Published in final edited form as:
Mol Cell. 2010 June 11; 38(5): 675–688. doi:10.1016/j.molcel.2010.03.019.

Short RNAs are transcribed from repressed Polycomb target genes and interact with Polycomb Repressive Complex-2

Aditi Kanhere[1], **Keijo Viiri**[1], **Carla C. Araújo**[1], **Jane Rasaiyaah**[1], **Russell D. Bouwman**[1], **Warren A. Whyte**[2,3], **C. Filipe Pereira**[4], **Emily Brookes**[4], **Kimberly Walker**[2], **George W. Bell**[2], **Ana Pombo**[4], **Amanda G. Fisher**[4], **Richard A. Young**[2,3], and **Richard G. Jenner**[1,5]

[1]Division of Infection and Immunity, University College London, London, W1T 4JF, United Kingdom

[2]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

[3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02141, USA

[4]MRC Clinical Sciences Centre, Imperial College London, London, W12 0NN, United Kingdom

SUMMARY

Polycomb proteins maintain cell identity by repressing the expression of developmental regulators specific for other cell types. Polycomb repressive complex-2 (PRC2) catalyses trimethylation of histone H3 lysine-27 (H3K27me3). Although repressed, PRC2 targets are generally associated with the transcriptional initiation marker H3K4me3 but the significance of this remains unclear. Here, we identify a new class of short RNAs, ~50-200 nucleotides in length, transcribed from the 5′-end of polycomb target genes in primary T-cells and embryonic stem cells. Short RNA transcription is associated with RNA polymerase II and H3K4me3, occurs in the absence of mRNA transcription and is independent of polycomb activity. Short RNAs form stem-loop structures resembling PRC2 binding sites in Xist, interact with PRC2 through SUZ12, cause gene repression in cis and are depleted from polycomb target genes activated during cell differentiation. We propose that short RNAs play a role in the association of PRC2 with its target genes.

INTRODUCTION

Polycomb group (PcG) proteins are essential for embryogenesis and for maintaining embryonic stem (ES) cell pluripotency and differentiated cell states (Boyer et al., 2006; Faust et al., 1995; Lee et al., 2006; O'Carroll et al., 2001; Pasini et al., 2007; van der Stoop et al., 2008). PcG proteins form at least two polycomb repressive complexes (PRCs) that are conserved across Metazoa. PRC2 catalyses trimethylation of lysine 27 of histone H3 (H3K27me3), forming a binding site for PRC1 (Cao et al., 2002; Wang et al., 2004).

Genome-wide measurements of PRC2 localisation and H3K27 methylation have revealed that PRC2 represses the expression of hundreds of developmental regulators in ES cells that would otherwise induce cell differentiation (Azuara et al., 2006; Boyer et al., 2006; Lee et al.,

[5]Corresponding author. r.jenner@ucl.ac.uk, Tel: +44 20 7679 9635, Fax: +44 20 7679 9652.

Please see Supplemental Data for more detailed experimental procedures.
Microarray data are available at ArrayExpress (accession E-TABM-844).

2006). In differentiated cells, activated genes important for the identity of those cells lose H3K27me3 whereas genes that regulate alternate cellular identities remain methylated and repressed (Bernstein et al., 2006; Lee et al., 2006; Mikkelsen et al., 2007; Roh et al., 2006).

Although repressed, PRC2 target genes are thought to adopt a poised state that allows their rapid upregulation upon ES cell differentiation (Boyer et al., 2006; Lee et al., 2006). This poised state is reflected by the association of PRC2 target genes with histone H3K4me3, a marker of transcriptional initiation (Azuara et al., 2006; Bernstein et al., 2006; Roh et al., 2006). The unphosphorylated and Ser-5 phosphorylated forms of RNA polymerase (pol) II have also been detected at some polycomb target genes but not the Ser-2 phosphorylated form that is associated with transcriptional elongation (Dellino et al., 2004; Lee et al., 2006; Stock et al., 2007). The block to RNA pol II elongation at these genes is due, at least in part, to the activity of PRC1, which binds to H3K27me3. PRC1 contains Ring1 that ubiquitinates H2A, blocking elongation by RNA pol II (Stock et al., 2007; Zhou et al., 2008).

PRC2 is often associated with CpG islands (Lee et al., 2006; Ku et al., 2008) but PRC2 has no known DNA sequence binding specificity and it is not clear how PRC2 associates with its target genes in mammals. Recent results have shown that PRC2 can interact with long non-coding (nc)RNA transcripts. PRC2 interacts with the 1.6 kb ncRNA RepA transcribed from *Xist* and this is necessary for H3K27me3 and X-chromosome inactivation (Zhao et al., 2008). Similarly, the ncRNA HOTAIR associates with PRC2 and induces methylation of the *HOXD* locus in trans (Rinn et al., 2007) and the antisense transcript Kcnq1ot1 associates with PRC2 and induces methylation of *Kcnq1* (Pandey et al., 2008).

Because polycomb target genes show evidence of stalled RNA polymerase II and do not transcribe appreciable amounts of mRNA, we hypothesised that short RNAs may be transcribed from these genes. We describe here the identification of such a set of short RNAs transcribed from the 5′ ends of polycomb target genes in primary human CD4+ T-cells and in murine ES cells and present evidence for their interaction with PRC2.

RESULTS

Detection of short RNAs associated with the 5′ end of genes

We designed a DNA microarray to identify short RNAs at the 5′ end of human genes. RNA from primary human CD4+ T-cells was fractionated by size into short (<200nt) and large RNAs, labeled with fluorescent dyes and hybridized to the microarrays in replicate experiments (Figure 1A and S1). The arrays contained multiple probes allowing detection of sense-strand RNA at the promoter, 5′-most exons, 1st intron and 3′-most exons of protein-coding genes (Supplemental Data and Table S1). These arrays also contained probes for known short nuclear (sn)RNA and short nucleolar (sno)RNA that acted as positive controls (Figures 1B and S1).

We investigated whether we could detect short RNAs transcribed from protein-coding genes. Applying a threshold derived from our snRNA probes, we found thousands of short RNAs transcribed from the sense-strands of promoters, exons and introns of protein-coding genes in replicate experiments (Figure 1C, Table S2). The RNAs were concentrated at the 5′ end of genes, within 700bp upstream and downstream of the mRNA TSS (Figure 1D). A substantial subset of the genes associated with short RNAs in primary human CD4+ T-cells also produced short RNAs in cell lines and ES cells (Kapranov et al., 2007; Seila et al., 2008) (Figure 1E), although often from different locations (Figure S1). These results indicate that short RNAs are generated in primary human T-cells, are most abundant in promoter-proximal sequences, and are a general feature of the transcriptome of normal somatic cells.

Transcription of short RNAs from genes that are otherwise repressed

The short RNAs described in previous reports are associated with transcriptionally active genes but not with repressed polycomb target genes (Affymetrix/CSHL ENCODE, 2009; Core et al., 2008; Kapranov et al., 2007; Seila et al., 2008; Taft et al., 2009). We therefore examined the transcriptional status of T-cell genes that produce short RNAs. We found that short RNAs could be detected both at genes that produced mRNA transcripts and at genes that do not (Figure 2A). Although the mRNA signals are quite different between these two sets of genes, the short RNA signals are similar (Figure 2B). This indicates that the transcription of short RNAs can occur in the absence of mRNA transcription.

We next used ChIP-Chip to determine whether RNA pol II was present at genes that transcribe short RNAs but not mRNA (Figure 2C). The antibody used preferentially recognises the non-phosphorylated form of RNA pol II that is recruited to chromatin (Lee et al., 2006). We plotted average enrichment for RNA pol II across the TSS of genes that transcribe short RNA but not mRNA and compared this to genes that transcribe neither RNA type. We found that RNA pol II was enriched at genes that produce short RNAs but not mRNA. This enrichment of RNA pol II at these genes was also apparent using data from independent ChIP-Seq experiments (Barski et al., 2007) (Figure S2).

At genes that produce short RNAs but not mRNA, the location of the short RNA correlates with the position of the initiating form of RNA pol II. At the set of genes where short RNA was detected at promoter regions, the RNA pol II peak was shifted upstream of the mRNA TSS (Figures 2D and S2). At the set of genes where short RNA was detected at exon or introns, RNA pol II occupancy was enriched downstream of the mRNA TSS. The positioning of RNA pol II at the site of short RNA production, rather than the mRNA TSS, could also be observed at individual genes (Figure 2E). At genes that produce both short RNA and mRNA, RNA pol II was concentrated at the mRNA TSS (Figure S2). These results argue that the RNA pol II observed at repressed genes is primarily involved in the transcription of short RNA and that short RNAs can be produced from TSS distinct from the mRNA TSS.

