting protein that is required for excluding PolII on the X-chromosome, we sought to determine how it might carry out this role. SHARP is a direct RNA binding protein[6,22] that was first identified in mammals based on its interaction with the SMRT co-repressor complex[6], which is known to interact with HDAC3 and is required for activating its deacetylation and transcriptional silencing activity *in vivo*[7]. Based on these previous observations, we hypothesized that Xist-mediated transcriptional silencing through SHARP would occur through SMRT and the silencing function of HDAC3. (We would not expect to identify these proteins by RAP-MS, which was designed to identify only direct RNA-protein interactions.)

To test this hypothesis, we knocked down either SMRT or HDAC3 and measured the expression of X-chromosome genes upon Xist induction. Knockdown of SMRT or HDAC3 in both male and female ES cells abrogated silencing of X-chromosome genes upon induction of Xist expression (Figure 2b, Extended Data 5, 7, and 8). To ensure that the observed silencing defect is specific for HDAC3 and not other class I HDAC proteins, we knocked down HDAC1 or HDAC2 and observed no effect on gene silencing (Extended Data 5). To further confirm the specificity of our results, we used independent siRNAs to knock down SMRT or HDAC3 and in all cases identified a similar silencing defect (Extended Data 7).

To determine whether this effect is similar to the effect produced by knock down of SHARP or a distinct defect in transcriptional silencing, we tested whether HDAC3, the silencing protein in this complex[7,23], is required for the exclusion of RNA PolII from the Xist-coated territory. We found that knock down of HDAC3 in both male and female ES cells eliminated the exclusion of RNA PolII from the Xist-coated compartment to a similar degree to that seen for knock down of SHARP (Figure 3, Extended Data 9).

These results suggest that SHARP silences transcription through SMRT and the HDAC3 silencing protein. This role for HDAC3 in Xist-mediated silencing would explain the long-standing observation of global hypoacetylation on the entire X-chromosome as one of the very first events that occur upon initiation of XCI[3,24].

One of the features of XCI is the recruitment of PRC2 and its associated H3K27me3 repressive chromatin modifications across the X-chromosome in an Xist-dependent manner[3,4,9]. While PRC2 is not required for the initiation of XCI[11,12] (Figure 3b), it or its associated H3K27me3 repressive chromatin modifications may be involved in establishing an epigenetically silenced state[25]. Yet, how Xist recruits the PRC2 complex across the X-chromosome is unknown. Since we failed to identify any PRC2 components by RAP-MS,

and various HDAC complexes are known to recruit PRC2[26], we hypothesized that PRC2-recruitment is mediated by SHARP and HDAC3.

To test this hypothesis, we looked at PRC2 recruitment to the Xist-coated territory. In wild-type cells, we observe a strong enrichment of EZH2, a component of PRC2, over the Xist-coated territory after 16 hours of induction (Figure 4a). Upon knock down of EED, a distinct component of the PRC2 complex that is required for its proper localization to chromatin[20], we observe no enrichment of EZH2 over the Xist cloud at this same time point (Figure 4a). Similarly, upon knock down of SHARP, we identified a loss of EZH2 over the Xist coated territory, of comparable magnitude to that observed in the absence of EED (Figure 4a). Conversely, upon knock down of LBR, we observed a strong enrichment of EZH2 over the Xist coated territory, of comparable magnitude to the levels of recruitment in wild-type conditions (Figure 4b). To determine whether HDAC3 is required for PRC2 recruitment, we knocked down HDAC3 and observed a loss of PRC2 recruitment (Figure 4a), of comparable magnitude to that observed upon loss of SHARP (Figure 4b). Knockdown of SHARP or HDAC3 led to the same PRC2-recruitment defect in female ES cells (Figure 4c, Extended Data 10).

These results argue that Xist-mediated recruitment of PRC2 across the X-chromosome is dependent on SHARP and HDAC3. Whether this occurs through an interaction with SHARP or HDAC3 (direct recruitment) or due to the HDAC3-induced silenced transcription state, chromatin modifications, or compact chromatin structure (indirect recruitment) remains unclear (Supplementary Note 5). Yet, our results are in contrast to a previous model that PRC2 is recruited through a direct interaction between EZH2 and the A-repeat of Xist[8]. The evidence for this PRC2-Xist interaction is based on *in vitro* binding and purifications in non-denaturing conditions[8]. Recently, the specificity of this interaction has been questioned because PRC2 appears to bind promiscuously to many RNAs, including bacterial RNAs, in these conditions[27]. Instead, our results are consistent with reports that deletion of the A-repeat, unlike knockdown of SHARP or HDAC3, has no significant effect on PRC2 recruitment to the Xist-coated territory[9] (Figure 4b).

Taken together, our data suggest a model for how Xist can orchestrate transcriptional silencing on the X-chromosome (Figure 4d). Upon initiation of Xist expression, Xist can localize to sites on the X-chromosome by binding to the SAF-A protein[10], which is known to interact directly with chromatin[28]. Xist directly interacts with SHARP to recruit SMRT[6] to these DNA sites across the inactive X-chromosome. This Xist-SHARP-SMRT complex either directly recruits HDAC3 to the X-chromosome or may act to induce the enzymatic activity of HDAC3[7] that may already be present at active genes across the X-chromosome[29]. Through HDAC3, Xist can direct the removal of activating histone acetylation marks on chromatin thereby compacting chromatin and silencing transcription[30]. Upon initiating the silenced state, Xist recruits PRC2 across the X-chromosome in an HDAC3-dependent manner, either through a direct interaction between PRC2 and HDAC3 or indirectly through HDAC3-induced transcriptional silencing or chromatin compaction (Supplemental Note 5). In this way, the same Xist interacting protein might achieve two essential roles in XCI – initiating the inactive state by recruiting transcriptional silencers (HDAC3) and maintaining the inactive state by recruiting stable epigenetic silencers (PRC2)[25]. Beyond Xist, RAP-MS

provides a critical tool that will accelerate the discovery of novel lncRNA mechanisms that have thus far proved elusive.

METHODS

Mouse ES cell culture

All mouse ES cell lines were cultured in serum-free 2i/LIF medium as previously described[13]. We used the following cell lines: (i) Wild-type male ES cells (*V6.5 line*); (ii) Male ES cells expressing Xist from the endogenous locus under control of a tet-inducible promoter (*pSM33* ES cell line) as previously described[13]. (iii) Male ES cells carrying a cDNA Xist transgene without the A-repeat integrated into the Hprt locus under control of the tet-inducible promoter (*A-repeat deletion*: kindly provided by A. Wutz)[31]. (iv) Female ES cells (*F1 2-1 line*). This wild-type female mouse ES cell line is derived from a 129 × castaneous F1 mouse cross as previously described[13].

Xist induction

For Dox inducible cells (pSM33 and A-repeat deletion), we induced Xist expression by treating cells with 2 μg/ml doxycycline (Sigma) for 6 hours, 16 hours, or 24 hours based on the application. For female ES cells (F1 2-1 line), we induced Xist expression by inducing differentiation; 2i was replaced with MEF media (DMEM, 10% Gemini Benchmark FBS, 1× L-glutamine, 1× NEAA, 1× Pen/Strep; Life Technologies unless otherwise indicated) for 24 hours followed by treatment with 1 μM retinoic acid (RA) (Sigma) for an additional 24 hours.

We measured the amount of Xist RNA in both the doxycycline-inducible cells (6 hours induction) and differentiating female ES cells (24 hour induction) by qRT-PCR. We normalized this level to various RNA housekeeping controls, 18S, 28S, and U6, in both cell populations and calculated the fold expression difference between male and female cells using the comparative Ct method. We observed a range of expression, with the male inducible system expressing from 5–20 fold (12-fold average) more Xist than the female cells. We note that this estimate likely represents an upper limit of the actual differences because the female ES cell system is known to be heterogeneous in Xist-induction, such that not every cell will induce Xist to the same level after 24 hours of retinoic acid treatment. Accordingly, we expect that the actual differences between the male inducible system and differentiating female ES cells are actually significantly lower. While the precise levels are hard to compare by single molecule FISH, the size and intensity of each Xist RNA cloud is similar in both systems at the time points used.

The male-inducible system is more sensitive for identifying proteins that affect silencing compared to a female system because Xist-mediated silencing in males will lead to loss of 100% of X-chromosome transcripts rather than only 50% in a female system, which still retains one active X.

UV crosslinking

Cells were washed once with PBS and then crosslinked on ice using 0.8 Joules/cm^2 (UV8k) of UV at 254 nm in a Spectrolinker UV Crosslinker. Cells were then scraped from culture dishes, washed once with PBS, pelleted by centrifugation at 1500 × g for 4 minutes, and flash frozen in liquid nitrogen for storage at − 80 °C.

SILAC ES cell culture

For SILAC experiments, we adapted our ES cell culture procedures to incorporate either light or heavy lysine and arginine amino acids. The 2i/LIF SILAC medium was composed as follows: custom DMEM/F-12 without lysine or arginine (Dundee Cell Products) was supplemented with either 0.398 mM heavy Arg10 (Sigma) or unlabeled arginine (Sigma) and either 0.798 mM heavy Lys8 (Cambridge Isotope Labs) or unlabeled lysine (Sigma), 0.5× B-27 (Gibco), 2 mg/mL bovine insulin (Sigma), 1.37 μg/mL progesterone (Sigma), 5 mg/mL BSA Fraction V (Gibco), 0.1 mM 2-mercaptoethanol (Sigma), 5 ng/mL murine LIF (GlobalStem), 0.1 μM PD0325901 (SelleckChem) and 0.3 μM CHIR99021 (SelleckChem). Cells in both heavy and light 2i/LIF SILAC medium were also supplemented with 0.2 mg/mL of unlabeled proline (Sigma) to prevent conversion of labeled arginine to proline. 2i inhibitors were added fresh with each medium change.

Adapting cells to SILAC conditions

Prior to mass spectrometry, ES cells were adapted to SILAC conditions over three passages. The heavy or light culture medium was replaced every 24–48 hours depending on cell density, and cells were passaged every 72 hours using 0.025% trypsin (Gibco), rinsing dissociated cells from the plates with DMEM/F12 containing 0.038% BSA Fraction V (Gibco). Cells were grown in two different types of medium: (i) 2i/LIF SILAC medium with light (unlabeled) lysine and arginine, or (ii) 2i/LIF SILAC medium with heavy isotope labeled lysine and arginine.

Measuring SILAC incorporation

To examine the efficiency of SILAC labeling in pSM33 cells, we tested for the incorporation of labeled amino acids after 10 days of growth (3 cell passages) in heavy 2i/LIF SILAC medium. Pellets of 2 million cells were boiled for 10 minutes in LDS Sample Loading Buffer (Invitrogen) and then proteins were separated by SDS-PAGE on a 4–12% Tris-Glycine polyacrylamide gel (Invitrogen). Total protein was stained with Colloidal Coomassie (Invitrogen) and gel slices were excised with a clean scalpel and transferred to microcentrifuge tubes for in-gel tryptic digest. Protein disulfide bonds were reduced with DTT then alkylated with iodoacetamide. Proteins were digested with trypsin overnight and then extracted using successive washes with 1% formic acid/2% acetonitrile, 1:1 acetonitrile/water, and 1% formic acid in acetonitrile. Peptides were collected, lyophilized, then resuspended in 1% formic acid for mass spectrometry analysis (described below in ***Mass Spectrum Measurements***). Peptides were identified from mass spectra using MaxQuant (described below in ***MS data analysis***). The incorporation rate of labeled amino acids was calculated based on the ratio of the intensity of heavy and light versions of each peptide identified. In cells used for subsequent assays, we confirmed that over 95% of

peptides from cellular proteins showed >95% incorporation of labeled amino acids (Extended Data 1b).

RNA Affinity Purification-Mass Spectrometry (RAP-MS)

Probe design and generation—To create the probes used to capture target RNAs, we designed and synthesized 90-mer DNA oligonucleotides (Eurofins Operon) that spanned the entire length of the target RNA. The sequence of each DNA oligonucleotide probes was antisense to the complementary target RNA sequence. Each DNA oligonucleotide probe was also modified with a 5' biotin in order to enable capture of DNA-RNA hybrids on streptavidin coated magnetic beads (described below). While we had previously used 120-mer probes, we found that 90-mer probes provided comparable stringency and yield in the conditions used. For Xist, we used 142 probes that covered the entire mature RNA sequence, with the exception of regions that match to other transcripts or genomic regions as previously described[13,32].

Total cell lysate preparation—For the 18S and U1 experiments we used total cellular lysates prepared in the following manner. We lysed batches of 20 million cells by completely resuspending frozen cell pellets in ice cold detergent-based Cell Lysis Buffer (10 mM Tris pH 7.5, 500 mM LiCl, 0.5% dodecyl maltoside (DDM, Sigma), 0.2% sodium dodecyl sulfate (SDS, Ambion), 0.1% sodium deoxycholate (Sigma)). Next, 1× Protease Inhibitor Cocktail (Set III, EDTA-free, Calbiochem) and 920 U of Murine RNase Inhibitor (New England Biolabs) were added and the sample was incubated for 10 minutes on ice to allow lysis to proceed. During this incubation period, the cell sample was passed 3–5× through a 26-gauge needle attached to a 1 mL syringe in order to disrupt the pellet and shear genomic DNA. Each sample was then sonicated using a Branson Digital Sonifier with a microtip set at 5 watts power for a total of 30 seconds in intermittent pulses (0.7 seconds on, 1.3 seconds off). During sonication the samples were chilled to prevent overheating of the lysate. The samples were then treated for 10 minutes at 37 °C with 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$, and 20 U of TURBO DNase (Ambion) to digest DNA. Samples were returned to ice and the reaction was immediately terminated by the addition of 10 mM EDTA and 5 mM EGTA. Disulfide bonds were reduced by addition of 2.5 mM Tris-(2-carboxyethyl) phosphine (TCEP) and samples were then mixed with twice the lysate volume of 1.5× LiCl/Urea Buffer (the final1 × Buffer contains 10 mM Tris pH 7.5, 5 mM EDTA, 500 mM LiCl, 0.5% DDM, 0.2% SDS, 0.1% deoxycholate, 4M urea, 2.5 mM TCEP). Lysates were incubated on ice for 10 minutes then cleared by centrifugation in an Eppendorf 5424R centrifuge for 10 minutes at 16,000 × *g*. Supernatants were pooled and flash frozen in liquid nitrogen for storage at −80 °C.

