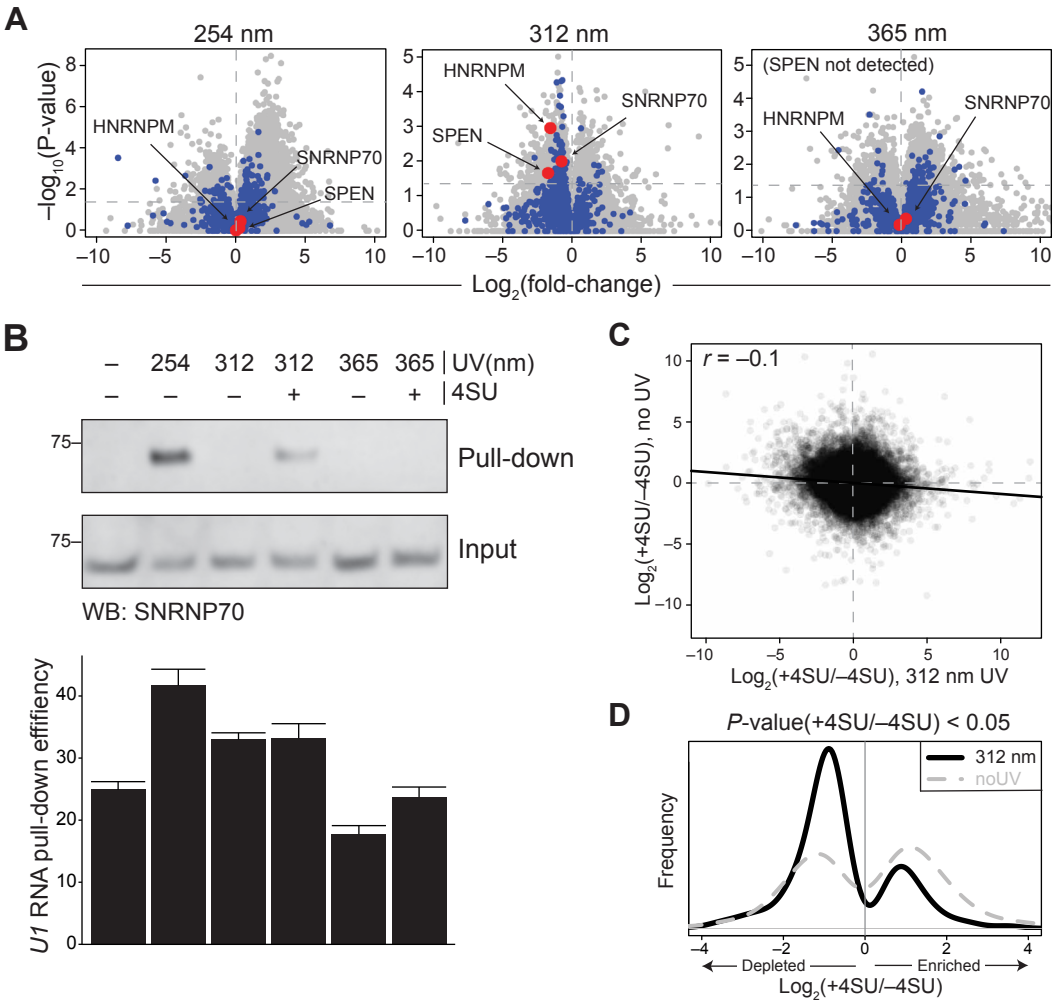

Figure S1. Comparison of Crosslinking Conditions, Related to Figure 1

(A) Volcano plots as in Figure 1D, highlighting the position of RRM-spanning peptides from three additional known RBPs.

(B) Mouse ESCs were treated or not with 4SU and irradiated with UV at the indicated wavelengths. After lysis U1 RNA-containing complexes were purified with specific U1 probes in denaturing conditions to only recover crosslinked protein, following the ChIRP protocol (Chu *et al*., 2015). Co-purifying SNRNP70 was detected by western blot (top panels) and comparable efficiencies of RNA pull-downs are shown by RT-qPCR, as a control (bottom panel). Bars represent the mean of three technical replicates + s.e.m.

(C) Scatter plot comparing 4SU-dependent depletion of peptides after 312 nm UV crosslinking (*x* axis) or no UV irradiation (*y* axis). The regression line is shown in black and the Pearson correlation score is indicated. Data is from six replicates.

(D) Density plot showing the distribution of log-fold changes in spectral intensity for peptides with consistent ($P < 0.05$) differences between +4SU and –4SU in the 312 nm irradiation conditions (black line) or no UV conditions (dashed gray line).