We next examined the transcriptional status of RNA pol II at short RNA loci. We used ChIP-Chip to map H3K4me3, a marker of transcriptional initiation, and H3K79me2, a marker of transcription through a gene (Guenther et al., 2007; Steger et al., 2008). Looking across all human genes, we found that H3K79me2 showed comparable enrichment to H3K4me3 (Figure 2F). In contrast, genes that produced short RNAs but not mRNAs were enriched for H3K4me3 but not H3K79me2 (Figure 2G and Figure S2). We conclude that these genes associated with short RNAs experience transcriptional initiation but the protein-coding portion of the gene is not transcribed.

Repressed genes producing short RNAs are targets for polycomb

Genes repressed by polycomb proteins are generally associated with nucleosomes containing H3K4me3, a marker of transcriptional initiation, and often RNA pol II (Azuara et al., 2006; Bernstein et al., 2006; Roh et al., 2006; Stock et al., 2007; Chopra et al., 2009). We therefore considered the possibility that genes that generate short transcripts but not mRNA could be associated with polycomb proteins. To test this, we used ChIP-Chip to measure H3K27me3 and compared the results to our short RNA data. We found that genes that transcribed short RNA but not mRNA were enriched for H3K27me3 (Figures 3A and S3). H3K27me3 could be detected more frequently at genes that transcribe short RNA but not mRNA than at any other category of gene (Figure 3B), demonstrating the close association between short RNA transcription and H3K27 methylation. Sites of short RNA transcription were also commonly associated with CpG islands (Figure 3C), previously linked to polycomb recruitment (Ku et al., 2008).

Transcription of short RNAs from polycomb target genes would explain the association of these genes with H3K4me3. We therefore sought to verify that the transcriptional machinery was present at H3K27-methylated genes that express short RNAs. Plotting enrichment of H3K27me3, H3K4me3 and RNA pol II across these genes confirms that transcriptional initiation occurs at polycomb target genes in CD4+ T-cells (Figures 3D and S3). Strikingly, H3K27-methylated nucleosomes flank the sites of RNA pol II and H3K4me3 occupancy, resembling the pattern of H3K27me3 around Drosophila polycomb response elements (Schwartz et al., 2006). PRC2 targets developmental regulators that must be repressed to maintain cell identity (Barski et al., 2007;Boyer et al., 2006;Lee et al., 2006). We therefore asked whether the set of genes that expressed short RNAs in the absence of mRNA were enriched for genes that play a role in development. Using Gene Ontology (GO), we found that functional categories such as multicellular development and cell-cell signalling were significantly enriched in the set of genes from which short RNA was transcribed in the absence of mRNA (Figure 3E). Therefore, consistent with polycomb-mediated silencing, genes from which short RNA is transcribed in the absence of mRNA tend to have functions in development and cell differentiation.

Short RNAs transcribed from polycomb target genes are ~50-200 nt in length

We next used Northern blotting to verify the transcription of short RNAs from polycomb target genes and to perform a more accurate determination of their size. We selected array probes that detected short RNAs at genes for which no mRNA could be detected and were also associated with H3K27me3. We then purified short RNA from peripheral blood mononuclear cells (PBMC), treated with DNase and performed northern blotting for short RNAs (Figure 3F). 17 of the 22 probes we tested over the course of this study (77%) detected a short RNA species (Table S3). Some probes identified a single RNA species of between ~50 and 200 nt, while others detected multiple RNA products. Short RNA transcripts could be detected from exons, introns and promoter regions, consistent with our array data.

Short RNA transcription is not dependent on polycomb activity

We considered that short RNA production might be a by-product of polycomb activity and thereby dependent on H3K27 methylation. To test this, we made use of the murine ES cell line Ezh2-1.3, in which deletion of the Ezh2 SET domain can be induced by tamoxifen (Pereira et al., unpublished data). We first sought to establish if short RNAs are transcribed from polycomb target genes in ES cells. We identified short RNA loci conserved between human and mouse and used histone methylation data (Boyer et al., 2006) to identify those targeted by polycomb in ES cells. Northern blotting detected a short RNA at each of these genes in murine ES cells (Figure 4A), indicating that the transcription of short RNAs from repressed polycomb target genes is common to different cell types and involves loci that are conserved between different mammalian species.

We next measured short RNA transcription in Ezh2-1.3 cells at timepoints after the addition of tamoxifen. Loss of full-length Ezh2 protein (together with the appearance of a truncated form lacking the SET domain) and a loss of H3K27me3 could be observed in these cells over the five-day timecourse (Figure 4B). Blotting for short RNAs at the genes Hes5, Msx1 and Ybx2 showed that loss of Ezh2 and H3K27me3 had no effect on the levels of short RNAs at these genes (Figure 4C). Although we did not observe an increase in mRNA expression from these genes upon Ezh2 deletion (Figure S4), our results show that the production of short RNAs at polycomb target genes is not dependent on H3K27 methylation.

The block to mRNA transcription at bivalent genes is dependent on the PRC1 component Ring1 that catalyzes the ubiquitination of histone H2A. Deletion of Ring1 causes activation of PRC1 target genes, including Msx1 (Stock et al., 2007; Figure S4). We blotted for short RNAs in the

murine ES cell line ES-ERT2 with and without addition of tamoxifen that induces deletion of *Ring1b* and loss of H2AK119ub (Stock et al., 2007). As we found for deletion of *Ezh2*, loss of Ring1b had no effect on short RNA transcription (Figure 4D). These data show that transcription of short RNAs is not dependent on H2A ubiquitination and, taken together with the Ezh2 deletion experiments, indicate that short RNA transcription is independent of polycomb activity.

Short RNAs encode stem-loop structures and interact with PRC2

If short RNAs are not a product of polycomb-mediated gene silencing they might instead act upstream. Supporting this hypothesis, the long ncRNAs HOTAIR and Xist RepA interact with PRC2 and this interaction is necessary for H3K27me3 of *HOXD* and the X-chromosome, respectively (Rinn et al., 2007; Zhao et al., 2008). We therefore considered the possibility that PRC2 may interact with the short ncRNAs identified here.

The PRC2 binding site within mouse Xist RepA appears to be a double stem-loop structure that is repeated 7 times (Zhao et al., 2008). We therefore first examined whether the short RNAs we have identified could form such a structure. We derived a general structural motif based on the RepA sequence and searched for the presence of this motif in the DNA sequences immediately surrounding probes that detect short RNAs from H3K27-methylated genes (Figure 5A). We found that 71% of these sequences encode this PRC2-binding structure, compared to 36% of control sequences not associated with short RNAs but with an equal distribution around the mRNA TSS. Examples of these structures are given in Figures 5B and S5. These data indicate that short RNAs transcribed from polycomb target genes have the potential to interact with PRC2. However, we found that these structures were not limited to short RNAs transcribed from polycomb-associated genes and that they were also present within short RNAs transcribed from genes not associated with H3K27me3 (Figure 5A). These RNAs may therefore also have the potential to bind PRC2.

To test for an interaction between short RNAs encoding the stem-loop structure and PRC2, we performed electrophoretic mobility shift assays (EMSA). Consistent with recent observations (Zhao et al., 2008), we found that incubation of T-cell lysate with radiolabeled RNA oligonucleotides encoding the Xist-RepA stem-loop produced a mobility shift and that mutation of the RNA stem-loop structure abolished this interaction (Figure 5C). We then repeated the experiment with RNA oligonucleotides corresponding to the stem-loops encoded by BSN, C20orf112, HEY1, MARK1 and PAX3 short RNAs. We observed a similar shift with RNA oligonucleotides corresponding to these short RNA stem-loop structures as we did for Xist-RepA and did not observe an interaction when the BSN short RNA structure was disrupted (Figure 5C). The BSN short RNA stem-loop was also able to compete with Xist-RepA for binding (Figure 5D), indicating that the interaction between PRC2 and Xist-RepA is similar to that between PRC2 and short RNAs.

We next sought to identify which PRC2 component was responsible for the interaction with short RNAs. We purified GST-tagged recombinant SUZ12, EZH2, EED and RBBP4 from E. coli and incubated each protein with radiolabeled Xist-RepA and BSN oligonucleotides (Figure 5E). We found that SUZ12 interacted strongly with Xist-RepA and BSN short RNA stem-loops. SUZ12 also interacted with other short RNA stem-loops (Figure 5F), displayed a weaker interaction with mutated Xist-RepA RNA (Figure 5G) and did not bind to the BSN sequence encoded by single-stranded DNA, double-stranded DNA or DNA:RNA duplexes (Figure 5H). These data demonstrate that SUZ12 specifically interacts with RNA stem-loop structures encoded by Xist-RepA and short RNAs.

