# The interaction of PRC2 with RNA or chromatin is mutually antagonistic

Manuel Beltran, Christopher M. Yates, Lenka Skalska, Marcus Dawson, Filipa P. Reis, Keijo Viiri, Cynthia L. Fisher, Christopher R. Sibley, Benjamin M. Foster, Till Bartke, Jernej Ule and Richard G. Jenner

### **SUPPLEMENTAL METHODS**

### **Mouse ESC culture**

Suz12<sup>+/+</sup> and Suz12<sup>-/-</sup> mESC (Lee et al., 2006) (used for CLIP) were cultured on mitomycintreated PEFs with KO-DMEM medium, 10% FCS, 5% knockout serum replacement, nonessential amino acids, L-glutamine, 2-mercaptoethanol, Penicillin/Streptomycin and leukaemia inhibitory factor (LIF; 1000 U/ml, Abcam). Ezh2<sup>fl/fl</sup> (Kanhere et al. 2010), JM8 (Jarid2<sup>+/+</sup>), Jarid2<sup>-/-</sup> (Landeira et al. 2010) and Suz12<sup>-/-</sup> (Pasini et al., 2007, used for iCLIP) mESC were maintained on 0.1% gelatin, as previously described. Ezh2fl/fl mESC were treated with 4hydroxy-tamoxifen (OHT) for 96 hours (800 nM, Sigma) to induce Cre-mediated deletion of the SET domain, generating  $Ezh2^{\Delta/\Delta}$  cells (Kanhere et al. 2010). cKO Eed mESC clone 4D2 (Ura et al. 2008) were cultured on 0.1% gelatin with DMEM medium, 10% FCS, nonessential amino acids, L-glutamine, 2-mercaptoethanol, antibiotics and LIF (1000 U/ml). Cells were treated with doxycycline (10 µg/ml, Sigma) for 96 hours to repress the Myc-tagged *Eed* transgene. NSCs were cultured as described (Sofueva et al., 2013) as an adherent monolayer in DMEM/Ham's F-12 media with L-glutamine (PAA), 0.5x N2 (PAA), 0.5x B27 (Invitrogen), 50 nM 2mercaptoethanol (Invitrogen), 120 µg/ml BSA, 4.5 mM HEPES buffer solution (PAA), 1x non-essential amino acids (PAA), 0.15% glucose (Sigma-Aldrich), 1xpenicillin/streptomycin (PAA), 10 ng/ml murine EGF (Peprotech), 10 ng/ml human FGF (Peprotech), and 10 ng/ml Laminin (Sigma). Rad21 was deleted by adding 200 nM OHT for 48 hrs. Stable Ezh2<sup>fl/fl</sup>-GFP and Suz12<sup>-/-</sup>-GFP cell lines were generated by electroporation of Ezh2<sup>fl/fl</sup> and Suz12<sup>-/-</sup> cells (Pasini et al., 2007), respectively, with pCAG-eGFP vector using the 4D-Nucleofector system (Lonza) and selected with 200 µg/ml puromycin.