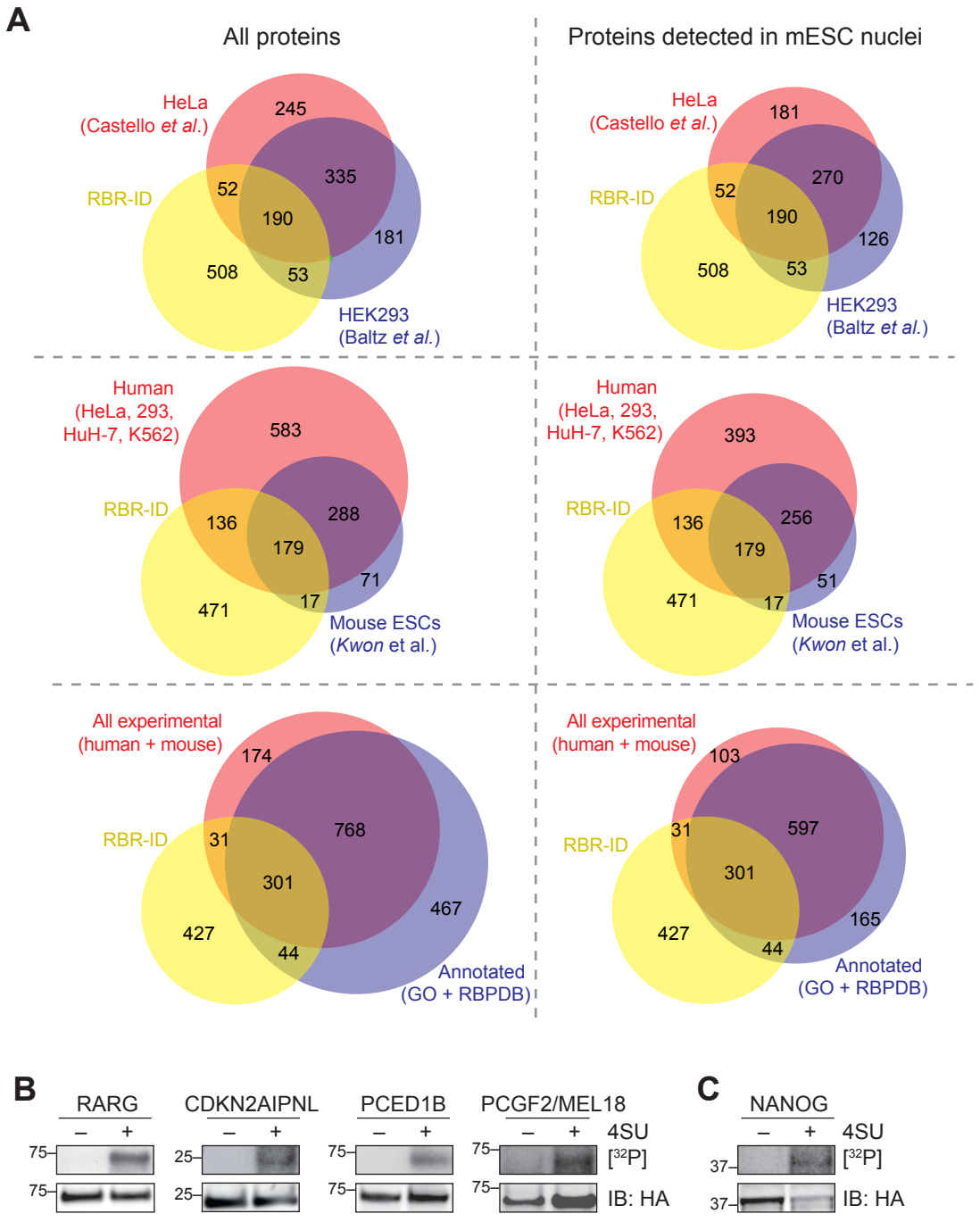

Figure S2. Analysis and Validation of RBP Candidates Predicted by RBR-ID, Related to Figure 2

(A) Venn diagrams for the overlaps are shown. On the left ("All proteins"), all protein identifiers were considered; on the right only those proteins that had at least one spectrum in all RBR-ID runs (including different UV wavelengths shown in Figure 1) were considered, with the rationale that no overlap could be calculated for proteins that were not expressed/nuclear/detectable in ESCs.

(B) Validation by PAR-CLIP of four unknown RBPs . Epitope-tagged candidates were expressed in HEK 293 cells and crosslinked to RNA using 365 nm UV. After pull-down, RNA was end-labeled, complexes resolved on SDS-PAGE and revealed by autoradiography (top) or western blot for loading control (bottom).

(C) Validation as in (B) but using 312 nm UV in mouse ESCs.