PRC2 interacts with short RNAs in cells

We next sought to verify that the interaction between PRC2 and short RNA occurred in living cells. We immunoprecipitated SUZ12 from a female T-cell line, isolated co-purifying RNA and subjected this to quantitative reverse-transcription (RT)-PCR (Figure 6A). Immunoprecipitation with the SUZ12 antibody, but not a non-specific rabbit control antibody, enriched for Xist RNA over Actin mRNA and mRNAs encoding GAPDH and HPRT, indicating we could specifically detect PRC2-RNA interactions in these cells. We then performed RT-PCR for short RNAs transcribed from polycomb target genes. We found that 4 of the 5 short RNAs we tested were enriched by SUZ12 IP, although to a lesser extent than Xist. These short RNAs also possessed a stem-loop structure (Figure S5). Amplification of short RNAs was not observed in control reactions lacking RT and enrichment was maintained under stringent wash conditions (Figure S6). We also examined short RNAs transcribed from genes that were not associated with H3K27me3. We found that these were also often enriched by SUZ12 IP, consistent with the presence of stem-loop structures in these transcripts (Figure 6A). Therefore, PRC2 interacts with short promoter-associated RNA but this interaction does not necessarily produce detectable levels of H3K27me3. The snRNAs U1, U2 and U3 and the short structured RNA pol III transcripts 7SK and 5S ribosomal RNA were not enriched by SUZ12 IP, demonstrating that enrichment of short RNA was specific to the set transcribed from the 5′ ends of genes.

PRC2-binding RNA stem-loops cause gene repression in cis

Our data suggest a model in which short RNAs transcribed at the 5′ end of polycomb target genes interact with PRC2, stabilising its interaction with chromatin in cis. This model suggests that addition of sequences encoding PRC2-binding stem-loop structures to the 5′ end of a gene would lead to ectopic PRC2 binding and consequent H3K27 methylation and gene repression. To test the model, we introduced sequences encoding PRC2-binding stem loops into the HIV long terminal repeat (LTR). The R portion of the LTR located immediately downstream of the TSS encodes the structured transactivation response element (TAR) RNA that is transcribed as a short RNA and as the 5′ UTR of HIV mRNA transcripts. We replaced the R and U5 portions of the LTR with the Xist-RepA stem-loop or the stem-loop from the short RNA at C20orf112, which also interacts with PRC2 (Figures 5 and 6A), and cloned the luciferase coding region downstream (Figure 6B).

We first examined whether the incorporation of these PRC2-binding RNA stem-loops into the HIV LTR would allow binding by PRC2. The different constructs were transfected into Hela cells and the enrichment of luciferase RNA in SUZ12 IP material relative to input RNA measured by quantitative RT-PCR (Figure 6C). We found that addition of Xist-RepA and short RNA stem-loops resulted in the specific enrichment of luciferase RNA in the SUZ12 IP fraction.

We next tested whether incorporation of PRC2-targeting RNA stem-loops affects expression of the luciferase gene (Figure 6D). We found that there was a significant drop ($p<0.05$) in luciferase activity in cells transfected with constructs containing wild-type Xist-RepA or short RNA stem-loops. Furthermore, mutation of the stem-loop sequences restored luciferase activity.

Finally, we asked whether incorporation of PRC2-binding RNA stem-loops caused an increase in H3K27 methylation. We performed ChIP for H3K27me3 and total H3 and measured enrichment of luciferase construct DNA by quantitative PCR (Figure 6E). We found that the constructs encoding PRC2-binding RNA stem-loops showed a small increase in H3K27 methylation compared to the wild-type LTR construct. This increase in H3K27me3 was not observed with constructs in which the RNA stem-loops had been mutated. Taken together with

our other experiments, these data show that the incorporation of Xist-RepA and short RNA stem-loop sequences at the 5′ end of genes allows PRC2 binding, H3K27me3 and gene repression in cis.

Short RNAs are lost from polycomb target genes active in other cell types

Differential H3K27 methylation allows for cell-type specific expression of developmental regulators and this underlies the differing identities of specialized cell types. We hypothesised that short RNAs may be specifically depleted from polycomb target genes that are derepressed in other cell types. To address this, we examined short RNA expression in the neuronal cell line SH-SY5Y. We choose neuronal cells because our GO analysis (Figure 3E) revealed that many genes that express short RNAs in T-cells play roles in neuronal development. We used gene expression and functional data to select polycomb target genes that are repressed in T-cells and active in neuronal tissue (Figure S7). We then purified short RNAs from both PBMC and the neuronal cultures and blotted for short RNAs identified at these genes in T-cells (Figure 7A). We found that the increased expression of mRNA from the genes *FOXN4*, *HEY1*, *MARK1*, *NKX2-2*, *BSN* and *HES5* in the neuronal cells was accompanied by a reduction in the expression of short RNAs. In contrast, short RNAs transcribed from *YBX2*, a gene expressed only in germ cells, or from the thyroid and lung-specific *NKX2-1*, were present equally in PBMC and neuronal cells. These results show that high-levels of short RNA transcription is specific to polycomb target genes silent in a given cell type.

We next asked whether activation of polycomb target genes during differentiation of ES cells into specialised cell types is accompanied by loss of short RNAs. To test this, we differentiated murine ES cells over a period of 4 days, through embryoid bodies and neural precursor cells to precursor motor neurons (Wichterle et al., 2002). During this process there is an increase in the expression of mRNA encoding the neuronal proteins Hes5 and Pcdh8 (Figure 7B). Blotting for short RNAs at these genes reveals that activation of *Hes5* and *Pcdh8* is accompanied by a progressive drop in the levels of the ~190 nt *Hes5* and ~100 nt *Pcdh8* short RNAs (Figure 7C). As the levels of the ~190 nt *Hes5* short RNA decrease, there is a concomitant increase in the levels of shorter species, implying that the longer RNA is degraded. We conclude that derepression of polycomb target genes during ES cell differentiation is accompanied by a decrease in short RNA transcripts.

DISCUSSION

We report the identification of a class of short RNAs transcribed from the 5′ end of polycomb target genes in primary CD4+ T-cells and ES cells. These short RNAs are ~50-200nt in length and are transcribed from promoters, introns and exons by RNA pol II, independently of polycomb activity. The short RNAs interact with PRC2 through a stem-loop structure, causing gene repression and H3K27 methylation in cis, and are depleted from polycomb target genes derepressed during cell differentiation. These results support a model in which short RNAs transcribed from PRC2 target genes stabilise the association of PRC2 with chromatin (Figure 7D).

Relationship with other short RNAs

Short RNAs of less than 200nt have previously been found to be transcribed from the 5′ end of active genes in ES cells and cell lines (Affymetrix/ CSHL Laboratory ENCODE 2009; Kapranov et al., 2007, Core et al., 2008; Seila et al., 2008; Taft et al., 2009) and our detection of short RNAs in primary differentiated human cells demonstrates that they are a core component of the transcriptome. Previous studies have not identified short RNAs at polycomb target genes for several likely reasons. Transcription run-on techniques (Core et al., 2008) only detect transcripts in the process of being synthesised, so the absence of polycomb-associated

RNAs from this dataset may be because the bulk of the short RNAs we detect are not associated with actively transcribing polymerase. Analysis of 18-30nt RNAs (Seila et al., 2008; Taft et al., 2009) will also miss polycomb-associated short RNAs that our northern blotting evidence indicates are primarily ~50 to 200 nt in size (Figure 3F). The appearance of smaller RNAs during *Hes5* activation suggests how the different sizes of RNA may be related.

Implications for bivalent chromatin

The transcription of short RNAs from polycomb target genes explains why these loci are often associated with H3K4me3 (Azuara et al., 2006; Bernstein et al., 2006; Roh et al., 2006). The association of genes with both H3K27me3 and H3K4me3 was first described in ES cells and was hypothesised to poise genes for activation during subsequent stages of embryogenesis (Azuara et al., 2006; Bernstein et al., 2006; Lee et al., 2006). The detection of RNA pol II phosphorylated at ser-5, but not ser-2, at polycomb target genes in ES cells suggested that the poised state reflected the stalling of RNA pol II during mRNA transcription, with this block being released upon ES cell differentiation (Stock et al., 2007). We find that short RNAs are often transcribed from sites independent of the mRNA TSS (Figure 2) and that this is independent of polycomb activity (Figure 4). These results argue that the bivalent state can reflect short RNA production *per se* and is not necessarily a consequence of RNA pol II stalling during mRNA transcription. This would explain why genes are associated with H3K27me3 and H3K4me3 in differentiated cells even though such genes are unlikely to be poised in preparation for subsequent activation (for example neuronal genes in T-cells). Gene activation may switch RNA pol II from a form that transcribes short RNA to a form that transcribes mRNA.