Nuclear lysate preparation—For the Xist versus U1 and 45S versus U1 comparisons, we used nuclear lysates prepared in the following manner. We lysed batches of 50 million cells by resuspending frozen pellets in 1 mL Lysis Buffer 1 (10 mM HEPES pH7.2, 20 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM EDTA, 1 mM Tris(2-carboxyethyl)phosphine (TCEP), 0.5 mM PMSF). Then the samples were centrifuged at 3,300 × *g* for 10 minutes to pellet cells. The cell pellets were resuspended in 1 mL Lysis Buffer 1 with 0.1% dodecyl maltoside (DDM) and dounced 20 times using a glass dounce homogenizer with the small clearance

pestle (Kontes). Nuclei released from the cells after douncing were pelleted by centrifugation at 3,300 × *g* then resuspended in 550 μl Lysis Buffer 2 (20 mM Tris pH 7.5, 50 mM KCl, 1.5 mM $MgCl_2$, 2 mM TCEP, 0.5 mM PMSF, 0.4% sodium deoxycholate, 1% DDM, and 0.1% N-lauroylsarcosine (NLS)). Samples were incubated on ice for 10 minutes, then each sample was sonicated using a Branson Sonifier at 5 watts power for a total of 1 minute in intermittent pulses (0.7 seconds on, 3.3 seconds off) to lyse nuclei and solubilize chromatin. During sonication the samples were chilled to prevent overheating of the nuclear lysate. Samples were then treated with 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$, and 330 U TURBO DNase (Ambion) for 12 minutes at 37 °C to further solubilize chromatin. After DNase treatment, lysates were mixed with equal volume of 2× Hybridization Buffer (the final 1× Buffer contains 10 mM Tris pH 7.5, 5 mM EDTA, 500 mM LiCl, 0.5% DDM, 0.2% SDS, 0.1% deoxycholate, 4M urea, 2.5 mM TCEP). Finally, lysates were cleared by centrifugation for 10 minutes at 16,000 × *g* in an Eppendorf 5424R centrifuge and the resulting supernatants were pooled and flash frozen in liquid nitrogen for storage at −80 °C.

RNA affinity purification of crosslinked complexes—Lysates from 200 million or 800 million cells were used for each capture. For 200 million cells the following protocol was used, and scaled appropriately for larger cell numbers. For each capture, a sample of heavy or light SILAC labeled frozen lysate was warmed to 37 °C. For each sample, 1.2 mL of Streptavidin Dynabeads MyOne C1 magnetic beads (Invitrogen) were washed 6 times with equal volume of hybridization buffer (10 mM Tris pH 7.5, 5 mM EDTA, 500 mM LiCl, 0.5% DDM, 0.2% SDS, 0.1% deoxycholate, 4M urea, 2.5 mM TCEP). Lysate samples were pre-cleared by incubation with the washed Streptavidin C1 magnetic beads at 37 °C for 30 minutes with intermittent shaking at 1100 rpm on a Eppendorf Thermomixer C (30 seconds mixing, 30 seconds off). Streptavidin beads were then magnetically separated from lysate samples using a Dynamag magnet (Life Technologies). The beads used for pre-clearing lysate were discarded and the lysate sample was transferred to fresh tubes twice to remove all traces of magnetic beads. Biotinylated 90-mer DNA oligonucleotide probes specific for the RNA target of interest (20 μg per sample, in water) were heat-denatured at 85 °C for 3 minutes and then snap-cooled on ice. Probes and pre-cleared lysate were mixed and incubated at 67 °C using an Eppendorf thermomixer with intermittent shaking (30 seconds shaking, 30 seconds off) for 2 hours to hybridize probes to the capture target RNA. Hybrids of biotinylated DNA probes and target RNA were then bound to streptavidin beads by incubating each sample with 1.2 mL of washed Streptavidin coated magnetic beads at 67 °C for 30 minutes on an Eppendorf Thermomixer C with intermittent shaking as above. Beads with captured hybrids were washed 6 times with LiCl/Urea Hybridization Buffer at 67 °C for 5 minutes to remove non-specifically associated proteins. Between 0.5 – 1% of the total beads were removed and transferred to a fresh tube after the final wash to examine RNA captures by qPCR (see “Elution and analysis of RNA samples”). The remaining beads were resuspended in Benzonase Elution Buffer (20 mM Tris pH 8.0, 2 mM $MgCl_2$, 0.05% NLS, 0.5 mM TCEP) for subsequent processing of the protein samples.

Elution of protein samples—Elution of captured proteins from streptavidin beads was achieved by digesting all nucleic acids (both RNA and DNA, double-stranded and single-stranded) using 125 U of Benzonase nonspecific RNA/DNA nuclease for 2 hours at 37 °C

(Millipore, #71206-3). Beads were then magnetically separated from the sample using a DynaMag magnet (Life Technologies) and the supernatant containing eluted Xist-specific proteins were precipitated overnight at 4 °C with 10% trichloroacetic acid (TCA). TCA treated protein elution samples were pelleted by centrifugation for 30 minutes at >20,000 × *g*, then washed with 1 mL cold acetone and recentrifuged. Final protein elution pellets were air dried to remove acetone and stored at − 20 °C until processing for mass spectrometry.

Elution and analysis of RNA samples—Beads with hybrids were magnetically separated using a 96-well DynaMag (Life Technologies) and the supernatant was discarded. Beads were then resuspended by pipetting in 20 μL NLS RNA Elution Buffer (20 mM Tris pH 8.0, 10 mM EDTA, 2% NLS, 2.5 mM TCEP). To release the target RNA, beads were heated for 2 minutes at 95 °C in an Eppendorf Thermomixer C. Beads were then magnetically separated using a 96-well DynaMag (Life Technologies) and the supernatants containing eluted target RNA were digested by the addition of 1 mg/mL Proteinase K for 1 hour at 55 °C to remove all proteins. The remaining nucleic acids were then purified by ethanol precipitation onto SILANE beads (Invitrogen) as previously described[13,32]. DNA probes were removed by digestion with TURBO DNase (Ambion). To quantify RNA yield and enrichment, qPCR was performed as previously described[13].

Mass Spectrometry Analysis

Preparation of proteins for mass spectrometry—Proteins from RAP-MS captures were resuspended in fresh 8 M urea dissolved in 40 μL of 100 mM Tris-HCl pH 8.5. Disulfide bonds were reduced by incubation with 3 mM TCEP for 20 minutes at room temperature, followed by alkylation with 11 mM iodoacetamide for 15 minutes at room temperature in the dark. Samples were then digested with 0.1 μg endoproteinase Lys-C for 4 hours at room temperature. After Lys-C digestion the samples were diluted to a final concentration of 2M urea by the addition of 100 mM Tris-HCl pH 8.5, and $CaCl_2$ was added to a final concentration of 1 mM. Tryptic peptides were generated by treatment with 0.1 to 0.5 μg of trypsin overnight at room temperature. Contaminating detergents were removed from peptides using HiPPR detergent removal columns (Thermo), and peptides were protonated by the addition of 5% formic acid before desalting on a Microm Bioresources C8 peptide MicroTrap column. Peptide fractions were collected and lyophilized, and dried peptides were resuspended in 0.2% formic acid with 5% acetonitrile.

Mass spectrum measurements—Liquid chromatography-mass spectrometry and data analyses of the digested samples were carried out as previously described[33] with the following modifications. All experiments were performed on a nanoflow LC system, EASY-nLC 1000 coupled to a hybrid linear ion trap Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nanoelectrospray ion source (Thermo Fisher Scientific). For the EASY-nLC II system, solvent A consisted of 97.8% H_2O, 2% ACN, and 0.2% formic acid and solvent B consisted of 19.8% H_2O, 80% ACN, and 0.2% formic acid. For the LC-MS/MS experiments, 200 ng of digested peptides were directly loaded at a flow rate of 500 nL/min onto a 16-cm analytical HPLC column (75 μm ID) packed in-house with ReproSil-Pur C_{18}AQ 3 μm resin (120 Å pore size, Dr. Maisch, Ammerbuch, Germany). The column was enclosed in a column heater operating at 30°C.

After 30 min of loading time, the peptides were separated with a 75 min gradient at a flow rate of 350 nL/min. The gradient was as follows: 0–2% Solvent B (5 min), 2–30% B (60 min), and 100% B (10 min). The Elite was operated in data-dependent acquisition mode to automatically alternate between a full scan (m/z=400–1600) in the Orbitrap and subsequent rapid 20 CID MS/MS scans in the linear ion trap. CID was performed with helium as collision gas at a normalized collision energy of 35% and 10 ms of activation time.

MS data analysis—Thermo RAW files were searched with MaxQuant (v 1.5.0.30)[34,35]. Spectra were searched against all UniProt mouse entries (43,565 entries, downloaded 02 Oct 14) and MaxQuant contaminant database (245 entries). Decoy sequences (reversed peptide sequences) were generated in MaxQuant to estimate the false discovery rate[36]. Search parameters included multiplicity of 2 with heavy Arg (+10.0083) and heavy Lys (+8.0142) as heavy peptide modifications. Variable modifications included oxidation of Met (+15.9949) and protein N-terminal acetylation (+42.0106). Carboxyamidomethylation of Cys (+57.0215) was specified as a fixed modification. Protein and peptide false discovery rates were thresholded at 1%. Precursor mass tolerance was 7 ppm (or less for individual peptides). Fragment mass tolerance was 0.5 Da. Requantify and match between runs were both enabled. Trypsin was specified as the digestion enzyme with up to 2 missed cleavages.

Identification of RNA interacting proteins—Proteins of interest from RAP-MS captures were identified based on several criteria. First, proteins were considered identified only if 2 or more unique peptides were found in the mass spectrum. Then proteins of interest were selected based on the SILAC ratio of capture versus control samples. SILAC ratios for each peptide were calculated based on the intensity ratios of heavy and light SILAC pairs. The protein ratio is the median of all calculated peptide ratios, with a minimum of two SILAC pairs required for a SILAC ratio to be assigned to a given protein. A SILAC ratio cutoff of ≥3.0 (fold enrichment over control sample) was used as a cutoff for further analysis. We excluded known contaminants, including human keratin and proteins introduced during the sample purification and preparation process (such as streptavidin, Benzonase, and trypsin), as well as naturally biotinylated proteins that contaminate the preparation by binding to streptavidin beads.

RAP-MS experiments and controls

18S rRNA versus U1 snRNA—To validate the RAP-MS method and identify proteins specifically interacting with 18S ribosomal RNA or U1 snRNA, we performed captures of each target RNA in parallel samples from heavy and light labeled lysates from wild-type V6.5 ES cells. The total protein quantity in elution samples from each RAP-MS capture was measured by comparing the median intensity of peptides identified in a single quantitation MS run for each sample. The heavy and light label swapped samples were then mixed equally based on total protein quantity and analyzed by mass spectrometry to identify the SILAC enrichment ratio of proteins originating from 18S rRNA or U1 snRNA captures. The experiment was performed twice and each experimental set contained two biological replicates of 18S and U1 captures (heavy and light labeling states).

Xist lncRNA versus U1 snRNA captures—To identify proteins specifically interacting with Xist lncRNA, we performed captures as described above with either 200M cells or 800M pSM33 cells treated with doxycycline for 6 hours. The total protein quantity in elution samples from each RAP-MS capture was measured by a single quantitation MS run for each sample. Heavy and light label swapped samples were mixed equally based on total protein quantity, and analyzed by mass spectrometry. SILAC ratios of Xist enriched proteins versus U1 enriched proteins were calculated and used to identify Xist-specific interacting proteins for further analysis. The experiment was performed twice and each experimental set contained two biological replicates of Xist and U1 captures, from heavy and light labeled samples. Proteins replicated well between samples, with a sole exception (LBR) that was missed only because its enrichment level (2-fold) fell below our enrichment cutoff (3-fold) in some replicate samples.

Xist lncRNA capture from non-crosslinked cells—As a control to ensure that purified proteins are not non-specifically associated or binding *in vitro* with target RNAs during capture, we performed RAP-MS captures of Xist from non-crosslinked cells otherwise treated in the same manner (i.e. doxycycline treated for 6 hours).

Xist lncRNA capture from cells where Xist is not expressed—To confirm that the identified proteins are not resulting from background proteins or probe association with other RNAs or proteins in the pSM33 cells, we performed RAP captures of Xist from pSM33 cells that were not treated with doxycycline, but which were otherwise treated identically.

45S pre-rRNA capture versus U1 capture—To ensure that the proteins enriched in Xist captures using RAP-MS are not simply due to increased protein capture as a consequence of long target RNA transcripts, we additionally performed captures of the 13,000 nucleotide long 45S pre-ribosomal RNA as a control. To ensure specific capture only of the 45S, and not the mature 18S and 28S, we designed probes that specifically targeted the internal transcribed spacer regions (ITS1 and ITS2) that are only present in the 45S pre-ribosomal RNA. The experiment was performed in the same manner and with the same conditions as the Xist lncRNA captures described above. To compare Xist protein enrichment to 45S protein enrichment, we used a SILAC approach based on direct comparison of two samples that share a common denominator (called spike-in SILAC[37]). Specifically, we calculated an overall Xist/45S SILAC ratio by multiplying the Xist/U1 ratio by the U1/45S ratio for each identified protein.

