G-tract RNA removes Polycomb repressive complex 2 from genes

Manuel Beltran1, Manuel Tavares1, Neil Justin2, Garima Khandelwal1, John Ambrose1,4, Benjamin M. Foster3, Kaylee B. Worlock1, Andrey Tvardovskiy3, Simone Kunzelmann2, Javier Herrero3, Till Bartke3, Steven J. Gamblin2, Jon R. Wilson2 and Richard G. Jenner1*

Polycomb repressive complex 2 (PRC2) maintains repression of cell-type-specific genes but also associates with genes ectopically in cancer. While it is currently unknown how PRC2 is removed from genes, such knowledge would be useful for the targeted reversal of deleterious PRC2 recruitment events. Here, we show that G-tract RNA specifically removes PRC2 from genes in human and mouse cells. PRC2 preferentially binds G tracts within nascent precursor mRNA (pre-mRNA), especially within predicted G-quadruplex structures. G-quadruplex RNA evicts the PRC2 catalytic core from the substrate nucleosome. In cells, PRC2 transfers from chromatin to pre-mRNA upon gene activation, and chromatin-associated G-tract RNA removes PRC2, leading to H3K27me3 depletion from genes. Targeting G-tract RNA to the tumor suppressor gene CDKN2A in malignant rhabdoid tumor cells reactivates the gene and induces senescence. These data support a model in which pre-mRNA evicts PRC2 during gene activation and provides the means to selectively remove PRC2 from specific genes.

Chromatin structure is responsive to changes in transcriptional state, but the mechanisms for this are unclear. Nascent pre-mRNA has primarily been considered to be a passive intermediary, but a potential regulatory role for nascent pre-mRNA may explain some of the changes in chromatin structure that occur upon gene expression.

The chromatin regulator PRC2 prevents inappropriate activation of genes specific for other cell types and other stages of cell differentiation. The PRC2 subunit EZH2 methylates histone H3 lysine 27 (H3K27) and, together with PRC1, induces formation of a repressive chromatin conformation. PRC2 is essential for cell differentiation, both during embryogenesis and throughout life. Dysregulation of PRC2 function occurs in a range of cancers and can drive cancer cell proliferation, invasion and metastasis. EZH2 methyltransferase inhibitors block proliferation of a range of cancer cell types, including malignant rhabdoid tumors (MRTs), germinal center B-cell diffuse large B-cell lymphoma and diffuse intrinsic pontine glioma (DIPG), and are currently being evaluated in clinical trials.

The binding of PRC2 to genes is dynamic. During cell differentiation, PRC2 is lost from genes that become activated and gained at genes that become repressed. Changes in PRC2 occupancy and trimethylated H3K27 (H3K27me3) are also observed during cell transformation and in cancer. The oncogenic effects of PRC2 have been linked to ectopic repression of particular genes, for example CDKN2A (encoding p16INK4A) in MRT and DIPG. However, rather than targeting these key genes specifically, EZH2 inhibition leads to the reactivation of Polycomb target genes across the genome, which may alter tumor cell identity and promote tumor progression.

PRC2 is recruited to chromatin through CpG islands (CGIs). Insertion of CGIs into the genome is sufficient to induce PRC2 recruitment. The recruitment of PRC2 to CGIs is consistent with the binding of the accessory factors PHF1 (PCL1) and MTF2 (PCL2) to nonmethylated CpG DNA and the binding of JARID2 to H2AK119ub, deposited by PRC1. Although recognition of CGIs offers an explanation for the spatial pattern in which PRC2 is associated with the genome, this mechanism does not account for changes in PRC2 occupancy that occur during cell differentiation or during cell transformation. Knowledge of the mechanisms responsible for these dynamic patterns of PRC2 chromatin binding is necessary to understand how cell differentiation programs are regulated and may allow the development of methods to inhibit Polycomb activity at specific genes.