Interaction with PRC2

We have shown that there is no change in short RNA levels after loss of H3K27 methylation or H2A ubiquitination (Figure 4). However, the same short RNAs decline upon differentiation of ES cells to neurons (Figure 7), indicating that their invariance upon Ezh2 and Ring1 loss does not merely reflect their stability. These data instead indicate that short RNAs could function upstream of polycomb and RNA binding data suggests that this function is to stabilize the association of PRC2 with chromatin. The interaction of PRC2 with ncRNAs was previously described at *HOXD*, the inactive X-chromosome and Kcnq1 (Rinn et al., 2007;Pandey et al., 2008;Zhao et al., 2008). The polycomb protein Cbx7 also interacts with chromatin in an RNA-dependent manner (Bernstein et al., 2006). A more recent study described a set of large intergenic ncRNAs that associate with PRC2 (Khalil et al., 2009) but how these may effect association of polycomb with protein-coding genes is unclear. The present study, demonstrating an interaction between PRC2 and short ncRNAs transcribed from polycomb-target genes themselves, presents a simple model for the association of PRC2 with its target genes.

We have found that short RNAs possess stem-loop structures similar to those present within Xist RNA. This argues that the repeated Xist-RepA stem-loop is a specialised form of an RNA motif found in many PRC2-interacting RNAs. Of all the PRC2 subunits, we found that SUZ12 interacted most strongly with both Xist-RepA and short RNA stem-loops (Figure 5E). The weak interaction observed between EED and RNA (Figure 5) suggests it too may play a role. A previous study (in which the interaction of SUZ12 with RNA was not assessed) suggested that Ezh2 interacted directly with the Xist-RepA stem-loop (Zhao et al., 2008) but we only observed a weak interaction using EZH2 purified from bacteria (Figure S5). It is possible that the use of baculovirus as a source of recombinant Ezh2 by Zhao and colleagues could have resulted in co-purification of endogenous Su(z)12 or that bacterially-produced EZH2 protein lacks an important post-translational modification.

Decline of short RNAs during gene activation

Short RNA levels are reduced at polycomb-target genes derepressed in neurons and this decrease can be observed directly as ES cells differentiate. Given the interaction between short RNAs and PRC2, a reduction in short RNA levels would seem likely to destabilize the association of PRC2 with chromatin, thereby allowing demethylation of H3K27 and mRNA transcription. Although this provides a mechanism for how polycomb-target genes become activated, it does not explain how other genes expressing short RNAs capable of interacting with PRC2 (Figure 6A) remain free of H3K27 methylation. At this set of genes, it is possible that PRC2 is inactive, counteracted by a H3K27 demethylase or other RNA-binding transcriptional regulators, such as pTEFb.

In summary, we have identified a class of short RNAs that interact with PRC2 through a stem-loop structure, causing gene repression and H3K27 methylation in cis, and that are depleted from polycomb target genes derepressed during cell differentiation. These results explain why polycomb target genes are associated with markers of transcriptional activation and provide a mechanism for the interaction of PRC2 with its target loci across the genome.

EXPERIMENTAL PROCEDURES

RNA purification and fractionation

CD4+ T-cells were isolated from PBMC by magnetic cell sorting (Miltenyi). Total RNA was purified with Trizol (Invitrogen) and short RNA (<200bp) was purified using Ambion's mirVana miRNA purification system and size fractionation confirmed using an Agilent Bioanalyzer.

Microarray analysis

Short RNA was poly-adenylated (Ambion), amplified and labeled with Agilent's Low RNA Input Linear Amplification protocol and hybridised with labeled mRNA or fractionated long RNA to custom microarrays. The signals for probes for snRNAs and snoRNAs was used to devise an algorithm for the prediction of other short RNAs ($\log_2$ ratio >1.5, p-value <0.01, normalized signal >50 in both experiments).

ChIP-Chip

DNA associated with RNA pol II, H3K4me3, H3K79me2, H3K27me3 and total H3 was enriched with the following antibodies: RNA pol II (8WG16, Abcam), H3 (Abcam ab1791), H3K4me3 (Abcam ab8580), H3K79me2 (Abcam ab3594) and H3K27me3 (Abcam ab6002) and analysed with DNA microarrays (Agilent) according to published protocols (Lee et al., 2006).

Northern blotting

Short RNA was resolved on 15% acrylamide-7M Urea TBE gels (Invitrogen) and electroblotted to Nytran membranes (Whatman). RNA was crosslinked to the membrane and exposed to probes designed from microarray elements and radioactively labeled using StarFire 3′-extension (IDT). Estimates of RNA length were made in comparison to Century and Decade markers (Ambion).

Conditional deletion of Ezh2 and Ring1

The Ezh2-1.3 ES cell line was derived from a cross between mice carrying the floxed SET domain of Ezh2 (Su et. al., 2003) and mice carrying tamoxifen-inducible Cre at the Rosa locus (Cre-ERT2). The conditional *Ezh2* mutation, present on both alleles, was induced by 4-hydroxy-tamoxifen (800nM). Cells were maintained on 0.1% gelatin using KO-DMEM, 10%

FCS, 5% knock-out serum replacement (Invitrogen) and 1000U/ml leukaemia inhibitory factor (Chemicon). Conditional Ring1b cell lines were cultured and Ring1b deletion induced as described (Stock et al., 2007).

RNA structure prediction

RNA structures were identified within 200nt of sequence surrounding probes that detected short RNAs using RNAmotif (Macke et al., 2001) and free-energy structures predicted using RNAfold (Hofacker and Stadler, 2006).

EMSA

SUZ12 was amplified from the IMAGE clone and cloned into pGEX-4T1 (GE). Plasmids encoding other PRC2 subunits were kindly provided by Dr.Y. Zhang. GST-tagged proteins were purified using glutathione-sepharose beads and 4μg used for EMSA with oligonucleotides end-labeled with [γ-^{32}P]ATP, as described (Zhao et al., 2008). Nuclear extracts were prepared from CEM cells using reagents from Active Motif.

RNA IP

RNA associated with PRC2 was enriched from CEM cells with an antibody to SUZ12 (Abcam) or unspecific rabbit antibody (Santa Cruz) following published protocols (Zhao et al., 2008). RNA was treated with DNase-turbo (Ambion), and reverse transcribed using random priming (Invitrogen). Enrichment of RNA species compared to input RNA was quantified using real-time PCR and normalized to Actin mRNA.

Luciferase assay

R and U5 regions of the HIV LTR (SF2) were replaced with DNA encoding the murine Xist-RepA stem-loop, the stem-loop from C20orf112 short RNA or their mutated forms, cloned in place of the CMV promoter in pIRESneo3 (Clontech) and luciferase from pCSFLW inserted downstream. Plasmid DNA was transfected into Hela cells in triplicate together with the Renilla luciferase plasmid phRL-null. Firefly and Renilla luciferase activities were measured after 48 hours (Promega) and cells harvested for RNA IP and ChIP after 72 hours. Significance was estimated with a one-sided paired T-test.

Neuronal cell culture

SH-SY5Y cells were cultured in DMEM with 10% FCS until 40-50% confluent and then differentiated in DMEM with 5% FCS and 10μM retinoic acid (Sigma) for 7 days.

ES cell differentiation

ES cell-derived motor neuron precursors were generated as described (Wichterle et al., 2002). v6.5 mouse ES cells were partially dissociated and cultured in ADFNK medium. After two days, 1μM retinoic acid and 0.5μg/ml Hedgehog agonist Hh-Ag1.3 were added and cells cultured for an additional 24hrs (for NPCs) or 48hrs (for PMNs).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Thanks to Anne Palser for advice on SH-SY5Y culture. This work was funded by a MRC Career Development Award to RGJ, an MRC Centre grant that supports ASK and CCA, a project grant to RGJ and core funding from the UCLH/

UCL Comprehensive Biomedical Research Centre and NIH grants HG002668 and NS055923 to RAY. KV acknowledges support from the Orion-Farmos Research Foundation and Academy of Finland.