Protein domain classification

We defined the conserved domain structures of proteins using the Protein Families database (Pfam[38]).

RNA Immunoprecipitation in UV-crosslinked cells

We crosslinked pSM33 cells after 6 hours of doxycycline-treatment with 0.4 Joules/cm^2 of UV_{254}. Cells were lysed and RNA was digested with RNase I to achieve a size range of 100–500 nucleotides in length. Lysate preparations were precleared by mixing with Protein

G beads for 1 hour at 4 °C. For each sample, target proteins were immunoprecipitated from 20 million cells with 10 μg of antibody (Supplementary Table 1) and 60 μl of Protein G magnetic beads (Invitrogen). The antibodies were pre-coupled to the beads for 1 hour at room temperature with mixing before incubating the precleared lysate to the antibody-bead complexes for 2 hours at 4 °C. After the immunoprecipitation, the beads were rinsed with a wash buffer of 1× PBS with detergents. After a dephosphorylation treatment, the RNA in each sample was ligated to a mixture of barcoded adapters in which each adapter had a unique barcode identifier. After ligation, beads were rinsed with 1× PBS and detergents and then 5× PBS (750 mM NaCl) and detergents prior to pooling 3–4 antibodies in a new tube. The proteins and RNA were then eluted from the Protein G beads with 6 M urea and 40 mM DTT at 60 °C. Protein-RNA complexes were separated away from free RNA by covalently coupling proteins to NHS-magnetic beads (Pierce) and washing 3 times in 6 M GuSCN (Qiagen RLT buffer) and heating in 1% NLS at 98 °C for 10 minutes. The proteins were then digested with Proteinase K and RNA was purified for subsequent analysis. From the barcoded RNA in each pool, we generated Illumina sequencing libraries as previously described[32]. We saved a small percentage (~1%) of starting material prior to immunoprecipitation and processed and sequenced this sample in parallel.

Analysis of crosslinked RNA Immunoprecipitation Data

We computed the enrichment for any RNA upon immunoprecipitation with a specific protein relative to its total levels in the cell. To do this, we counted the total number of reads overlapping the RNA in either the immunoprecipitation sample or the input control. To account for differences in read coverage between samples, each of these numbers was normalized to the total number of reads within the same experiment. This generates a normalized score, per RNA, within each sample. We then computed an enrichment metric by taking the ratio of these normalized values (IP/Input). We then compared these enrichment levels across different proteins and controls (i.e. IgG). To enable direct comparison across proteins for a given gene, we need to account for differences in the protein specific background level, which may occur to differences in IP efficiency or non-specific binding of each antibody. To do this, we computed a normalized enrichment ratio by dividing the ratio for each gene by the average ratio across all genes for a given protein, as previously described[1].

To exclude the possibility of promiscuous binding to all RNAs, we considered various mRNA controls, which are not expected to bind to these proteins, including Oct4, Nanog, Stat3, and Suz12. These mRNAs were selected as examples because they are expressed in ES cells, although many mRNAs show similar results. To account for the possibility that the Xist RNA non-specifically binds to any RBP, we evaluated Xist with other RBPs that we did not identify as interacting with Xist by RAP-MS (Pum1 and hnRNP-H). To ensure that a negative result (i.e. no enrichment for Xist) is meaningful and does not reflect a failed immunoprecipitation experiment, we evaluated Neat1-1, which we previously found immunoprecipitates with hnRNPH[1]. To further evaluate the level of enrichment on other lncRNAs, we considered several lncRNAs including Malat1, Firre, and Tug1. These lncRNAs were selected as examples because they are well-known and expressed in ES cells, although many ES lncRNAs show similar results.

Immunoprecipitation and RT-qPCR

Female ES cells were differentiated then crosslinked with UV4k as described above. Pellets of 20M cells were lysed and treated with TURBO DNase (Ambion) to destroy DNA by incubation for 10 minutes at 37 °C in an Eppendorf Thermomixer C. The lysate was pre-cleared by incubation with 180 μL of Dynabeads Protein G magnetic beads (Life Technologies). Meanwhile, 10 μg of antibody for immunoprecipitation (SHARP antibody, Novus NBP1-82952 or IgG antibody, Cell Signaling 2729S) was coupled to 60 μl Protein G magnetic beads. After pre-clearing was completed, the lysate was then mixed with the appropriate antibody-coupled Protein G magnetic beads and incubated for 2 hours at 4 °C on a Hulamixer sample mixer (Life Technologies) for protein capture. After immunoprecipitation, beads were washed with a wash buffer of 1× PBS with detergents and then captured nucleic acids were eluted by digesting all proteins with 5.6 U proteinase K (New England Biolabs). Eluted RNA was purified using the RNA Clean and Concentrator-5 Kit (Zima Research) and RT-qPCR was performed as described previously[13] to evaluate RNA enrichment.

V5-epitope tagged protein expression

For V5-tagged protein expression and immunoprecipitation, mouse ES cells were electroporated using the Neon transfection system (Invitrogen) with an episomally-replicating vector (pCAG-GW-V5-Hygro) encoding expression of a C-terminal V5 tagged ORF driven by a CAG promoter. ORFs were obtained from the DNASU plasmid repository as Gateway entry clones and inserted into pCAG-GW-V5-Hygro using an LR recombination reaction (Invitrogen). Transfected cells were selected on 125ug/mL Hygromycin B (Invitrogen) to generate stably expressing lines.

siRNA Transfections

For siRNA knockdown experiments, 20 nM siRNAs were transfected using the Neon transfection system (settings: 1200V, 40ms width, 1 pulse). For each transfection, two 10 μL transfections with the same siRNA were carried out in succession using 100,000 cells each, mixed, and plated equally between two poly-L-lysine or poly-D-lysine (Sigma) and 0.2% gelatin (Sigma)-coated #1.5 coverslips placed into wells of a 24-well plate containing 2i media. After 48 hours, 2i media was replaced and cells on one coverslip of each pair were treated with 2 μg/mL doxycycline (Sigma) for 16hr to induce Xist expression. Coverslips were then fixed in Histochoice (Sigma) for 5 min, washed thoroughly in PBS, and dehydrated in ethanol for storage until FISH staining.

For all proteins we used siRNAs pool from Dharmacon (ON-TARGETplus SMARTpool siRNAs). For each of these, we tested whether the siRNA successfully reduced the targeted mRNA expression by >70%. For SAF-A and SMRT, the siRNAs failed to achieve this level of mRNA reduction, so we purchased additional siRNAs (and their associated controls) for SAF-A and SMRT from Qiagen and Ambion respectively, and selected siRNAs that successfully reduced on-target mRNA levels. siRNA against GFP was purchased from Qiagen. For additional independent siRNAs, the siRNAs were purchased as a pool from Dharmacon, Qiagen, and Ambion, or as each individual siRNA deconvoluted from the pool from Dharmacon and Qiagen (Supplementary Table 2).

In addition to the proteins identified by RAP-MS, we knocked down several proteins previously reported to associate with Xist – including EED (a component of PRC2)[8], YY1[39], Satb1[40], SRSF1[41], hnRNPC[42], and Atrx[43].

siRNA experiments in female ES cells

Female ES F1 2-1 cells were similarly transfected. To initiate differentiation and Xist expression for these cells, 2i was replaced with MEF media (DMEM, 10% Gemini Benchmark FBS, 1× L-glutamine, 1× NEAA, 1× Pen/Strep; Life Technologies unless otherwise indicated) at 12 hours post-transfection. Forty-eight hours after transfection, 1uM retinoic acid (Sigma) was administered for 24 hours and cells were fixed as described above. For cells not undergoing differentiation, 2i was replaced 12hr and 48hr after transfection.

Single molecule RNA FISH

Single molecule RNA Fluorescence *in situ* hybridization (FISH) experiments were done using QuantiGene ViewRNA ISH Cell Assay (Affymetrix) and QuantiGene ViewRNA ISH Cell 740 Module (Affymetrix) according to manufacturer's protocol. Cells fixed on coverslips were first permeabilized with Detergent Solution QC at room temperature for 5 min, and then incubated with desired mixture of probe set (Affymetrix) in Probe Set Diluent QF at 40°C for 3 h, followed by incubated with PreAmplifier Mix at 40°C for 30 min, Amplifier Mix at 40°C for 30 min, and Label Probe Mix at 40°C for 30 min sequentially. For DAPI staining, coverslips were incubated in 30 nM DAPI in PBS at room temperature for 15–20 min. Probe set and conjugated fluorophore for FISH were TYPE 1-XIST (550 nm), TYPE 4-GPC4, RBMX, SMC1A, MECP2 (488 nm), TYPE 10-ATRX (740 nm), and TYPE 6-EED1, SHARP, LBR, SAFA, RBM15, MYEF2, PTBP1, HNRNPC, HNRNPM, CELF1, RALY, HDAC3, NCOR2, MID1, PIR (650 nm).

Immunofluorescence and RNA FISH

For immunofluorescence (IF), cells were fixed on coverslips and permeabilized with 0.1% Triton-X in PBS at room temperature for 10 min, and blocked with 5% normal goat serum in PBS at room temperature for 10 min. Cells were then incubated with primary antibodies at room temperature for 1 h, followed by incubating with secondary antibodies at room temperature for 1 h. The samples were then processed using the RNA FISH protocol, as described above. Primary antibodies and the dilution used for IF were anti-RNA polymerase II CTD repeat YSPTSPS (phospho S2) (Abcam; ab5095) (1:100), anti-Nanog (Abcam; ab80892) (1:100), and anti-EZH2 (Active Motif; 39933) (1:100). Secondary antibodies and the dilution used for IF were Alexa Fluor® 405 goat anti-rabbit IgG (H+L) (Life Technology; 1575534) (1:100) and Alexa Fluor® 488 F(ab')$_2$ fragment of goat anti-rabbit IgG (H+L) (Life Technology; 1618692) (1:100).

Microscopic Imaging

FISH and IF/FISH samples were imaged using a Leica DMI 6000 Deconvolution Microscope with the Leica HC PL APO 63×/1.30 GLYC CORR CS2 objective. Samples stained with TYPE 10-ATRX (740 nm) were imaged using Nikon Ti Eclipse with the Nikon

CFI Plan Apochromat λ DM 60×/1.40 oil objective. Images were projected with maximum projection (3 μm; step size, 0.5 μm).

X-chromosome Silencing Assay

Cells were stained for Xist RNA, Gpc4 mRNA, Atrx mRNA and siRNA-targeted mRNA by FISH and imaged. In addition, in some siScramble and siSHARP samples, we used probes against Rbm15, Mecp2, Smc1a, Mid1 or Pir mRNA. Images were then analyzed using Matlab R2013b (described below). Cells were selected if the copy number of the targeted mRNA was less than 30% of the level of the no siRNA treated cells and if they induced Xist expression. Within these cells, the copy number of Gpc4 mRNA and Atrx mRNA were quantified using a peak finding method (described below) and compared across conditions. We quantified mRNA levels for 50 individual cells. We also evaluated Xist expression in siRNA-treated cells, and observed no difference in the percentage of cells that induced Xist expression in any of the siRNA conditions relative to untreated cells.

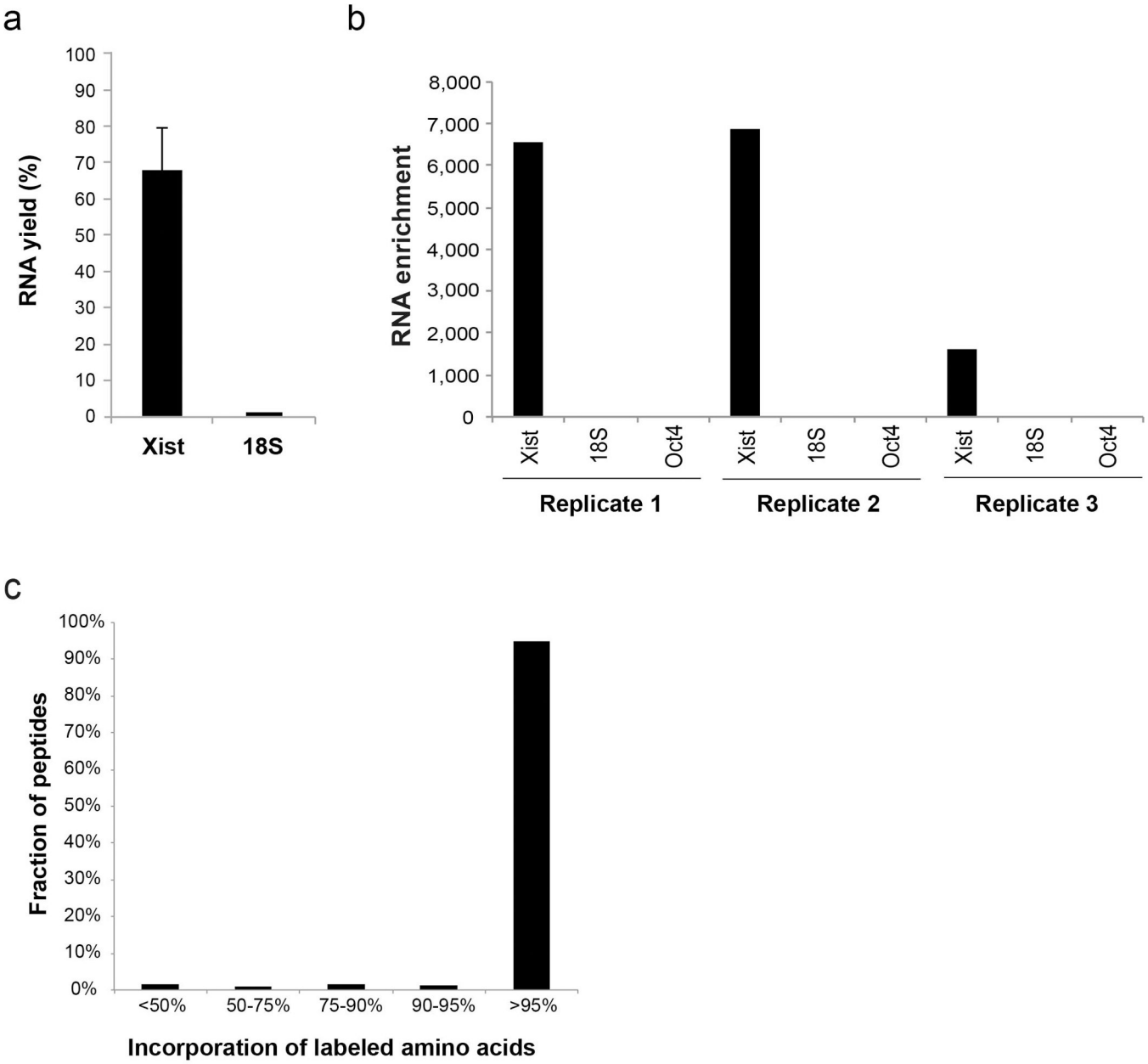
Quantifying mRNAs by single molecule FISH

All image analysis was carried out using Matlab (version R2013b) utilizing built-in functions from the Image Processing toolbox. Images were first filtered using a two-dimensional median filter to remove background. Cell boundaries were outlined manually, guided by DAPI staining, to create a binary mask and applied to the various channels from the same field of view. Top-hat morphological filtering, a background subtraction method that enhances the individual focal spots, was applied to the images[44]. The spots were then identified using a 2D peak finding algorithm that identifies local maximal signals within the cell. Once regional maxima were identified, the number of spots was counted for each cell.

