doi:10.1038/nature14443

Note 1: RAP-MS identifies known U1 snRNA interacting proteins.

We validated the RAP-MS approach by defining the proteins that interact with two well-characterized non-coding RNAs: (i) U1 small nuclear RNA, a core component of the spliceosome[1] and (ii) 18S ribosomal RNA, a component of the small ribosomal subunit[2]. In the U1 purifications, we identified 9 enriched proteins, all of which are known to interact with the U1 snRNA. The list includes 7 of the 10 proteins that comprise the core U1 snRNP complex (U1-A, U1-C, U1-70K, Sm-B, Sm-D2, Sm-D3, Sm-E)[3] as well as the Gemin5 processing factor involved in U1 snRNP biogenesis[4] **(Extended Data 2)**. The ninth enriched protein, SF3a1, had not previously been identified as a U1-interacting protein but was recently shown to interact directly with the U1 snRNA *in vivo*[5].

Note 2: Validation of the Xist interacting proteins identified by RAP-MS

To confirm that the identified proteins reflect specific interactions with Xist, and are not due to background proteins or non-specific purification of other RNAs, we performed RAP using the same Xist probes in uninduced cells in which Xist is not expressed. Furthermore, to confirm that the identified interactions represent proteins that are crosslinked to Xist *in vivo* rather than interactions that form in solution, we purified Xist from cells that were not crosslinked (no UV light). In both cases, we identified none of these 10 Xist-interacting proteins, nor any other specifically enriched proteins in either of these control samples (**Methods**). Together, these results demonstrate that the identified proteins are direct RNA binding proteins that are covalently crosslinked with Xist in cells.

To confirm that these proteins do not merely associate non-specifically with any nuclear long ncRNA, we compared Xist to the 45S pre-ribosomal RNA, which is of comparable length to Xist (~13,000 vs. ~17,000 nucleotides, respectively) (**Methods, Extended Data 2**). Importantly, each of the 10 proteins that were enriched when comparing Xist to U1 was still enriched when comparing Xist to 45S. In three cases (hnRNPC, RALY, and LBR) the enrichment level was only ~2-fold because these proteins had higher levels in the 45S purification, consistent with the fact that they are known to be present in the

nucleolus[6]. These results demonstrate that these 10 proteins associate with Xist specifically and not merely with any long RNA in the nucleus.

Finally, we independently validated these interactions by testing whether we could enrich the Xist RNA upon immunoprecipitation of the identified proteins. To do this, we obtained high-quality IP-grade antibodies or epitope-tagged proteins for 8 of the 10 proteins (Ptbp1, hnRNPC, CELF1, Myef2, Rbm15, LBR, RALY, and SHARP) and purified protein-RNA complexes from UV-crosslinked lysate (**Methods**). In all cases, we observed a strong enrichment for the Xist RNA relative to total input RNA levels (>4-fold, **Extended Data 3, Methods**). In contrast, we did not observe similar enrichment for other control RNAs – including mRNAs (i.e. Oct4, Nanog, or Stat3) or lncRNAs (i.e. Neat1, Malat1, Tug1, or Firre) (**Extended Data 3**). For the remaining 2 proteins, we were unable to identify antibodies or generate affinity reagents that could be used to independently validate their interactions. In one case (SAF-A), the protein has been previously shown to directly interact with Xist in human cells[7] – providing independent confirmation. For the remaining protein (hnRNPM), future work will be needed to provide independent confirmation of its interaction with Xist.

Note 3: Xist silencing defects are not due to ES cell differentiation

Since previous studies have shown that Xist can no longer initiate transcriptional silencing after a certain critical window during differentiation[8], we wanted to ensure that the loss of Xist silencing upon knock down of SHARP, LBR, and SAF-A was not merely due to cellular differentiation. To address this, we performed single molecule FISH for Gpc4 mRNA along with immunofluorescence for Nanog, a marker of the pluripotent state that is rapidly lost upon differentiation[9]. We confirmed that knockdown of SHARP, LBR, or SAF-A also abolished gene silencing on the X-chromosome in Nanog-positive cells (**Extended Data 5c**).

Note 4: Additional characterization of SHARP

Having identified a critical role for SHARP in PolII exclusion, we sought to further confirm the functional importance of SHARP. (i) We confirmed that SHARP is required for silencing additional X-chromosome genes by selecting three additional genes (Rbmx, Mecp2, and Smc1a) that are silenced at different times during the induction of XCI[10]. Knockdown of SHARP abolished the silencing of all 3 additional X-chromosome genes. In contrast, knockdown of SHARP had no impact on the expression of two X-chromosome genes (Mid1 and Pir) that are known to escape XCI[11] (**Extended Data 6)**. (ii) We confirmed that SHARP similarly interacts with Xist in differentiating female ES cells. To do this, we purified SHARP from lysates of UV-crosslinked RA-treated female ES cells and identified a strong enrichment for Xist (>45-fold) but not for Neat1 or 45S (<1-fold) relative to levels in IgG (**Methods, Extended Data 3b).** (iii) We confirmed that SHARP, but not LBR, is similarly required for PolII exclusion in differentiating female ES cells. We knocked down SHARP, LBR, and several controls in female ES cells and induced Xist expression through RA-treatment and identified higher levels of PolII localization over the Xist-coated territory upon knock down of SHARP, but not the other proteins (**Figure 3c, Extended Data 9**).