REFERENCES

Affymetrix/CSHL ENCODE Transcriptome Project. Post-transcriptional processing generates a diversity of 5′-modified long and short RNAs. Nature 2009;457:1028–1032. [PubMed: 19169241]

Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, et al. Chromatin signatures of pluripotenT-cell lines. NaT-cell Biol 2006;8:532–538. [PubMed: 16570078]

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. Cell 2007;129:823–837. [PubMed: 17512414]

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 2006;125:315–326. [PubMed: 16630819]

Bernstein E, Duncan EM, Masui O, Gil J, Heard E, Allis CD. Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. Mol Cell Biol 2006;26:2560–2569. [PubMed: 16537902]

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 2006;441:349–353. [PubMed: 16625203]

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 2002;298:1039–1043. [PubMed: 12351676]

Chopra VS, Hong JW, Levine M. Regulation of Hox gene activity by transcriptional elongation in Drosophila. Curr Biol 2009;19:688–693. [PubMed: 19345103]

Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 2008;322:1845–1848. [PubMed: 19056941]

Dellino GI, Schwartz YB, Farkas G, McCabe D, Elgin SC, Pirrotta V. Polycomb silencing blocks transcription initiation. Mol Cell 2004;13:887–893. [PubMed: 15053881]

Faust C, Schumacher A, Holdener B, Magnuson T. The eed mutation disrupts anterior mesoderm production in mice. Development 1995;121:273–285. [PubMed: 7768172]

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. Cell 2007;13:77–88. [PubMed: 17632057]

Hofacker IL, Stadler PF. Memory efficient folding algorithms for circular RNA secondary structures. Bioinformatics 2006;22:1172–1176. [PubMed: 16452114]

Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermuller J, Hofacker IL, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 2007;316:1484–1488. [PubMed: 17510325]

Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci USA 2009;106:11667–11672. [PubMed: 19571010]

Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, Presser A, Nusbaum C, Xie X, Chi AS, et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet 2008;10:e1000242. [PubMed: 18974828]

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 2006;125:301–313. [PubMed: 16630818]

Macke TJ, Ecker DJ, Gutell RR, Gautheret D, Case DA, Sampath R. RNAMotif, an RNA secondary structure definition and search algorithm. Nucleic Acids Res 2001;29:4724–4735. [PubMed: 11713323]

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 2007;448:553–560. [PubMed: 17603471]

O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T. The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol 2001;21:4330–4336. [PubMed: 11390661]

Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell 2008;32:232–246. [PubMed: 18951091]

Pasini D, Bracken AP, Hansen JB, Capillo M, Helin K. The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol Cell Biol 2007;27:3769–3779. [PubMed: 17339329]

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007;129:1311–1323. [PubMed: 17604720]

Roh TY, Cuddapah S, Cui K, Zhao K. The genomic landscape of histone modifications in human T-cells. Proc Natl Acad Sci. USA 2006;103:15782–15787. [PubMed: 17043231]

Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, Pirrotta V. Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat Genet 2006;38:700–705. [PubMed: 16732288]

Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, Young RA, Sharp PA. Divergent transcription from active promoters. Science 2008;322:1849–1851. [PubMed: 19056940]

Steger DJ, Lefterova MI, Ying L, Stonestrom AJ, Schupp M, Zhuo D, Vakoc AL, Kim JE, Chen J, Lazar MA, et al. DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. Mol Cell Biol 2008;28:2825–2839. [PubMed: 18285465]

Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koseki H, Brockdorff N, Fisher AG, Pombo A. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. NaT-cell Biol 2007;9:1428–1435. [PubMed: 18037880]

Su I, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT, Tarakhovsky A. Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat Immunol 2002;4:124–131. [PubMed: 12496962]

Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, et al. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci. USA 2004;101:6062–6067. [PubMed: 15075390]

Taft RJ, Glazov EA, Cloonan N, Simons C, Stephen S, Faulkner GJ, Lassmann T, Forrest AR, Grimmond SM, Schroder K, et al. Tiny RNAs associated with transcription start sites in animals. Nat Genet 2009;41:572–578. [PubMed: 19377478]

van der Stoop P, Boutsma EA, Hulsman D, Noback S, Heimerikx M, Kerkhoven RM, Voncken JW, Wessels LF, van Lohuizen M. Ubiquitin E3 ligase Ring1b/Rnf2 of polycomb repressive complex 1 contributes to stable maintenance of mouse embryonic stem cells. PloS One 2008;3:e2235. [PubMed: 18493325]

Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS. Hierarchical recruitment of polycomb group silencing complexes. Mol Cell 2004;14:637–646. [PubMed: 15175158]

Wichterle H, Lieberam I, Porter JA, Jessell TM. Directed differentiation of embryonic stem cells into motor neurons. Cell 2002;110:385–397. [PubMed: 12176325]

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]

Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J, Glass CK, Rosenfeld MG. Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. Mol Cell 2008;29:69–80. [PubMed: 18206970]

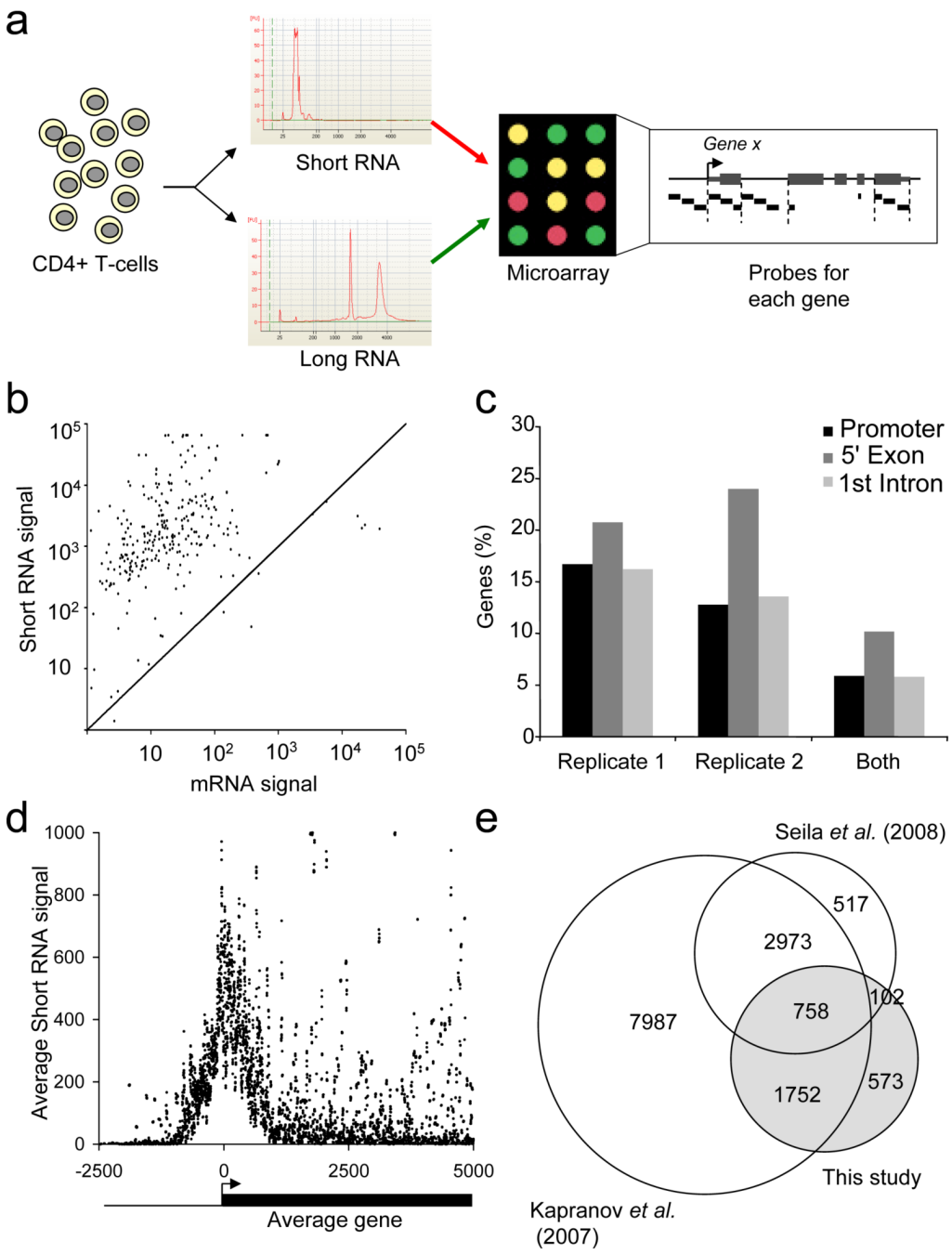

Figure 1. Detection of promoter-associated short RNAs in primary human T-cells

A. Experimental strategy. Short RNA was purified from human CD4+ T-cells, poly-adenylated, amplified and labeled with Cy5 and hybridized together with Cy3-labeled mRNA (experiment 1) or total long RNA (experiment 2) to a custom microarray containing probes for the promoters, 5′ exons, first intron and 3′ exons of human RefSeq genes, together with control probes for snRNA and snoRNAs.

B. Background-subtracted array signals for control snRNA and snoRNA probes in the mRNA channel (x-axis) and the short RNA channel (y-axis).

C. Fraction of promoter regions (~1 kb upstream of mRNA transcription start site), 5′ exons (up to 1kb downstream of mRNA 5′ end) and first introns that give rise to short RNAs.