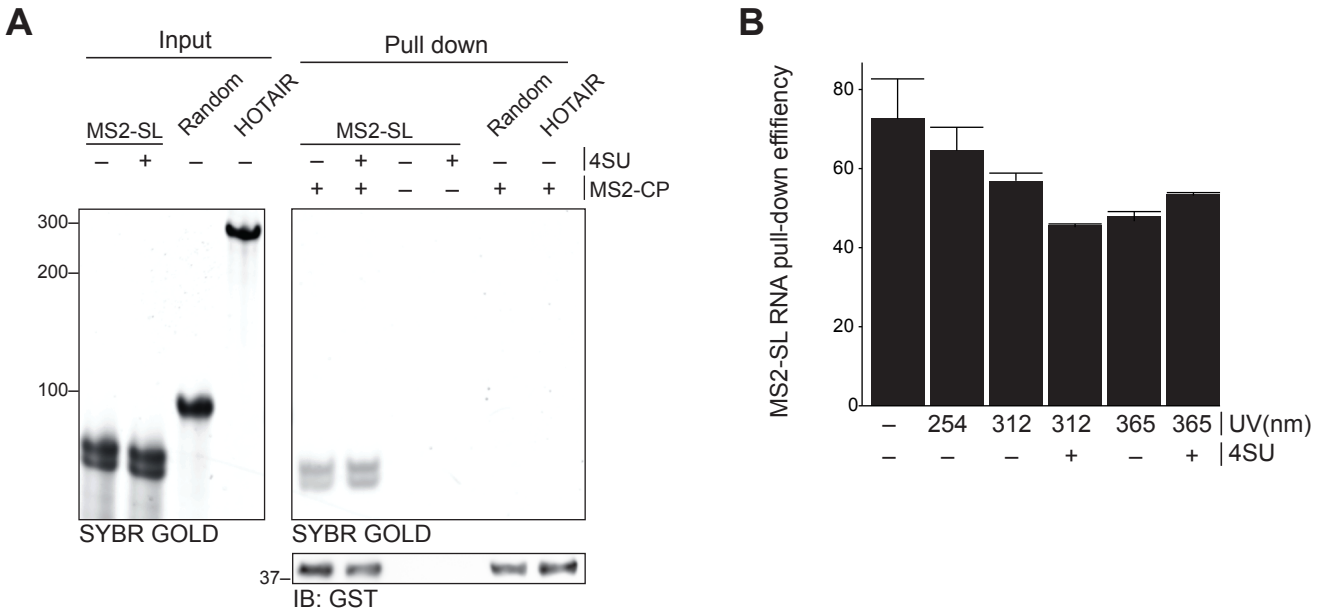

Figure S3. Controls for *In Vitro* MS2 Pull-Downs, Related to Figure 3

(A) MS2-CP protein pull-down with glutathione beads in non-denaturing conditions, showing its specific interaction with MS2-SL, regardless of the incorporation of 4SU. A random 100 nts RNA as well as a 5'-terminal fragment of the lncRNA HOTAIR were used as specificity controls. Inputs are shown to the left for RNA and at the bottom for the GST-MS2-CP fusion protein.

(B) RT-qPCR for *in vitro*-transcribed MS2-SL RNA after crosslinking to MS2-CP in the indicated condition and hybridized with biotinylated MS2-SL DNA probe followed by streptavidin pull-down from the experiment shown in Figure 3. Data is shown as an RNA pull-down efficiency control. Bars represent the mean of three technical replicates + s.e.m.

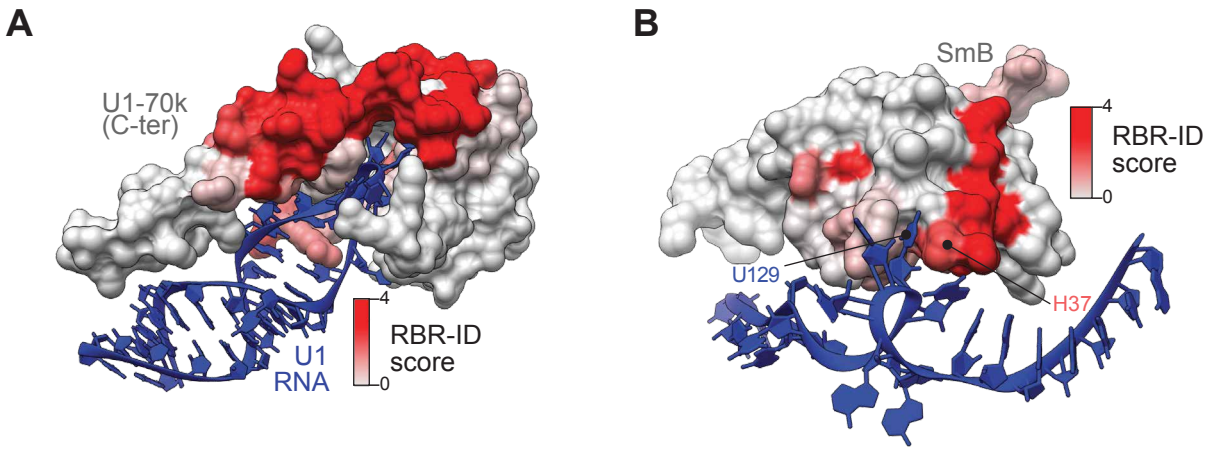

Figure S4. Additional Examples of RBR Mapping to Known Crystal Structures, Related to Figure 4

(A) Zoomed-in regions of the crystal structure of the C-terminal RRM of U1-70K bound to stem loop II of U1 snRNA (PDB ID: 4PKD (Kondo *et al*., 2015)) showing protein surfaces color-coded according to their single-residue RBR-ID score.

(B) Same as (A) but for the SmB protein (PDB ID: 4PJO).