### **CLIP**

CLIP was performed as described (Huppertz et al. 2014). Briefly, cells were irradiated with 0.15 J/cm<sup>2</sup> of 254 nm UV light in a Stratalinker 2400 (Stratagene). 5x10<sup>6</sup> cells were used per IP and were lysed in 1 ml of lysis buffer with Complete (Roche) protease inhibitor. Lysates were passed through a 27 G needle, 4 U/ml of DNase Turbo (Ambion) and RNase I (Ambion, range between 1-20 U/ml) added, and incubated in a Thermomixer at 37 °C and 1100 rpm for 3 minutes. Lysates were then cleared by centrifugation and Proteus clarification spin column, according to the manufacturer's instructions. 5  $\mu$ l of  $\alpha$ -SUZ12 (Cell Signaling #3737) or 5  $\mu$ g  $\alpha$ -EZH2 (Cell Signaling #3147), 5 μg α-FLAG (Sigma #F1804), α-MYC (Cell Signaling, #2276), α-FUS (Novus Biologicals #100-565), α-HNRNPC (Santa Cruz #sc-32308), non-specific IgG (Santa Cruz #5279) or α-GFP (Abcam #ab290) antibody was used per experiment and were bound to 50 µl of pre-washed protein G beads (Dynal) for 1 hour at RT. Antibody-bound beads were then incubated with lysate for 5 hours at 4 °C. Beads were washed 3 times with 900 µl of high-salt buffer and twice with 900 µl of wash buffer (both supplemented with 1 M Urea). We also tested whether addition of 5M Urea (Supplemental Fig. 1B) or 2% lauryl-dimethylbetaine (Supplemental Fig. 1C) could be used to prevent co-precipitation of EZH2 by SUZ12. After transfer, the membrane was washed twice with 1x PBS and exposed overnight to film (Fuji) at -80 °C. Immunoblotting was performed using α-SUZ12 (Santa Cruz #46264), α-EZH2 (Cell Signaling #3147), α-JARID2 (Abcam #48137), α-FLAG (Sigma #8592), α-MYC (Cell Signaling, #2276), α-GFP (Abcam #ab290), α-HNRNPC (Santa Cruz #sc-32308) or α-FUS (Novus Biologicals #100-565) antibodies and the relevant HRP-coupled secondary antibody (DAKO). Crosslinked RNA was quantified using a Typhoon phosphorimager (GE) and ImageQuantTL (GE). Protein was quantified using an ImageQuantLAS 4000 imager and ImageQuantTL (GE).

## Over-digestion with RNase I or DNase Turbo

After radiolabelling of the IP, the radioactive mix was discarded and a second nuclease treatment was performed by incubating beads in 1 ml of PNK buffer supplemented with 10 U of DNase Turbo and/or 200 U of RNase I for 10 minutes at 37 °C (Supplemental Fig. 1E).

#### PAR-CLIP

PAR-CLIP was performed as described (Hafner et al. 2010). Cells were treated with 100  $\mu$ M 4-thiouridine (Sigma) for 14 hours and irradiated with 0.15 J/cm² of 365 nm UV light in a Stratalinker 2400 (Stratagene).  $10^6$  U/ml of RNase T1 was used, unless indicated.  $5x10^6$  cells were used per IP and were lysed in 0.25 ml of NP40 lysis buffer. 5  $\mu$ g of antibody was incubated with pre-washed protein G beads (Dynal) for 1 hour at RT. Antibody-bound beads were then incubated with lysate for 5 hours at 4 °C.

Ezh2 PAR-CLIP data were downloaded from GEO (GSE49435) and processed as described (Kaneko et al. 2013) retaining only those reads containing a T>C transition. A list of ezRNA genes was obtained from Table S2 of Kaneko et al. 2013.

#### **iCLIP**

iCLIP was performed as described (Huppertz et al. 2014) with the following variations. Cells were irradiated with 0.15-0.2 J/cm<sup>2</sup> of 254 nm UV light in a Stratalinker 2400 (Stratagene). 2x10<sup>8</sup> cells were used per IP for SUZ12, EED-MYC, IgG and GFP and 2.5x10<sup>7</sup> cells per IP for FUS and HNRNPC. Cells were lysed in 1 ml of lysis buffer and lysates were passed through a 27 G needle. For SUZ12, IgG, EED and GFP iCLIP, 200 U/ml of DNase Turbo (Ambion), and 40 U/ml of RNase I (Ambion) were added. For FUS and HNRNPC iCLIP, 20 U/ml of DNase Turbo (Ambion), and 5 U/ml of RNase I (Ambion) were added. Lysates were then incubated in a thermomixer at 37 °C and 1100 rpm for 3 minutes and cleared by centrifugation and Proteus clarification spin columns, according to the manufacturer's instructions. 10 µg of antibody was used for each CLIP, as described in the CLIP section. After SDS-PAGE and transfer to membrane, cross-linked RNPs running 10-30 (SUZ12), 10-45 (FUS), 10-40 (HNRNPC) and 5-30 (GFP) kDa above the sizes of their respective proteins were isolated. For IgG, we isolated RNPs of the same molecular weight as for SUZ12. RNA was purified and after reverse transcription, cDNA was fractionated by running samples on a precast 6% TBE-urea gel at 180 V for 40 min and cDNA bands running between 250-180 nt (top), 120-180 nt (high), 85-120 nt (medium) and 70-85 nt (low) isolated. One µl of each fraction was pooled to optimize the number of cycles in the PCR, determined by the minimum number of cycles that produced detectable amplicon in gels stained with SYBR Green I. Once the optimal number of optimal cycles was established, the library PCR was performed separately for each fraction, checked by gel electrophoresis and pooled the following proportions: 1:5:5:1. Library concentration was determined using the KAPA Universal Library Quantification kit, according to the manufacturer's instructions and library concentration was corrected by multiplication by 0.38 to improve cluster density. Libraries were sequenced using Illumina HiSeq 2500 sequencers (50 bp single-end reads).