Ezh2 Recruitment and PolII Exclusion

Cells were stained for Xist RNA and the siRNA-targeted mRNA (FISH) along with Ezh2 or PolII (IF) as described above. For image acquisition, the exposure time for each individual channel was kept the same across all samples. Images were then analyzed and selected for XIST-induced and cells showing knock down of the target mRNA, as described above. Specifically, the nuclei of individual cells were identified manually using the DAPI staining. We identified the Xist cloud by using an intensity-based threshold to partition the image within the nucleus and find contiguous 2-dimensional regions of high intensity. The threshold was determined based on Otsu method as previously described[45], which splits the image into 2 bins – high and low – and identifies a threshold that minimizes the variance within the partition. This creates a binary mask on the image. We visually confirmed that this binary mask accurately reflected the Xist cloud. We then applied this binary mask to all other images in that field of view (PolII or Ezh2) for all images. We then quantified the intensity of fluorescence signal by taking the average intensity of all the pixels within the region (i.e. PolII or Ezh2, respectively). We computed this average intensity (1 number per cell) across all conditions and compared them using a 2-same unpaired t-test relative to the scramble sample across 50 single cells.

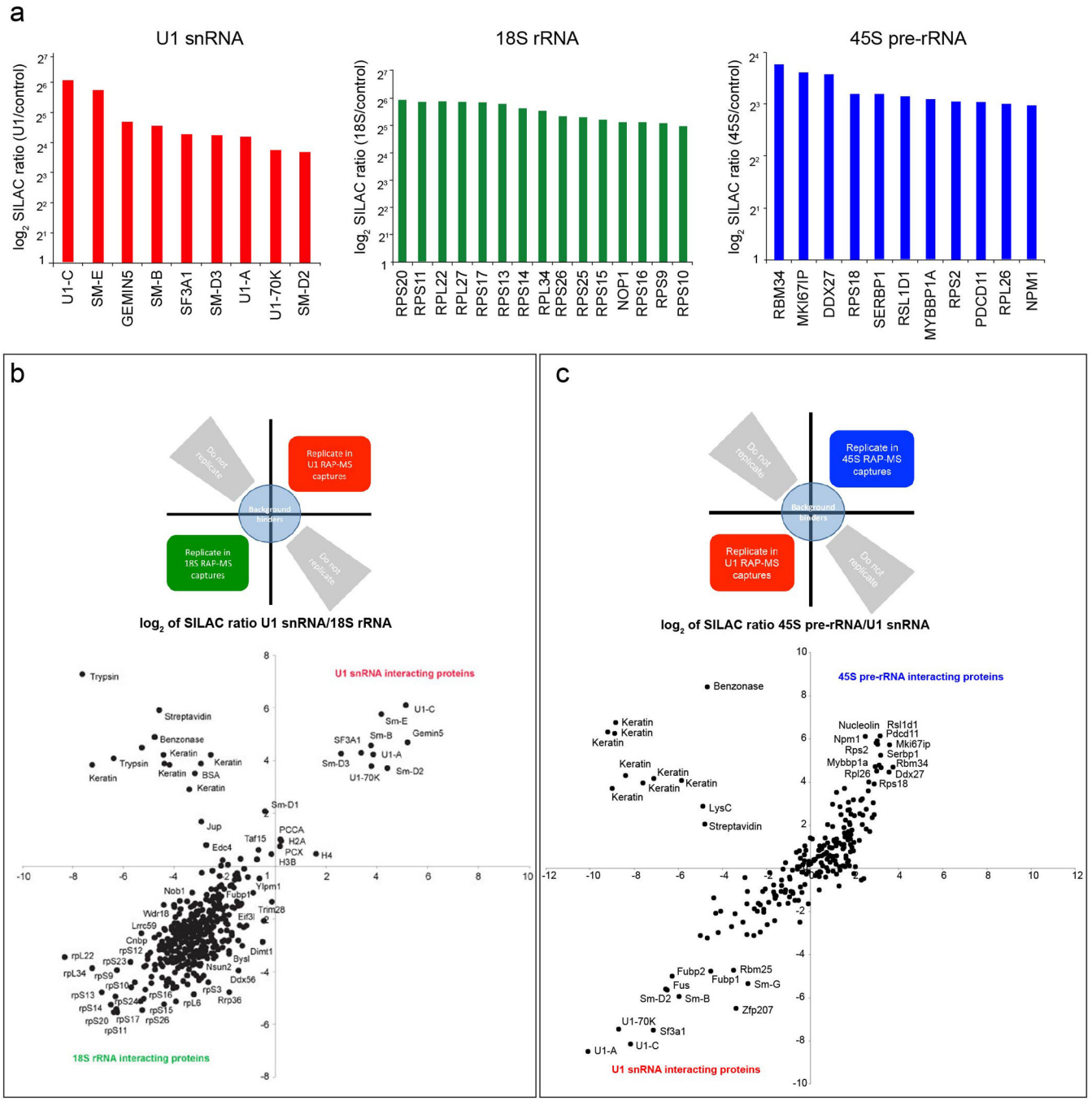
Extended Data



Extended Data Figure 1. RAP-MS recovers and enriches the majority of Xist RNA from mouse ES cells, and these cells can be efficiently labeled with SILAC

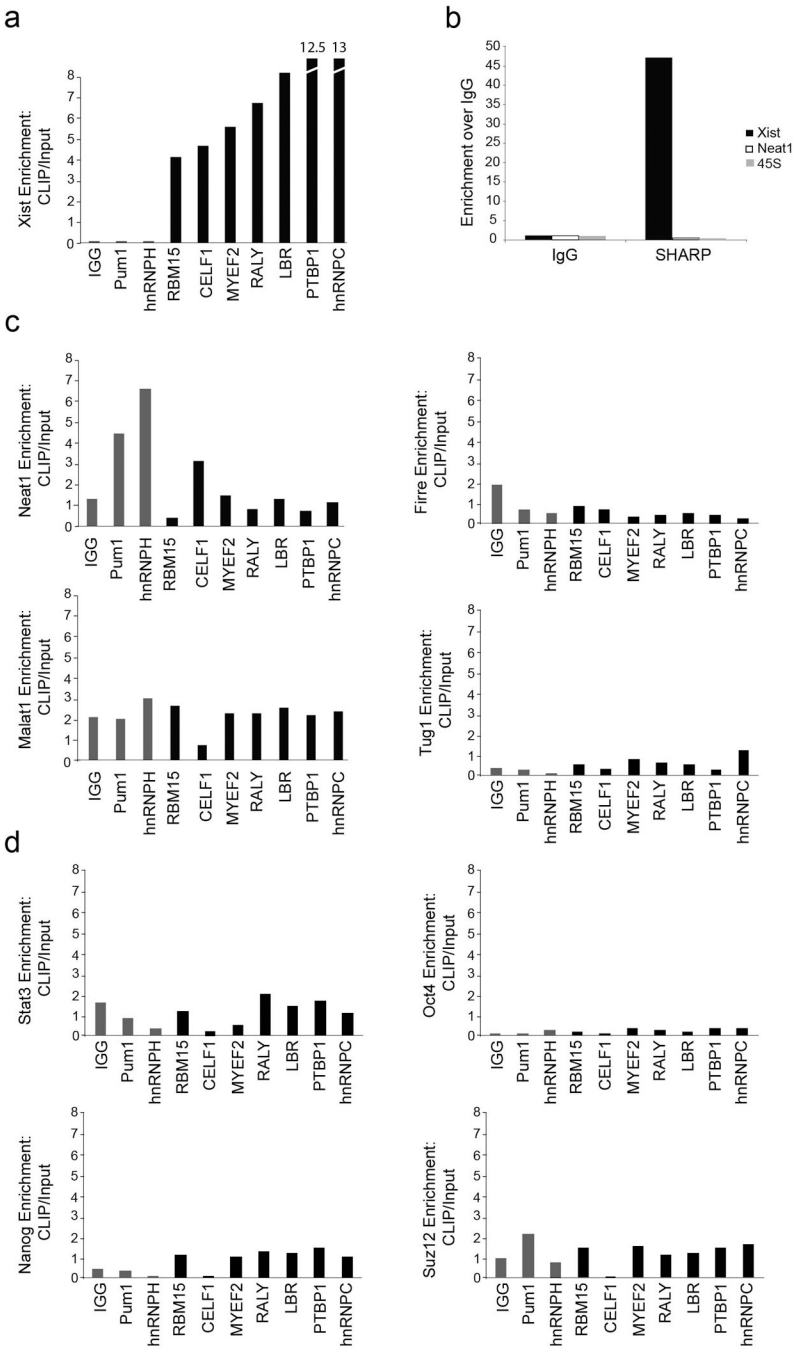
(a) RT-qPCR measuring the percentage of the total cellular Xist or 18S recovered after RAP-MS of Xist. Values are computed as the amount of each RNA in the elution divided by the amount of RNA in the starting ("input") lysate material. Error bars represent the standard error of the mean from 5 biological replicates. (b) Enrichment of Xist after RAP-MS captures from pSM33 cells as measured by qPCR. Bars indicate RNA levels of Xist, 18S, and Oct4 after purification of Xist, normalized to RNA in input sample. Each bar represents the RNA levels of Xist, 18S, and Oct4 after purification of Xist, normalized to RNA in input

sample, from 3 biological replicates. (c) SILAC labeling efficiency of a representative culture of pSM33 mouse ES cells after 10 days of growth (3 cell passages) in SILAC medium. Peptides were analyzed by mass spectrometry, and values indicate the fraction of identified peptides with heavy-label incorporation with different levels of peptide labeling (shown in bins).



Extended Data Figure 2. RAP-MS identifies proteins that are known to directly interact with specific ncRNAs, and separates specific RNA interacting proteins from background proteins

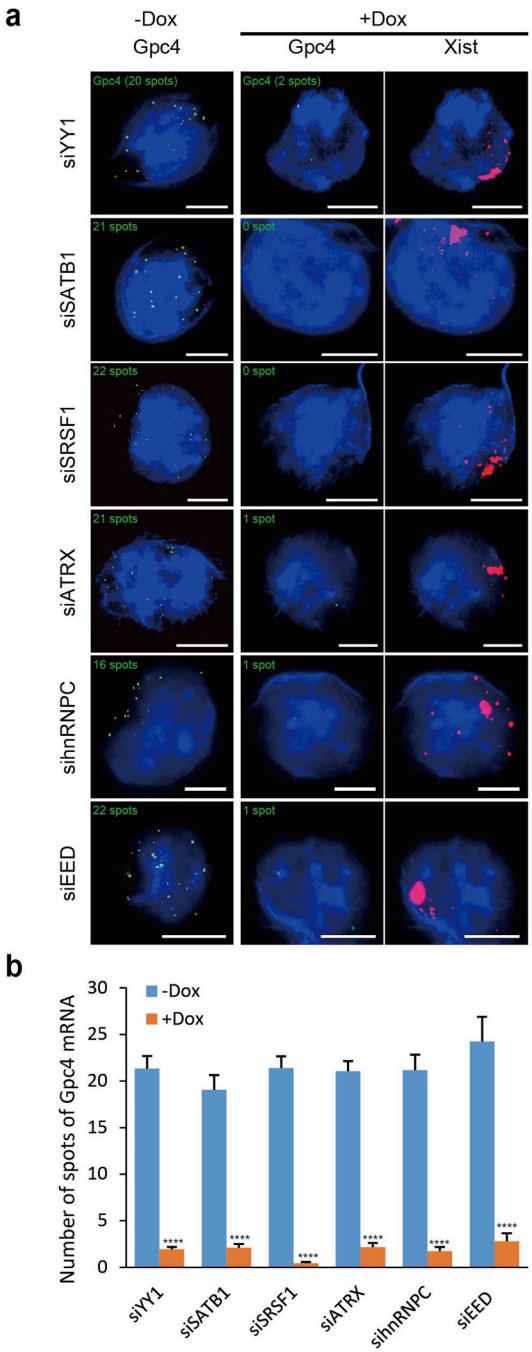
(a) SILAC ratios of top proteins enriched in the RAP-MS U1 snRNA, 18S rRNA, and 45S pre-rRNA experiments. (b) SILAC ratio plot of replicate captures of U1 snRNA versus 18S rRNA from one of two biologically independent label-swap experiments. Proteins associated with U1 are consistently found in U1 samples, both light and heavy labeled (top right quadrant), and proteins specifically associated with 18S are consistently identified in 18S, both light and heavy (lower left quadrant). Background contaminant proteins have low enrichments (center of panel) or are consistently found in the light channel and do not replicate between experiments (i.e. keratin, streptavidin). (c) SILAC ratio plot of replicate captures of U1 snRNA versus 45S pre-rRNA from one label-swap experiment. Proteins that are known to associate with 45S pre-rRNA are consistently identified in 45S captures.