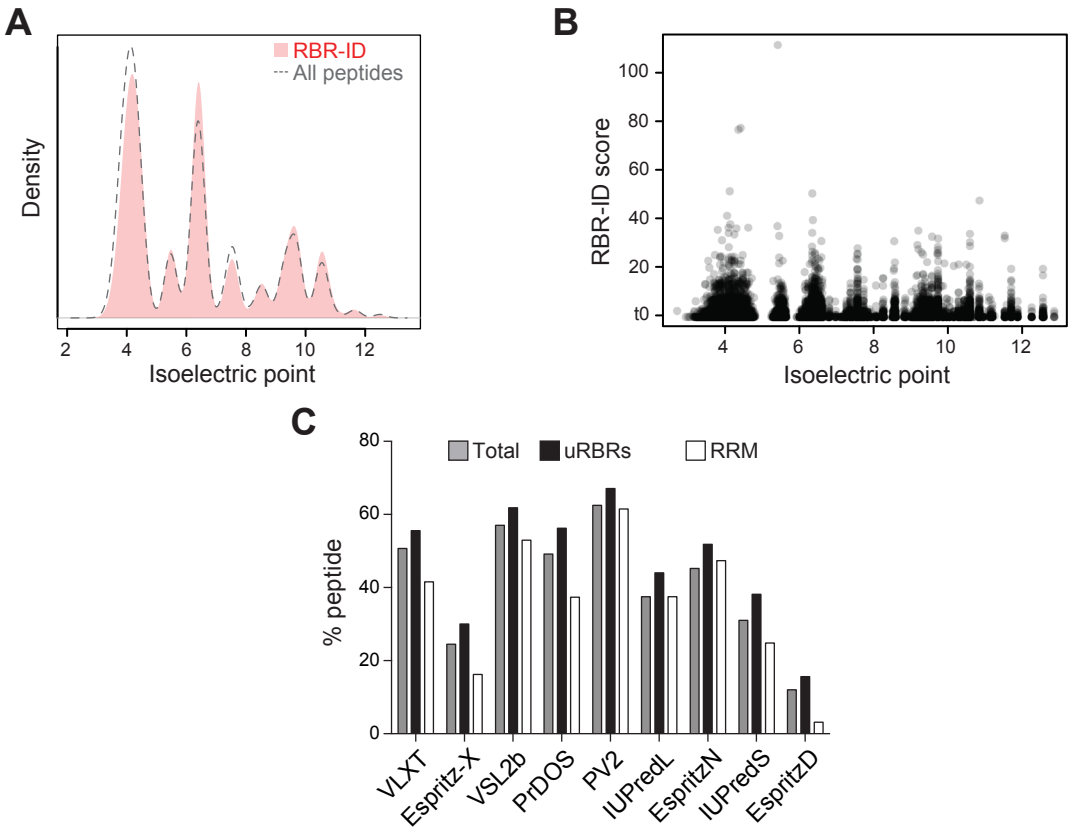

Figure S5. Isolectric Point and Disorder Predictions for RNA-Binding Peptides, Related to Figure 5

(A) Density plot of the isoelectric point distributions for all peptides (dashed) or newly identified peptides within RBRs (red fill). Values are the same as for Figure 5F.

(B) Correlation plot for the isoelectric point and the RBR-ID score for each detected peptide.

(C) Percentage of peptides overlapping with disordered regions from various predictors, compiled by Oates et al. in the D2P2 database (Oates *et al*., 2013). Values are shown for all detected peptides (tot), all top-tier RBR-ID peptides not mapping to a known RNA-binding domain (uRBRs), and all peptides overlapping RRM domains.

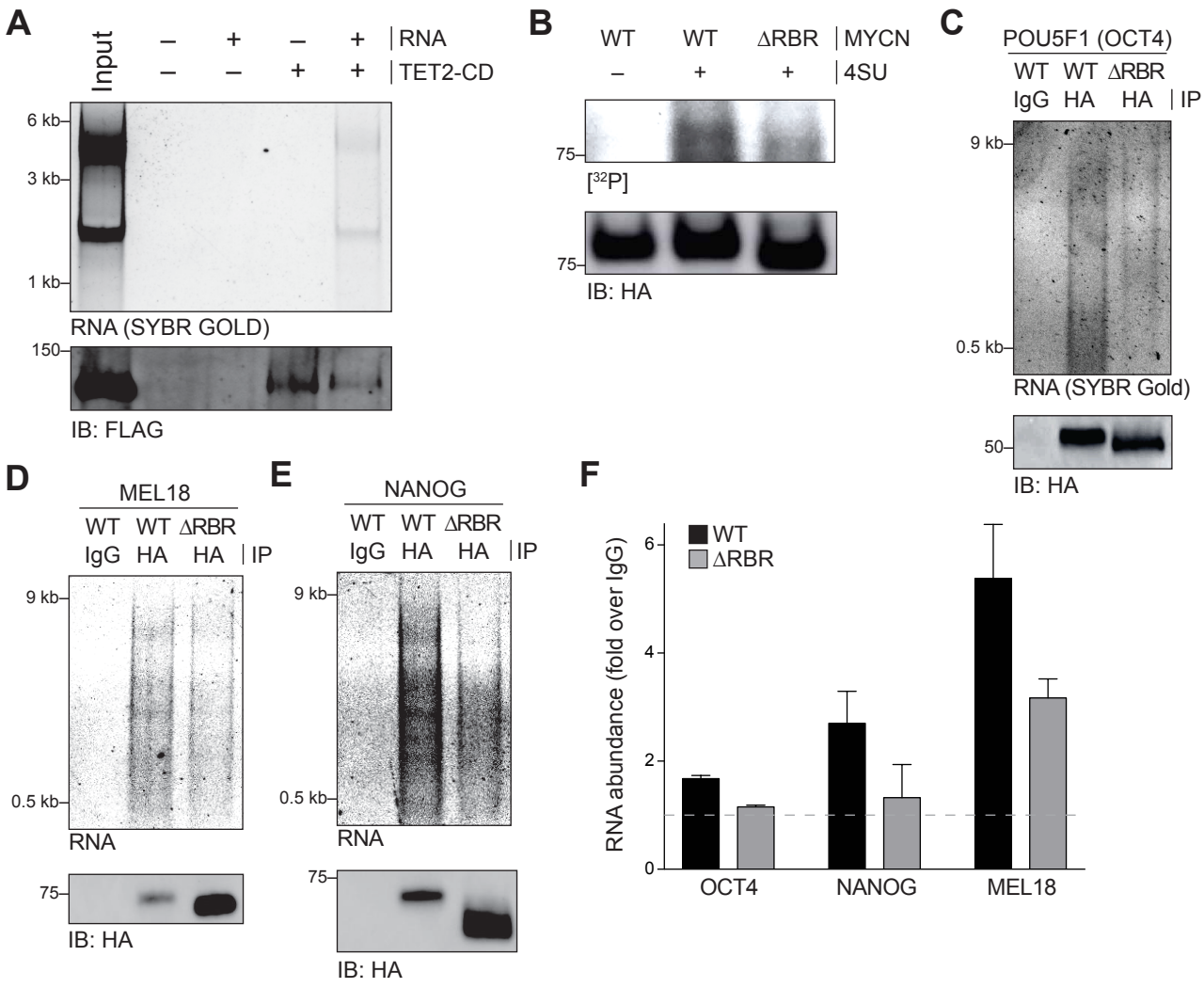

Figure S6. Additional *In Vitro* and *In Vivo* Validations for RBR-ID Candidates, Related to Figure 6