## iCLIP data processing

iCLIP data were processed with iCount (http://icount.biolab.si/) and filtered to remove any crosslinks overlapping a RepeatMasker feature or ncRNAs under 200nt in length or annotated as a snoRNA in RefSeq, GENCODE (vM1) or Ensembl (v59). High-confidence clusters of crosslink sites were identified using the low FDR function in iCount (FDR<0.05), with a 50 nt flank (Konig et al. 2010). The background was obtained by randomising reads within the entire gene ("region as one" setting in iCount). Bedtools was used to assign crosslinks to gene segments annotated in Ensembl. Where a crosslink could map to multiple segments, it was assigned as either intron, UTR or CDS, in that order, with protein-coding genes taking precedence over lincRNA (using Ensembl definitions of gene types). Crosslink density profiles were generated with ngsplot (Shen et al. 2014), with the resultant heatmaps filtered to remove the top/bottom 1% of genes before plotting the mean read density. For visualizing iCLIP data at individual genes, bigwig files were generated using bedGraphToBigWig and uploaded to the UCSC Genome Browser. Reads were split by the DNA strand and each track was auto-scaled.

To assess the number of reads at spliced and unspliced exons, reads were realigned using TopHat2 (Kim et al. 2013), using the very-sensitive Bowtie2 setting, up to two multi-hits, up to one mismatch per segment, and a segment length of 18. Overlaps between the aligned reads and exon-exon junctions were obtained using bedtools intersect, requiring at least 5 bp each side of the splice site for an unspliced read to be counted.

To identify genes with significantly more PRC2 crosslinks compared to FUS and HNRNPC, EDASeq (Risso et al. 2011) was used to normalise for GC content and gene length, both of which can bias RNA-seq experiments, and the normalisation factors from EDASeq fed into DESeq2 (Love et al. 2014) to identify significant genes (FDR<0.05).

To identify crosslinks to the *CreERT2* transgene, the sequence of this transgene was downloaded from https://www.addgene.org/14797/. iCLIP reads were then aligned to a genome consisting of this sequence and the mouse genome (mm9) and processed as above.

## RNA-seq

Total RNA was purified with TRizol and treated with DNase-Turbo (Ambion). Ribosomal RNA was depleted with Ribo-Zero Gold (EpiCentre). Libraries were constructed using the Illumina Directional mRNA-Seq Sample Prep and the NEBNext Multiplex Small RNA Library Prep kits and sequenced on an Illumina HiSeq 2500 (50bp paired-end). Nuclear RNA-seq data were downloaded from GEO (GSE57092 (Bulut-Karslioglu et al. 2014)).

Reads were trimmed to remove adaptors using fastq-mcf, requiring a minimum remaining read length of 15 nt and a 0.0001% threshold for trimming. Low-quality bases were removed using seqtk, removing anything with a Phred score under 0.01. Reads passing these quality-control steps were aligned to the mouse genome (build mm9) using TopHat2, using the very-sensitive setting on Bowtie2 and allowing up to two multi-hits. Reads were then filtered for repeats, snRNAs and snoRNAs as for iCLIP. Reads overlapping genes were quantified using bedtools intersect, requiring both reads in a pair to have been successfully mapped.

EDASeq and DESeq2 were used to quantify fold-changes in expression between  $Ezh2^{fl/fl}$  and  $Ezh2^{\Delta/\Delta}$  cells and to identify genes enriched for RNA crosslinks compared to nuclear RNA abundance (FDR<0.05).