Extended Data Figure 3. Immunoprecipitation of the identified Xist-interacting proteins confirms Xist RNA interaction

RNA immunoprecipitation experiments were performed for seven Xist-interacting proteins (black bars), two control RNA binding proteins that were not identified by RAP-MS and IgG (gray bars) in UV-crosslinked cell lysate after 6 hours of Xist induction by doxycycline addition (**Methods**). The RNA associated with each protein was measured and enrichment levels were computed relative to the level of the RNA in total cellular input and normalized to the total efficiency of capture in each sample to allow for direct comparison across all IP experiments (**Methods**). (a) Enrichment of the Xist lncRNA after immunoprecipitation from

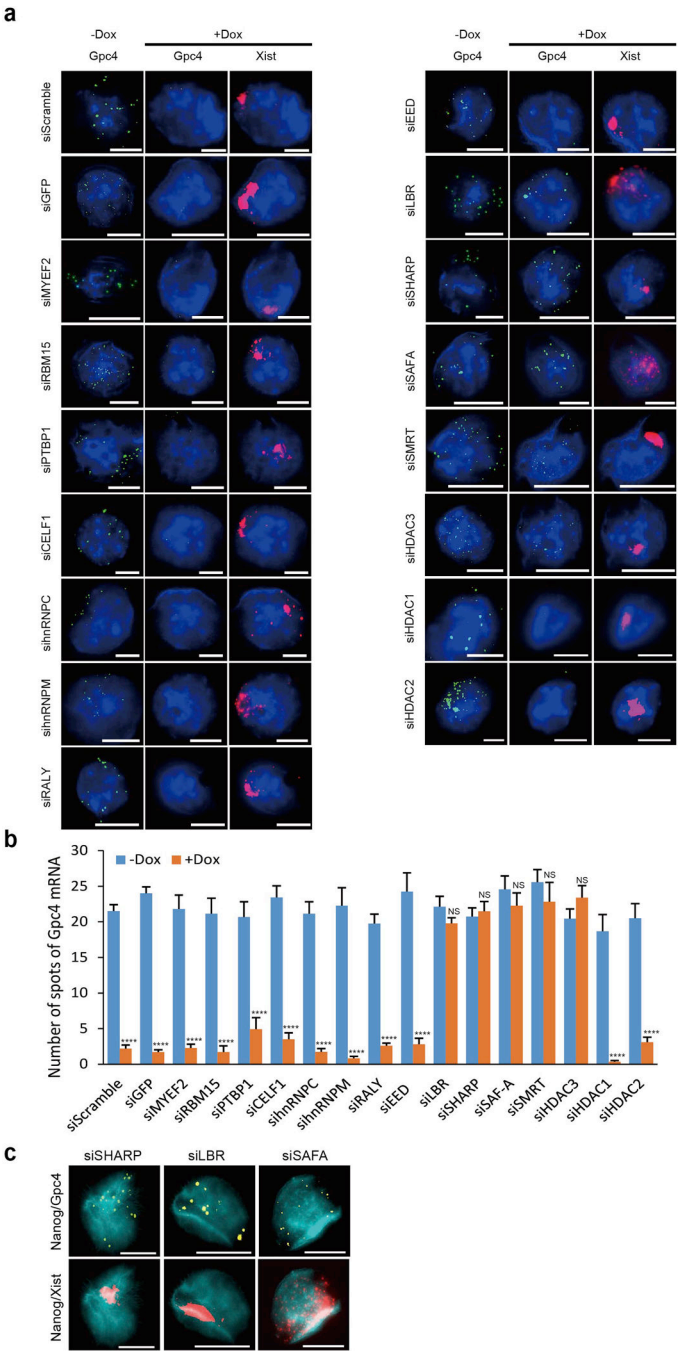
a sample of pSM33 male cells (b) Immunoprecipitation of SHARP was performed from a sample of UV-crosslinked females ES cells that were treated with retinoic acid for 24 hours. The levels of recovered Xist lncRNA (black bars), Neat1 lncRNA (white bars), and 45S pre-ribosomal RNA (gray bars) were measured by RT-qPCR. Enrichment of each RNA after capture with anti-SHARP antibody was calculated relative to the level of RNA captured with IgG control antibody. (c) The enrichment of various lncRNAs after immunoprecipitation in pSM33 male cells – including Neat1, Malat1, Firre, and Tug1 – are shown. (d) The enrichment of various mRNA controls after immunoprecipitation in pSM33 male cells – including Oct4, Nanog, Stat3, and Suz12 – are shown.



Extended Data Figure 4. Previously identified proteins associated with XCI are not required for Xist-mediated transcriptional silencing

(a) To confirm the specificity of our assay, we tested the function of several proteins that were previously identified to associate with Xist, but not to silence transcription, for their role in transcriptional silencing in our inducible male ES cells prior to Xist induction (−Dox; left) or after Xist induction for 16 hours (+Dox; middle and right). Representative images are shown after knockdown of each protein – DAPI (blue), Xist (red), and Gpc4 (green). (b) Quantification of the copy number of Gpc4 before and after Xist induction upon treatment with different siRNAs. Error bars represent the standard error of the mean across 50

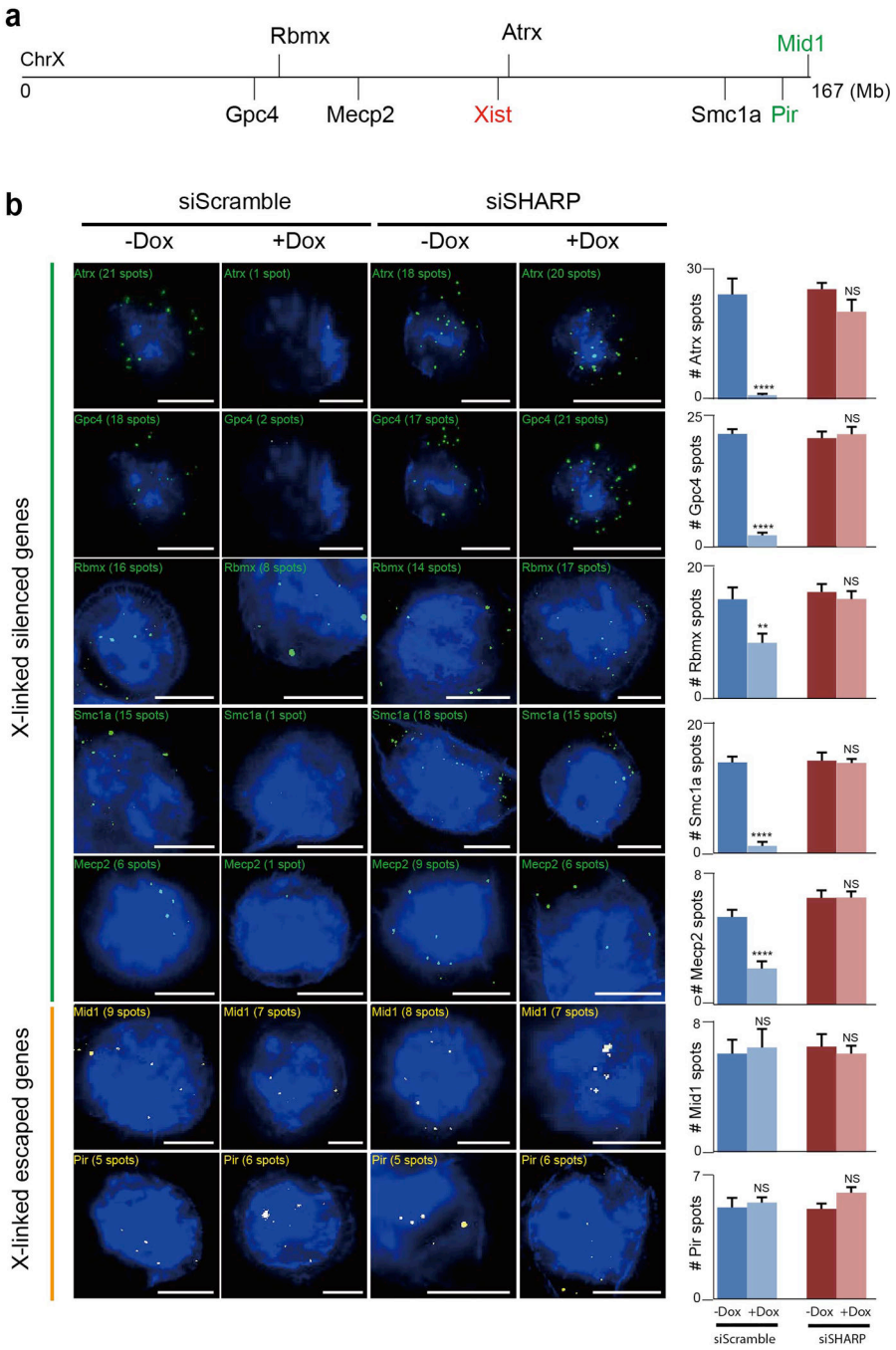
individual cells from one experiment. **** represents values with a p-value<0.001 between +Dox and −Dox cells based on an unpaired two-sample t-test. Scale bars on the images represent 5 μm. Importantly, while these proteins do not have a role in the initiation of transcriptional silencing, we do not mean to imply that they do not have other roles in XCI.



Extended Data Figure 5. SHARP, LBR, SAF-A, SMRT, and HDAC3 are required for Xist-mediated transcriptional silencing

(a) Representative images showing staining of DAPI (blue), Xist (red), and Gpc4 (green) for different siRNA knockdown in male ES cells prior to Xist induction (−Dox; left) or after

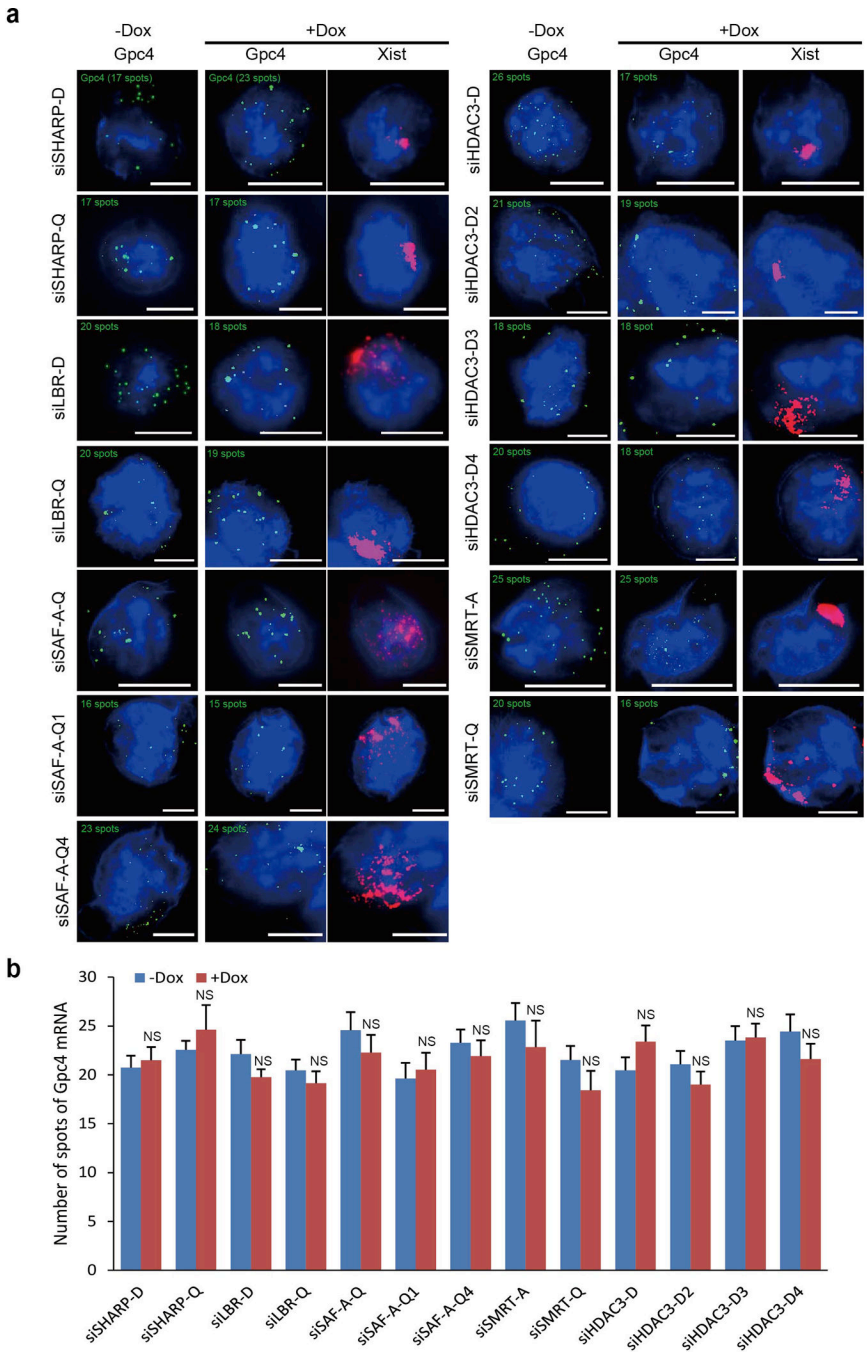
Xist induction for 16 hours (+Dox; middle and right). (b) Quantification of the copy number of Gpc4 in −Dox and +Dox cells after knockdown with siRNAs targeting different mRNAs. Error bars represent the standard error of the mean across 50 individual cells from one experiment. NS: not significantly different between +Dox and −Dox cells; **** represents values with a p-value<0.001 between +Dox and –Dox cells based on an unpaired two-sample t-test. Scale bars on the images represent 5 μm. (c) Knockdown of SHARP, LBR, or SAF-A abrogates Xist-mediated gene silencing without causing pluripotency defects. Representative images showing staining of Nanog (cyan), Xist (red), and Gpc4 (green) upon knockdown of SHARP, LBR or SAF-A after 16 hours of Xist induction with doxycycline. Scale bars on the images represent 5 μm.