(A) A recombinant FLAG-tagged fragment encompassing the TET2 catalytic domain and RBR (TET2-CD) was incubated with total RNA from mouse ESCs and immunoprecipitated with anti-FLAG-conjugated agarose beads. The RNA fraction was purified and detected on a denaturing gel with SYBR gold (top). The protein pull-down control is shown at the bottom by western blot.

(B) PAR-CLIP for MYCN in presence or absence of 4SU for the WT sequence or a mutant lacking the predicted RBR (ΔRBR). Autoradiography of crosslinked RNA (top) and western blot for loading control (bottom) are shown.

(C–E) The indicated proteins either WT or lacking the RBR predicted by RBR-ID were fused to HA and overexpressed in mouse ESCs. After RNA-IP in native conditions, the RNA (top panel) and protein fractions (bottom panels) were separated and resolved on gels.

(F) Densitometry on RIP signal from gels shown in (C–E) and additional replicates. The intensity of the SYBR Gold signal in the WT and ΔRBR was normalized to the respective background intensity (IgG lane), shown as a dashed line. Bars show average of two biological replicates + s.e.m.

SUPPLEMENTAL TABLE LEGENDS

Table S1. List of peptides depleted after 4SU and 312 nm UV crosslinking, Related to Figure 2

Table S2. List of RBR-ID proteins, Related to Figure 2

Table S3. Gene ontology analysis for primary RBR-ID candidates, Related to Figure 2

Table S4. Gene ontology analysis of known RBPs not detected by RBR-ID, Related to Figure 2

Table S5. Gene ontology analysis of unknown RBPs from RBR-ID, Related to Figure 2

Table S6. Interpro analysis of primary RBR-ID hits, Related to Figure 2

Table S7. Interpro analysis of unknown RBPs, Related to Figure 2

Table S8. Oligonucleotide sequences, Related to Figures 3, 6, S1, S3, S6

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cells

E14Tg2A.4 mESC lines (E14 mESCs) and HEK 293 cells were cultured as described previously (Kaneko et al., 2010). KH2 ESCs expressing the reverse tetracycline-controlled transactivator (rtTA) (Hochedlinger et al., 2005) were maintained in standard mouse ESC (mESC) culture conditions. For RNA immunoprecipitation experiments, stable KH2 lines were generated by transfection of the relevant pINTA-N3 constructs and selecting with 50 μg/ml Zeocin (InvivoGen, CA). Transgene expression was induced with 2 μg/ml doxycycline for 24 h.

Plasmids and sequences

The construction of the backbones for pGEX-6P1 was described previously (Kaelin et al., 1992). The DNA sequence for MS2-CP was synthesized by IDT and subcloned into the pGEX-6P1 expression vector. pINTO-N3 vector was based on the pINTO system (Gao et al., 2012) containing three N-terminal epitope tags (FLAG, HA, and Twin-Strep-Tag). For expression in

HEK 293 cells, we cloned into pINTO-N3 *L1td1*, *Mycn*, *Rarg*, *Cdkn2aipnl*, and *Pced1b* from mouse cDNA and the catalytic domain of *Tet2* from a plasmid kindly provided by Rahul Kohli (University of Pennsylvania) (Crawford et al., 2016). *Pou5f1*, *Pcgf1*, *Pcgf2*, and *Nanog* were cloned from mouse cDNA into the pINTA-N3 vector, which is based on the Tet-On 3G system (Clontech, CA) and encodes three different N-terminal epitope tags: Flag, HA, and Twin-Strep-Tag (Kaneko et al., 2013). Truncations were obtained by PCR. All oligonucleotide and synthetic DNA sequences used are in **Table S8**.

Antibodies information

The following antibodies were used for Western blots: SNRNP70 (#sc-9571 Santa Cruz Biotechnologies, TX), GST (#sc-33613; Santa Cruz Biotechnologies), FLAG (#F1804; Sigma-Aldrich, MO), HA (#901501 BioLegend, CA). Antibody against HA (#ab9110 Abcam, UK) was used for PAR-CLIP experiment and RNA immunoprecipitation.

Recombinant protein expression and purification

GST fusion MS2-CP were induced with 0.1 mM IPTG and expressed in BL21(DE3) cells for 24 h at 16°C and purified using glutathione-sepharose 4 Fast Flow beads (GE Healthcare Life Sciences, PA). The beads were washed with PBS in a column, and the proteins were eluted in the presence of 10 mM glutathione. The purified proteins were dialyzed against 20 mM Tris pH 7.5, 100 mM KCl and 10% glycerol. Recombinant, FLAG-fused TET2-CD was kindly provided by Rahul Kohli (Crawford et al., 2016).