Note 5: Possible models for SHARP-mediated recruitment of PRC2

While it is clear that SHARP and HDAC3 are required for the recruitment of PRC2, whether this is due to direct recruitment or indirect recruitment remains unclear. Previous studies suggest several possible mechanisms: (i) PolII exclusion has been shown to be sufficient to trigger PRC2 recruitment in other contexts[12] and because SHARP is required for PolII exclusion on the X-chromosome, this might indirectly lead to PRC2 recruitment. (ii) Previous studies have shown that the PRC2 complex can interact with various HDAC complexes[13] and accordingly PRC2 might be recruited directly by Xist through the HDAC3 complex. (iii) Chromatin compaction has been shown to be sufficient to mediate PRC2 recruitment[14] because HDAC3 can lead to chromatin compaction this may indirectly lead to PRC2 recruitment. (iv) SHARP has been shown to interact *in vitro* with RbAp48[15], a component of several chromatin regulatory complexes including the PRC2[16,17] and HDAC3 complexes[18,19] and therefore Xist might directly recruit PRC2

through an interaction between SHARP or HDAC3 and RbAp48 and the PRC2 complex. Future work will be needed to distinguish between these possible mechanisms.

References

1 Mount, S. M., Pettersson, I., Hinterberger, M., Karmas, A. & Steitz, J. A. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. *Cell* **33**, 509-518 (1983).

2 Steitz, T. A. A structural understanding of the dynamic ribosome machine. *Nat Rev Mol Cell Biol* **9**, 242-253 (2008).

3 Pomeranz Krummel, D. A., Oubridge, C., Leung, A. K., Li, J. & Nagai, K. Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. *Nature* **458**, 475-480 (2009).

4 Yong, J., Kasim, M., Bachorik, J. L., Wan, L. & Dreyfuss, G. Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol Cell* **38**, 551-562 (2010).

5 Sharma, S., Wongpalee, S. P., Vashisht, A., Wohlschlegel, J. A. & Black, D. L. Stem-loop 4 of U1 snRNA is essential for splicing and interacts with the U2 snRNP-specific SF3A1 protein during spliceosome assembly. *Genes Dev* **28**, 2518-2531 (2014).

6 Jarboui, M. A., Wynne, K., Elia, G., Hall, W. W. & Gautier, V. W. Proteomic profiling of the human T-cell nucleolus. *Molecular immunology* **49**, 441-452 (2011).

7 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* **19**, 469-476 (2010).

8 Wutz, A. & Jaenisch, R. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Molecular cell* **5**, 695-705 (2000).

9 Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-642 (2003).

10 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* **20**, 2223-2237 (2006).

11 Yang, F., Babak, T., Shendure, J. & Disteche, C. M. Global survey of escape from X inactivation by RNA-sequencing in mouse. *Genome Res* **20**, 614-622 (2010).

12 Riising, E. M. *et al.* Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol Cell* **55**, 347-360 (2014).

13 van der Vlag, J. & Otte, A. P. Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. *Nat Genet* **23**, 474-478 (1999).

14 Yuan, W. *et al.* Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. *Science* **337**, 971-975 (2012).

15 Shi, Y. *et al.* Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev* **15**, 1140-1151 (2001).

16 Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343-349 (2011).
17 Simon, J. A. & Kingston, R. E. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol* **10**, 697-708 (2009).
18 Nicolas, E., Ait-Si-Ali, S. & Trouche, D. The histone deacetylase HDAC3 targets RbAp48 to the retinoblastoma protein. *Nucleic Acids Res* **29**, 3131-3136 (2001).
19 Zhang, Y. *et al.* Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev* **13**, 1924-1935 (1999).

Supplementary Table 1: A list of all antibodies used for immunoprecipitation experiments.

Protein	Epitope	Vendor	Catalog #
RALY	V5-tagged clone	Sigma	V8137
LBR	V5-tagged clone	Sigma	V8137
hnRNPC	V5-tagged clone	Sigma	V8137
RBM15	Endogenous	Santa Cruz	sc-366873
PTBP1	Endogenous	Abcam	ab5642
CELF1	Endogenous	Abcam	ab129115
PUM1	Endogenous	Santa Cruz	sc-135049
MYEF2	Endogenous	Santa Cruz	sc-102031
hnRNPH	Endogenous	Bethyl	A300-511A
IgG	None	Cell Signaling	2729S
SHARP	Endogenous	Novus	NBP1-82952

Supplementary Table 2: A list of all siRNAs used for knockdown experiments.

