

Interactions with RNA direct the Polycomb group protein SCML2 to chromatin where it represses target genes

Roberto Bonasio^{1†‡}, Emilio Lecona^{1†§}, Varun Narendra¹, Philipp Voigt¹, Fabio Parisi^{2,3}, Yuval Kluger^{2,3}, Danny Reinberg^{1*}

¹Department of Biochemistry and Molecular Pharmacology, Howard Hughes Medical Institute, New York University School of Medicine, New York, United States;

²Department of Pathology, Yale University School of Medicine, New Haven, United States; ³Yale Cancer Center, Yale University School of Medicine, New Haven, United States

*For correspondence: danny.reinberg@nyumc.org

†These authors contributed equally to this work

Present address: †Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, United States; ‡Genome Instability Group, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

Competing interests: See page 18

Funding: See page 18

Received: 25 February 2014

Accepted: 27 May 2014

Published: 01 July 2014

Reviewing editor: Ali Shilatifard, Stowers Institute for Medical Research, United States

© Copyright Bonasio et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract Polycomb repressive complex-1 (PRC1) is essential for the epigenetic regulation of gene expression. SCML2 is a mammalian homolog of *Drosophila* SCM, a Polycomb-group protein that associates with PRC1. In this study, we show that SCML2A, an SCML2 isoform tightly associated to chromatin, contributes to PRC1 localization and also directly enforces repression of certain Polycomb target genes. SCML2A binds to PRC1 via its SPM domain and interacts with ncRNAs through a novel RNA-binding region (RBR). Targeting of SCML2A to chromatin involves the coordinated action of the MBT domains, RNA binding, and interaction with PRC1 through the SPM domain. Deletion of the RBR reduces the occupancy of SCML2A at target genes and overexpression of a mutant SCML2A lacking the RBR causes defects in PRC1 recruitment. These observations point to a role for ncRNAs in regulating SCML2 function and suggest that SCML2 participates in the epigenetic control of transcription directly and in cooperation with PRC1.

DOI: [10.7554/eLife.02637.001](https://doi.org/10.7554/eLife.02637.001)

Introduction

Polycomb group (PcG) genes are key epigenetic regulators in *Drosophila*, mammals, and plants (Schuettengruber et al., 2007). PcG mutants fail to maintain transcriptional repression of differentiation genes during development, which results in their ectopic expression, compromised patterning of the embryo, and homeotic transformations (Lewis, 1978; Struhl, 1981; Simon et al., 1992). Most PcG genes encode proteins that recognize chromatin or modify its structure, often by forming multi-subunit protein complexes such as Polycomb repressive complex 1 (PRC1) (Simon and Kingston, 2009). Genetic screens and biochemical studies have identified several additional PcG proteins that are not stable components of known Polycomb complexes. One of these proteins is sex comb on midleg (SCM), which is required for appropriate body patterning during embryonic development in *Drosophila* (Simon et al., 1992; Bornemann et al., 1996, 1998). SCM and its human homologs were recovered in sub-stoichiometric amounts after biochemical purifications of the PRC1 complex (Shao et al., 1999; Levine et al., 2002; Gao et al., 2012), suggesting a functional link between SCM and PRC1. SCM also belongs to a small family of proteins characterized by the presence of a conserved malignant brain tumor domain (MBT), which functions as a binding module for methylated lysines on histones (Bonasio et al., 2010a), further suggesting that SCM might function on chromatin. In addition to the MBT domain, SCM contains a C-terminal SPM domain, through which it is thought to interact with PRC1 (Peterson et al., 2004).

eLife digest Almost every cell in our bodies has the same genes but at any one time different genes will be switched on in different cells. Much of the DNA in a cell is wrapped around proteins called histones to form a compact structure called chromatin. The DNA in chromatin is often so tightly packed that the genes in the DNA cannot to be accessed or switched on.

A complex of proteins called the Polycomb Repressive Complex 1 (or PRC1 for short) pack DNA into chromatin to switch genes off (and keep them off) in both plants and animals. Other proteins are known to weakly bind to this complex, including one called SCM in fruit flies. However, this protein has not been extensively studied.

Bonasio, Lecona et al. have now looked at a related version of this protein that binds to chromatin in humans. These experiments revealed that this protein, which is called SCML2A, also binds to molecules of RNA, and Bonasio, Lecona et al. also identified a previously unrecognized domain within SCML2A that interacts with these molecules.

Like other domains in this protein that bind to PRC1 and histones, SCML2A needs its RNA-binding region to be able to bind to chromatin in order to target and switch off certain genes. As such, the findings of Bonasio, Lecona et al. support models whereby RNA molecules can regulate the expression of genes, and suggest that some RNA molecules do this by interacting with SCML2A.

Future work is needed to address whether RNA molecules serve as guides that direct SCML2A to specific genes that need to be switched off, or whether these molecules' roles are more complex.

DOI: [10.7554/eLife.02637.002](https://doi.org/10.7554/eLife.02637.002)

Despite the crucial importance of PcG proteins in gene regulation, development, and disease (**Sparmann and van Lohuizen, 2006**), a coherent model that explains how these proteins are recruited and regulate different target genes in different cell lineages is lacking (**Simon and Kingston, 2009**). In *Drosophila*, arrays of DNA motifs recognized by a variety of transcription factors function as recruitment 'hubs' for PRCs and are known as *Polycomb* responsive elements (PREs) (**Müller and Kassisi, 2006**); however, PREs alone are not sufficient to explain or predict the genome-wide occupancy patterns of PcG complexes (**Simon and Kingston, 2009**). In mammals, our understanding of PRC targeting is even more limited, given that only a few examples of PRE-like DNA elements have been described to date (**Bengani et al., 2013; Cuddapah et al., 2012; Sing et al., 2009; Woo et al., 2010, 2013**) and several of the transcription factors that bind and recruit *Drosophila* PRCs are absent or poorly conserved (**Schuettengruber et al., 2007**). Moreover, there are several mammalian versions of the two best characterized *Drosophila* PcG complexes, PRC1 and PRC2, which differ in subunit composition and chromatin localization (**Kuzmichev et al., 2004; Margueron et al., 2008; Gao et al., 2012**).

The mammalian PRC1 closest in subunit composition to their *Drosophila* homolog (PRC1.2 and PRC1.4 [**Gao et al., 2012**]) have conserved functions in development and disease (**Richly et al., 2011**) and contain subunits that recognize trimethylated lysine 27 of histone H3 (H3K27me3), the catalytic product of PRC2 (**Cao et al., 2002; Kuzmichev et al., 2002**). However, the mechanistic details of PRC1 repression in mammals are poorly understood. Although H3K27me3 may contribute to stabilizing some forms of PRC1 on chromatin, the genome-wide distribution of PRC1 is largely unchanged in embryonic stem cells that do not have H3K27me3 (**Tavares et al., 2012**), and many PRC1 complexes do not contain the subunit that binds to H3K27me3 (**Gao et al., 2012**). Various factors, including RUNX1 (**Yu et al., 2012**) and the RNA helicase MOV10 (**El Messaoudi-Aubert et al., 2010**), have also been proposed to guide PRC1 recruitment in specific contexts, but an overarching model remains lacking. Given that increasing evidence points to a role for noncoding RNAs (ncRNAs) in mediating recruitment of chromatin complexes (**Koziol and Rinn, 2010; Wang and Chang, 2011**) and that at least one ncRNA has been shown to make contact with PRC1 (**Yap et al., 2010**), we hypothesized that an RNA-binding chromatin protein may be involved in PRC1 recruitment in mammals.

The potential role of SCM and its mammalian homologs in the regulation of PRC1 and function remain unexplored. A recent study showed that *Drosophila* SCM is recruited to the PRE upstream of *Ultrabithorax* independent of PRC1 or PRC2, whereas PRC1 and PRC2 binding requires the presence of SCM (**Wang et al., 2010b**). The human genome comprises four genes with homology to *Scm*, but only two that contain conserved MBT, SPM, and DUF3588 domains, *SCMH1* and *SCML2*. *SCMH1* functions in spermatogenesis (**Takada et al., 2007**), whereas little is known about *SCML2*, except that

it is ubiquitously expressed and that deletions of its coding sequence are found in a subset of medulloblastomas (Northcott et al., 2009), suggesting it may have tumor-suppressive activity.

In this study, we report that SCML2 interacts with ncRNAs through an RNA-binding region (RBR), contributes to the recruitment of PRC1 to target genes and also directly collaborates in gene repression. Furthermore, we show that the chromatin localization of SCML2 is achieved through the combined action of the RBR, MBT, and SPM domains and that the expression of Δ RBR mutants of SCML2 causes mislocalization of PRC1.

Results

SCML2 interacts with RNA through a novel RNA-binding region

Interaction with RNA regulates recruitment and assembly of several chromatin-associated factors, including members of the *Polycomb* and *trithorax* groups (Rinn et al., 2007; Khalil et al., 2009; Kanhere et al., 2010; Tsai et al., 2010; Wang et al., 2011; Kaneko et al., 2013, 2014). To find novel RNA-interacting proteins among potential epigenetic regulators, we performed a candidate-based screen using RNA immunoprecipitation (RIP) from HeLa S3 nuclear extracts. We observed that the PcG protein SCML2, a homolog of *Drosophila* SCM, consistently co-purified with a relatively large amount of RNA, especially when compared to EZH2 and MLL, which are part of complexes known to interact with ncRNAs (Rinn et al., 2007; Kaneko et al., 2010, 2013; Wang et al., 2011; Figure 1A). SCML2 did not co-precipitate with either EZH2 or MLL or vice versa (data not shown).

RNA binding assays with total HeLa S3 RNA and recombinant fragments of SCML2 (Figure 1B) revealed that amino acids 256–330 were necessary and sufficient for RNA binding in vitro (Figure 1C). Although the most prominent bands recovered by pull-down of SCML2 fragments corresponded to ribosomal RNA, this was most likely due to their abundance in the input and not to a specific interaction. Other chromatin-associated proteins that bind RNA are rather promiscuous in vitro (Davidovich et al., 2013; Kaneko et al., 2014); therefore, this type of assay is only suitable to identify the protein region that harbors an affinity for RNA. Because the region spanning 256–330 bears no resemblance to other known RNA-binding domains, we tentatively refer to it as the RNA binding region (RBR) of SCML2. The SCML2 RBR exhibited little specificity for RNA sequence, but discriminated between RNA and DNA in vitro (Figure 1D), and strongly preferred single-stranded to double-stranded RNA (Figure 1D, left). SCML2 fragments lacking the RBR (Δ RBR) failed to bind any of the nucleic acids tested (Figure 1D, right).

The RBR is sufficient for RNA binding in vitro

To obtain a more quantitative description of the nucleic acid binding preferences of SCML2, we performed electromobility shift assays (EMSAs) with a recombinant SCML2 fragment encompassing the RBR alone (aa 256–330). The RBR shifted both *HOTAIR* RNA and DNA (Figure 2A), and so did the larger MBT-DUF fragment (Figure 2—figure supplement 1A), but only *HOTAIR* RNA and not DNA competed efficiently for binding to the RBR (Figure 2B) or to the MBT-DUF fragments (Figure 2—figure supplement 1B).

Because a basic patch was reported to mediate SCML2–DNA and SCML2–nucleosome interactions in vitro (Nady et al., 2012) and partially overlaps with the RBR, we also tested the ability of this fragment to bind nucleosomes by EMSA. The RBR displayed an affinity for nucleosome particles (Figure 2A), but much lower than for RNA as demonstrated by the fact that nucleosomes could not compete with *HOTAIR* RNA for binding to the RBR (Figure 2B). When both nucleosomes and RNA were present in limiting amounts, we noticed the appearance of a super-shifted band that was not observed when either species were added separately (Figure 2C, asterisk), suggesting that the RBR—and therefore SCML2—is capable of forming a ternary complex with nucleosomes and RNA simultaneously (Figure 2—figure supplement 2A). The appearance of the super-shifted band was not due to GST-mediated dimerization, as it also occurred after the removal of the GST moiety from the RBR (Figure 2—figure supplement 2B). The possibility of ternary complex formation is further supported by the observation that nucleosomes co-precipitated in vitro with both the SCML2 RBR and *HOTAIR* simultaneously (Figure 2—figure supplement 2C), whereas *HOTAIR* alone did not bind nucleosomes (Figure 2—figure supplement 2D).