## RNase A treatment and cell fractionation

RNase A treatment and cell fractionation were performed as described (Zoabi et al. 2014). ESC were trypsinized, washed twice with PBS, permeabilized with 0.05% Tween-20 in PBS for 10 min on ice, washed once, resuspended with PBS and mock-treated or treated with 1 mg/ml RNase A (Sigma) for 30 min at RT. Cells were centrifuged at 1200 rpm, washed twice with PBS and crosslinked for ChIP or resuspended in 1 ml of buffer A (10 mM HEPES (pH 7.9), 10 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.34 M sucrose, 10% glycerol, 1 mM DTT with Complete protease inhibitor). Triton X-100 (0.1%) was added, and the cells were incubated for 5 min on ice. Nuclei

were collected by low-speed centrifugation (4 min, 1,300g, 4 °C). The supernatant (cytoplasmic fraction) was further clarified by high-speed centrifugation (15 min, 20,000g, 4 °C). Nuclei were washed twice in buffer A, and then lysed in buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, Complete protease inhibitor). Insoluble chromatin was collected by centrifugation (4 min, 1,700g, 4 °C), and the supernatant (nucleoplasmic fraction) was combined with the cytoplasmic fraction. The final chromatin pellet (chromatin fraction) was washed twice with buffer B and resuspended in Laemmli buffer and sonicated (Diagenode Bioruptor). Samples were blotted for SUZ12 (Santa Cruz #46264), EZH2 (Cell Signaling #3147), alpha-tubulin (Cell Signaling #2144), beta-actin (Cell Signaling #4967), BRD4 (Bethyl #985A100), histone H3 (Abcam #ab1791), RNA pol II (8WG16; Millipore #05-952) and RAD21 (Abcam #ab992).

To test if re-addition of RNA could titrate PRC2 back off chromatin, following 15 min RNase A treatment, cells were washed twice with ice-cold PBS, resuspended in 100  $\mu$ l of room temperature PBS containing 25, 12.5, 5, 1 and 0.2  $\mu$ g/ $\mu$ l of yeast tRNA (Life Technologies) and incubated 15 min at RT. Cells were collected by 5 min centrifugation at 4 °C at 300g and chromatin fraction separated as before. BSA (2.5  $\mu$ g/ $\mu$ l) was used as a control instead of tRNA.

To test if RNase A treatment altered chromatin accessibility, mock- or RNase-treated nuclei, were resuspended in 100  $\mu$ l of Buffer A supplemented with 100  $\mu$ g/ml BSA, 5 mM CaCl<sub>2</sub> and 0, 5, 10, 50 and 250 gel units of Micrococcal Nuclease (NEB M0247S). After 10 min incubation at 37 °C, the reaction was stopped with 0.25 volumes of 4% SDS and 0.1 M EDTA. Samples were treated with RNase A and proteinase K, DNA purified by phenol-chloroform extraction and ethanol precipitation and resolved by agarose electrophoresis .

### **Chromatin IP**

ChIP was performed for SUZ12 (Abcam #ab12073), EZH2 (Cell Signaling #3147), JARID2 (Cell Signalling #13594), H3K27me3 (Abcam #ab6002), H3K27me1 (Active Motif #61015), EP300 (Bethyl #A300-358) and total H3 (Abcam #ab1791) as described (Kanhere et al. 2012), except that cell lysates were sonicated using a Diagenode Bioruptor. For ChIP-seq, cells were crosslinked with 1% formaldehyde immediately after mock and RNase A treatment. Libraries were sequenced using an Illumina HiSeq 2500 (50-bp single-end reads). ChIP-seq reads were

trimmed to remove adaptors using fastq-mcf with a 0.0001% threshold for trimming and low-quality bases were removed as for RNA-seq reads. Reads passing QC were aligned to the mouse genome (mm9) using Bowtie2 with the very-sensitive setting (Langmead and Salzberg, 2012). SUZ12 and EP300 binding at active and inactive genes were plotted from duplicate experiments using ngsplot (Shen et al. 2014). Active genes were selected as any genes with a nuclear RNA and total RNA RPKM of greater than 1, with inactive genes having both values equal to 0. CpG islands were downloaded from the UCSC Genome Browser and genes with CpG islands at their TSS identified. MACS14 (Zhang et al. 2008) was used to identify binding sites in the RNase A and mock treated ChIP samples. MAnorm (Shao et al. 2012) was used to identify sites at which binding significantly changed (p<0.05) between the conditions, requiring a gene to have a SUZ12 binding peak within 1 kb of its TSS in each RNase treatment replicate, with this peak showing significantly increased binding in at least one replicate. The significance of the difference in PRC2 RNaseA/mock ratio between genes that gain PRC2 upon triptolide treatment with those at which PRC2 remains constant was estimated using a Wilcoxon signed-rank test.