Description	Company	Catalog number
SMARTpool: ON-TARGETplus Hnrnpc siRNA	Dharmacon	L-044147-01-0005
SMARTpool: ON-TARGETplus Ncor2 siRNA	Dharmacon	L-044147-01-0005
SMARTpool: ON-TARGETplus Hdac3 siRNA	Dharmacon	L-043553-02-0005
SMARTpool: ON-TARGETplus Spen siRNA	Dharmacon	L-062019-01- 0005
SMARTpool: ON-TARGETplus Rbm15 siRNA	Dharmacon	L-048728-01- 0005
SMARTpool: ON-TARGETplus Lbr siRNA	Dharmacon	L-051330-01- 0005
SMARTpool: ON-TARGETplus Ptbp1 siRNA	Dharmacon	L-042865-01- 0005
SMARTpool: ON-TARGETplus Hnrnpu siRNA	Dharmacon	L-051574-01- 0005
SMARTpool: ON-TARGETplus Myef2 siRNA	Dharmacon	L-058553-01- 0005
SMARTpool: ON-TARGETplus YY1 siRNA	Dharmacon	L-050273-00-0005
SMARTpool: ON-TARGETplus Celf1 siRNA	Dharmacon	L-064577-01-0005
SMARTpool: ON-TARGETplus Raly siRNA	Dharmacon	L-044852-02-0005
SMARTpool: ON-TARGETplus Hnrnpm siRNA	Dharmacon	L-044465-01-0005
SMARTpool: ON-TARGETplus Atrx siRNA	Dharmacon	L-046292-01-0005
SMARTpool: ON-TARGETplus Satb1 siRNA	Dharmacon	L-045547-01-0005
SMARTpool: ON-TARGETplus Eed siRNA	Dharmacon	L-049898-00-0005
SMARTpool: ON-TARGETplus Srsf1 siRNA	Dharmacon	L-040886-01-0005
ON-TARGETplus Non- targeting Pool	Dharmacon	D-001810-10- 05
Silencer Select Pre-Designed siRNA: Ncor2	Ambion/Life Technologies	Assay Id s74030
Silencer Select Pre-Designed siRNA: Spen	Ambion/Life Technologies	Assay Id s80456
FlexiTube GeneSolution GS56381 for Spen	Qiagen	GS56381
FlexiTube GeneSolution GS98386 for Lbr	Qiagen	GS98386
FlexiTube GeneSolution GS51810 for Hnrnpu	Qiagen	GS51810
FlexiTube GeneSolution GS20602 for Ncor2	Qiagen	GS20602
GFP siRNA (1 nmol)	Qiagen	SI04380467

Individual siRNA deconvoluted from the pool

Description	Company	Catalog number
SMARTpool: ON-TARGETplus Spen siRNA Upgrade	Dharmacon	LU-062019-01-0002
D1: CGAGAGGGAGAGACGAAUA D2: CUAAAGAGCCGGAGCCGAA D3: CCUAAAAUCACGUCGGUUA D4: GGAAACACCUCAAGGCCGA		
SMARTpool: ON-TARGETplus Lbr siRNA Upgrade	Dharmacon	LU-051330-01-0002
D1: UGUUGAAGCCGUUCGGAAA D2: AUACAAAGAUGGCACCGAA D3: AUAAACACAUAGACGACUU		

D4: GUACUAGUGAGGUUGGAUA		
SMARTpool: ON-TARGETplus Hdac3 siRNA Upgrade	Dharmacon	LU-043553-02-0002
D1: GGGAAUGUGUUGAAUAUGU D2: CGGCAGACCUCCUGACGUA D3: GCACCCGCAUCGAGAAUCA D4: UAUAAGAAGAUGAUCGUCU		
FlexiTube GeneSolution for Spen	Qiagen	GS56381
Q1: CAGGAGCATTTGATCGGACAA Q2: CCCGAGCATCGTCACCACCAA Q3: CAGGAAGACTAACAAGAGCAA Q4: CACGAGGGAAGGTAACCCTAA		
FlexiTube GeneSolution for Lbr	Qiagen	GS98386
Q1: CAGCCCAATATTGGTGTTATA Q2: TACAATTCATTTAAAGTTAAA Q3: AAAGTGCATTTGGAAATATAA Q4: GAGGGAAATTCTGCAAGTTAA		
FlexiTube GeneSolution for Hnrnpu	Qiagen	GS51810
Q1: CACAGTGTCTTGGCAAGTTTA Q2: CACCAAGGATATTATTGAATA Q3: ACGGGTATTCTCTGAAAGGAA Q4: CAGTAAGACACTTATATACAA		
FlexiTube GeneSolution for Ncor2	Qiagen	GS20602
Q1: CGCATTTGGAACCAAAGTCTA Q2: CAGGCCTTATGACCTGTAGAA Q3: CCGGAACGAGCCAGAATACAA Q4: TACGATGAGAACCGGAAGAAA		
Silencer Select Pre-Designed siRNA: Ncor2	Ambion/Life Technologies	Assay Id s74030
Sense: CGUUCUCUGGGUUACCACAtt Antisense: UGUGGUAACCCAGAGAACGga		