***In vitro* RNA pull-down assays**

RNA fragments were *in vitro* transcribed using the HiScribe kit (New England Biolabs, MA) and purified by TRIzol (Thermo Fisher). The 5'-terminal HOTAIR RNA fragment used was as previously described (Bonasio et al., 2014); template information for MS2-SL and random 100 nts RNA fragments is in **Table S8.** For the binding assays, recombinant proteins were incubated with total E14 RNAs or *in vitro* transcripts in 1 ml RIP buffer (20 mM Tris pH 8, 0.2 mM EDTA, 100 mM KCl, 3 mM $MgCl_2$, 0.05% IGEPAL CA-630) with the addition of 2 u/µl murine RNAse inhibitor (New England Biolabs) for 30 min at 4°C. Protein-RNA complexes were pulled down using glutathione-sepharose 4 Fast Flow beads (GE) or Flag-M2 beads (Sigma). After three washes with RIP buffer, proteins were eluted from the beads in Laemmli sample buffer and nucleic acid with TRIzol. RNAs were resolved on polyacrylamide/urea gels and visualized with SYBR gold (Thermo Fisher).

RNA immunoprecipitation

Nuclear extracts were obtained using an established protocol (Dignam et al., 1983) with minor modifications to minimize RNAse activity. Briefly, cells were washed with PBS and with Buffer A (10 mM Tris pH $7.9_{4ºC}$, 1.5 mM $MgCl_2$, 10 mM KCl, protease inhibitors, phosphatase inhibitors) and lysed in Buffer A plus 0.2% IGEPAL CA-630 for 5 min on ice. Nuclei were isolated by centrifugation at 2,500g for 5 min and lysed in Buffer C (20 mM Tris pH $7.9_{4ºC}$, 25% glycerol, 400 mM NaCl, 1.5 mM $MgCl_2$, 10 mM EDTA, 0.4 u/µl murine RNAse inhibitor, protease inhibitors, phosphatase inhibitors) for 30 min at 4ºC.

Lysates were cleared at 18,000g for 30 min then incubated with HA antibody for 3 h at 4ºC. Immunocomplexes were recovered by adding 7 µl of protein G-coupled Dynabeads (Thermo Fisher) per µg of antibody used and incubating for 1 h at 4ºC. Beads were washed in RIP-W buffer (20 mM Tris pH $7.9_{4ºC}$, 1 mM $MgCl_2$, 200 mM KCl, 0.05% IGEPAL CA-630) twice and incubated with 2 u of TURBO DNase (Thermo Fisher) in 20 µl RIP-W buffer for 10 min at room temperature, to eliminate potential bridging effects of protein–DNA and DNA–RNA interactions. After two additional washes in RIP-W buffer RNA was eluted from the beads with TRIzol and collected by precipitation with isopropanol. Residual DNA was removed with TURBO DNAse for 20 min at 37ºC.

To quantify the RNA abundance after immunoprecipitations, we measured the intensity of the smears using ImageJ and normalized to the background observed in the IgG pull-down.

PAR-CLIP

HEK 293 cells were transiently transfected with plasmids using Lipofectamine 3000 reagent (Thermo Fisher) and pulsed with 100 µM 4-SU (Sigma) for 24 h. Cells were crosslinked with 400 mJ/cm^2 UVA (365 nm) using a Spectrolinker (Spectroline, NY) and lysed in CLIP buffer (20 mM HEPES pH 7.4, 5 mM EDTA, 150 mM NaCl, 2% Empigen) with protease inhibitors (Roche), 20 U/ml Turbo DNase (Thermo Fisher), and 200 U/ml murine RNase inhibitor (New England Biolabs). Tagged proteins were first bound to BSA-blocked Strep-Tactin beads (IBA, Germany) in CLIP buffer for 3 h at 4°C. Beads were washed 5 times using CLIP buffer and eluted in CLIP buffer with 2 mM biotin (Sigma), protease inhibitors (Roche Life Science, IN), and murine RNase inhibitor (New England Biolabs). Eluted proteins were incubated with HA antibody (Abcam) in CLIP buffer overnight at 4°C. Immunocomplexes were recovered with protein G-coupled dynabeads for 45 min at 4°C. DNA was removed with TURBO DNase (2 U in 20 µl). Crosslinked RNA was labeled by incubations with 5 U Antarctic phosphatase and 5 U T4 PNK

(both from New England Biolabs) in presence of 10 µCi [γ-^{32}P] ATP (PerkinElmer, MA). Labeled material was resolved on 8% Bis-Tris gels, transferred to nitrocellulose membrane, and exposed to autoradiography films for 1–24 hr.

For NANOG PAR-CLIP, E14 ESCs were transiently transfected with plasmids using Lipofectamine 3000 reagent (Thermo Fisher) and pulsed with 500 µM 4SU (Sigma) for 2 h. Cells were crosslinked with 400 mJ/cm^2 UVB (312 nm) using a Spectrolinker.