Not all RNA sequences were bound by the RBR to the same extent. In vitro-transcribed RNA from the 601 nucleosome-positioning template competed less efficiently for binding to the RBR as compared to *HOTAIR* (Figure 2B). We also tested the binding specificity of the SCML2 RBR for isolated stem-loops within the 5' fragment of the long noncoding RNA (lncRNA) *HOTAIR* that interacts with

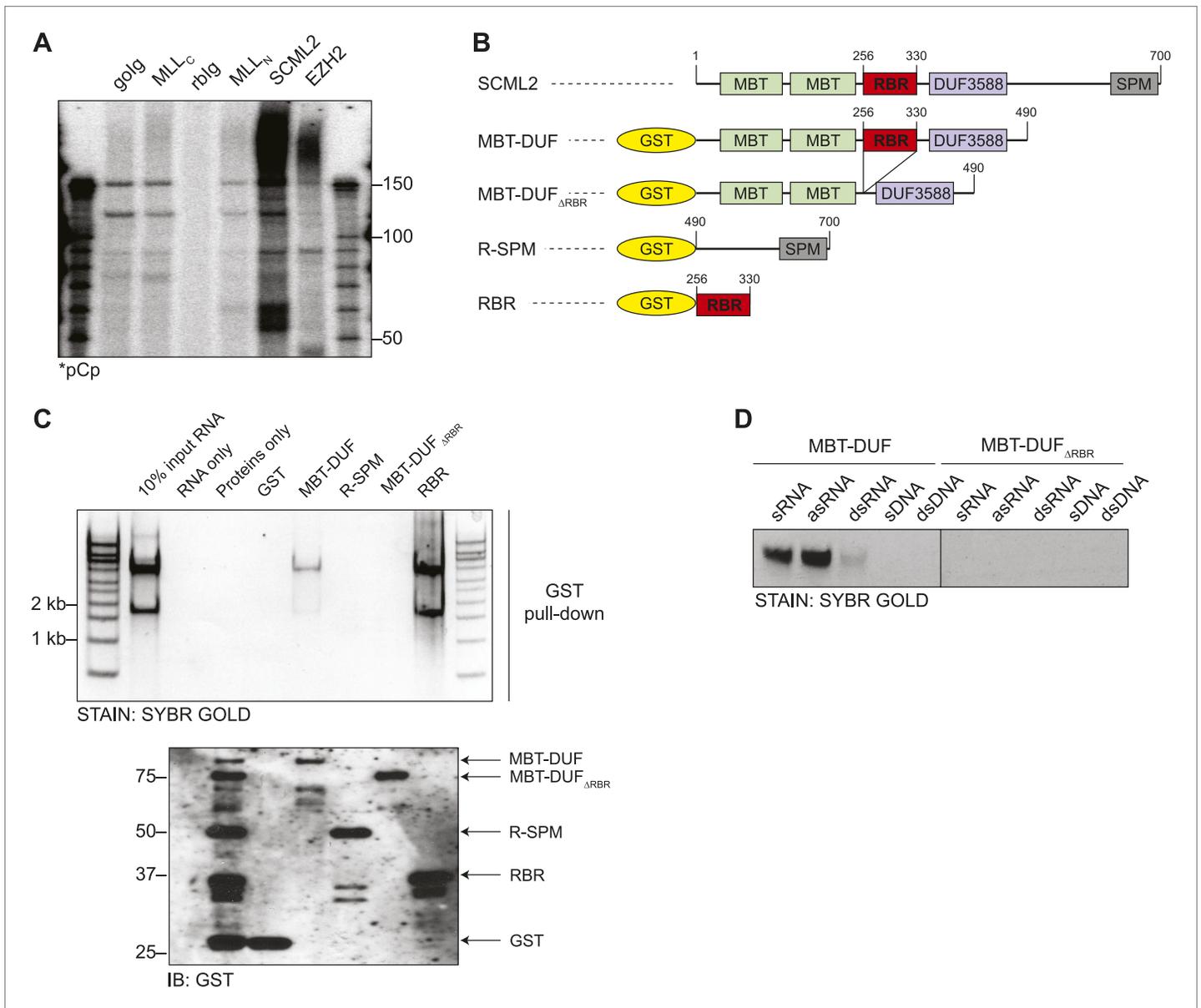


Figure 1. Identification of the RNA-binding region of SCML2. **(A)** Endogenous RIP from HeLaS3 nuclear extract for the indicated chromatin proteins or with goat (golg) and rabbit (rblg) control antibodies. Co-precipitating RNA was labeled in 3' and resolved by denaturing PAGE. **(B)** Domain organization of SCML2 and scheme of the GST-fused truncations. **(C)** In vitro pull-down of total RNA with the indicated SCML2 fragments. RNA fraction (top) and protein fraction (bottom) are shown. **(D)** The indicated sense RNA (sRNA) and antisense RNA (asRNA) were transcribed in vitro from a synthetic random 100 nts DNA template, PAGE-purified and annealed to form double-stranded RNA (dsRNA). The sense DNA oligonucleotide (sDNA) and the double-stranded DNA template (dsDNA) were used as controls.

DOI: 10.7554/eLife.02637.003

PRC2 (Tsai et al., 2010). Only the full 1–300 fragment bound to the SCML2 RBR with sufficient affinity to co-precipitate, whereas none of the isolated stem-loops did (data not shown), suggesting that the RBR is capable of recognizing complex structures and/or sequences and that it does not just form non-specific interactions with any type of RNA.

RBR, MBT, and SPM domains coordinate recruitment of SCML2A to chromatin

Having found that SCML2 binds to RNA in vivo and in vitro, we sought to determine the consequences of these binding events on the cellular and molecular functions of SCML2.

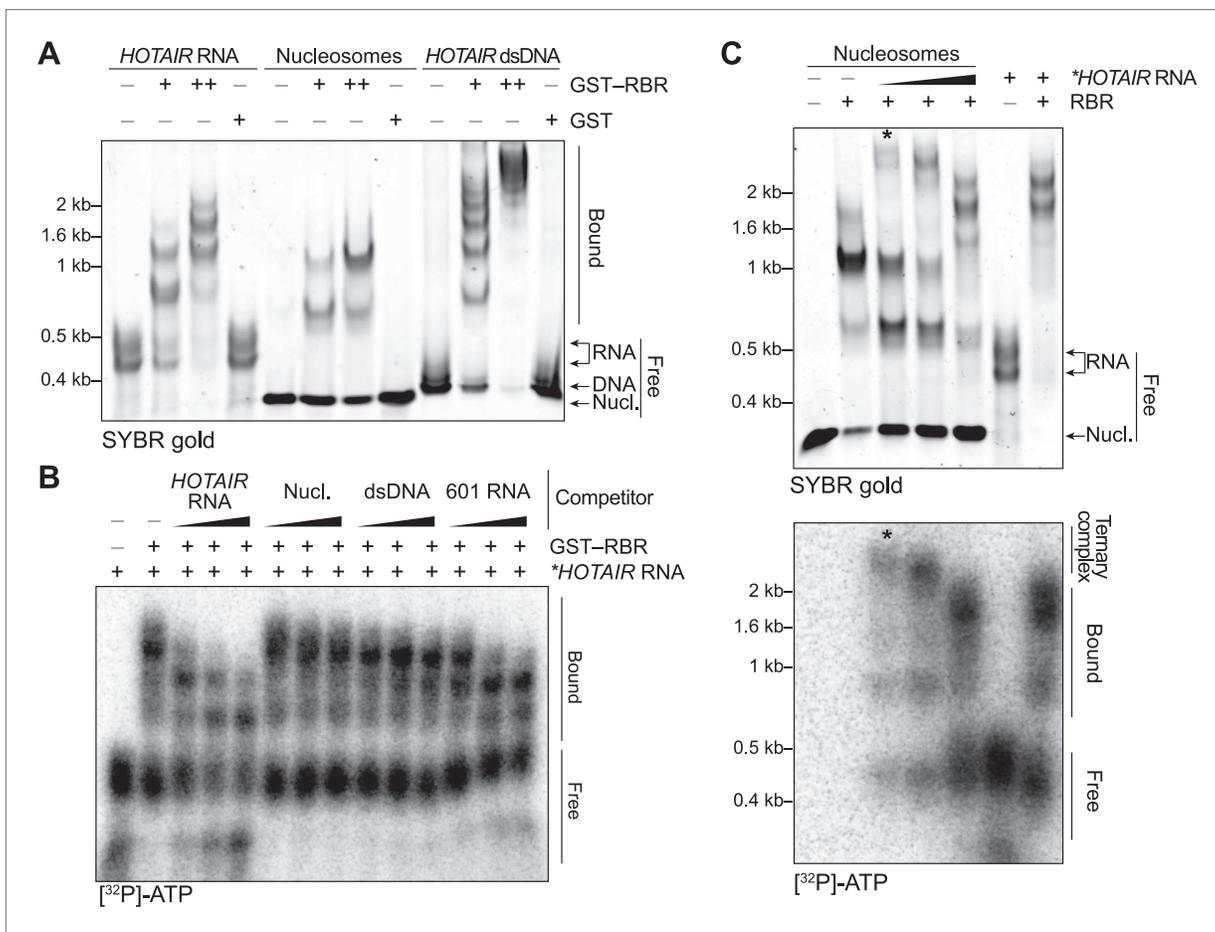


Figure 2. In vitro characterization of the binding preferences of the SCML2 RBR. **(A)** EMSA with 2 (+) and 4 (++) pmol GST-fused RBR and 520 fmol *HOTAIR* RNA 1–300, 690 fmol nucleosomes, or 275 fmol dsDNA encoding *HOTAIR*_{1–300}. 5.2 pmol GST were used as a control. Complexes were separated on native gels and detected with SYBR gold stain. Data are representative of ≥ 4 experiments. **(B)** EMSA with 3.4 pmol GST–RBR and 520 fmol of labeled *HOTAIR* RNA was performed as in **(A)** with the addition of, from left to right, 260, 520, and 1040 fmol of unlabeled *HOTAIR* RNA, nucleosome particles, dsDNA, and 601 RNA. Bands were visualized by autoradiography. Data are representative of ≥ 4 experiments. **(C)** EMSA of RBR with nucleosomes and labeled *HOTAIR* RNA. Assay was performed as in **(B)** varying the amount of labeled RNA. Nucleic acid stain (top) and autoradiography (bottom) are shown. The asterisk indicates the putative ternary complex.

DOI: 10.7554/eLife.02637.004

The following figure supplements are available for figure 2:

Figure supplement 1. Additional EMSAs.

DOI: 10.7554/eLife.02637.005

Figure supplement 2. Ternary complex EMSAs and pull-downs.

DOI: 10.7554/eLife.02637.006

As we reported previously, the human *SCML2* locus gives rise to two alternative isoforms: one encodes the full-length version (*SCML2A*) and the other a C-terminal truncation, generated by alternative splicing (*SCML2B*; **Figure 3—figure supplement 1A**) (Lecona et al., 2013). *SCML2A* has a domain organization similar to that of *Drosophila* SCM: two N-terminal MBT repeats, a DUF3588 domain, and a C-terminal SPM domain. Although both *SCML2* isoforms are predominantly nuclear, *SCML2A* is enriched in the chromatin fraction, whereas the bulk of *SCML2B*, which lacks the SPM domain is found in the soluble nuclear fraction, where it interacts with the cell cycle machinery and participates in the G1/S checkpoint (Lecona et al., 2013). Therefore, *SCML2A* is the isoform more likely to have a direct function on chromatin and transcription and is the focus of the remainder of this study.

Multiple lines of evidence point to a role for ncRNAs in recruiting epigenetic regulators, in particular members of the PcG and trxB, to chromatin (Rinn et al., 2007; Koziol and Rinn, 2010; Bonasio et al.,

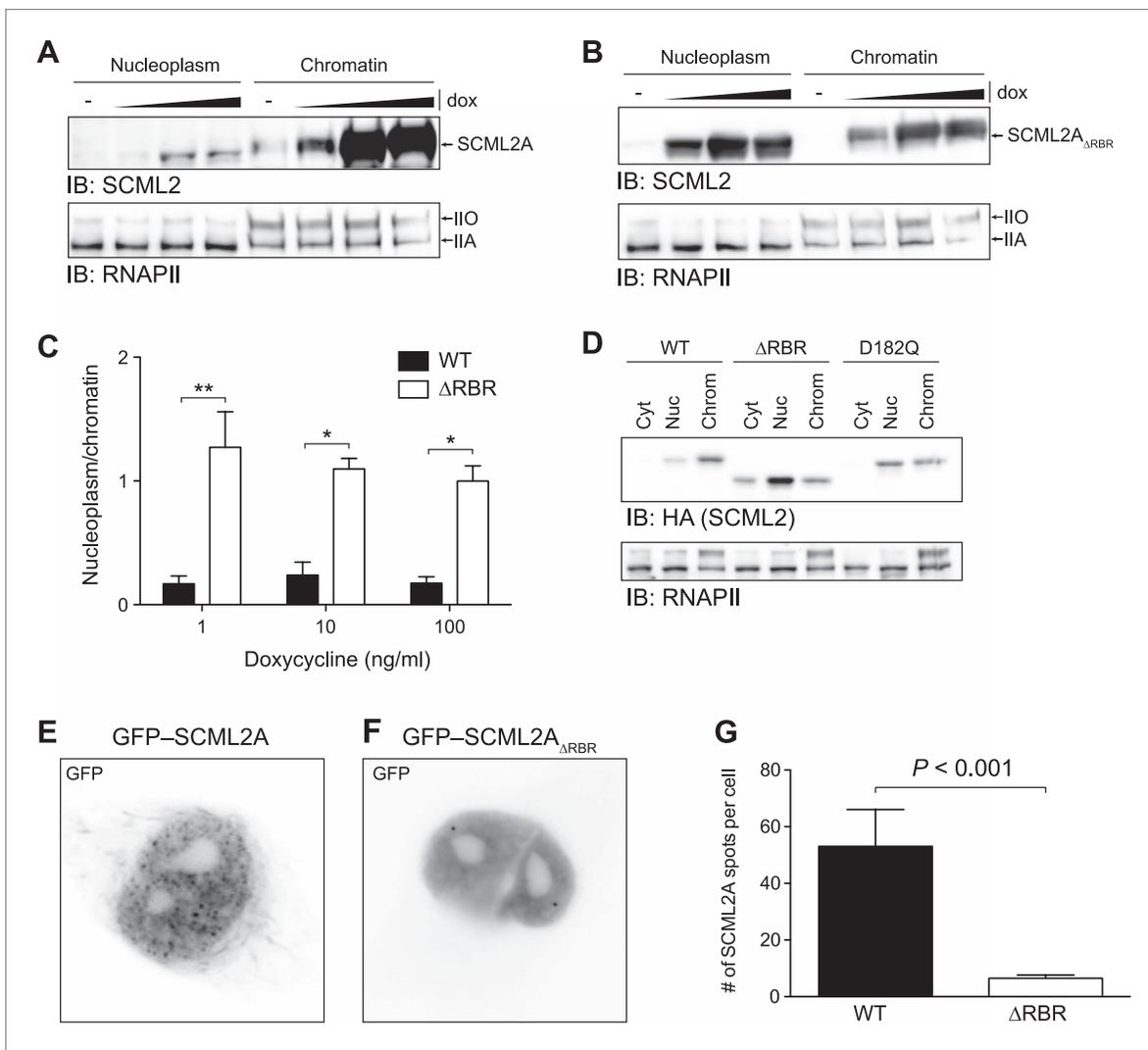


Figure 3. Loss of the RBR alters the subnuclear distribution of SCML2A. **(A and B)** Salt fractionation of 293T-REx induced to express SCML2A WT **(A)** and Δ RBR **(B)** with increasing amounts of doxycycline (dox). The distribution of RNAPIIA and RNAPIIO are shown as fractionation and loading control. **(C)** Densitometry-based quantification of the ratio of nucleoplasmic vs chromatin-associated SCML2A WT (black bars) and Δ RBR (white bars). Bars show mean measurements from two independent clones at increasing levels of inductions. p-values were calculated with 2-way ANOVA. **(D)** Salt fractionation of 293T-REx induced to express SCML2A WT, Δ RBR, and the MBT-domain mutant D182Q. **(E and F)** Fluorescence microscopy of HeLa cells transiently transfected with vectors encoding GFP-SCML2A_{WT} **(E)** or GFP-SCML2A _{Δ RBR} **(F)**. **(G)** Quantification of SCML2 spots in multiple cells transfected as in **(E)** and **(F)**. Bars represent mean + SEM. N = 11 for SCML2A; N = 17 for SCML2A _{Δ RBR}. p-value was calculated with an unpaired t test.