D. Distribution of short RNAs relative to the TSS of the average gene, plotted as a moving average with a window of 10 bp.

E. Venn diagram showing the overlap between genes with associated short RNAs detected in this study compared with two previous studies (Kapranov et al., 2007; Seila et al., 2008).

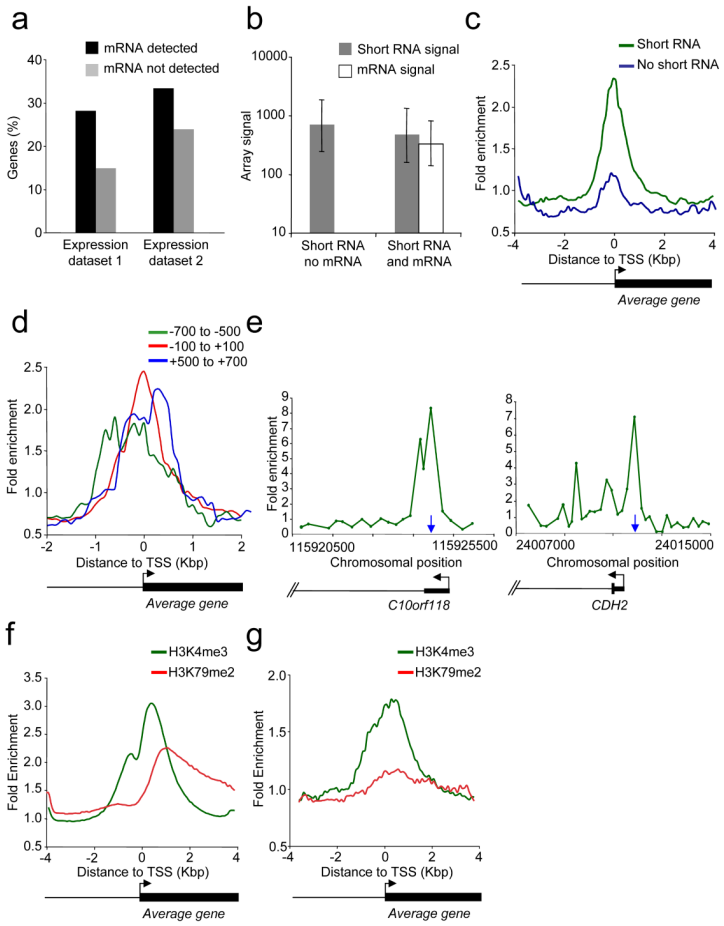

Figure 2. Transcriptional status of short RNA genes

A. Percentage of genes with promoter-associated short RNAs that produce detectable mRNA (black bars) or no detectable RNA (grey bars) according to our mRNA expression data (1) or published data (2, (Su et al., 2004)).

B. Array signals (median and inter-quartile range) for short RNAs (grey) and mRNA (white) at genes that produce detectable mRNA or no detectable mRNA.

C. Composite enrichment profile of RNA pol II at genes for which no mRNA can be detected. The plot shows average fold enrichment (normalized signal from RNA pol II ChIP vs. input DNA) and genes are divided into those that are associated with short RNA (green) and those not associated with short RNA (blue). The start and direction of transcription of the average gene is indicated by an arrow.

D. Composite enrichment profile of RNA pol II at genes that express short RNAs in the absence of detectable RNA. Genes are divided into those that detect a short RNA with probes positioned −100 to +100 (red), −700 to −500 (green) or +500 to +700 (blue) relative to the mRNA TSS.

E. Examples of RNA pol II ChIP signals at short RNA loci. The plots show unprocessed enrichment ratios for all probes within a genomic region (RNA Pol II ChIP vs. whole genomic DNA). Chromosomal positions are from Build 35. Genes are shown to scale below and aligned with the plots by chromosomal position. The positions of short RNAs are indicated by the vertical blue arrow.

F. Composite enrichment profile of H3K4me3 (green) and H3K79me2 (red) across all RefSeq genes. The plot shows average fold-enrichment (normalized signal from H3 methylation ChIP vs. total H3 ChIP).

G. As F., except at the set of genes associated with short RNA but not mRNA.

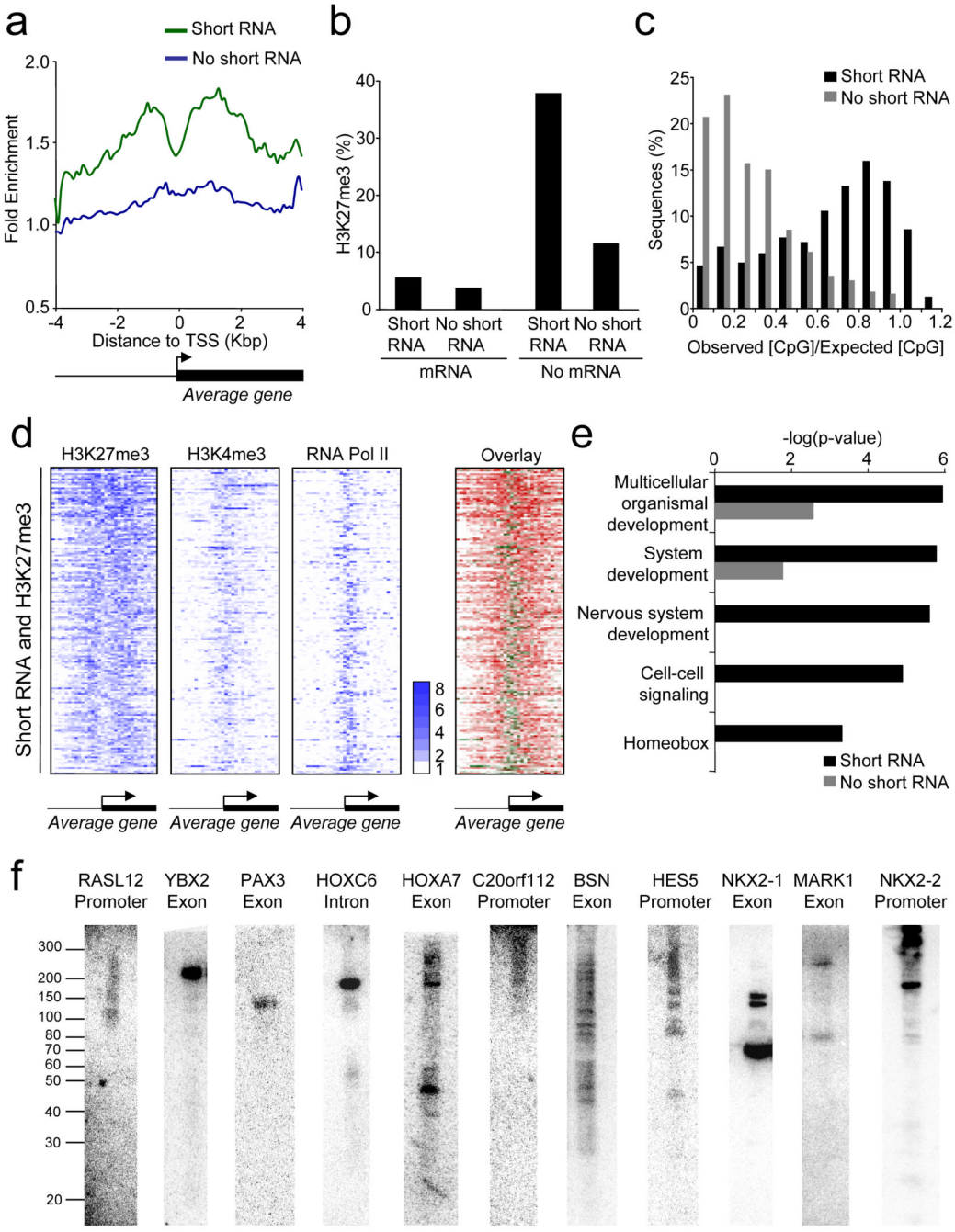

Figure 3. Short RNA loci are associated with H3K27me3

A. Composite enrichment profile of H3K27me3 at genes for which no mRNA can be detected, divided into those that are associated with short RNA (green) and those not associated with short RNA (blue). The plot shows average fold-enrichment (normalized signal from H3K37me3 ChIP vs. total H3 ChIP).

B. Percentage of genes in different transcriptional categories that are associated with H3K27me3. Genes are first divided into those that produce detectable mRNA and those that do not and then into those that produce detectable short RNA and those that do not.