In addition to interacting with chromatin, PRC2 also binds RNA; however, the impact of this binding on PRC2 function remains unclear. Although first identified to bind specific noncoding RNAs (ncRNAs), UV cross-linking-based methods have revealed that PRC2 directly interacts with most nascent pre-mRNAs and nascent ncRNAs in embryonic stem cells (ESCs). Binding short RNA oligonucleotides in vitro, recombinant forms of PRC2 display a preference for repeated G tracts, especially when folded into G-quadruplex (G4) structures; however, the relevance of this for PRC2 RNA function in cells is unknown. Though it was first postulated that RNA promotes the recruitment of PRC2 to chromatin, it has recently been found that RNA blocks the interaction of PRC2 with nucleosomes and inhibits its methyltransferase activity. Potentially consistent with this, global inhibition of RNA polymerase II or global RNA degradation triggers PRC2 recruitment to chromatin at active genes in cells. Similarly, insertion of premature poly(A) signals or promoter or enhancer inactivation increases PRC2 binding and H3K27me3 in cis. However, it is unknown whether these results reflect loss of antagonistic RNA, loss of RNA polymerase II or depletion of antagonizing chromatin modifications, such as H3K4me3.

We hypothesized that nascent RNA plays a role in the temporal regulation of PRC2 occupancy at its target genes. Specifically, we
considered that nascent, chromatin-associated RNA may remove PRC2 from chromatin. To address this possibility, we sought to identify the RNA sequences preferentially bound by PRC2 in cells and determine the impact of these RNA elements on PRC2 occupancy at genes. Our results support a model in which chromatin-associated G-tract RNA evicts PRC2 from chromatin during gene activation and provides the means to remove PRC2 from specifically targeted genes. These data also support the broader consideration of nascent pre-mRNA as a regulatory molecule that modulates chromatin state at active genes.

**Results**

PRC2 binds G tracts within nascent RNA in cells. Using individual-nucleotide-resolution UV cross-linking and immunoprecipitation (iCLIP), we previously found that PRC2 directly interacts with most nascent pre-mRNA and ncRNA in mouse ESCs\(^5\). We sought to determine whether PRC2 favored any particular sequence within nascent transcripts. To ensure identified sequences were specific for PRC2, we also mapped background protein cross-link sites on input RNA. Comparing PRC2 and input RNA cross-link sites, we identified a strong enrichment of G tracts at PRC2 RNA binding sites (Fig. 1a and Supplementary Fig. 1a). Enrichment of these sequences was also observed at cross-link sites for the RNA binding protein FUS, as has been observed previously in mouse brain\(^5\), but was not apparent for HNRNPC, which binds poly(U) sequences\(^6\).

In vitro, the binding of PRC2 to G-tract sequences has been reported to increase when these RNAs are folded into a G4 structure\(^7\); our findings also confirmed this (Supplementary Fig. 1b–d). We therefore explored whether PRC2 maintained this binding preference in cells. Calculating the propensity for G4 formation across all genes using G4Hunter\(^8\) revealed a peak in predicted G4 formation 50 nucleotides (nt) into the first intron (Fig. 1b). This finding is consistent with previous reports of G-tract enrichment at the 5' end of introns\(^9,10\). Although PRC2 bound these sequences near the first 5' splice site, we did not observe any effects of PRC2 on splicing (Supplementary Fig. 2a).

To explore whether the potential for G4 formation increased PRC2 RNA binding, we measured PRC2 cross-link site density at the first 5' splice sites predicted to be able to form G4 structures versus those that were not. The first 5' splice sites predicted to form G4 structures showed significantly higher PRC2 RNA binding (Fig. 1c). This increased binding was localized at the site of predicted G4 formation (Fig. 1d and Supplementary Table 1). This increased binding was observed even when normalizing for G content (Supplementary Fig. 2c). The presence of a predicted G4 structure was also associated with increased FUS RNA binding at the beginning of the first intron, consistent with previous reports of FUS binding to G4 RNA in vitro\(^6\), but no change in HNRNPC binding (Fig. 1c). We found that PRC2 bound across the range of predicted G4 structures but showed a slight but significant preference for those