DOI: 10.7554/eLife.02637.007

The following figure supplements are available for figure 3:

Figure supplement 1. Nuclear distribution of SCML2A_{D182A} and SCML2B.

DOI: 10.7554/eLife.02637.008

2010b; Wang et al., 2011). Because SCML2A exhibited stronger affinity for RNA than for DNA, we hypothesized that interactions between RNA and the RBR may contribute to the recruitment of SCML2 to chromatin. Consistent with this hypothesis, an SCML2A mutant lacking the RBR (SCML2A _{Δ RBR}) was displaced from chromatin, as demonstrated by the fact that a larger fraction of this protein was released with low-salt extraction when compared to the wild-type (WT) form, regardless of the amount of transgenic protein being expressed (**Figure 3A–C**).

We then assessed the potential role of the MBT domains in binding to chromatin. Mutation of the acidic residue that coordinates the methylated lysine within the aromatic cage impairs its binding (**Sathyamurthy et al., 2003; Grimm et al., 2007; Santiveri et al., 2008**) and also resulted in loss of SCML2A from the chromatin fraction, but not to the extent observed with the RBR deletion (**Figure 3D**,

Figure 3—figure supplement 1B). The SCML2 isoform lacking the SPM domain, SCML2B, displayed a predominantly diffuse distribution within the nucleus, which was not affected by the removal of the RBR; however the residual amount of SCML2B that could be recovered from the chromatin fraction was further decreased upon deletion of the RBR (**Figure 3—figure supplement 1C–G**). Together, these results show that the affinity of SCML2 for chromatin is regulated by multiple molecular mechanisms, including binding to PRC1 via the SPM domain, histone marks via the MBT domains, and RNA via the RBR.

We confirmed the altered distribution of SCML2A_{ΔRBR} by fluorescence microscopy. Similar to the endogenous SCML2 protein (**Figure 3—figure supplement 1H**), GFP-SCML2A accumulated in the nucleus and exhibited a punctate distribution, suggestive of recruitment to specific sub-nuclear compartments, presumably target sites on chromatin (**Figure 3E**). In contrast, GFP-SCML2A_{ΔRBR} exhibited a diffuse nuclear distribution (**Figure 3F**) consistent with its looser association with chromatin revealed by biochemical fractionations (**Figure 3A–C**). The number of intense GFP spots decreased from a median value of 39 per cell for SCML2A to only 5 per cell for SCML2A_{ΔRBR} (**Figure 3G**). Deletion of the RBR did not affect the ability of SCML2A to be imported into the nucleus, nor its ability to interact with other protein factors, including the core subunits of the PRC1 complex (see below).

SCML2 interacts with PRC1 and represses transcription

To further explore the functional relationship between SCML2 and PRC1, we purified epitope-tagged SCML2A and SCML2B from the chromatin fraction of 293T-REx cells. Both SCML2A_{WT} and SCML2A_{ΔRBR} co-purified with all the components of the canonical PRC1, including RING1A, RING1B, CBX2/8, PHC2/3, and BMI1/PCGF4 (a subunit specific for PRC1.4 [**Gao et al., 2012**]), suggesting that deletion of the RBR does not affect the ability of SCML2A to interact with PRC1. Only small traces of PRC1 components could be detected among the SCML2B-associated proteins, despite the fact that a comparable number of SCML2 peptides were detected in the two purifications (**Figure 4A**). SCML2A also co-purified with SCMH1 (the other SCM homolog in humans) and L3MBTL3, both of which comprise C-terminal SPM domains (**Figure 4A**). As a control, both SCML2A and SCML2B recovered comparable amounts of USP7, in good agreement with previous reports (**Sowa et al., 2009**) and as we also independently confirmed (Lecona et al. unpublished). In keeping with these observations, immunoprecipitation (IP) of BMI1 from the chromatin fraction of HeLa cells recovered preferentially the chromatin-associated and SPM-containing SCML2A isoform (**Figure 4B**, compare top and bottom band in lane 4 vs lane 1 or lane 3).

Consistent with the classification of SCML2 as a PcG (and therefore repressive) gene, artificial tethering of both SCML2A and SCML2B to an integrated reporter using the DNA-binding domain of *Saccharomyces cerevisiae* Gal4p (**Vaquero et al., 2004**) resulted in a dose-dependent repression of transcription (**Figure 4C**). However, only tethering of SCML2A but not SCML2B caused an increase in BMI1 recruitment to the UAS-containing reporter (**Figure 4—figure supplement 1A,B**), confirming that the SPM domain is required for interacting with BMI1 and suggesting that SCML2A contributes to PRC1 targeting. On the other hand, recruitment of BMI1 was not necessary for SCML2-mediated repression in this context. Consistent with this observation, both the MBT domains and the SPM domain contribute to the repressive activity of SCM in *Drosophila* (**Roseman et al., 2001; Peterson et al., 2004; Grimm et al., 2007**).

These results show that RNA binding and the interaction with PRC1 are independent events and that they require different protein regions (the RBR and SPM domain, respectively). Our data support a model in which the RBR contributes to the localization of SCML2A to chromatin, where it interacts with PRC1 and creates a chromatin environment that is refractory to transcription.

SCML2 is enriched at PRC target sites on chromatin

Because SCM is a PcG protein and binds to at least one PcG target in *Drosophila* (**Wang et al., 2010b**), we reasoned that we could gain further insight on the function of SCML2 and its relationship with the Polycomb pathway by determining its genome-wide localization using ChIP sequencing (ChIP-seq). Following a strategy we previously described (**Gao et al., 2012; Kaneko et al., 2013**), we generated 293T-REx lines carrying stably integrated, inducible transgenes for human SCML2 fused to a triple N-terminal affinity tag (FLAG-HA-Twin-Strep-tag, henceforth 'N3') and shRNAs against the endogenous SCML2. In these cells, more than 30% (781 out of 2479 total) of SCML2-enriched regions (ERs) were located in the vicinity (<10 kb) of the transcription start site (TSS) of protein-coding genes (**Figure 4D**),

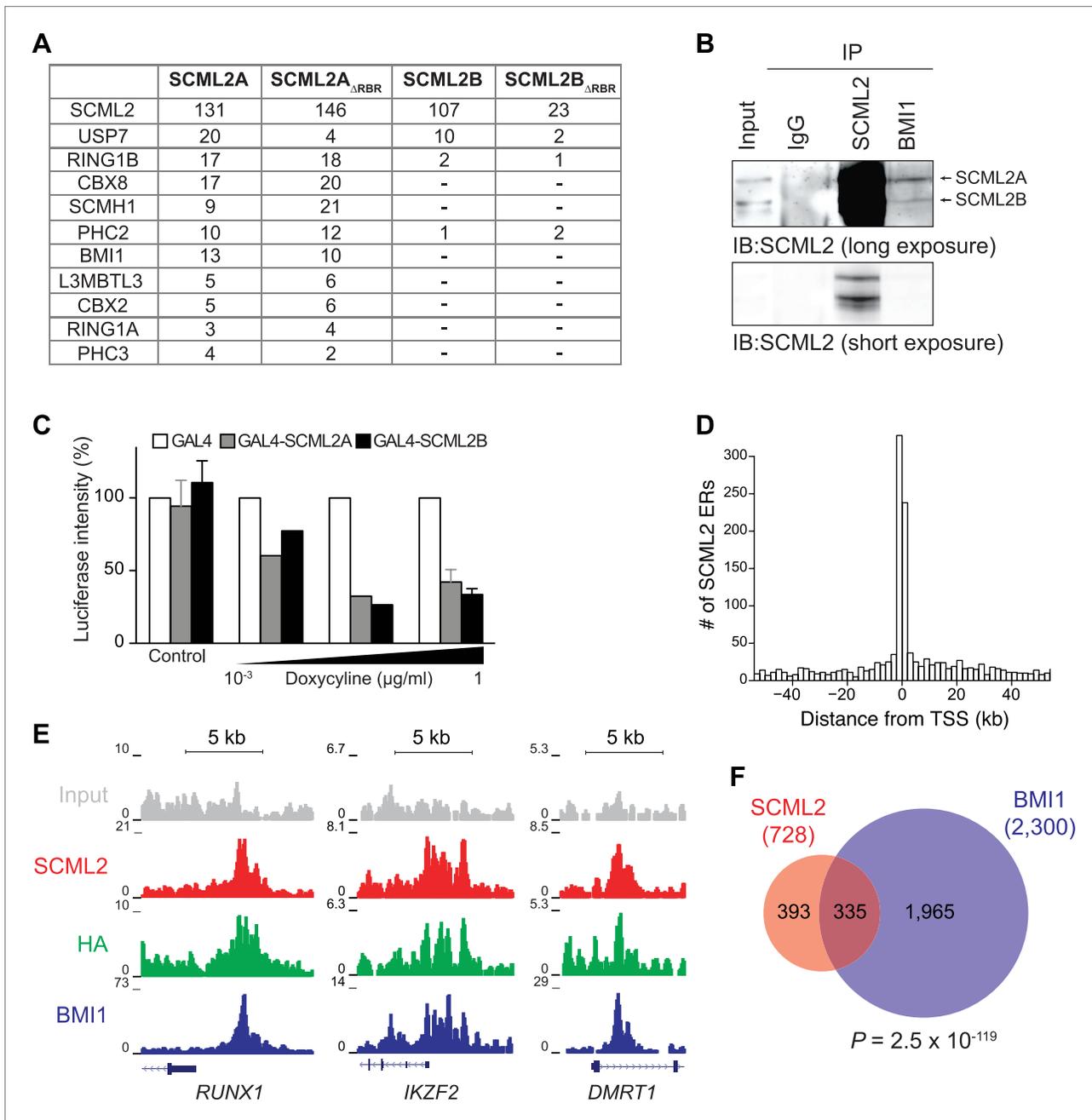


Figure 4. SCML2 interacts and shares target genes with BMI1. **(A)** Proteins detected by mass spectrometry after affinity purification of SCML2A, SCML2A_{ΔRBR}, SCML2B and SCML2B_{ΔRBR} from the chromatin fraction of 293T-REx cells. The number of peptides identified for each protein is indicated. **(B)** IP of SCML2 and BMI1 from the chromatin fraction of HeLa cells. Two exposures are shown. **(C)** 293T-REx cells containing an integrated luciferase reporter preceded by UAS repeats were induced to express the indicated transgenes with increasing amounts of doxycycline. 24 hr after induction the activity of the luciferase reporter was assayed. Bars represent the luciferase activity normalized to the non-induced GAL4-only control. **(D)** Distribution of SCML2 ERs around TSSs of protein coding genes. **(E)** Normalized ChIP-seq read densities for SCML2 (red), HA (green), and BMI1 (blue) as compared to input (gray) at three representative loci. **(F)** Overlap of gene targets for SCML2 and BMI1, as defined by having an ER for the respective protein within 10 kb of the TSS. p-value was calculated according to the hypergeometric distribution.

DOI: [10.7554/eLife.02637.009](https://doi.org/10.7554/eLife.02637.009)

The following figure supplements are available for figure 4:

Figure supplement 1. Artificial recruitment of SCML2 and ChIP-seq analysis.

DOI: [10.7554/eLife.02637.010](https://doi.org/10.7554/eLife.02637.010)

consistent with a role for SCML2 in gene regulation. SCML2 targets were significantly enriched for gene ontology (GO) terms related to transcriptional regulation, development, and differentiation (**Figure 4—figure supplement 1C, Supplementary file 1A**), as observed for targets of the *Polycomb* repressive machinery in other systems (*Schuettengruber et al., 2007*). To determine whether SCML2 shared target genes with PRC1, we also performed ChIP-seq for BMI1 (also known as PCGF4), the key component of canonical PRC1.4 (*Gao et al., 2012*). SCML2 and BMI1 co-occupied several promoters, including those of transcription factors *RUNX1*, *IKZF2*, and *DMRT1* (**Figure 4E**). Overall, SCML2 shared ~46% of its target genes with BMI1 (**Figure 4F, Supplementary file 1B**), almost four times more than expected by chance ($p=2.5 \times 10^{-119}$, hypergeometric distribution). This overlap became more pronounced as a function of decreasing p-values for the ERs (**Figure 4—figure supplement 1D,E**), suggesting that many of the SCML2-only or BMI1-only sites were false positives due to noise in the ChIP-seq signal. We made similar observations for endogenous SCML2 in non-transgenic K562 cells (**Figure 4—figure supplement 1F,G**), demonstrating that our conclusions can be generalized and do not depend on the presence of the N-terminal tag or on exogenous expression.