Extended Data Figure 6. SHARP is required for silencing many genes across the X-chromosome

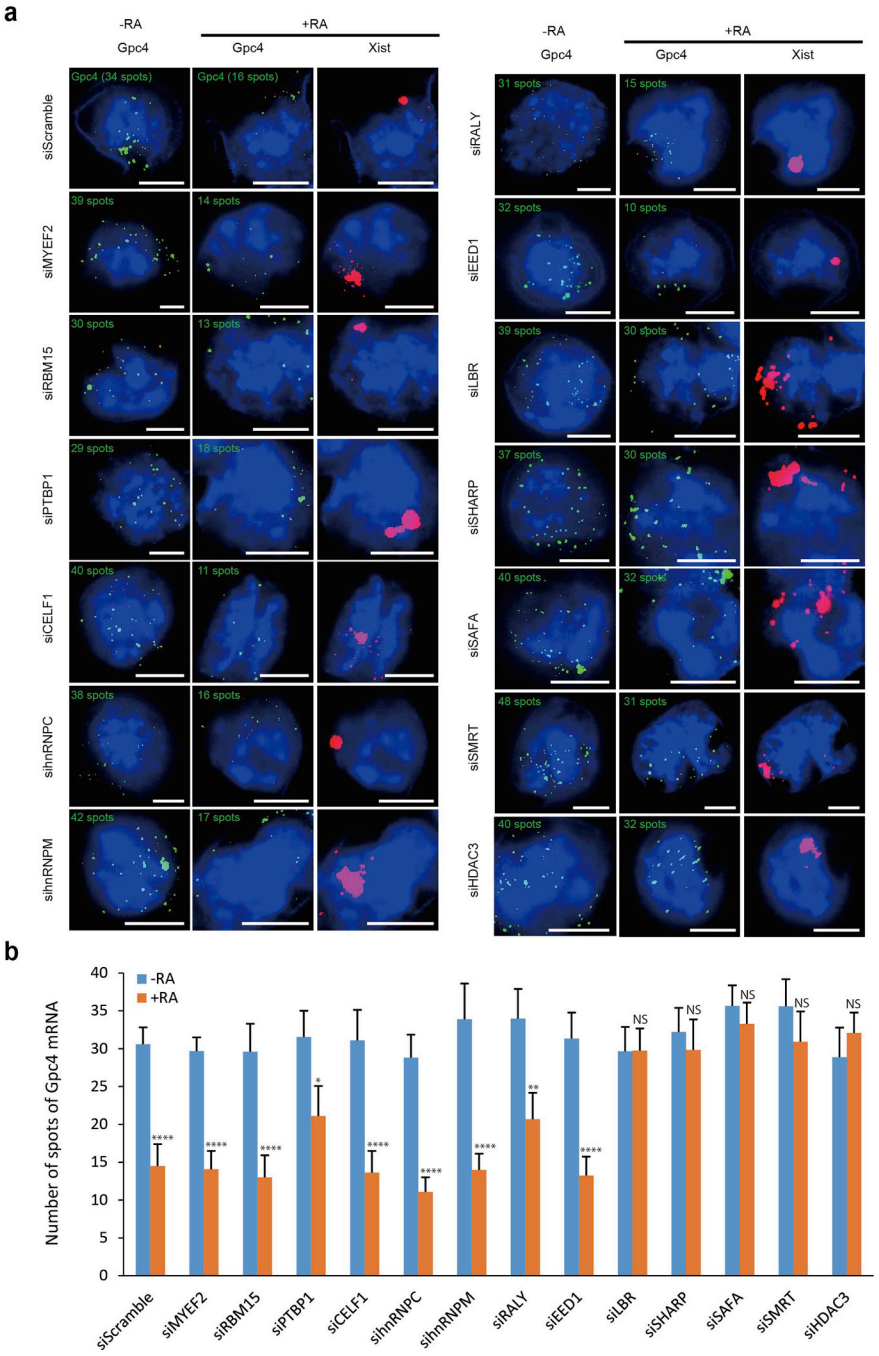
(a) A diagram showing the locations of Xist (red), X-linked silenced genes (black), and X-linked escaped genes (green) along the X-chromosome. (b) Representative images showing staining of DAPI (blue), Xist (red), X-linked silenced genes (green), and X-linked escaped genes (yellow) upon knockdown of SHARP or control male ES cells prior to Xist induction (−Dox) or after Xist induction for 16 hours (+Dox). Knock of SHARP abolishes the silencing of Atrx, Gpc4, Rbmx, Smc1a and Mecp2, which are normally silenced upon Xist expression, but has no effect on Mid1 and Pir, which normally escape Xist-mediated

silencing. The bar graphs show the quantification of the copy number of the mRNA for each gene for −Dox and +Dox cells upon transfection with SHARP siRNA or control siRNA; Error bars represent the standard error of the mean across 50 individual cells from one experiment, NS: not significantly different, **** represents values with a p-value<0.001, and ** represents values with a p-value<0.01 between +Dox and −Dox cells based on an unpaired two-sample t-test. Scale bars on the images represent 5 μm.



Extended Data Figure 7. Multiple independent siRNAs targeting SHARP, LBR, SAF-A, HDAC3, or SMRT demonstrate the same silencing defect

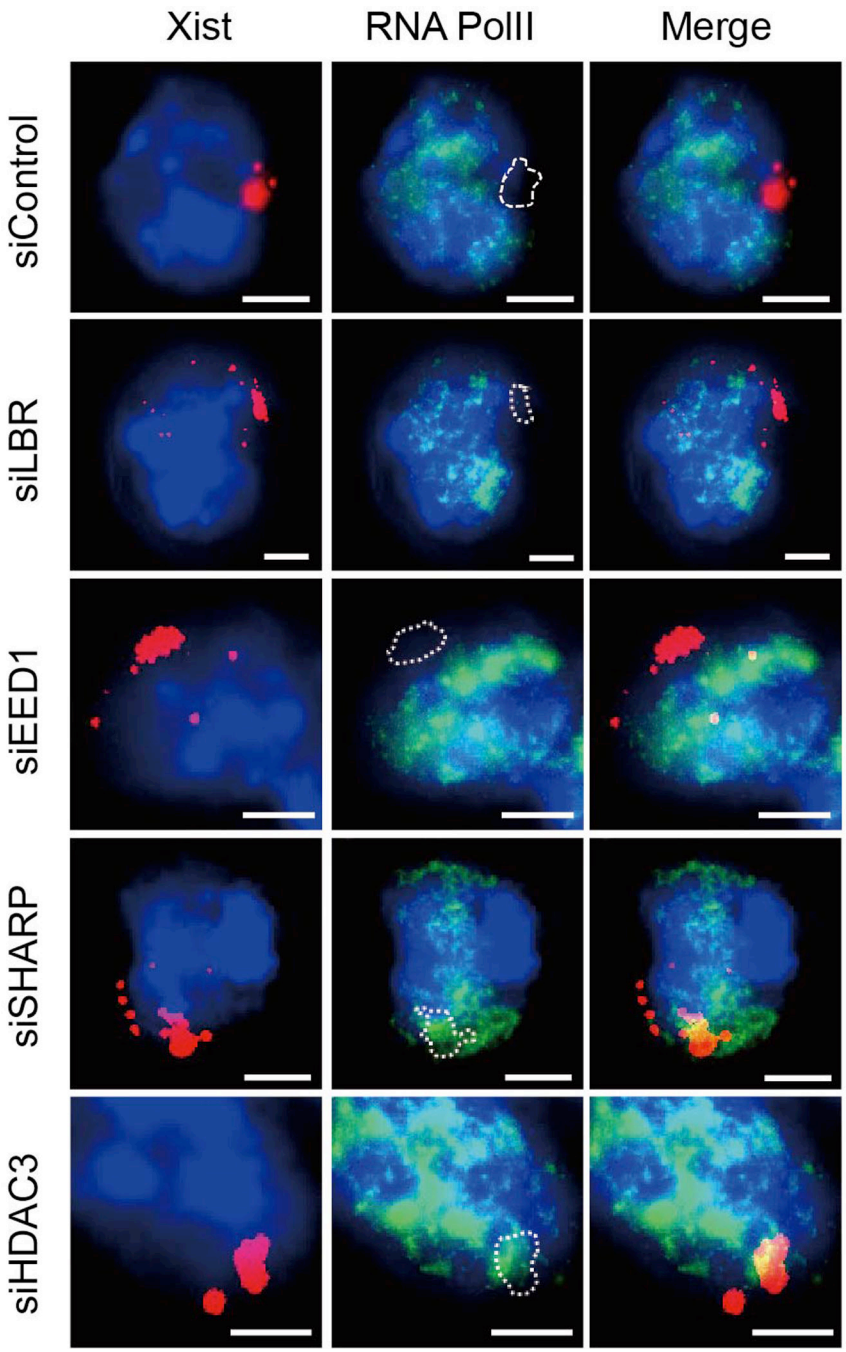
(a) Representative images showing staining of DAPI (blue), Xist (red), and Gpc4 (green) after knockdown of proteins using independent, non-overlapping, siRNA pools, or individual siRNA deconvoluted from the pool prior to Xist induction (−Dox; left) or after Xist induction for 16 hours (+Dox; middle and right). Cells were either transfected with the siRNA pool from Dharmacon (siRNA-D), Qiagen (siRNA-Q) or Ambion/Life Technologies (siRNA-A), or each individual siRNA deconvoluted from the pool from Dharmacon (siRNA-D1, 2, 3, 4) or Qiagen (siRNA-Q1, 2, 3, 4). Error bars represent the standard error of the mean across 50 individual cells from one experiment. NS: not significantly different between +Dox and −Dox cells based on an unpaired two-sample t-test. Scale bars on the images represent 5 μm. We excluded all siRNAs that did not reduce the targeted mRNA level by >70% (**Methods**). The sequences of deconvoluted siRNAs are shown in Supplementary Table 2.



Extended Data Figure 8. SHARP, LBR, SAF-A, SMRT, and HDAC3 are required for transcriptional silencing in differentiating female ES cells

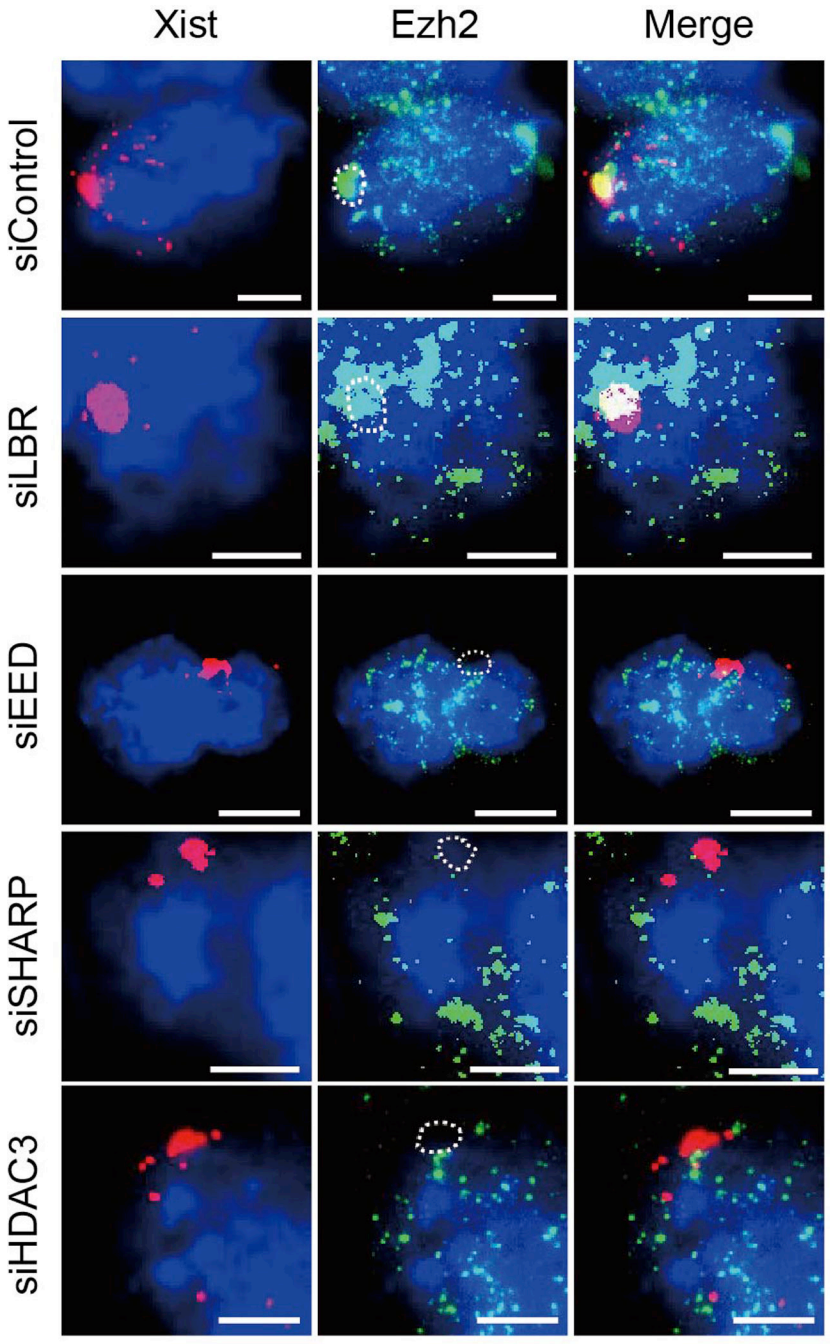
(a) Representative images showing staining of DAPI (blue), Xist (red), and Gpc4 (green) upon knockdown of specific proteins using different siRNAs in female ES cells prior to differentiation (−RA; left) or after differentiation for 24 hours (+RA; middle and right). (b) Quantification of the copy number of Gpc4 for −RA and +RA cells upon transfection with different siRNAs. Error bars represent the standard error across 50 individual cells from one experiment. NS: not significantly different between +RA and −RA cells; **** represents values with a p-value<0.001, ** represents values with a p-value<0.01, and * represents

values with a p-value<0.05 between +RA and −RA cells based on an unpaired two-sample t-test. Scale bars on the images represent 5 μm.