***In vitro* RBR-ID**

GST-fused MS2-CP was incubated with MS2-SL RNA in binding buffer (1 mM ATP, 10 mM HEPES pH 7.2, 3 mM $MgCl_2$, 5% glycerol, 1 mM DTT, 100 mM KCl) for 30 min at 4°C. RNA-protein complex were crosslinked with 1 J/cm^2 UVB (312 nm) using a Spectrolinker. The complexes were treated with RNase A (1 µg/µl) for 30 min at 37°C and the protein digested with trypsin or chymotrypsin. For trypsin digestion, proteins were diluted in 50 mM ammonium bicarbonate (NH_4HCO_3, pH 8) and incubated with 5 mM dithiothreitol (DTT) for 45 min at 56°C for disulfide bond reduction. This was followed by 20 mM iodoacetamide (IAA) incubation for 30 min in the dark for alkylation of the free cysteines. Trypsin was then added at an enzyme:sample ratio of 1:20, overnight at 37°C. For chymotrypsin digestion, proteins were diluted in 100 mM Tris pH 8, 10 mM $CaCl_2$, and incubated with 5 mM dithiothreitol (DTT) for 45 min at 56°C for disulfide bond reduction. This was followed by 20 mM iodoacetamide (IAA) incubation for 30 min in the dark for alkylation of the free cysteines. Samples were then digested using chymotrypsin at an enzyme:sample ratio of 1:20, overnight at 25°C. Reactions were blocked by adding 1% trifluoroacetic acid. Desalting was performed by using in-house packed Stage tips made of C_{18} material. Eluted peptides were dried and resuspended in 0.1% formic acid prior to nanoLC-MS analysis.

***In vivo* RBR-ID**

Cells were pulsed with 500 µM 4SU (Sigma) for 2 h and then crosslinked with 1 J/cm^2 UVA (365 nm), 1 J/cm^2 UVB (312 nm), or 800 mJ/ cm^2 UVC (254 nm) using a Spectrolinker. We verified that 2 hours was sufficient to incorporate 4SU in virtually all coding and noncoding transcripts by 4SU-sequencing (data not shown). Cells were lysed in Buffer A (10 mM Tris pH 8, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 0.2 mM PMSF) with 0.2% IGEPAL CA-630 for 5 min on ice to isolate nuclei. Nuclei were washed with Buffer A and lysed in denaturing lysis buffer (9 M urea, 100 mM Tris pH 8). Lysate was diluted in 50 mM ammonium bicarbonate (NH_4HCO_3, pH: 8.0) and incubated with 5 mM dithiothreitol (DTT) for 60 min at 25°C for disulfide bond reduction.

This was followed by 20 mM iodoacetamide (IAA) incubation for 30 min in the dark for alkylation of the free cysteines. Samples were then digested using trypsin at an enzyme:sample ratio of 1:100, overnight at 37°C. Reaction was blocked by adding 1% trifluoroacetic acid. Desalting was performed by using in-house packed Stage tips made of C_{18} material. Eluted peptides were dried and resuspended in 0.1% formic acid prior MS analysis. Crosslinked RNA was removed with Benzonase (250 U in 20 µl).

In vitro denaturing protein–RNA pull-downs

GST fusion MS2-CP were incubated with MS2-SL RNA in binding buffer (1 mM ATP, 10 mM HEPES pH 7.2, 3 mM $MgCl_2$, 5% glycerol, 1 mM DTT, 100 mM KCl) for 30 min at 4°C. RNA-protein complex were crosslinked with 1 J/cm^2 UVA (365 nm), 1 J/cm^2 UVB (312 nm), or 800 mJ/ cm^2 UVC (254 nm) using a Spectrolinker. Biotin-labeled DNA probes were hybridized to crosslinked RNA in hybridization buffer (500 mM NaCl, 1% SDS, 50 mM Bis-Tris pH6.7, 10 mM EDTA, 10% formamide) with protease inhibitor for 4 h at 37°C. DNA probes and protein–RNA complexes were recovered by incubating with streptavidin-conjugated dynabeads (Thermo Fisher) for 30 min at 37°C. After three washes with wash buffer (2X SSC, 0.5% SDS, 0.4 mM PMSF), proteins were eluted by PBS with 0.5 µg/µl RNase A.

In vivo denaturing protein–RNA pull-downs

Cells were pulsed with 4SU (Sigma) for 2 h and then crosslinked with 1 J/cm^2 UVA (365 nm), 1 J/cm^2 UVB (312 nm), or 800 mJ/ cm^2 UVC (254 nm) using a Spectrolinker. Cells were lysed in lysis buffer (50 mM Bis-Tris pH 6.7, 10 mM EDTA, 1% SDS) with protease inhibitor. A further sonication step with a Bioruptor (Diagenode, NJ) was performed to homogenize the cell lysates. Biotin-labeled DNA probes (**Table S8**) were hybridized to crosslinked RNA in hybridization buffer (500 mM NaCl, 1% SDS, 50 mM Bis-Tris pH 6.7, 10 mM EDTA, 10% formamide) with protease inhibitor for 4 h at 37°C. DNA probes and protein–RNA complexes were recovered by incubating with streptavidin-conjugated dynabeads for 30 min at 37°C. After three washes with wash buffer (2X SSC, 0.5% SDS, 0.4 mM PMSF), proteins were eluted by PBS with 0.5 µg/µl RNase A.