Together, our genome-wide analysis of SCML2 distribution on chromatin supports the notion that human SCML2, like its *Drosophila* homolog, belongs to the PcG and regulates the expression of developmental genes and other *Polycomb* targets.

The RBR is required for efficient binding of SCML2A to a subset of PRC1-regulated loci

Next, we sought to determine whether the large-scale redistribution of SCML2A $_{\Delta RBR}$ that we observed by microscopy and biochemical fractionation (**Figure 3**) was due to the loss of recruitment of SCML2A to target genes. To this end, we compared the distribution of N3-tagged SCML2A $_{WT}$ and SCML2A $_{\Delta RBR}$ in transgenic 293T-REx cells. In the absence of doxycycline, endogenous SCML2 levels were equal in all transgenic lines and equivalent to those measured in the parental 293T-REx cell line (data not shown). Upon doxycycline treatment, the endogenous copy of SCML2 was partially silenced by the shRNA and HA-tagged SCML2A $_{WT}$ and SCML2A $_{\Delta RBR}$ were expressed at comparable levels (**Figure 5—figure supplement 1A**). Pull-down of these epitope-tagged proteins did not recover untagged SCML2 (**Figure 5—figure supplement 1B**) and could therefore be used to determine the specific localization of SCML2 WT and ΔRBR without confounding effects due to interactions with endogenous SCML2.

Visual inspection of ChIP-seq profiles revealed that SCML2A $_{\Delta RBR}$ was lost from several target genes (**Figure 5A**), suggesting that the RBR, and presumably RNA-protein interactions might participate in the recruitment or stabilization of SCML2 on chromatin. A smaller number of SCML2 ERs, including the ER near the *RUNX1* TSS, exhibited SCML2A $_{\Delta RBR}$ levels comparable to WT (**Figure 5A**), consistent with the existence of RNA-independent pathways of SCML2A recruitment at some loci. Nonetheless, RBR-independent ERs constituted a small proportion of the total SCML2 binding profile, because the N3-SCML2A $_{\Delta RBR}$ ChIP-seq signal was much weaker than WT when all significant ($p < 10^{-5}$ according to MACS 1.4.0rc2) ERs for SCML2 WT were combined in a single meta-ER plot (**Figure 5B**).

Consistent with our observations that SCML2A colocalizes with BMI1 (**Figure 4E,F**) and can recruit it to chromatin (**Figure 4—figure supplement 1B**), partial replacement of endogenous SCML2 with SCML2A $_{\Delta RBR}$ resulted in a decreased, albeit not abolished, accumulation of BMI1 at several of the target sites (**Figure 5A,C**). The residual recruitment of BMI1 may be explained either by the fact that the knockdown of endogenous SCML2 is far from complete (**Figure 5—figure supplement 1A**) or by the contribution of RBR-independent loci. When inspected at the level of individual genomic regions, out of 148 high-confidence ERs obtained from the triple intersection of SCML2, HA, and BMI1 ERs, a majority displayed losses in SCML2 binding in the absence of the RBR and many also had decreased BMI1 occupancy (**Figure 5E**). We confirmed these findings with two additional ChIP-seq data sets from independent biological replicates (**Figure 5—figure supplement 1C**) and ChIP-qPCR analyses of a sample of ERs (**Figure 5—figure supplement 1D**).

From this set of experiments, we conclude that SCML2 is recruited to a subset of target genes through a process that requires its RNA-binding activity and that the presence of mutant SCML2A $_{\Delta RBR}$ that is not chromatin-bound causes a partial release of BMI1 and, presumably, PRC1 from chromatin.

Genome-wide characterization of SCML2-bound ncRNAs by RIP-seq

Given that SCML2 binds RNA through an RBR that is required for proper genome-wide targeting and downstream recruitment of PRC1, we sought to identify the ncRNAs bound to SCML2 in vivo. To this

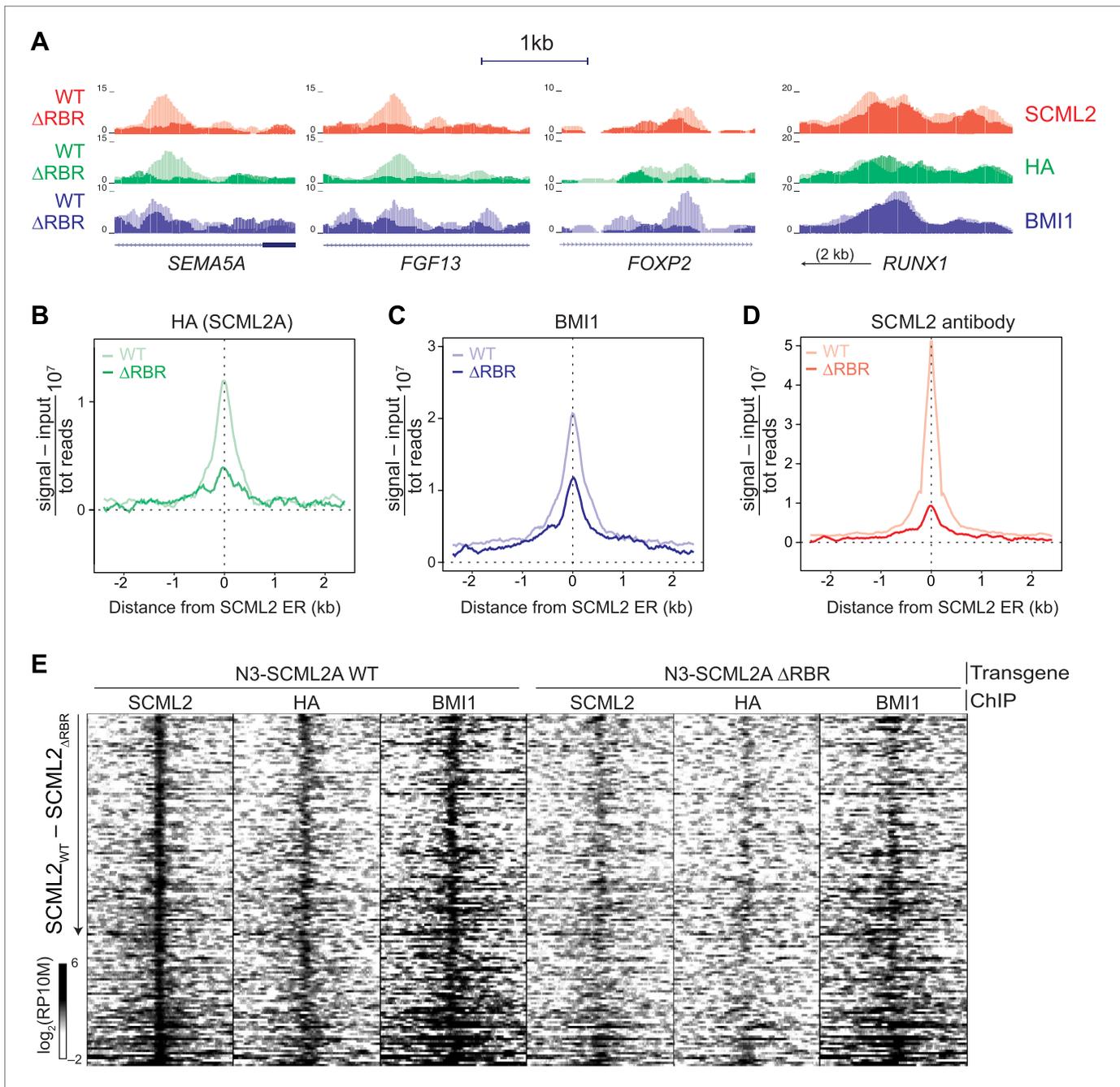


Figure 5. The role of the RBR in the localization of SCML2A and PRC1 to target genes. **(A)** Normalized read density profiles for SCML2 (red), HA (green), and BMI1 (blue) in cells expressing N3-SCML2A WT (light colors) or ΔRBR (dark colors). Three RBR-dependent and one RBR-independent sites are shown. **(B–D)** Normalized ChIP-seq enrichment (signal reads–input reads) for HA **(B)**, BMI1 **(C)**, and SCML2 **(D)**, centered around 2479 SCML2 ERs in 293T-REx::N3-SCML2A. **(E)** Heatmap of read densities for the indicated ChIP-seq samples centered on 148 overlapping ERs for SCML2, HA, and BMI1 in the WT sample. Each row represents a genomic region spanning ± 2.5 kb from the center of the ER. For each sample the log-converted normalized read density was calculated in 25 bp bins and is represented by color intensity. Regions were sorted by difference in SCML2 read densities in WT vs ΔRBR transgenic lines.

DOI: [10.7554/eLife.02637.011](https://doi.org/10.7554/eLife.02637.011)

The following figure supplements are available for figure 5:

Figure supplement 1. Validation and replication of ChIP-seq for SCML2 WT and ΔRBR.

DOI: [10.7554/eLife.02637.012](https://doi.org/10.7554/eLife.02637.012)

end, we performed HA IPs in 293T-REx expressing N3-tagged SCML2A WT, Δ RBR, or the N3 tag alone. Although comparable amounts of WT and Δ RBR protein were recovered, only in the case of WT SCML2A, we could detect co-IP of RNA (**Figure 6A,B**), which we subjected to deep sequencing (RIP followed by sequencing, RIP-seq).

Because increasing experimental evidence points to a role for long intergenic ncRNAs (lincRNAs) in the function of chromatin-associated epigenetic regulators (*Koziol and Rinn, 2010; Wang and Chang, 2011*), we measured the abundance of these RNAs in the IP'ed material by referring to a published lincRNA database (*Cabili et al., 2011*). Although overall abundant lincRNAs were more likely to be detected by RIP-seq, several specific interactions distinguished the set of RNAs recovered in RIPs for SCML2A WT from all other controls (**Figure 6C**), demonstrating the specificity of RNA binding by SCML2 in vivo.

Next, we tested whether the set of ncRNAs associated with SCML2 was fixed in a given cell type or could be altered by processes such as differentiation. To this end, we performed RIP-seq for endogenous SCML2 in K562 before and after treatment with hemin or PMA, which stimulate their differentiation toward the erythroid or megakaryocytic lineage, respectively (*Jacquel et al., 2006*). Indeed, cell

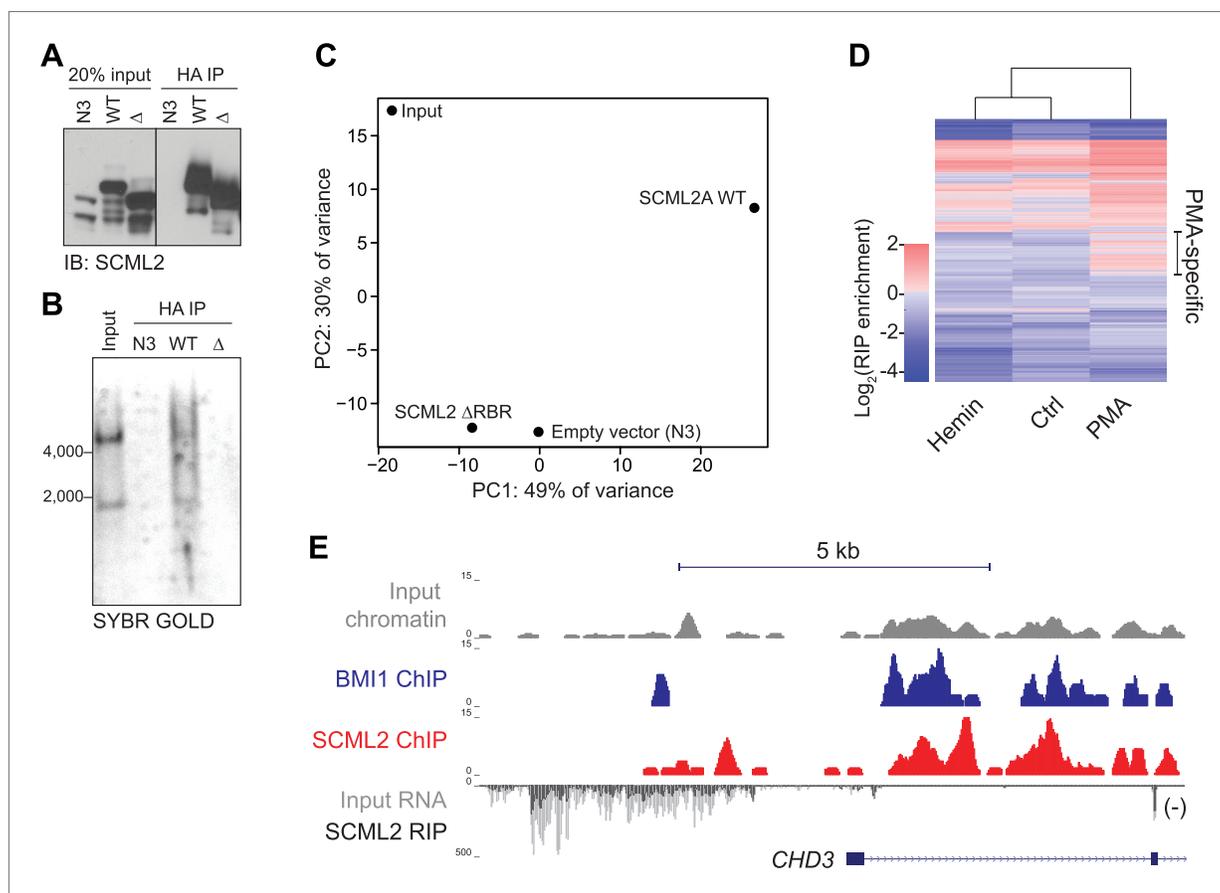


Figure 6. RIP-seq for SCML2 in 293T-REx and K562 cells. (**A** and **B**) HA IP from 293T-REx expressing the empty tag alone (N3), N3-SCML2A WT (WT) or N3-SCML2A Δ RBR (Δ). After washing, the bound fraction was divided into two parts; one was analyzed by western blot (**A**) and one by denaturing RNA electrophoresis (**B**). (**C**) Principal component analysis of the lincRNA abundances for the four samples. (**D**) RIP enrichment heatmap for 668 lincRNAs in K562 stimulated to differentiate with hemin or PMA. Colors indicate \log_2 enrichment in SCML2 RIP-seq compared to the respective inputs. A set of lincRNAs specifically associated with SCML2 in PMA-treated K562 cells is indicated. (**E**) ChIP and RIP profiles at the CHD3 promoter. Normalized read densities are shown for input chromatin (gray), BMI1 ChIP (blue), SCML2 ChIP (red), input RNA (dark gray), and SCML2 RIP (light gray). For simplicity, only the (-) strand of the RNA is shown (antisense to the CHD3 transcript).