Enrichment of specific gene sequences was measured by qPCR (Applied Biosystems) using QuantiTect SYBR Green PCR Kit (Qiagen) and the following oligos:

## Primers used for ChIP-qPCR.

Gene	Forward primer (5' to 3')	Reverse primer (5' to 3')
Hoxd11	GGCACAGCGCCTGTCCAACA	TCTTCCCTGCAGAGCCTACCG
Втр6	AGCCGCCTCTGAGGGTTC	GCCAGGTGTGTCCTAGGCAG
Gata4	CTCTCCCGAGCTCACTTCAAGG	GGAGAAGGTGACCTCGCACAC
Utp6	TCTATGGCCTTACCCACTGC	TGACACGTTTCTGCTTCCAG
Rab7A	GCTCCACTTGAAGGTCTTGG	ATACACAGAAGGCCCACAGG
Sal1	AGAAGAACCAGGCCTCAGA	AGAAGAACCAGGCCTCAGA
Actb	CCGTGCCCCAGAATTATCAG	GCCGTCCATATCCACCTAAGAA
Pou5f1	TCGGATGCCCCATCGCA	GGCTCCCAGAGGATGGCTGAG
Intergenic	GCCGTCCATATCCACCTAAGAA	CCGTGCCCCAGAATTATCAG

ChIP-seq data were also downloaded from GEO: H3K4me3 and H3 from GSE12241 (Mikkelsen et al. 2007), RNA Pol II (8WG16) and input from GSE34518 (Brookes et al. 2012), SUZ12 with/without addition of triptolide from GSE58023 (Riising et al. 2014), and SUZ12 from WT ESC and cells overexpressing RNase H from GSE67584 (Chen et al. 2015). Average profiles were plotted using ngsplot. Input was subtracted from Suz12 and RNA pol II and histone modifications were divided by total H3. SUZ12 peaks changing with the addition of triptolide were identified using MACS14 and MAnorm (p<0.05) (Shao et al. 2012). PRC2 binding sites in differentiated cells were taken from the ENCODE project (Yue et al. 2014).

## Co-immunoprecipitation

Co-immunoprecipitations were performed using 5 µg of antibody (SUZ12; Cell Signaling #3737 and EZH2; Cell Signalling #3147) or control IgG (Santa Cruz #sc2027)) cross-linked to 50 µl of Protein G Dynabeads (Life Technologies). Cells were lysed in iCLIP lysis buffer and IPs washed in wash buffer with 1M Urea. Immunoblotting was performed using anti-SUZ12 (Santa Cruz #sc46264), anti-EZH1 (Abcam #ab13665), anti-EZH2 (Cell Signaling #4905) or anti-beta actin (ACTB; Cell Signaling #4967S).

### Biotinylated nucleosome pull-downs

Modified histone H3 protein was prepared by native chemical ligation and refolded into octamers together with recombinant human histone H2A, H2B and H4 expressed and purified from *E. coli*, as described (Bartke et al., 2010). Modified octamers were then assembled into nucleosomes by salt deposition dialysis using a biotinylated 185 bp DNA fragment containing the 601 nucleosome positioning sequence, as described (Bartke et al., 2010). 100 nM nucleosomes were incubated with 50 nM recombinant PRC2 (Active Motif #31387), 10 μl MyOne Streptavidin T1 Dynabeads (Thermo scientific) and with purified nuclear RNA or yeast tRNA, titrated from 0.5-65 ng/μl (0.02-2.6 μM or 0.38-48 nM, respectively), in 10 mM HEPES pH 7.9, 100 mM NaCl, 0.25 mM EDTA (pH 8.0), 1 mM DTT, 5% Glycerol, 0.05% IGEPAL CA-630, 33 ng/μl BSA, plus Complete protease inhibitor (Roche) with a final volume of 100 μl. BSA at 330 ng/μl (5 μM) was used to control for potential non-specific effects of the RNA occluding the interaction with nucleosomes. The binding reaction was allowed to proceed for 3 hours at 4 °C, beads were

then washed 3 times and resuspended in protein loading buffer for SDS-PAGE and immunoblotting.