Extended Data Figure 9. SHARP is required for exclusion of RNA Polymerase II from the Xist-coated territory in differentiating female ES cells

Images of individual cells that are labeled with Xist (red), RNA Polymerase II (green), and DAPI (blue) across different siRNA conditions (rows) in female ES cells after 24 hours of retinoic acid treatment. The dashed white region represents the outlined Xist coated territory.



Extended Data Figure 10. SHARP is required for PRC2 recruitment across the Xist coated territory in differentiating female ES cells

Images of individual cells that are labeled with Xist (red), Ezh2 (green) and DAPI (blue) across different siRNA conditions (rows) in female ES cells after 24 hours of differentiation. The dashed white region represents the outlined Xist coated territory.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank Jesse Engreitz for extensive discussions, help in adapting the RAP method, and critical comments on the manuscript; Andi Gnirke, Steve Carr, Jake Jaffe, and Monica Schenone for initial discussions about the RAP-MS method; Andres Collazo, Eric Lubek, and Long Cai for microscopy help; Anton Wutz for providing transgenic cell lines; Roxana Eggleston-Rangel for assistance with mass spectrometry; Shari Grossman, Ido Amit, Manuel Garber, and John Rinn for comments on the manuscript and helpful suggestions; and Sigrid Knemeyer for illustrations. CAM is supported by a post-doctoral fellowship from Caltech. CKC is supported by an NIH NRSA training grant (T32GM07616). Imaging was performed in the Biological Imaging Facility, with the support of the Caltech Beckman Institute and the Arnold and Mabel Beckman Foundation. This work was funded by the Gordon and Betty Moore Foundation (GBMF775), the Beckman Institute, and NIH (1S10RR029591-01A1 to SH), an NIH Director's Early Independence Award (DP5OD012190), the Rose Hills Foundation, Edward Mallinckrodt Foundation, Sontag Foundation, Searle Scholars Program, and funds from the California Institute of Technology.

REFERENCES

1. Guttman M, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature. 2011; 477:295–300. [PubMed: 21874018]
2. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012; 81:145–166. [PubMed: 22663078]
3. Wutz A. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat Rev Genet. 2011; 12:542–553. [PubMed: 21765457]
4. Lee JT. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. Genes Dev. 2009; 23:1831–1842. [PubMed: 19684108]
5. McHugh CA, Russell P, Guttman M. Methods for comprehensive experimental identification of RNA-protein interactions. Genome biology. 2014; 15:203. [PubMed: 24467948]
6. Shi Y, et al. Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. Genes Dev. 2001; 15:1140–1151. [PubMed: 11331609]
7. You SH, et al. Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 in vivo. Nat Struct Mol Biol. 2013; 20:182–187. [PubMed: 23292142]
8. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science. 2008; 322:750–756. [PubMed: 18974356]
9. Plath K, et al. Role of histone H3 lysine 27 methylation in X inactivation. Science. 2003; 300:131–135. [PubMed: 12649488]
10. Hasegawa Y, Brockdorff N, Kawano S, Tsutui K, Nakagawa S. The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. Dev Cell. 2010; 19:469–476. [PubMed: 20833368]
11. Schoeftner S, et al. Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. The EMBO journal. 2006; 25:3110–3122. [PubMed: 16763550]
12. Kalantry S, Magnuson T. The Polycomb group protein EED is dispensable for the initiation of random X-chromosome inactivation. PLoS Genet. 2006; 2:e66. [PubMed: 16680199]
13. Engreitz JM, et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science. 2013; 341:1237973. [PubMed: 23828888]
14. Darnell RB. HITS-CLIP: panoramic views of protein-RNA regulation in living cells. Wiley interdisciplinary reviews. RNA. 2010; 1:266–286. [PubMed: 21935890]
15. Ong SE, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nature protocols. 2006; 1:2650–2660. [PubMed: 17406521]
16. Ariyoshi M, Schwabe JW. A conserved structural motif reveals the essential transcriptional repression function of Spen proteins and their role in developmental signaling. Genes Dev. 2003; 17:1909–1920. [PubMed: 12897056]
17. Raffel GD, et al. Ott1(Rbm15) has pleiotropic roles in hematopoietic development. Proc Natl Acad Sci U S A. 2007; 104:6001–6006. [PubMed: 17376872]
18. Haas S, Steplewski A, Siracusa LD, Amini S, Khalili K. Identification of a sequence-specific single-stranded DNA binding protein that suppresses transcription of the mouse myelin basic protein gene. J Biol Chem. 1995; 270:12503–12510. [PubMed: 7539003]

19. Olins AL, Rhodes G, Welch DB, Zwerger M, Olins DE. Lamin B receptor: multi-tasking at the nuclear envelope. Nucleus. 2010; 1:53–70. [PubMed: 21327105]
20. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. Nature. 2011; 469:343–349. [PubMed: 21248841]
21. 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 & development. 2006; 20:2223–2237. [PubMed: 16912274]
22. 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. 2014; 42:6742–6752. [PubMed: 24748666]
23. Li J, Lin Q, Wang W, Wade P, Wong J. Specific targeting and constitutive association of histone deacetylase complexes during transcriptional repression. Genes Dev. 2002; 16:687–692. [PubMed: 11914274]
24. Keohane AM, O'Neill LP, Belyaev ND, Lavender JS, Turner BM. X-Inactivation and histone H4 acetylation in embryonic stem cells. Dev Biol. 1996; 180:618–630. [PubMed: 8954732]
25. Riising EM, et al. Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. Mol Cell. 2014; 55:347–360. [PubMed: 24999238]
26. van der Vlag J, Otte AP. Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. Nat Genet. 1999; 23:474–478. [PubMed: 10581039]
27. Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. Nat Struct Mol Biol. 2013; 20:1250–1257. [PubMed: 24077223]
28. Fackelmayer FO, Dahm K, Renz A, Ramsperger U, Richter A. Nucleic-acid-binding properties of hnRNP-U/SAF-A, a nuclear-matrix protein which binds DNA and RNA in vivo and in vitro. European journal of biochemistry / FEBS. 1994; 221:749–757. [PubMed: 8174554]
29. Wang Z, et al. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell. 2009; 138:1019–1031. [PubMed: 19698979]
30. Kuo MH, Allis CD. Roles of histone acetyltransferases and deacetylases in gene regulation. Bioessays. 1998; 20:615–626. [PubMed: 9780836]

References

31. Wutz A, Rasmussen TP, Jaenisch R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nature genetics. 2002; 30:167–174. [PubMed: 11780141]
32. Engreitz JM, et al. RNA-RNA Interactions Enable Specific Targeting of Noncoding RNAs to Nascent Pre-mRNAs and Chromatin Sites. Cell. 2014; 159:188–199. [PubMed: 25259926]
33. Kalli A, Hess S. Effect of mass spectrometric parameters on peptide and protein identification rates for shotgun proteomic experiments on an LTQ-orbitrap mass analyzer. Proteomics. 2012; 12:21–31. [PubMed: 22065615]
34. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized pp**b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008; 26:1367–1372. [PubMed: 19029910]
35. Cox J, et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. Journal of proteome research. 2011; 10:1794–1805. [PubMed: 21254760]
36. Elias JE, Gygi SP. Target-decoy search strategy for mass spectrometry-based proteomics. Methods in molecular biology. 2010; 604:55–71. [PubMed: 20013364]
37. Geiger T, et al. Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. Nature protocols. 2011; 6:147–157. [PubMed: 21293456]
38. Finn RD, et al. Pfam: the protein families database. Nucleic Acids Res. 2014; 42:D222–D230. [PubMed: 24288371]
39. Jeon Y, Lee JT. YY1 Tethers Xist RNA to the Inactive X Nucleation Center. Cell. 2011; 146:119–133. [PubMed: 21729784]
40. Agrelo R, et al. SATB1 defines the developmental context for gene silencing by Xist in lymphoma and embryonic cells. Developmental cell. 2009; 16:507–516. [PubMed: 19386260]

41. Royce-Tolland ME, et al. The A-repeat links ASF/SF2-dependent Xist RNA processing with random choice during X inactivation. Nat Struct Mol Biol. 2010; 17:948–954. [PubMed: 20657585]

42. Brown CJ, Baldry SE. Evidence that heteronuclear proteins interact with XIST RNA in vitro. Somatic cell and molecular genetics. 1996; 22:403–417. [PubMed: 9039849]

43. Sarma K, et al. ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. Cell. 2014; 159:869–883. [PubMed: 25417162]

44. Theodosiou Z, et al. Automated analysis of FISH and immunohistochemistry images: a review. Cytometry. Part A : the journal of the International Society for Analytical Cytology. 2007; 71:439–450. [PubMed: 17559119]

45. Fumagalli M, et al. Telomeric DNA damage is irreparable and causes persistent DNA-damage-response activation. Nat Cell Biol. 2012; 14:355–365. [PubMed: 22426077]

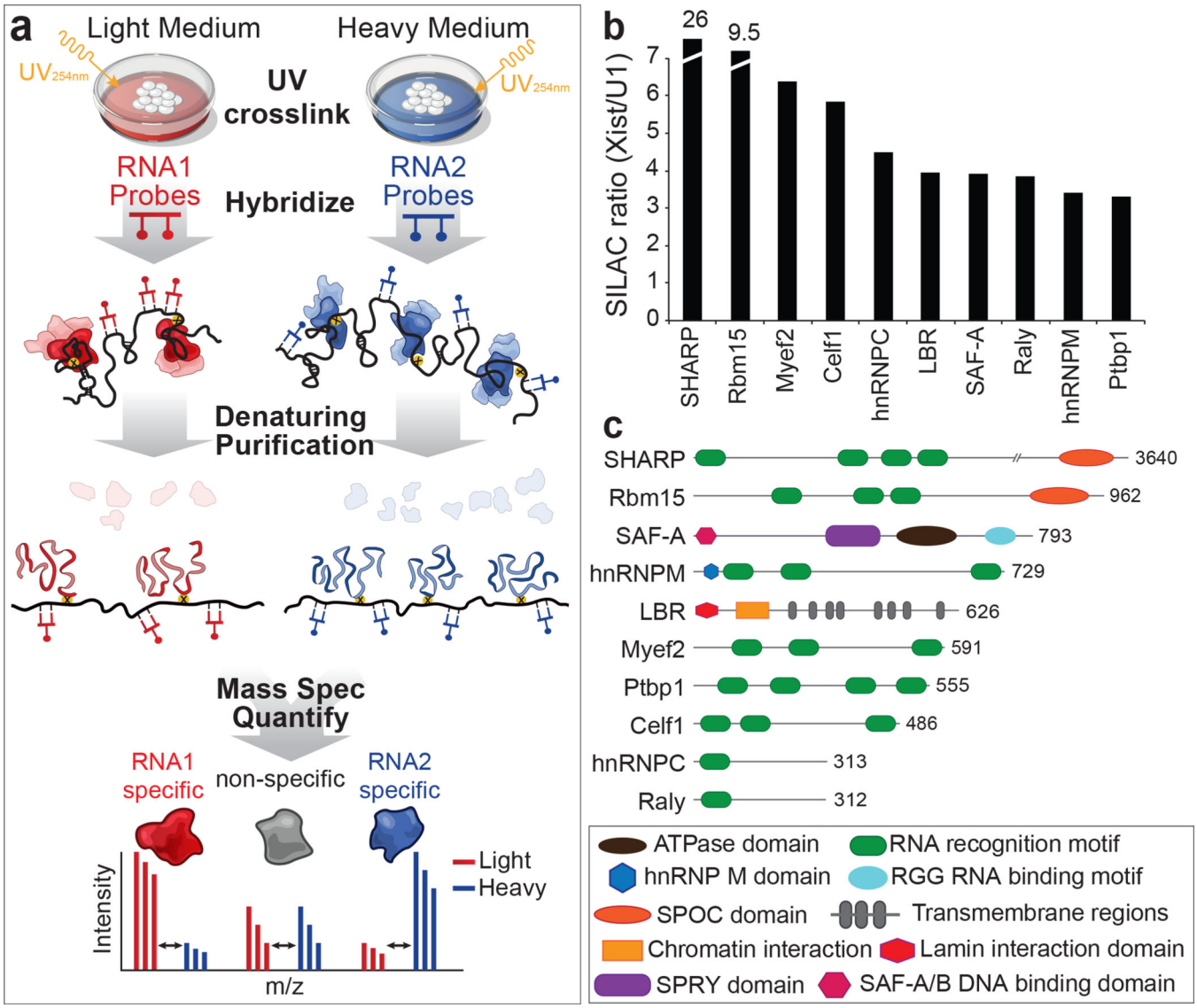


Figure 1. RAP-MS identifies direct Xist-interacting proteins

(a) A schematic overview of the RAP-MS method. (b) The SILAC ratio (Xist/U1) for each Xist-enriched protein identified by RAP-MS for one representative sample of four biological replicates. For SHARP and RBM15, the enrichment values are indicated above their bars. (c) Each Xist-interacting protein is shown (scaled to protein length). The locations of functional domains are shown.