DOI: [10.7554/eLife.02637.013](https://doi.org/10.7554/eLife.02637.013)

The following figure supplements are available for figure 6:

Figure supplement 1. Examples of SCML2-associated RNAs.

DOI: [10.7554/eLife.02637.014](https://doi.org/10.7554/eLife.02637.014)

differentiation caused a measurable shift in the pool of lincRNAs associated with SCML2 (**Figure 6D**), suggesting that in addition to intrinsic binding preferences, cellular context may also determine which RNAs bind to SCML2 and, in turn, how its function and distribution on chromatin is affected by these interactions. This phenomenon was not restricted to lincRNAs. RIP for SCML2 also recovered a variety of mRNAs, some of which were specific for certain states (**Figure 6—figure supplement 1A**).

We also found several examples of unannotated ncRNAs that were enriched in SCML2 RIPs compared to input RNA and were transcribed near SCML2 ERs (**Figure 6E**, **Figure 6—figure supplement 1B–E**), which is consistent with the notion that some ncRNAs interacting with SCML2 might act in *cis*. We found different types of ncRNAs that bound SCML2 and originated near SCML2 target sites on chromatin: (1) divergently transcribed antisense ncRNAs (**Figure 6E**, **Figure 6—figure supplement 1B,D**); (2) ncRNAs that originated downstream of the TSS, were transcribed in the antisense direction, and overlapped with portions of the coding transcript (**Figure 6—figure supplement 1C**); and (3) ncRNAs transcribed upstream of the TSS but from the same strand as the downstream coding gene (**Figure 6—figure supplement 1E**).

Thus, SCML2 associates with annotated and unannotated lincRNAs via the RBR in vivo. These interactions are specific and they are also dynamic, likely reflecting the need to adjust SCML2 localization or regulation as the cells transition from one transcriptional and/or epigenetic state to another.

SCML2 contributes to gene repression independently of BMI1

Having observed that SCML2 is sufficient to repress a reporter gene after artificial tethering to chromatin (**Figure 4C**), we tested whether its presence is also necessary for the repression of endogenous target genes. To this end, we depleted endogenous SCML2 by siRNA-mediated knockdown for 3 or 5 days with two different siRNAs (**Figure 7A**). We generated genome-wide expression profiles by RNA-seq in two biological replicates and combined all data sets to identify differentially expressed genes (DEGs). Using a loose cutoff of $p < 0.2$, we identified a total of 530 DEGs of which a significant fraction overlapped ($p = 1.4 \times 10^{-4}$, hypergeometric distribution) with the set of genes associated with an SCML2 ER (**Figure 7B**). In keeping with the postulated repressive function of SCML2, a majority (341) of these genes were upregulated (**Supplementary file 1C**) and this subset accounted for ~86% of the DEGs associated with SCML2 occupancy on chromatin (**Figure 7C**). These observations remained valid when we restricted our analysis to the smaller set of genes with an SCML2 ER within 10 kb from the TSS and only considered DEGs with a p-value < 0.05 (**Figure 7—figure supplement 1**), although the small number of genes obtained with these more stringent filters limited the power of our statistical analysis.

Given that SCML2A can recruit BMI1, and therefore PRC1, to an artificial reporter and that PRC1 is generally thought as the effector of *Polycomb*-mediated repression (**Simon and Kingston, 2009**), we hypothesized that the changes in transcriptional activity observed above might be secondary to a loss of BMI1 from SCML2-bound genes. Unexpectedly, this was not the case. In fact, the levels of BMI1 on chromatin were unaffected even when we analyzed only the 520 sites that lost SCML2 occupancy after SCML2 knockdown (**Figure 7D,E**). Together with the observation that SCML2B retains the ability to repress a reporter independent of BMI1 recruitment (**Figure 4C**), these findings reinforce the notion that SCML2 represses transcription by itself, and that, in its absence, PRC1 alone is not sufficient to maintain repression at all target genes.

Discussion

Proteins belonging to the *Polycomb* and *trithorax* group are epigenetic regulators and maintain specialized chromatin structures that are refractory or permissive to transcription, respectively (**Schuettengruber et al., 2007**). Unlike conventional transcription factors, these proteins do not contain a DNA binding domain that directs them to targets on chromatin; instead, there is accumulating evidence that recruitment and regulation of PcG proteins are coordinated by a number of molecular pathways that act in a combinatorial fashion to repress specific genes in different cell types (**Simon and Kingston, 2013**).

Drosophila SCM and its human homologs, including SCML2, associate with the PRC1 complex (**Shao et al., 1999; Levine et al., 2002; Gao et al., 2012**), but their role in PRC1 function has remained elusive. Here, we have unveiled two mechanisms by which SCML2A contributes to PRC1 function: (1) SCML2 and PRC1 cooperate in chromatin binding, as evidenced by loss of BMI1 upon overexpression of SCML2A_{ΔRBR}, recruitment of BMI1 by SCML2A at artificial tethering sites, and the weak association

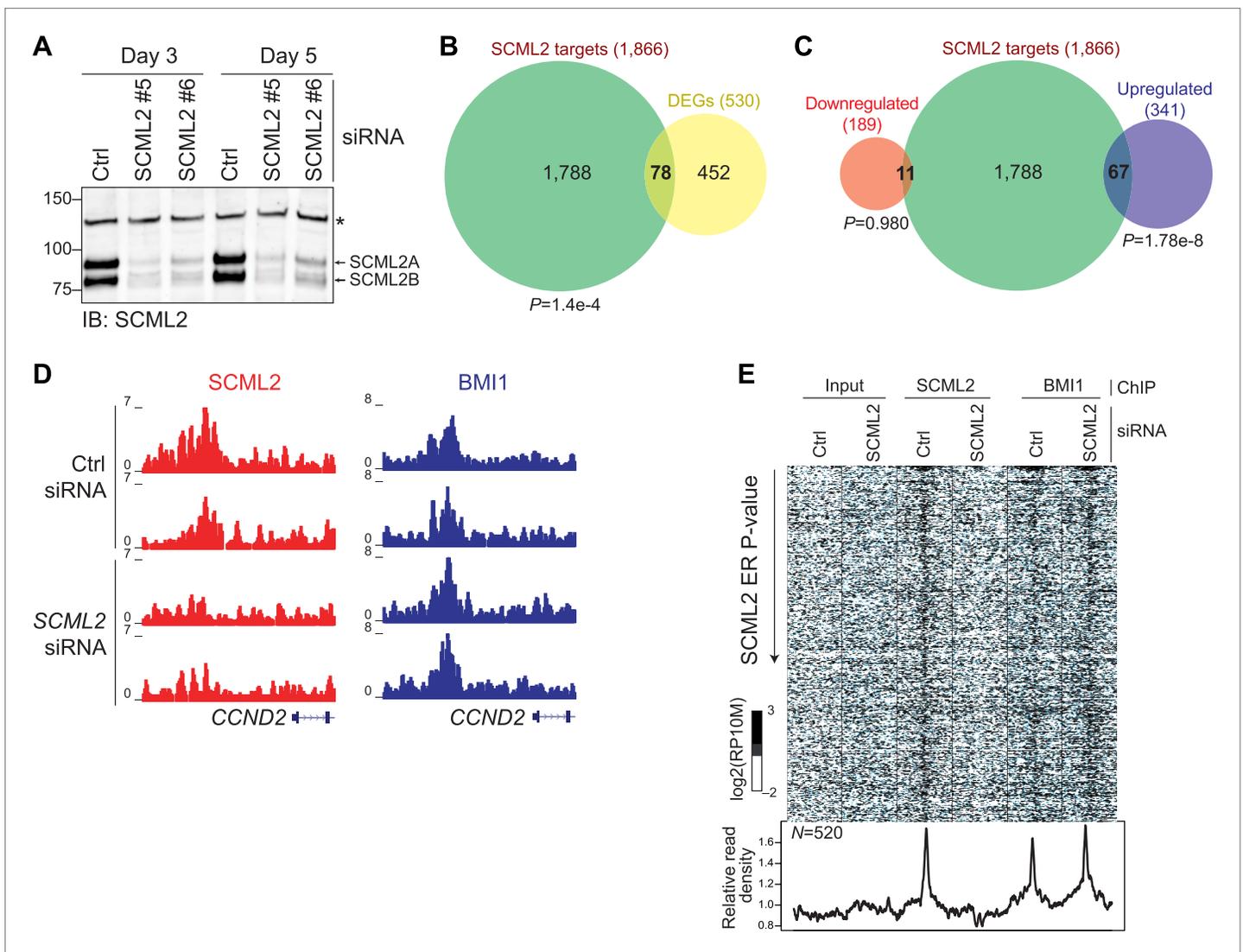


Figure 7. Functional consequences of SCML2 depletion. **(A)** Western blot for SCML2 in 293T-REx treated for 3 or 5 days with control siRNAs (Ctrl) or two different siRNAs against SCML2 (#5 and #6). The band marked with an asterisk corresponds to a cross-reacting protein and serves as loading control. **(B)** Overlap of all DEGs with $p < 0.2$ and genes closest to an SCML2 ER. **(C)** Same as **(B)** but DEGs were divided in downregulated and upregulated. **(D)** Normalized ChIP-seq read densities for SCML2 (red) and BMI1 (blue) in two biological replicates after treating 293T-REx with control (Ctrl) or SCML2 siRNAs. **(E)** Heatmap for all SCML2 ERs showing decreased occupancy ($\log_2(\text{SCML2}_{\text{ctrl}}/\text{SCML2}_{\text{kd}}) > 0.5$) after SCML2 KD ($n = 520$). Profiles for input chromatin, SCML2 and BMI1 ChIP-seq are shown. Read densities were normalized and log-converted. Windows span 2.5 kb on each side of the ER summit and were divided into 25 bp bins.

DOI: [10.7554/eLife.02637.015](https://doi.org/10.7554/eLife.02637.015)

The following figure supplements are available for figure 7:

Figure supplement 1. Additional RNA-seq analyses.

DOI: [10.7554/eLife.02637.016](https://doi.org/10.7554/eLife.02637.016)

of SCML2B to chromatin; and (2) SCML2 enforces repression at some target sites independent of the presence of PRC1. In addition, our results show that RNA binding also plays a role in SCML2 function, as it is required for its localization to chromatin.

In *Drosophila*, the interaction of SCM and PRC1 relies on the C-terminal SPM domain (Peterson et al., 2004) and our results show that a similar mechanism is conserved in humans. Genome-wide analyses revealed an extensive overlap of SCML2A target sites with those of BMI1 (Figure 4), suggesting that this interaction is relevant in a chromatin context. Despite the considerable overlap some sites were occupied by SCML2 or BMI1 alone. The former might reflect the complexity of PRC1 subtypes in mammals and the fact that SCML2 is also part of PRC1.2, which does not contain BMI1 (Gao et al., 2012); the

latter is likely due to the higher sensitivity of the BMI1 antibody, although we cannot exclude that some BMI1 target sites are devoid of SCML2. Together with genetic evidence in *Drosophila* (Simon et al., 1992) and mice (Takada et al., 2007), these observations firmly establish SCML2 as an important player in Polycomb-mediated epigenetic regulation. Consistent with this, the SCML2B isoform, which lacks the SPM domain, does not bind to BMI1 and is unable to recruit PRC1 to chromatin (Figure 4—figure supplement 1B); however, when artificially tethered to a chromatinized reporter, SCML2B is still capable to repress it (Figure 4C), suggesting that SCML2 (and possibly other SCM homologs) might directly enforce gene repression. This notion is further supported by the observation that knockdown for SCML2 causes derepression of target genes without affecting BMI1 levels (Figure 7). This PRC1-independent silencing may be mediated by the MBT domain, which binds to methylated histone tails (Grimm et al., 2007; Santiveri et al., 2008) and, in the case of L3MBTL1, represses transcription via chromatin compaction (Trojer et al., 2007). In fact, the MBT domain of SCM in *Drosophila* is important for repression (Grimm et al., 2007), and our results show that the MBT domain contributes to the chromatin localization of SCML2 (Figure 3D).

In vitro binding assays and domain mapping experiments revealed the existence of a previously unknown RBR within SCML2. Despite the fact that the 74-residue fragment containing the RBR is both necessary and sufficient for RNA binding in vitro (Figures 1 and 2), prediction algorithms failed to detect known protein folds or the propensity to form secondary structures in this region, suggesting that, in isolation, the RBR remains unstructured. This is not entirely unexpected, considering that induced fit is a common feature of RNA-protein interactions (Williamson, 2000). For example, the fragile X syndrome protein FMR1/FMRP contains an arginine-rich stretch that is disordered in solution but becomes ordered in presence of an RNA ligand (Phan et al., 2011). Although the RBR was dispensable for the interaction with PRC1, its disruption resulted not only in a delocalization of SCML2A from chromatin (Figure 3), but also in decreased PRC1 occupancy at several target genes (Figure 5). This observation is intriguing in view of the fact that short-term SCML2 knockdown did not affect BMI1 localization, and suggests that nucleoplasmic SCML2 $_{\Delta RBR}$, being still capable of interacting with BMI1 through the SPM domain, might function in a dominant negative fashion, titrating PRC1 away from chromatin.