Bottom-up nanoLC-MS/MS

Samples were analyzed by using a nanoLC-MS/MS setup. NanoLC was configured with a 75 µm ID x 17 cm Reprosil-Pur C_{18}-AQ (3 µm; Dr. Maisch GmbH, Germany) nano-column using an EASY-nLC nanoHPLC (Thermo Fisher). The HPLC gradient was 0-30% solvent B (A = 0.1%

formic acid; B = 95% acetonitrile, 0.1% formic acid) over 120 min for the nuclear proteome experiments and over 45 min for the recombinant protein (MS2-CP) analysis. The gradient proceeded from 30% to 85% solvent B in 5 minutes and 10 min isocratic at 85% B. The flow rate was set to 300 nL/min. NanoLC was coupled with an Orbitrap Fusion mass spectrometer (Thermo Fisher) for the proteome experiments or with an Orbitrap Elite (Thermo Fisher) for the single protein analysis. Spray voltage was set at 2.3 kV and capillary temperature was set at 275 °C. Full scan MS spectrum (*m/z* 350−1200) was performed in the Orbitrap with a resolution of 120,000 (at 200 *m/z*) with an AGC target of $5x10^5$. For the proteome experiment in the Orbitrap Fusion the Top Speed MS/MS option was set to 2.5 sec, and the most intense ions above a threshold of 50,000 counts were selected for fragmentation. Fragmentation was performed with higher-energy collisional dissociation (HCD) with normalized collision energy of 32, an AGC target of 10^4 and a maximum injection time of 120 msec. For the single protein experiment in the Orbitrap Elite the top 10 most intense ions above a threshold of 10,000 counts were selected for fragmentation. Fragmentation was performed with collisional induced dissociation (CID) with normalized collision energy of 35, an AGC target of 10^4 and a maximum injection time of 150 msec. MS/MS data for both experiment types were collected in centroid mode in the ion trap mass analyzer (normal scan rate). Only charge states 2-5 were included.

MS analysis

All MS/MS spectra were processed through the MaxQuant program (Cox and Mann, 2008). Parameters for MS/MS database searching included the following: precursor mass tolerance 4.5 ppm; product mass tolerance 0.5 Da; enzyme trypsin; missed cleavages allowed 2; static modifications carbamidomethyl (C); variable modifications none; label-free quantification method iBAQ (for protein reports); database used was *Mus musculus* (Uniprot, September 2015, including not reviewed proteins) for the proteome searches, and a custom database including MS2-CP and RNase A for *in vitro* RBR-ID. PSMs and protein false discovery rate was filtered for < 0.01. Match between runs was enabled using a tolerance of 1 min to extend the peptide identification to MS signals not identified in some of the replicates.

RBR-ID analysis

For each peptide, the maximum intensity of the corresponding extracted chromatogram calculated by MaxQuant was considered and inter-run variability was accounted for by normalizing for the sum of all peptide intensities in each MS run. To calculate the extent of crosslinking-induced depletion we calculated the $\log_2$-converted ratio of the mean intensity of

each peptide in the +4SU samples divided by the mean intensity of the same peptide in -4SU samples. The list of primary hits contains all peptides showing depletion (i.e. $\log_2$(fold-change) < 0 with a *P*-value < 0.05 (Student's *t* test). For the extended list we relaxed the *P*-value requirement to 0.1. For both the primary and extended list we removed peptides that passed the same cutoffs when comparing signals for +4SU and -4SU in absence of UV.

RBR-ID scores were calculated by combining the extent of depletion and the *P*-value according to the following formula:

score = $-\log_2$(normalized +4SU intensity/normalized -4SU intensity) * $(\log_{10}(\text{P-value}))^2$.

For residue level RBR-ID scores, we summed the RBR-ID score of each peptide overlapping any given amino acid and smoothed the resulting curve using Friedman's 'super smoother' (Friedman, 1984).

GO and Interpro enrichment

For protein list comparisons, all proteins identifiers were converted to official mouse symbols using the Biomart database (version 84). One-to-one human-mouse orthologs were mapped directly, whereas one-to-many and many-to-many homologs were reduced to one-to-one by considering the protein with highest percentage of homology, according to the Biomart database (version 84). For GO and Interpro annotation, tables were downloaded from the Uniprot and Interpro websites directly. Enrichment values and statistics were obtained using the DAVID web server (Huang da et al., 2009) either using the unique Uniprot accession identifiers or the converted symbols, when needed.

Data accessibility

MS raw data are available at the Chorus database (https://chorusproject.org) under project number 1128. Peptide lists and RBR-ID score plots are available at http://rbrid.bonasiolab.org.

SUPPLEMENTAL REFERENCES

Bonasio, R., Lecona, E., Narendra, V., Voigt, P., Parisi, F., Kluger, Y., and Reinberg, D. (2014). Interactions with RNA direct the Polycomb group protein SCML2 to chromatin where it represses target genes. eLife *3*, e02637.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology *26*, 1367-1372.

Crawford, D.J., Liu, M.Y., Nabel, C.S., Cao, X.J., Garcia, B.A., and Kohli, R.M. (2016). Tet2 Catalyzes Stepwise 5-Methylcytosine Oxidation by an Iterative and de novo Mechanism. J Am Chem Soc *138*, 730-733.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res *11*, 1475-1489.

Friedman, J.H. (1984). A variable span smoother (DTIC Document).

Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., and Reinberg, D. (2012). PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. Mol Cell *45*, 344-356.

Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell *121*, 465-477.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols *4*, 44-57.

Kaelin, W.G., Jr., Krek, W., Sellers, W.R., DeCaprio, J.A., Ajchenbaum, F., Fuchs, C.S., Chittenden, T., Li, Y., Farnham, P.J., Blanar, M.A.*, et al.* (1992). Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. Cell *70*, 351-364.

Kaneko, S., Li, G., Son, J., Xu, C.F., Margueron, R., Neubert, T.A., and Reinberg, D. (2010). Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. Genes Dev *24*, 2615-2620.

Kaneko, S., Son, J., Shen, S.S., Reinberg, D., and Bonasio, R. (2013). PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. Nat Struct Mol Biol *20*, 1258-1264.