---

**Fig. 1** PRC2 binds G tracts with the potential to form G4 structures in nascent RNA. a. Enrichment of 8-mer sequences at PRC2, FUS and HNRNPC RNA cross-link sites identified by iCLIP (vs input controls). Gs per 8-mer are indicated by color. The ten 8-mers with the highest z scores are labeled. b. Average G4 prediction score (G4-forming sequences (G4FS)) for the coding (dark blue, above x axis) and noncoding (cyan, below x axis) strands around mouse gene splice sites. c. RNA cross-link density for PRC2, FUS, HNRNPC and their input controls at the set of first 5' splice sites that are predicted (red, n = 942) or not predicted (blue, n = 760) to be able to form G4 structures (PRC2, P < 2.2 × 10\(^{-16}\); FUS, P < 2.2 × 10\(^{-16}\), Wilcoxon rank-sum test). d. Left, heat map (blue) showing the position of sequences predicted to be able to form G4 structures −30 to +300 nt around the first 5' splice site of nascent RNAs expressed in mouse ESCs. Right, heat maps (red) showing the position of PRC2 and input RNA cross-link sites at the same 5' splice sites. The number of cross-link sites per 5-nt window is indicated by color.
G4 RNA blocks interaction of the PRC2 catalytic core with the substrate core nucleosome particle. We next sought to explore the basis for the antagonism between RNA and nucleosomes for PRC2 binding. We reasoned that because the PRC2 catalytic core has been reported to be competent for G4 RNA binding, G4 RNA may block the interaction of the PRC2 core with nucleosomes. We purified a recombinant catalytic core complex comprising EZH2, EED and the SUZ12 VEFS domain, and, using fluorescence anisotropy, found that it bound to an archetypal G4-forming RNA ((G4A4)₄); it also bound to the endogenous G4-forming sequence within PIM1 RNA in KCl, but with no substantial binding in LiCl buffer (Fig. 3a,b).

In order to measure PRC2 binding to the nucleosome, we engineered a nucleosome with a fluorescence-tagged histone H3 and 147 bp DNA. After confirming that robust binding to (G₄A₄)₄ and PIM1 G4 RNA was also observed in the low-salt nucleosome binding buffer required for the fluorescence assay (16.7 ± 1.2 nM and 22.5 ± 1.8 nM, respectively; Supplementary Fig. 4a,b), we then measured the effect of the RNA on the binding of the catalytic core to nucleosomes. In the absence of G4 RNA, the PRC2 catalytic core interacted with nucleosomes with high affinity (25.9 ± 10.7 nM); in the presence of 500 nM (G₄A₄)₄ RNA, PRC2 binding to the nucleosome was effectively blocked (Supplementary Fig. 4c).

The PRC2 catalytic core interacts with the substrate H3 tail through EZH2 and the K27-methylated H3 tail of the allostERIC nucleosome through EED. To specifically test the effect of G4 RNA on the binding of the PRC2 catalytic core to its substrate, we used nucleosomes containing a fluorescence-labeled H3K27M-modified tail, which engages EZH2 but not EED. In the absence of RNA, the PRC2 catalytic core interacted with this obligate substrate core nucleosome particle with high affinity (29.9 ± 3.9 nM). In the presence of (G₄A₄)₄ or PIM1 G4 RNA, the interaction was blocked, whereas a non-G4-forming portion of PIM1 RNA had no effect (Fig. 3c,d and Supplementary Fig. 4d). Given this antagonistic effect of G4 RNA on the interaction of the PRC2 core with its substrate, we considered that the RNA may also be able to displace PRC2 from the nucleosome. Strikingly, we found that both (G₄A₄)₄ and PIM1 G4 RNA blocks interaction of the PRC2 catalytic core with the nucleosome. Strikingly, we found that both (G₄A₄)₄ and PIM1 G4 RNA blocks interaction of the PRC2 catalytic core with the nucleosome.

**Fig. 2 | G4 structures within longer RNAs block PRC2 binding to nucleosomes. a,** Immunoblot for SUZ12 after pulldown of recombinant PRC2 (EZH2–SUZ12–EED–RBBP4 or RBBP7) with pre-folded biotinylated PIM1 RNA or control PIM1 RNA lacking the G4-forming sequence (ΔG4) in KCl- or LiCl-containing buffer. Streptavidin beads were incubated with 500, 50 or 5 ng/ul of RNA, washed, and then incubated with PRC2. Representative of three independent experiments (others shown in Supplementary Fig. 3d). b, Immunoblot for SUZ12, EZH2, JARID2 and ACTB after pulldown of PRC2 from ESC nuclear extract with ten-fold dilutions of biotinylated wild-type PIM1 RNA, ΔG4 RNA, G-to-H RNA (G4-forming G nucleotides mutated to non-G nucleotides) and G-rich RNA (G-to-H RNA with an equal number of non-G-to-G mutations outside of the G4-forming region). Representative of two independent experiments. c, Immunoblot for SUZ12 and H3 after pulldown of recombinant PRC2 with biotinylated nucleosomes (reconstituted with 185 bp DNA) in the presence of PIM1 or ΔG4 RNA (2, 20 or 200 ng/μl). Representative of three independent experiments (others shown in Supplementary Fig. 3e). d, Immunoblot for SUZ12, EZH2, JARID2, ACTB and H3 after pulldown of PRC2 from ESC nuclear extract with biotinylated nucleosomes in the presence of biotinylated wild-type PIM1, ΔG4, G-to-H or G-rich PIM1 RNA (2, 20 or 200 ng/μl). Representative of two independent experiments. Uncropped blot images are shown in Supplementary Data Set 1.
G4 RNA, but neither control non-G4 PIM1 RNA nor poly(A) RNA, was also able to remove PRC2 from a pre-formed core-PRC2–substrate nucleosome complex (Fig. 3e,f).