RIP-seq identified several species of RNAs that associate with SCML2, including annotated lincRNAs, unannotated RNAs, and protein-coding mRNAs. Despite promiscuous RNA binding in vitro, SCML2-RNA interactions appeared to be specific in vivo, as demonstrated by the following observations: (1) the profile of lincRNAs enriched in the SCML2A WT RIP-seq was unique and distinct from input and control samples (Figure 6C); (2) the lincRNAs associated with SCML2 changed upon cellular differentiation (Figure 6D); and (3) several divergently transcribed ncRNAs were enriched in SCML2 RIPs although the coding mRNAs originating from the same locus were not (Figure 6E, Figure 6—figure supplement 1B–E). The latter also demonstrates that physical proximity on chromatin alone was not the main determinant of SCML2-RNA interactions. Furthermore, the RBR is likely required to mediate the interactions with these RNAs, given that immunoprecipitation of the ΔRBR mutant recovered less RNA than the WT (Figure 6B) and that the few detected ΔRBR -associated RNAs were undistinguishable from those obtained in empty tag-expressing cells (Figure 6C). We propose that the RBR of SCML2A contributes to its recruitment either by interacting with *trans*-acting RNAs transcribed from distant intergenic loci, or by recognizing *cis*-acting ncRNAs. The latter may include divergent and overlapping antisense ncRNAs that originate downstream of the TSS, some of which have a demonstrated role in gene regulation (Faghihi and Wahlestedt, 2009; Modarresi et al., 2012), or promoter-associated ncRNAs that originate near the TSS of coding genes and are transcribed from the same strand (Kapranov et al., 2007), similar to the ncRNAs reported to bind PRC2 via SUZ12 (Kanhere et al., 2010). We note that despite the examples shown (Figure 6E, Figure 6—figure supplement 1B–E) and several others, we could not detect a significant spatial association between the RIP-seq and ChIP-seq signals at the genome-wide level (data not shown). This might be due to the fact that only a subset of the SCML2-associated ncRNAs act in *cis* or to the intrinsically noisy nature of the RIP-seq signal. Future studies utilizing improved technologies and/or different cellular and genetic tools will be needed to fully characterize the set of SCML2-associated ncRNAs and their mechanism of action.

Using different experimental approaches, we have dissected the molecular interplay of SCML2 and PRC1. We have identified a novel RBR in SCML2 that contributes to its chromatin localization. Along with the recognition of methylated lysines by the MBT domain, the RNA-binding activity of the RBR is

required to recruit or stabilize SCML2 on chromatin. We propose that SCML2A is then retained at target genes through the interaction with PRC1 and that, in turn, SCML2A contributes to stabilizing PRC1. We also report that, at a subset of genes, SCML2 appears to enforce repression directly, independent of PRC1 recruitment, via mechanisms that remain to be elucidated.

The identification of SCML2 as a PcG protein that binds RNA and contributes to repression of *Polycomb* targets will help shed new light on the interplay between ncRNAs and epigenetic regulation and on the combinatorial factors that drive the specificity of the diverse PcG complexes in different cellular contexts.

Materials and methods

Cell culture, transfections, and fractionation

HeLa (ATCC), HeLaS3 (ATCC), K562 (ATCC), and 293T-REx (Invitrogen, Carlsbad, CA) cells were grown in DMEM with 10% FBS, penicillin (100 IU/ml), streptomycin (100 µg/ml), and L-glutamine (300 µg/ml). Cytosolic and nuclear extracts were prepared as previously described ([Lecona et al., 2008](#)). The chromatin fraction was either extracted with 8 M urea and 1% Chaps in 50 mM Tris pH 7.9 or solubilized with ammonium sulfate.

Stable transfection of 293T-REx and 293T-REx-luc cells was carried out with polyethylenimine (PEI). The cells were seeded and grown overnight in 10-cm diameter dishes. The next day, 5 µg of plasmid were diluted in 250 µl of 150 mM, and 30 µl of PEI were diluted in 250 µl of 150 mM NaCl. Diluted plasmids and PEI were mixed and incubated for 15 min at room temperature. The mixture was added to the cells and incubated overnight. Next, fresh medium was added to the cells, and they were grown for one additional day. The next day 5 µg/ml blasticidin, 2 µg/ml puromycin (InvivoGen, San Diego, CA), 100 µg Zeocin (Invitrogen), or 1 mg/ml G418 (Sigma-Aldrich, St. Louis, MO) were added individually or in combination, according to the resistance encoded by the plasmid(s) to select resistant clones. Expression of recombinant proteins or shRNAs was induced by adding different amounts of doxycycline for 24 or 48 hr. Lines expressing N3-SCML2A WT, ΔRBR, or empty N3 control were established in two steps. First 293T-REx were transfected with pTRIPZ-V2THS_69827 (Thermo Fisher Scientific, Waltham, MA), and clones were selected with puromycin. The clone displaying the best knockdown efficiency was chosen for further processing and named 'shL2'. The shL2 clone was then transfected with pINTO-N3, pINTO-N3::SCML2A_{WT} or pINTO-N3::SCML2A_{ΔRBR} and sub-clones harboring both the shRNA and the transgene were selected with puromycin and Zeocin and expanded to yield the final lines utilized for the ChIP-seq and RIP-seq experiments.

Transient transfections of HeLa cells with pINTO-GFP plasmids with the different versions of SCML2A and SCML2B were performed with polyfect (Qiagen, Venlo, Netherlands), following manufacturer's instructions. Transfection of the siRNA for human SCML2 (#5 5'-CCAAACGATCTCCTCAGCAAA, #6 5'-CAGTATGTATTGCTACGGTTA) was performed using lipofectamine RNAimax (Invitrogen) according to manufacturer's instructions.

In some experiments K562 were treated for 48 hr with 160 nM PMA in DMSO (Sigma), 100 µM hemin (Sigma-Aldrich), or vehicle (DMSO) alone as control.

Plasmids and sequences

The construction of the backbones for pcDNA4/TO-GAL4 and pINTO plasmids has been described previously ([Vaquero et al., 2004](#); [Gao et al., 2012](#)). SCML2A and SCML2B were cloned from human cDNA; truncations and deletions were obtained by PCR.

Details on the oligonucleotides used for this study are reported in [Supplementary file 1D](#).

Antibodies

Rabbit antibody against SCML2 was generated using an antigen consisting of a central region of SCML2 fused to GST and affinity-purified using the same antigen fused to a 6xHis tag. The following antibodies were used for IP and ChIP experiments: HA (ab9110; Abcam, Cambridge, England), BMI1 (A301-694A; Bethyl, Montgomery, TX), RING1B (A302-869A; Bethyl), GAL4 (06-262; Millipore). Antibodies against GST (sc-33613; Santa Cruz Biotechnology, Santa Cruz, CA) or RBP1 (Reinberg lab) were used for Western Blots.

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_{4°C}, 1.5 mM MgCl₂, 10 mM KCl, protease inhibitors, phosphatase inhibitors) and lysed in Buffer A with 0.2% IGEPAL CA-630 for 5 min on ice. Nuclei were isolated by centrifugation at 2,500×g for 5 min and lysed in Buffer C (20 mM Tris pH 7.9_{4°C}, 25% glycerol, 400 mM NaCl, 1.5 mM MgCl₂, 10 mM EDTA, 0.4 μl/μl murine RNase inhibitor, protease inhibitors, phosphatase inhibitors) for 30 min at 4°C. Lysates were cleared at 20,000×g for 30 min.

For IP, lysates were diluted to 1 mg/ml in RIP buffer (20 mM Tris pH 7.9_{4°C}, 200 mM KCl, 0.05% IGEPAL CA-630, 10 mM EDTA), cleared by centrifugation at 20,000×g for 10 min, and incubated with a depleting amount of antibody (predetermined with titration experiments) for 3 hr at 4°C. Immunocomplexes were recovered by incubating for 1 hr at 4°C with 7 μl of protein G-coupled Dynabeads (Invitrogen) per μg of antibody used. Beads were washed in RIP-W buffer (RIP buffer without EDTA and with 1 mM MgCl₂) twice and incubated with 2 U of TURBO DNase (Life Technologies, Carlsbad, CA) in 20 μl RIP-W buffer for 10 min at room temperature, to eliminate potential bridging effects of protein–DNA and RNA–DNA 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 37°C.

Recombinant protein expression and purification

GST fusion proteins were expressed in BL21(DE3) cells for 16 hr at 16°C and purified using a glutathione–sepharose column. The column was washed extensively with 50 mM Tris pH 7.5, 50 mM NaCl, and the proteins were eluted in the presence of 10 mM glutathione. The purified proteins were dialyzed against 50 mM tris pH 7.5, 50 mM NaCl and 10% glycerol.

In vitro RNA binding assays

Single-stranded DNA fragments were commercially synthesized (IDT) and annealed to generate double-stranded species. Single-stranded RNA were in vitro transcribed with the HiScribe kit (New England Biolabs, Ipswich, MA) and purified by size exclusion on mini Quick Spin columns (Roche, Penzberg, Germany) followed by TRIzol (Invitrogen) extraction and isopropanol precipitation. *HOTAIR* RNA fragments were obtained from a construct kindly gifted by H Chang (Stanford).

For the binding assays, GST-tagged proteins and protein fragments were incubated with RNA or DNA in 100 μl RIP-W buffer (20 mM Tris pH 7.9_{4°C}, 1 mM MgCl₂, 0.05% IGEPAL CA-630) with the addition of 2 U/μl murine RNase inhibitor (New England Biolabs) for 1 hr at 4°C. Protein–RNA complexes were precipitated by adding 10–20 μl of glutathione-coupled beads (Qiagen) for 30 min at 4°C. After three washes with RIP-W buffer, proteins were eluted from the beads with Laemmli sample buffer and nucleic acid with TRIzol or phenol/chloroform/isoamyl alcohol. RNA and DNA were resolved on polyacrylamide/urea gels and visualized with SYBR gold (Invitrogen).

Electrophoretic mobility shift assays (EMSAs)

The GST-tagged RBR corresponding to residues 256–330 of human SCML2 was purified as described above. In some cases, the GST tag was cleaved off with prescission protease (GE Healthcare, Little Chalfont, UK) and residual uncleaved protein removed with a GST column. The 1–300 fragment of *HOTAIR* and a 200-nts construct spanning the 601 nucleosome positioning sequence were obtained by in vitro transcription. In vitro-transcribed *HOTAIR* was end-labeled with [³²P]-γ-ATP and purified by phenol–chloroform extraction. RNA was refolded prior to EMSA assays by heating at 95°C for 5 min in BTE buffer (10 mM bis-tris pH 6.7, 1 mM EDTA) then incubating on ice and room temperature for 5 min each. Double-stranded DNA encoding *HOTAIR* 1–300 was generated by PCR. Nucleosome core particles were generated by salt dialysis of recombinant histone octamers and a 147-bp fragment corresponding to the 601 nucleosome positioning sequence, as described (Voigt *et al.*, 2012). Dialysis was performed as a gradient of 2–0.4 M NaCl overnight at 4°C, followed by dialysis into TE buffer. Recombinant wild-type *Xenopus laevis* core histones were purified from *E. coli*.

Protein and nucleic acid were co-incubated for 30 min on ice in 10 mM HEPES pH 7.9, 100 mM KCl, 0.5 mM EDTA, 5% glycerol, 0.25 mg/ml BSA, and 400 U/ml murine RNase inhibitor (New England Biolabs). 50 ng of RNA, double-stranded DNA, or nucleosomes were incubated with 70–140 ng of GST or GST-fused RBR, or 17–35 ng of RBR without GST in 15-μl binding reactions. Samples were loaded onto 4% acrylamide mini gels pre-run for 1 hr at 100 V in 0.5x TGE (12.5 mM Tris pH 8.8, 95 mM glycine, 0.5 mM EDTA). Electrophoresis was performed for 2 hr at 100 V at 4°C. Gels were stained with SYBR gold (Invitrogen) and documented with a LAS-4000 imaging system (GE Healthcare). Radiolabeled RNA was detected using a phosphoimaging screen (GE Healthcare).

Nucleosome pull-downs

Nucleosome core particles were reconstituted on biotinylated 147-bp 601 DNA. Streptavidin agarose beads (Millipore, Billerica, MA) were blocked with 1% BSA in pull-down buffer (10 mM HEPES pH 7.9, 100 mM KCl, 0.5 mM EDTA, 5% glycerol, 0.01% NP-40, 400 U/ml murine RNase inhibitor) for 4 hr at 4°C. Binding reactions containing 1 µg of nucleosomes, 1 µg of RBR, and varying amounts of *HOTAIR* RNA (100–400 ng) were incubated on ice for 30 min and then with the blocked beads for 3 hr at 4°C. After two washes, beads from each reaction were divided for western blot and RNA analysis. RNA was purified with TRIzol (Invitrogen) and analyzed by non-denaturing acrylamide gel electrophoresis. Proteins were eluted by boiling in 1x SDS loading buffer and analyzed by Western blot with histone H3 and GST antibodies.

Luciferase assay

5 × 10⁴ 293T-REx-luc cells stably transfected with pcDNA4/TO::GAL4-SCML2A or SCML2B were seeded in 24-well plates. The cells were incubated for 24 hr in the presence of different doxycycline concentrations or the equivalent amount of ethanol as a control. After washing twice with PBS, the cells were resuspended in 100 mM potassium phosphate pH 7.8 containing 0.2% triton X-100, and lysed for 15 min at 4°C with agitation. Cell debris was removed by centrifugation at 20,000×g for 10 min, and the supernatant was assayed for luciferase activity using a luciferase assay system according to manufacturer's instructions (Promega, Madison, WI).