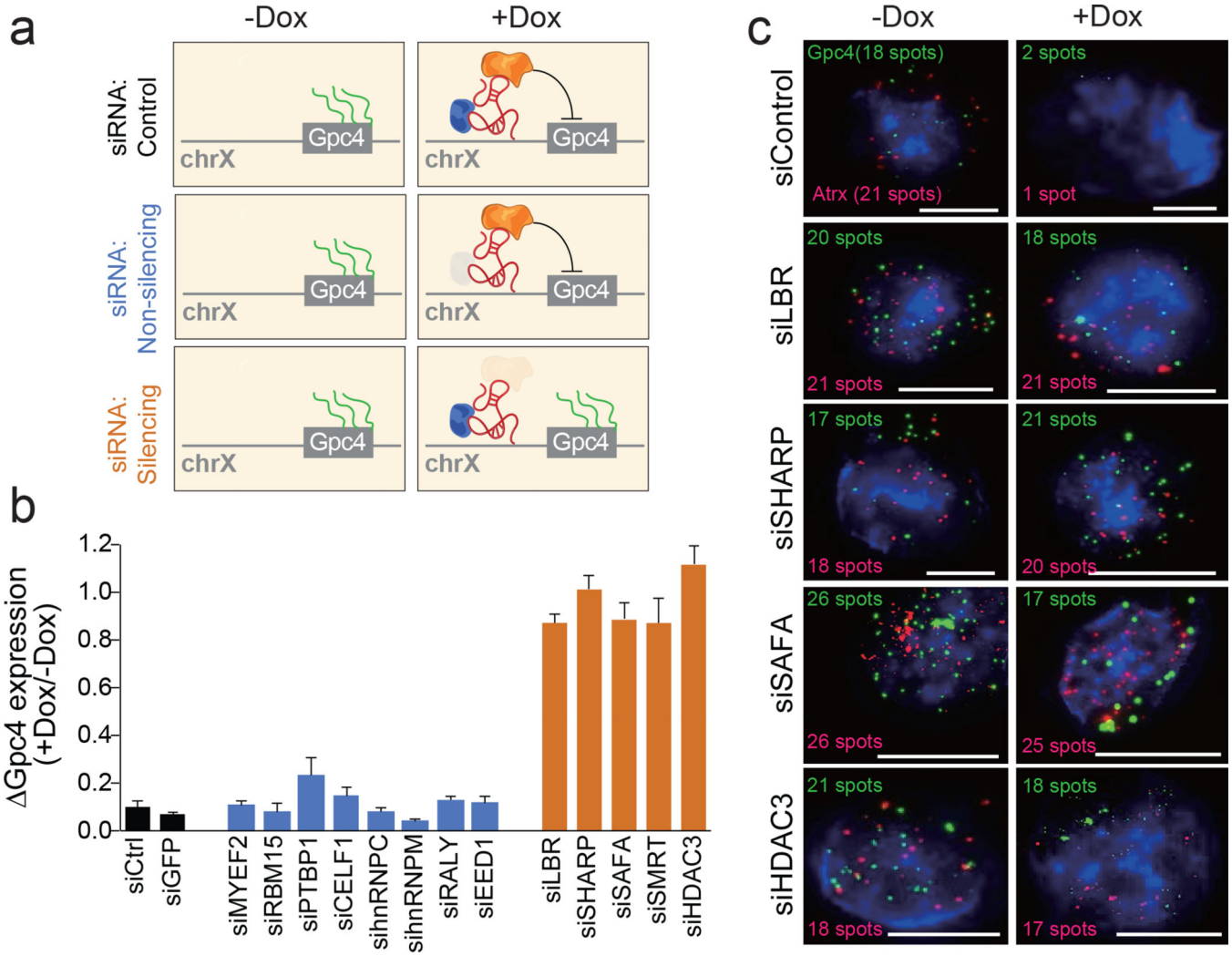


Figure 2. SHARP, LBR, and SAF-A are required for Xist-mediated gene silencing

(a) Screen for Xist-mediated gene silencing for knockdown of control (top), non-silencing proteins (middle), or silencing proteins (bottom). (b) Gpc4 mRNA levels after induction of Xist (+dox) normalized to Gpc4 levels before Xist induction (−dox). Error bars: standard error of the mean across 50 cells from one experiment. siCtrl: scrambled siRNA control. (c) Images of individual cells for two X-linked mRNAs, Gpc4 (green) and Atrx (red), and DAPI (blue) after treatment with different siRNAs (rows). The number of identified mRNAs is shown. Scale bars, 5 micrometers.

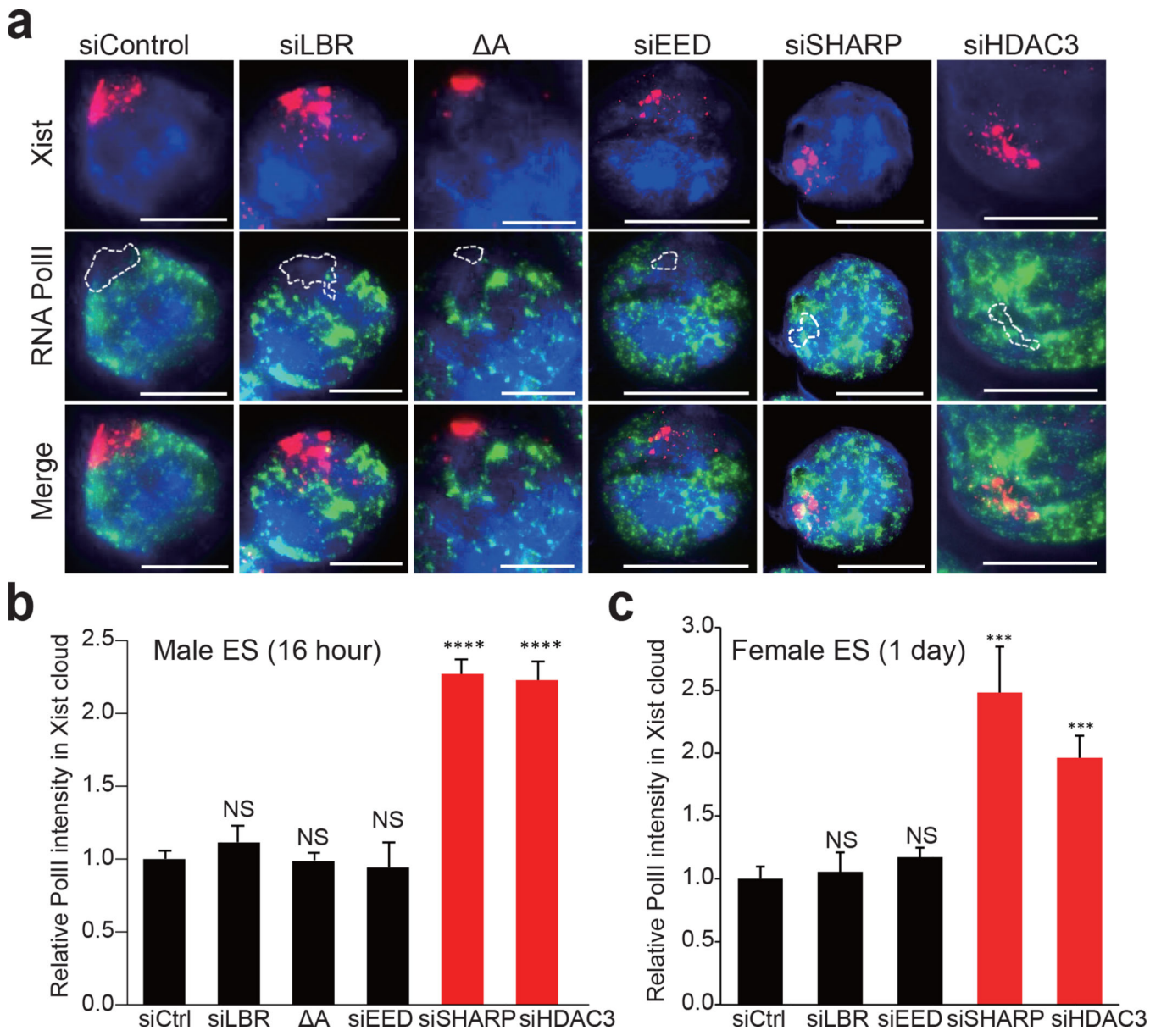


Figure 3. SHARP is required for exclusion of PolII from the Xist-coated territory

(a) Xist (red), PolII (green), and DAPI (blue) across different siRNA conditions (rows). Quantification of fluorescence intensity of PolII within Xist territory normalized to control siRNA levels for (b) male ES cells after 16 hours of doxycycline treatment and (c) female ES cells after 1 day of retinoic acid induced differentiation. Error bars: standard error of the mean across 50 cells from one experiment. NS: not significant, **** *p*-value<0.001 relative to siControl by unpaired two-sample t-test. ΔA: Genetic deletion of A-repeat. Scale bars, 5 micrometers.

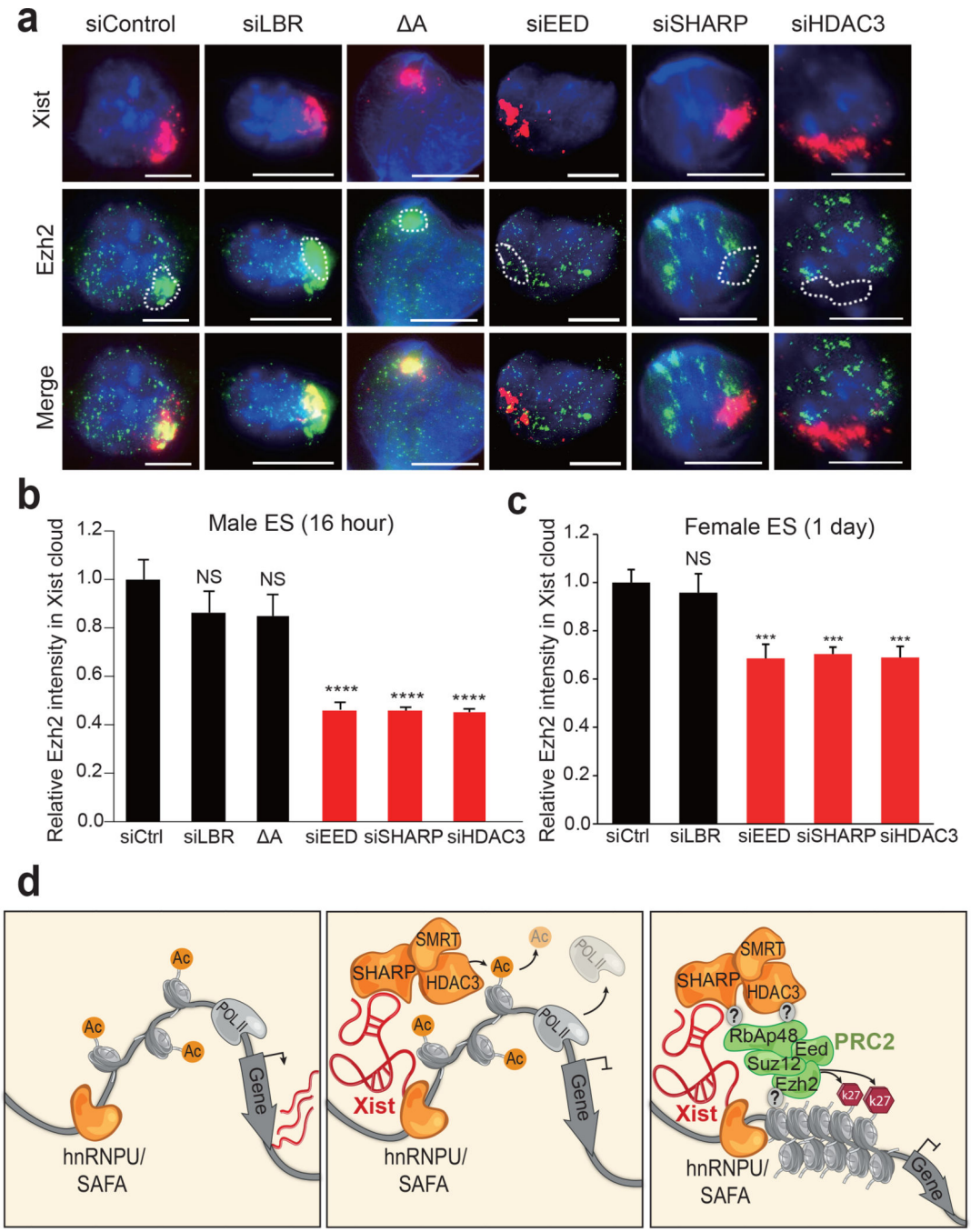


Figure 4. SHARP is required for PRC2 recruitment across the Xist-coated territory

(a) Xist (red), Ezh2 (green) and DAPI (blue) across siRNA conditions (rows). Quantification of Ezh2 levels within the defined Xist territory normalized to the levels in the control siRNA sample for (b) male ES cells and (c) differentiating female ES cells. Error bars: standard error of the mean across 50 cells from one experiment. NS: not significant, ***, *p*-value<0.005, **** *p*-value<0.001 relative to siControl by an unpaired two-sample t-test.

Scale bars, 5 micrometers. (d) A model for Xist-mediated transcriptional silencing and recruitment of PRC2 across the X-chromosome.