## **Biotinylated RNA pull-downs**

Nascent biotinylated RNA was generated by incubating mESC with 0.5 mM 5-ethynyl uridine for one hour, RNA was extracted using TRIzol and biotinylated with the Click-iT Nascent RNA Capture Kit (Thermo Scientific). Biotinylation was verified by dot-blot with streptavidin-HRP (Thermo). 2.5 μg biotinylated RNA bound to MyOne Streptavidin T1 Dynabeads (Thermo) from the final step of the Click-iT protocol was added to a reaction containing 50 nM recombinant PRC2 (Active Motif) and nucleosomes, titrated from 0.024 nM to 2.4 μM, in 10 mM HEPES pH 7.9, 100 mM NaCl, 0.25 mM EDTA (pH 8.0), 1 mM DTT, 5% Glycerol, 0.05% IGEPAL CA-630, 33 ng/μl BSA, plus Complete protease inhibitor (Roche). The binding reaction was allowed to proceed for 3 hours at 4 °C, beads were then washed 3 times and resuspended in protein loading buffer for SDS-PAGE and immunoblotting.

#### **UV-RIP**

2x10<sup>8</sup> Ezh2<sup>fl/fl</sup>, Ezh2<sup>fl/fl</sup>-GFP or Suz12<sup>-/-</sup>-GFP cells were UV irradiated at 0.4 J/cm2 using a Stratalinker 1800 (Stratagene), and nuclei lysed in polysomal lysis buffer (10 mM Hepes pH 7.0, 100 mM KCl, 5 mM MgCl<sub>2</sub>, 0.5% NP-40). *In vitro* transcribed GFP RNA was generated from linearized pcDNA-eGFP using MAXIScript Sp6/T7 (Ambion) and purified according to the manufacturer's instructions. 5 µg, 0.5 µg or 0.05 µg of in vitro transcribed GFP RNA were spiked-into the Ezh2<sup>fl/fl</sup> cell lysate (equivalent to 36,000, 3600 and 360 RNA copies per cell, respectively). RIP was performed as previously described (Zhao et al., 2008), except using 10 µg of anti-SUZ12 antibody (Cell Signalling), and including 0.5 M NaCl and 0.5 M urea in the washes. Beads were resuspended in TRIzol and purified RNA reverse-transcribed using random hexamer primers and SuperScript III (Invitrogen). Enrichment of cDNA compared to input control was measured by qPCR (Applied Biosystems) using QuantiTect SYBR Green PCR kit (Qiagen) with primers for GFP 5'-CGTAAACGGCCACAAGTTCAG-3' CAGGGTCAGCTTGCCGTAG-3' or for Srsf1 5'-CTTTTCACCCCCGGAGCG-3' and 5'-ATCTTCCGCGTCTCTGCG-3'. Primers against the untranscribed vector sequence were used as a negative control.

#### **Immunofluorescence**

mESC were cultured as described on gelatin-coated glass coverslips for 24 hours and then fixed using a 4% formaldehyde solution in PBS for 15 min at room temperature. After 2 hours incubation in blocking buffer (1X PBS, 5% BSA, 0.3% Triton<sup>TM</sup> X-100), samples were incubated overnight in antibody dilution buffer (1X PBS, 1% BSA, 0.3% Triton<sup>TM</sup> X-100) with the following antibodies: α-SUZ12 (1/800, Cell Signaling #3737) and α-EZH2 (1/250, Cell Signalling #3147). After washing, cells were incubated with α-mouse Alexa Fluor 546, α-rabbit Alexa Fluor 488 and Hoechst stain (Life Technologies) for 2 hours. For nuclear GFP staining cells were fixed and incubated in blocking buffer supplemented with Hoechst stain. Coverslips were washed and mounted using ProLong® anti-fade reagents (Life Technologies) and images acquired using a Perkin Elmer Spinning Disk Confocal microscope and Volocity software.