Fluorescence microscopy

48 hr after transfection, cells were washed three times and fixed with 4% paraformaldehyde in PBS for 10 min at room temperature. After washing twice with PBS and twice with water, the cells were mounted using SlowFade with DAPI (Invitrogen) and imaged using an Axiovert 200 M inverted microscope (Zeiss, Oberkochen, Germany).

Purification of TwinStrepTag-tagged proteins

30 mg of solubilized nuclear pellet from 293T-REx expressing FS-SCML2A or SCML2B, wild-type or ΔRBR were diluted in 10 ml of 50 mM Tris pH 7.5 with 200 mM NaCl and incubated with 600 µl of Strep-Tactin resin (IBA, Göttingen, Germany) for 1 hr at 4°C with constant agitation. The resin was allowed to settle and the flow-through was collected. The columns were washed with 10 ml of 50 mM Tris pH 7.5, 300 mM NaCl, and 0.1% Igepal CA630 (Sigma), and then with 4 ml of 50 mM Tris pH 7.5, 200 mM NaCl. The proteins bound to the resin were eluted in denaturing loading buffer and analyzed by mass spectrometry.

Chromatin immunoprecipitation

Cells were fixed in DMEM containing 10 mM HEPES, pH 7.6, 1% formaldehyde, 15 mM NaCl, 0.15 mM EDTA and 0.075 mM EGTA, for 10 min at room temperature. The reaction was quenched with 0.125 M glycine for 5 min at room temperature. After washing with PBS, cells were collected, and the pellet was lysed in 50 mM HEPES, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% Igepal CA630, and 0.25% Triton X-100. After centrifugation, isolated nuclei were washed once in 10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA and 0.5 mM EGTA, then resuspended in 10 mM Tris pH 8.0, 1 mM EDTA, 0.5 mM EGTA and 0.5% N-Lauryl Sarcosine, and sonicated using a Diagenode Bioruptor to an average chromatin size of 200 bp. Chromatin was diluted to 150–200 µg/300 µl in 10 mM Tris pH 8.0, 1 mM EDTA, 600 mM NaCl and 3% Triton X-100, the antibodies were added (4–8 µg) and incubated at 4°C overnight. Protein G Dynabeads (Invitrogen) were blocked with BSA, and 20–40 µl of beads were added to each IP. After 1 hr of incubation at 4°C, the beads were washed five times with 750 µl of 25 mM HEPES pH 7.6, 1 mM EDTA, 0.1% N-Lauryl Sarcosine, 1% NP-40, and 0.5 M LiCl. After a final wash with 750 µl of 10 mM Tris pH 8.0, 1 mM EDTA and 50 mM NaCl, the beads were resuspended in 100 mM sodium bicarbonate, 200 mM NaCl, and 1% SDS. 1 µg of proteinase K was added to each sample, the chromatin was incubated 15 min at room temperature and the cross-linking was reversed at 65°C for 16 hr. 10% of the input was treated in parallel. The DNA was extracted using a PCR purification kit (Qiagen).

Library construction

Libraries for ChIP-seq were prepared according to manufacturer's instructions (Illumina, San Diego, CA) and as described (Asp *et al.*, 2011). Briefly, IP'd DNA (~5 ng) was end-repaired using End-It Repair Kit (Epicenter, Madison, WI), tailed with deoxyadenine using Klenow exo⁻ (New England Biolabs),

and ligated to custom adapters with LigaFast (Promega). Fragments of 300 ± 50 bp were size-selected and subjected to ligation-mediated PCR amplification using Phusion DNA polymerase (New England Biolabs). Libraries were quantified by qPCR using primers annealing to the adapter sequence and sequenced at a concentration of 7 pM on an Illumina Genome Analyzer IIx or 10 pM on an Illumina HiSeq. In some cases barcoding was utilized for multiplexing.

For RIP-seq libraries, polyA+ RNA was isolated using Dynabeads Oligo(dT)₂₅ (Invitrogen) beads and constructed into strand-specific libraries using the dUTP method (*Parkhomchuk et al., 2009*). Once dUTP-marked double-stranded cDNA was obtained, the remaining library construction steps followed the same protocol as described above for ChIP-seq libraries.

ChIP-seq, RNA-seq, and RIP-seq analysis

ChIP-seq analysis was performed as described before with modifications (*Gao et al., 2012*). Sequenced reads from ChIP-seq experiments were mapped with Bowtie using parameters $-v2 -m4$ (*Langmead et al., 2009*). Normalized genome-wide read densities were visualized on the UCSC genome browser (<http://genome.ucsc.edu>). Enriched regions (ERs) were identified using MACS 1.40rc2 (*Zhang et al., 2008*) using default parameters requiring at least 10 reads per ER and an unadjusted p-value $<10^{-5}$. GO analyses were performed with DAVID (*Huang da et al., 2009*) and GREAT (*McLean et al., 2010*). Heatmaps were generated in R.

RIP-seq and RNA-seq reads were mapped to the reference genome with Bowtie using parameters $-v2 -m40$. Normalized genome-wide read densities for each strand separately were visualized on the UCSC genome browser. Reads were assigned to genes using DEGseq (*Wang et al., 2010a*) and either the ENSEMBL annotation or the lincRNA catalog (*Cabili et al., 2011*) made available by the Broad Institute. DEG identification was performed with the R package DESeq (*Anders and Huber, 2010*).

Sequencing data

All sequencing data is available in the NCBI GEO as SuperSeries GSE38275.

Acknowledgements

We wish to thank Moran Cabili and John Rinn for their help with lincRNA analyses. This work was supported by grants from the National Institute of Health (GM-64844 and R37-37120) and the Howard Hughes Medical Institute (to DR). RB was supported by a Helen Hay Whitney Foundation postdoctoral fellowship and by the Helen L and Martin S Kimmel Center for Stem Cell Biology postdoctoral fellow award. EL was supported by an International Outgoing Fellowship from the European Commission (Marie-Curie Actions, FP7). PV was supported by a fellowship from the Deutsche Akademie der Naturforscher Leopoldina (LPDS 2009-5) and subsequently by a fellowship from the Empire State Training Program in Stem Cell Research (NYSTEM, contract no. C026880). FP and YK were supported by a grant from the National Cancer Institute and National Institute of Health (CA-16359 to YK).

Additional information

Competing interests

DR: Reviewing editor, *eLife*. The authors declare that no competing interests exist.

Funding

Funder	Grant reference number	Author
National Institutes of Health (NIH)	GM-64844, R37-37120	Danny Reinberg
Howard Hughes Medical Institute (HHMI)		Danny Reinberg
National Institutes of Health (NIH)	CA-16359	Fabio Parisi, Yuval Kluger

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

RB, EL, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; VN, PV, Acquisition of data, Analysis and interpretation of data; FP, YK, Analysis

and interpretation of data, Drafting or revising the article; DR, Conception and design, Drafting or revising the article

Additional files

Supplementary file

• Supplementary file 1. (A) GO enrichment analysis for SCML2 targets in 293T-Rex. (B) SCML2 target genes in 293T-REX. (C) Differentially expressed genes after SCML2 KD. (D) Oligonucleotide sequences. DOI: [10.7554/eLife.02637.017](https://doi.org/10.7554/eLife.02637.017)

Major dataset

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset ID and/or URL	Database, license, and accessibility information
Bonasio R, Lecona E, Reinberg D	2014	ChIP-seq and RIP-seq for SCML2 and SCML2 mutants in 293T-REX and K562 cells	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE38275	Publicly available at NCBI Gene Expression Omnibus.

References

- Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biology* **11**:R106. doi: [10.1186/gb-2010-11-10-r106](https://doi.org/10.1186/gb-2010-11-10-r106).
- Asp P, Blum R, Vethantham V, Parisi F, Micsinai M, Cheng J, Bowman C, Kluger Y, Dynlacht BD. 2011. Genome-wide remodeling of the epigenetic landscape during myogenic differentiation. *Proceedings of the National Academy of Sciences of the United States of America* **108**:E149–E158. doi: [10.1073/pnas.1102223108](https://doi.org/10.1073/pnas.1102223108).
- Bengani H, Mendiratta S, Maini J, Vasanthi D, Sultana H, Ghasemi M, Ahluwalia J, Ramachandran S, Mishra RK, Brahmachari V. 2013. Identification and validation of a putative polycomb responsive element in the human genome. *PLOS ONE* **8**:e67217. doi: [10.1371/journal.pone.0067217](https://doi.org/10.1371/journal.pone.0067217).
- Bonasio R, Lecona E, Reinberg D. 2010a. MBT domain proteins in development and disease. *Seminars in Cell & Developmental Biology* **21**:221–230. doi: [10.1016/j.semcdb.2009.09.010](https://doi.org/10.1016/j.semcdb.2009.09.010).
- Bonasio R, Tu S, Reinberg D. 2010b. Molecular signals of epigenetic states. *Science* **330**:612–616. doi: [10.1126/science.1191078](https://doi.org/10.1126/science.1191078).
- Bornemann D, Miller E, Simon J. 1996. The Drosophila Polycomb group gene Sex comb on midleg (Scm) encodes a zinc finger protein with similarity to polyhomeotic protein. *Development* **122**:1621–1630.
- Bornemann D, Miller E, Simon J. 1998. Expression and properties of wild-type and mutant forms of the Drosophila sex comb on midleg (SCM) repressor protein. *Genetics* **150**:675–686.
- Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & Development* **25**:1915–1927. doi: [10.1101/gad.17446611](https://doi.org/10.1101/gad.17446611).
- Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**:1039–1043. doi: [10.1126/science.1076997](https://doi.org/10.1126/science.1076997).
- Cuddapah S, Roh TY, Cui K, Jose CC, Fuller MT, Zhao K, Chen X. 2012. A novel human polycomb binding site acts as a functional polycomb response element in Drosophila. *PLOS ONE* **7**:e36365. doi: [10.1371/journal.pone.0036365](https://doi.org/10.1371/journal.pone.0036365).
- Davidovich C, Zheng L, Goodrich KJ, Cech TR. 2013. Promiscuous RNA binding by Polycomb repressive complex 2. *Nature Structural & Molecular Biology* **20**:1250–1257. doi: [10.1038/nsmb.2679](https://doi.org/10.1038/nsmb.2679).
- Dignam JD, Lebovitz RM, Roeder RG. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Research* **11**:1475–1489. doi: [10.1093/nar/11.5.1475](https://doi.org/10.1093/nar/11.5.1475).
- El Messaoudi-Aubert S, Nicholls J, Maertens GN, Brookes S, Bernstein E, Peters G. 2010. Role for the MOV10 RNA helicase in polycomb-mediated repression of the INK4a tumor suppressor. *Nature Structural & Molecular Biology* **17**:862–868. doi: [10.1038/nsmb.1824](https://doi.org/10.1038/nsmb.1824).
- Faghihi MA, Wahlestedt C. 2009. Regulatory roles of natural antisense transcripts. *Nature Reviews Molecular Cell Biology* **10**:637–643. doi: [10.1038/nrm2738](https://doi.org/10.1038/nrm2738).
- Gao Z, Zhang J, Bonasio R, Strino F, Sawai A, Parisi F, Kluger Y, Reinberg D. 2012. PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Molecular Cell* **45**:344–356. doi: [10.1016/j.molcel.2012.01.002](https://doi.org/10.1016/j.molcel.2012.01.002).
- Grimm C, de Ayala Alonso AG, Rybin V, Steuerwald U, Ly-Hartig N, Fischle W, Müller J, Müller CW. 2007. Structural and functional analyses of methyl-lysine binding by the malignant brain tumour repeat protein Sex comb on midleg. *EMBO Reports* **8**:1031–1037. doi: [10.1038/sj.embor.7401085](https://doi.org/10.1038/sj.embor.7401085).
- Huang da W, Sherman BT, Zheng X, Yang J, Imamichi T, Stephens R, Lempicki RA. 2009. Extracting biological meaning from large gene lists with DAVID. *Current Protocols in Bioinformatics* Chapter 13, Unit 13.11. doi: [10.1002/0471250953.bi1311s27](https://doi.org/10.1002/0471250953.bi1311s27).