C. CpG content of short RNA loci not associated with mRNA. The data are plotted as a histogram of the observed CpG content divided by the expected CpG content and are compared to control genes that do not produce detectable short RNA or mRNA.
D. Heat maps showing enrichment of H3K27me3, H3K4me3 and RNA pol II at genes associated with short RNAs and H3K27me3. Each row represents one gene and each column represents the data from one probe, ordered by their position relative to the TSS. For the first three panels, fold enrichment is indicated by color, according to the scale on the right. The last panel overlays the RNA pol II (green) and H3K27me3 (red) to show their relative locations.
E. P-values for the enrichment of GO categories in the set of genes that are not associated with mRNA and transcribe short RNA (black bars) compared with those that do not produce detectable short RNA (grey bars).
F. Northern blotting for short RNAs transcribed from polycomb target genes in PBMC. The position of single-stranded RNA size markers are shown to the left. The genes and relative locations from which the short RNAs are transcribed are indicated above each blot.

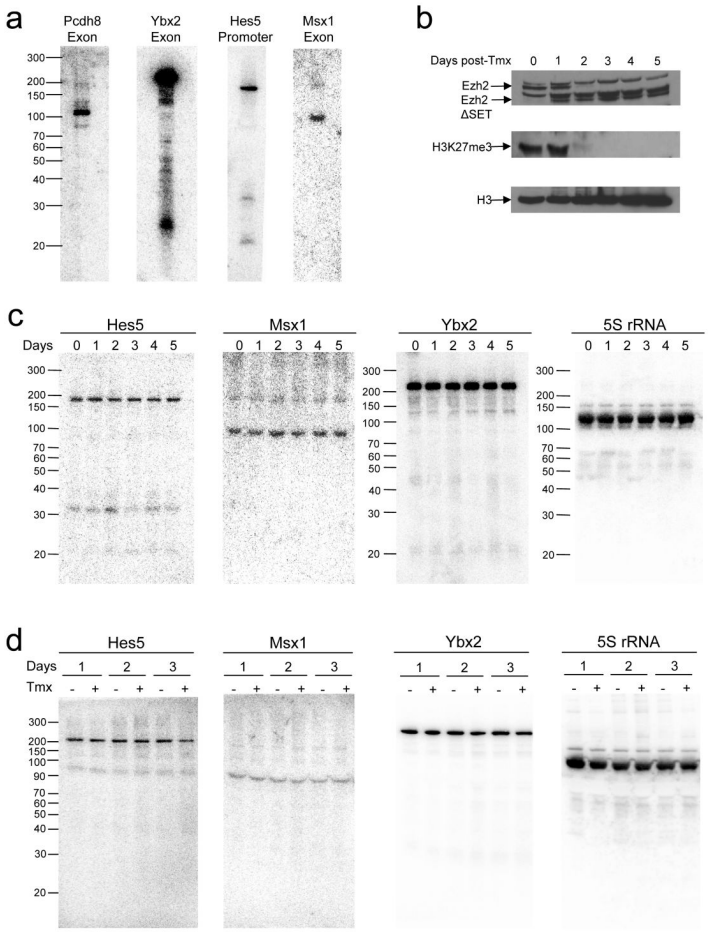

Figure 4. Transcription of short RNAs in murine ES cells deficient for Ezh2 and Ring1

A. Northern blotting for short RNAs transcribed from polycomb target genes in murine ES cells.

B. Western blotting for Ezh2, histone H3K27me3, and total histone H3 in Ezh2-1.3 ES cells over a 5-day treatment with tamoxifen that induces genetic deletion of the Ezh2 SET domain. Specific Ezh2 bands are marked.

C. Northern blotting for Hes5, Msx1 and Ybx2 short RNAs and 5S rRNA during tamoxifen treatment of Ezh2-1.3 cells.

D. Northern blotting in ES-ERT-cells, 1, 2 and 3 days after the addition of tamoxifen that induces genetic deletion of *Ring1b*, compared with untreated cells at the same timepoints.

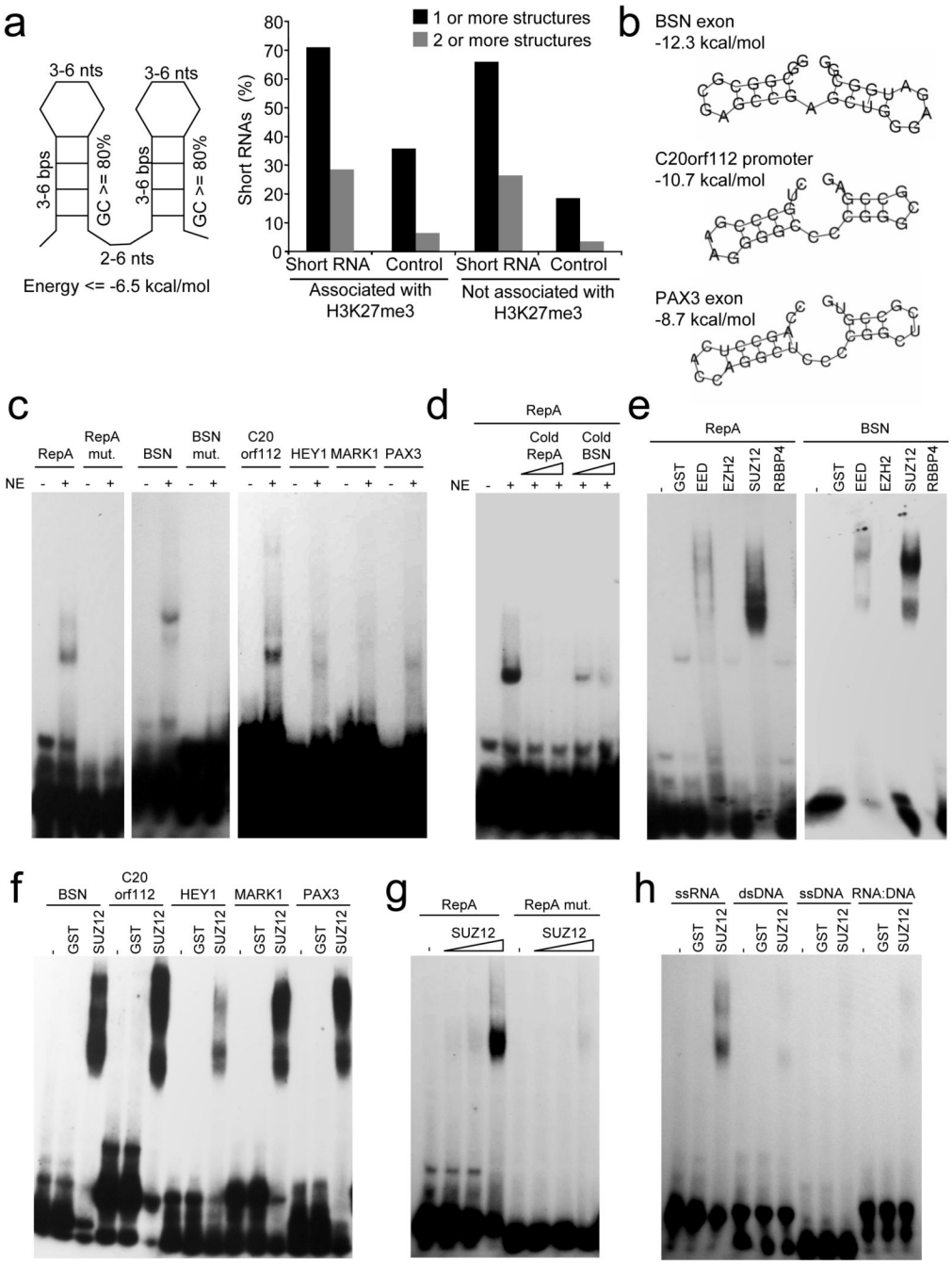

Figure 5. PRC2 interacts with short RNAs in vitro

A. Left: A stem-loop structure motif based on the repeat region of Xist-RepA RNA known to bind PRC2. Right: Percentage of short RNAs transcribed from genes associated with H3K27me3 that contain one or more (black bars) or two or more (grey bars) of the RNA structure motifs. The same data for genomic sequences showing identical distributions around TSS but that do not produce detectable short RNAs are shown as controls.

B. Example energy structures predicted in short RNAs transcribed from repressed genes.

C. EMSA for the interaction of ^{32}P-labeled RNA probes corresponding Xist-RepA, mutated Xist-RepA, BSN short RNA, mutated BSN short RNA, and C20orf112, HEY1, MARK1 and PAX3 short RNA probes with T-cell nuclear extract (NE).

D. Competition for Xist-RepA binding with 400 or 800-fold excess of cold Xist-RepA or cold BSN short RNA.
E. EMSA showing the interaction of PRC2 subunits expressed as GST-fusion proteins with Xist-RepA and BSN short RNA probes. GST alone was used a negative control.
F. EMSA showing the interaction of GST-SUZ12 with short RNA probes.
G. Affinity of GST-SUZ12 (0.25, 1 and 5μg) for wild-type and mutant Xist-RepA probes.
H. Binding specificity of SUZ12 defined by incubating GST-fusion protein with the BSN stem-loop sequence encoded by ssRNA, ssDNA, dsDNA and a RNA-DNA duplex.

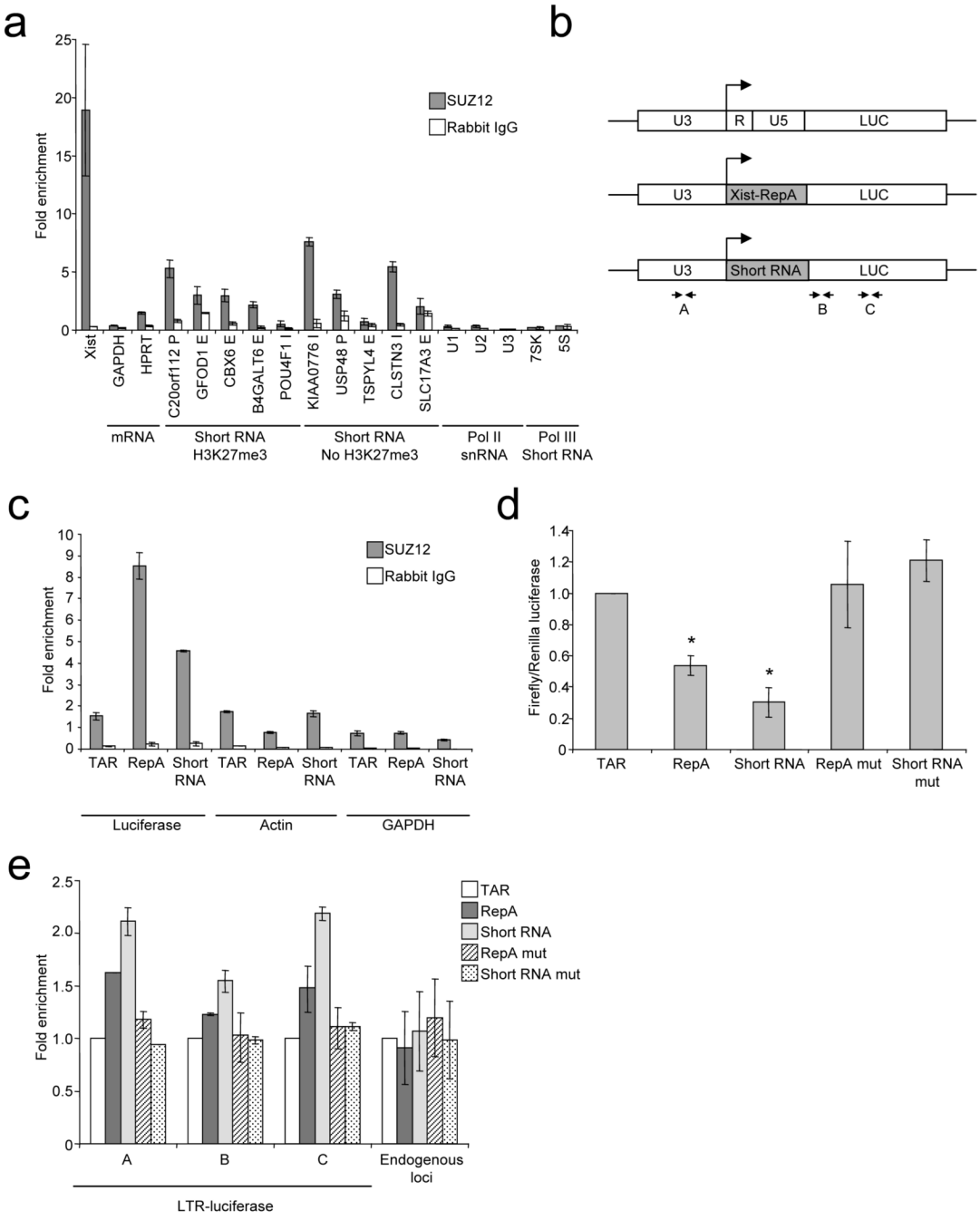

Figure 6. PRC2 interacts with short RNA in cells

A. Fold enrichment (mean and SD, n=3) of different RNA species in SUZ12 IP (grey) and control IP (white) compared with input RNA and normalized to Actin measured by quantitative reverse-transcription PCR. The genes and relative locations from which the short RNAs are transcribed are indicated above (E=exon, I=intron, P=promoter).

B. Schematic showing the wild-type HIV LTR-luciferase construct and modified forms in which R and U5, encoding TAR, are replaced with the Xist-RepA stem-loop or the C20orf112 short RNA stem-loop. The position of primers used for RNA IP and ChIP experiments are marked.

C. Fold enrichment (mean and SD, n=3) of luciferase RNA containing different RNA stem-loops, actin RNA and GAPDH RNA in SUZ12 IP (grey) and control IP (white) compared with input RNA.

D. Luciferase activity in Hela cells transfected with plasmids encoding firefly luciferase downstream of wild-type or modified HIV LTRs. Firefly luciferase activity is plotted relative to co-transfected Renilla luciferase and normalized to wild-type LTR (mean and SD of 7 experiments (performed with 2 clones), each experiment comprising 3 measurements. * comparisons that gave significance at $p<0.05$.

E. Fold enrichment of H3K27me3 over H3 (mean and SD, n=3) across the different luciferase constructs (primers marked in **B**), normalized to the ratio measured at the wild-type LTR construct. Data for 3 endogenous loci (ACCN2, HAPLN2 and HMX2) are shown in comparison.

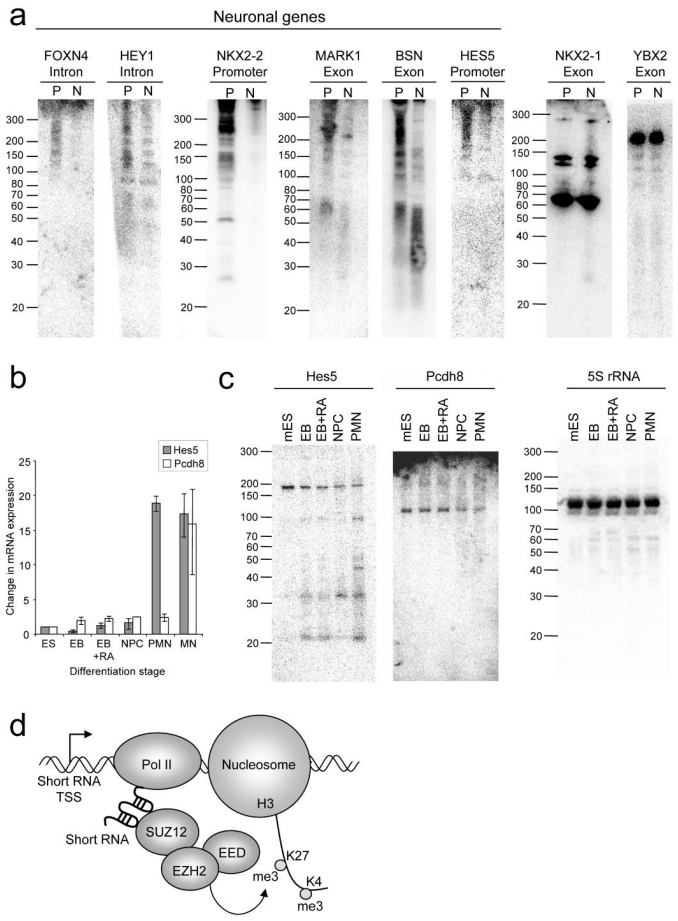

Figure 7. Loss of short RNAs from polycomb target genes activated in neuronal cells

A. Northern blotting for short RNAs transcribed from genes targeted by polycomb in CD4+ T-cells in RNA from PBMC (P) and from differentiated SH-SY5Y neuronal cells (N). Genes active in neuronal cells are labeled.

B. Quantitative PCR showing the expression (mean and SD, n=3) of Hes5 (grey) and Pcdh8 (white) mRNA relative to actin during the step-wise differentiation to motor neurons (PMN). EB, embryoid bodies (day 2); EB+RA, EB treated with retinoic acid (day 2, 8 hours later); NPC, neuronal precursor cells (day 3), PMN, precursor motor neuron (day 4); MN, motor neuron (day 7).

C. Northern blotting for Hes5 and Pcdh8 short RNAs and 5S rRNA in ES cells during the step-wise 4-day differentiation to PMN.

D. Our data support a model in which short RNAs transcribed from the 5′ end of polycomb target genes interact with PRC2, stabilizing its interaction with chromatin and allowing H3K27 methylation and the repression of mRNA transcription in cis.