- Jacquel A**, Herrant M, Defamie V, Belhacene N, Colosetti P, Marchetti S, Legros L, Deckert M, Mari B, Cassuto JP, Hofman P, Auburger P. 2006. A survey of the signaling pathways involved in megakaryocytic differentiation of the human K562 leukemia cell line by molecular and c-DNA array analysis. *Oncogene* **25**:781–794. doi: [10.1038/sj.onc.1209119](https://doi.org/10.1038/sj.onc.1209119).
- Kaneko S**, Bonasio R, Saldana-Meyer R, Yoshida T, Son J, Nishino K, Umezawa A, Reinberg D. 2014. Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Molecular Cell* **53**:290–300. doi: [10.1016/j.molcel.2013.11.012](https://doi.org/10.1016/j.molcel.2013.11.012).
- Kaneko S**, Li G, Son J, Xu CF, Margueron R, Neubert TA, Reinberg D. 2010. Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. *Genes & Development* **24**:2615–2620. doi: [10.1101/gad.1983810](https://doi.org/10.1101/gad.1983810).
- Kaneko S**, Son J, Shen SS, Reinberg D, Bonasio R. 2013. PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nature Structural & Molecular Biology* **20**:1258–1264. doi: [10.1038/nsmb.2700](https://doi.org/10.1038/nsmb.2700).
- Kanhere A**, Viiri K, Araujo CC, Rasaiyaah J, Bouwman RD, Whyte WA, Pereira CF, Brookes E, Walker K, Bell GW, Pombo A, Fisher AG, Young RA, Jenner RG. 2010. Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Molecular Cell* **38**:675–688. doi: [10.1016/j.molcel.2010.03.019](https://doi.org/10.1016/j.molcel.2010.03.019).
- Kapranov P**, Cheng J, Dike S, Nix DA, Dutttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermüller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt G, Ganesh M, Ghosh S, Piccolboni A, Sementchenko V, Tammana H, Gingeras TR. 2007. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316**:1484–1488. doi: [10.1126/science.1138341](https://doi.org/10.1126/science.1138341).
- Khalil AM**, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL. 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **106**:11667–11672. doi: [10.1073/pnas.0904715106](https://doi.org/10.1073/pnas.0904715106).
- Koziol MJ**, Rinn JL. 2010. RNA traffic control of chromatin complexes. *Current Opinion in Genetics & Development* **20**:142–148. doi: [10.1016/j.gde.2010.03.003](https://doi.org/10.1016/j.gde.2010.03.003).
- Kuzmichev A**, Jenuwein T, Tempst P, Reinberg D. 2004. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Molecular Cell* **14**:183–193. doi: [10.1016/S1097-2765\(04\)00185-6](https://doi.org/10.1016/S1097-2765(04)00185-6).
- Kuzmichev A**, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. *Genes & Development* **16**:2893–2905. doi: [10.1101/gad.1035902](https://doi.org/10.1101/gad.1035902).
- Langmead B**, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**:R25. doi: [10.1186/gb-2009-10-3-r25](https://doi.org/10.1186/gb-2009-10-3-r25).
- Lecona E**, Barrasa JI, Olmo N, Llorente B, Turnay J, Lizarbe MA. 2008. Upregulation of annexin A1 expression by butyrate in human colon adenocarcinoma cells: role of p53, NF- κ B, and p38 mitogen-activated protein kinase. *Molecular and Cellular Biology* **28**:4665–4674. doi: [10.1128/MCB.00650-07](https://doi.org/10.1128/MCB.00650-07).
- Lecona E**, Rojas LA, Bonasio R, Johnston A, Fernandez-Capetillo O, Reinberg D. 2013. Polycomb protein SCML2 regulates the cell cycle by binding and modulating CDK/CYCLIN/p21 complexes. *PLOS Biology* **11**:e1001737. doi: [10.1371/journal.pbio.1001737](https://doi.org/10.1371/journal.pbio.1001737).
- Levine SS**, Weiss A, Erdjument-Bromage H, Shao Z, Tempst P, Kingston RE. 2002. The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. *Molecular and Cellular Biology* **22**:6070–6078. doi: [10.1128/MCB.22.17.6070-6078.2002](https://doi.org/10.1128/MCB.22.17.6070-6078.2002).
- Lewis EB**. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**:565–570. doi: [10.1038/276565a0](https://doi.org/10.1038/276565a0).
- Margueron R**, Li G, Sarma K, Blais A, Zavadil J, Woodcock CL, Dynlacht BD, Reinberg D. 2008. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Molecular Cell* **32**:503–518. doi: [10.1016/j.molcel.2008.11.004](https://doi.org/10.1016/j.molcel.2008.11.004).
- McLean CY**, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology* **28**:495–501. doi: [10.1038/nbt.1630](https://doi.org/10.1038/nbt.1630).
- Modarresi F**, Faghihi MA, Lopez-Toledano MA, Fatemi RP, Magistri M, Brothers SP, van der Brug MP, Wahlestedt C. 2012. Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. *Nature Biotechnology* **30**:453–459. doi: [10.1038/nbt.2158](https://doi.org/10.1038/nbt.2158).
- Müller J**, Kassis J. 2006. Polycomb response elements and targeting of Polycomb group proteins in *Drosophila*. *Current Opinion in Genetics & Development* **16**:476–484. doi: [10.1016/j.gde.2006.08.005](https://doi.org/10.1016/j.gde.2006.08.005).
- Nady N**, Krichevsky L, Zhong N, Duan S, Tempel W, Amaya MF, Ravichandran M, Arrowsmith CH. 2012. Histone recognition by human malignant brain tumor domains. *Journal of Molecular Biology* **423**:702–718. doi: [10.1016/j.jmb.2012.08.022](https://doi.org/10.1016/j.jmb.2012.08.022).
- Northcott PA**, Nakahara Y, Wu X, Feuk L, Ellison DW, Croul S, Mack S, Kongkham PN, Peacock J, Dubuc A, Ra YS, Zilberberg K, McLeod J, Scherer SW, Sunil Rao J, Eberhart CG, Grajkowska W, Gillespie Y, Lach B, Grundy R, Pollack IF, Hamilton RL, Van Meter T, Carlotti CG, Boop F, Bigner D, Gilbertson RJ, Rutka JT, Taylor MD. 2009. Multiple recurrent genetic events converge on control of histone lysine methylation in medulloblastoma. *Nature Genetics* **41**:465–472. doi: [10.1038/ng.336](https://doi.org/10.1038/ng.336).
- Parkhomchuk D**, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A. 2009. Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Research* **37**:e123. doi: [10.1093/nar/gkp596](https://doi.org/10.1093/nar/gkp596).

- Peterson AJ, Mallin DR, Francis NJ, Ketel CS, Stamm J, Voeller RK, Kingston RE, Simon JA. 2004. Requirement for sex comb on midleg protein interactions in *Drosophila* polycomb group repression. *Genetics* **167**:1225–1239. doi: [10.1534/genetics.104.027474](https://doi.org/10.1534/genetics.104.027474).
- Phan AT, Kuryavyi V, Darnell JC, Serganov A, Majumdar A, Ilin S, Raslin T, Polonskaia A, Chen C, Clain D, Darnell RB, Patel DJ. 2011. Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nature Structural & Molecular Biology* **18**:796–804. doi: [10.1038/nsmb.2064](https://doi.org/10.1038/nsmb.2064).
- Richly H, Aloia L, Di Croce L. 2011. Roles of the Polycomb group proteins in stem cells and cancer. *Cell Death & Disease* **2**:e204. doi: [10.1038/cddis.2011.84](https://doi.org/10.1038/cddis.2011.84).
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. 2007. Functional Demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**:1311–1323. doi: [10.1016/j.cell.2007.05.022](https://doi.org/10.1016/j.cell.2007.05.022).
- Roseman RR, Morgan K, Mallin DR, Roberson R, Parnell TJ, Bornemann DJ, Simon JA, Geyer PK. 2001. Long-range repression by multiple polycomb group (PcG) proteins targeted by fusion to a defined DNA-binding domain in *Drosophila*. *Genetics* **158**:291–307.
- Santiveri CM, Lechtenberg BC, Allen MD, Sathyamurthy A, Jaulent AM, Freund SMV, Bycroft M. 2008. The malignant brain tumor repeats of human SCML2 bind to peptides containing monomethylated lysine. *Journal of Molecular Biology* **382**:1107–1112. doi: [10.1016/j.jmb.2008.07.081](https://doi.org/10.1016/j.jmb.2008.07.081).
- Sathyamurthy A, Allen MD, Murzin AG, Bycroft M. 2003. Crystal structure of the malignant brain tumor (MBT) repeats in Sex Comb on Midleg-like 2 (SCML2). *The Journal of Biological Chemistry* **278**:46968–46973. doi: [10.1074/jbc.M306469200](https://doi.org/10.1074/jbc.M306469200).
- Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. 2007. Genome regulation by polycomb and trithorax proteins. *Cell* **128**:735–745. doi: [10.1016/j.cell.2007.02.009](https://doi.org/10.1016/j.cell.2007.02.009).
- Shao Z, Raible F, Mollaaghababa R, Guyon JR, Wu CT, Bender W, Kingston RE. 1999. Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* **98**:37–46. doi: [10.1016/S0092-8674\(00\)80604-2](https://doi.org/10.1016/S0092-8674(00)80604-2).
- Simon J, Chiang A, Bender W. 1992. Ten different Polycomb group genes are required for spatial control of the *abdA* and *AbdB* homeotic products. *Development* **114**:493–505.
- Simon JA, Kingston RE. 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature Reviews Molecular Cell Biology* **10**:697–708. doi: [10.1038/nrm2763](https://doi.org/10.1038/nrm2763).
- Simon JA, Kingston RE. 2013. Occupying chromatin: polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Molecular Cell* **49**:808–824. doi: [10.1016/j.molcel.2013.02.013](https://doi.org/10.1016/j.molcel.2013.02.013).
- Sing A, Pannell D, Karaiskakis A, Sturgeon K, Djabali M, Ellis J, Lipshitz HD, Cordes SP. 2009. A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell* **138**:885–897. doi: [10.1016/j.cell.2009.08.020](https://doi.org/10.1016/j.cell.2009.08.020).
- Sowa ME, Bennett EJ, Gygi SP, Harper JW. 2009. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**:389–403. doi: [10.1016/j.cell.2009.04.042](https://doi.org/10.1016/j.cell.2009.04.042).
- Sparmann A, van Lohuizen M. 2006. Polycomb silencers control cell fate, development and cancer. *Nature Reviews Cancer* **6**:846–856. doi: [10.1038/nrc1991](https://doi.org/10.1038/nrc1991).
- Struhl G. 1981. A gene product required for correct initiation of segmental determination in *Drosophila*. *Nature* **293**:36–41. doi: [10.1038/293036a0](https://doi.org/10.1038/293036a0).
- Takada Y, Isono K-i, Shinga J, Turner JMA, Kitamura H, Ohara O, Watanabe G, Singh PB, Kamijo T, Jenuwein T, Burgoyne PS, Koseki H. 2007. Mammalian Polycomb Scmh1 mediates exclusion of Polycomb complexes from the XY body in the pachytene spermatocytes. *Development* **134**:579–590. doi: [10.1242/dev.02747](https://doi.org/10.1242/dev.02747).
- Tavares L, Dimitrova E, Oxley D, Webster J, Poot R, Demmers J, Bezstarosti K, Taylor S, Ura H, Koide H, Wutz A, Vidal M, Elderkin S, Brockdorff N. 2012. RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* **148**:664–678. doi: [10.1016/j.cell.2011.12.029](https://doi.org/10.1016/j.cell.2011.12.029).
- Trojer P, Li G, Sims RJ, Vaquero A, Kalakonda N, Bocconi P, Lee D, Erdjument-Bromage H, Tempst P, Nimer SD, Wang YH, Reinberg D. 2007. L3MBTL1, a histone-methylation-dependent chromatin lock. *Cell* **129**:915–928. doi: [10.1016/j.cell.2007.03.048](https://doi.org/10.1016/j.cell.2007.03.048).
- Tsai MC, Manor O, Wan Y, Mosammamparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. 2010. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**:689–693. doi: [10.1126/science.1192002](https://doi.org/10.1126/science.1192002).
- Vaquero A, Scher M, Lee D, Erdjument-Bromage H, Tempst P, Reinberg D. 2004. Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin. *Molecular Cell* **16**:93–105. doi: [10.1016/j.molcel.2004.08.031](https://doi.org/10.1016/j.molcel.2004.08.031).
- Voigt P, LeRoy G, Drury WJ III, Zee BM, Son J, Beck DB, Young NL, Garcia BA, Reinberg D. 2012. Asymmetrically modified nucleosomes. *Cell* **151**:181–193. doi: [10.1016/j.cell.2012.09.002](https://doi.org/10.1016/j.cell.2012.09.002).
- Wang KC, Chang HY. 2011. Molecular mechanisms of long noncoding RNAs. *Molecular Cell* **43**:904–914. doi: [10.1016/j.molcel.2011.08.018](https://doi.org/10.1016/j.molcel.2011.08.018).
- Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA, Chang HY. 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**:120–124. doi: [10.1038/nature09819](https://doi.org/10.1038/nature09819).
- Wang L, Feng Z, Wang X, Zhang X. 2010a. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**:136–138. doi: [10.1093/bioinformatics/btp612](https://doi.org/10.1093/bioinformatics/btp612).
- Wang L, Jähren N, Miller EL, Ketel CS, Mallin DR, Simon JA. 2010b. Comparative analysis of chromatin binding by Sex Comb on Midleg (SCM) and other polycomb group repressors at a *Drosophila* Hox gene. *Molecular and Cellular Biology* **30**:2584–2593. doi: [10.1128/MCB.01451-09](https://doi.org/10.1128/MCB.01451-09).

- Williamson JR.** 2000. Induced fit in RNA-protein recognition. *Nature Structural Biology* **7**:834–837. doi: [10.1038/79575](https://doi.org/10.1038/79575).
- Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE.** 2010. A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* **140**:99–110. doi: [10.1016/j.cell.2009.12.022](https://doi.org/10.1016/j.cell.2009.12.022).
- Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE.** 2013. Variable requirements for DNA-binding proteins at polycomb-dependent repressive regions in human HOX clusters. *Molecular and Cellular Biology* **33**:3274–3285. doi: [10.1128/MCB.00275-13](https://doi.org/10.1128/MCB.00275-13).
- Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM.** 2010. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Molecular Cell* **38**:662–674. doi: [10.1016/j.molcel.2010.03.021](https://doi.org/10.1016/j.molcel.2010.03.021).
- Yu M, Mazor T, Huang H, Huang HT, Kathrein KL, Woo AJ, Chouinard CR, Labadorf A, Akie TE, Moran TB, Xie H, Zacharek S, Taniuchi I, Roeder RG, Kim CF, Zon LI, Fraenkel E, Cantor AB.** 2012. Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors. *Molecular Cell* **45**:330–343. doi: [10.1016/j.molcel.2011.11.032](https://doi.org/10.1016/j.molcel.2011.11.032).
- Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, Liu XS.** 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biology* **9**:R137. doi: [10.1186/gb-2008-9-9-r137](https://doi.org/10.1186/gb-2008-9-9-r137).