To validate these findings, we measured the effect of RNA degradation on the binding of endogenous PRC2 in nuclear extract to wild-type mononucleosomes either lacking linker DNA (reconstituted with 147 bp DNA) or containing linker DNA (reconstituted with 183 bp DNA). We found that RNA depletion increased PRC2 binding to nucleosomes independently of linker DNA and independently of the DNA-binding accessory factors MTF2, AEBP2 and JARID2 (Fig. 3g and Supplementary Fig. 4e,f). Together, these data show that G4 RNA evicts PRC2 from the substrate core nucleosome particle via interactions with the PRC2 catalytic core, independently of accessory factors.

Chromatin-associated G-tract RNA removes PRC2 from genes. The binding of PRC2 to G tracts within nascent pre-mRNA in cells and the ability of G4 RNA to evict PRC2 from nucleosomes suggested that G tracts within nascent RNAs remove PRC2 from chromatin at genes. We considered that if this hypothesis is correct, then mimicking chromatin-associated, nascent RNA by tethering G-tract RNA to the 5′-end of genes with dCas9 should remove PRC2 from chromatin (Fig. 4a). To test this concept, we generated a doxycycline (dox)-inducible HA-dCas9 NIH-3T3 cell line and co-expressed short guide RNA (sgRNA) to recruit dCas9 to the first intron of the PRC2 target gene Fgf11 (Fig. 4b and Supplementary Fig. 5a,b). To the 3′ end of the sgRNA-155, we appended a 220-nt sequence composed of repeated G tracts; an equal-length sequence and co-expressed short guide RNA (sgRNA) to recruit dCas9 to the first intron of the PRC2 target gene Fgf11 (Fig. 4b and Supplementary Fig. 5a,b). To the 3′ end of the sgRNA-155, we appended a 220-nt sequence composed of repeated G tracts; an equal-length sequence with the same overall G content (50%) but lacking sequential sequence composed of repeated G tracts; an equal-length sequence (Supplementary Fig. 1a). We then performed chromatin immunoprecipitation (ChIP) for HA-dCas9, SUZ12, H3K27me3, total H3 and non-specific IgG control, before and after induction of dCas9 (Fig. 4b and Supplementary Fig. 5c). As predicted, dCas9 induction led to specific recruitment of the dCas9-G-tract RNA, dCas9-G-rich RNA, and the dCas9-A-tract RNA ribonucleoproteins to Fgf11. dCas9-tethered G-tract RNA significantly reduced PRC2 binding and H3K27me3 at Fgf11 but not at other genes. No change was observed in total histone H3 occupancy. In contrast, dCas9-tethered G-rich or A-tract RNAs had no effect on PRC2 chromatin binding or H3K27me3. The loss of PRC2 occupancy was not caused indirectly by induction of Fgf11 transcription (Supplementary Fig. 5d). We conclude that chromatin-associated G-tract RNA is sufficient to remove PRC2 and deplete H3K27me3 from genes.

We sought to determine whether the effect of chromatin-associated G-tract RNA was specific to PRC2. No changes were observed in the levels of H3K27ac or H2AK119ub (Fig. 4b and Supplementary Fig. 5c). As predicted, dCas9 induction led to specific recruitment of the dCas9-G-tract RNA, dCas9-G-rich RNA, and the dCas9-A-tract RNA ribonucleoproteins to Fgf11. dCas9-tethered G-tract RNA significantly reduced PRC2 binding and H3K27me3 at Fgf11 but not at other genes. No change was observed in total histone H3 occupancy. In contrast, dCas9-tethered G-rich or A-tract RNAs had no effect on PRC2 chromatin binding or H3K27me3. The loss of PRC2 occupancy was not caused indirectly by induction of Fgf11 transcription (Supplementary Fig. 5d). We conclude that chromatin-associated G-tract RNA is sufficient to remove PRC2 and deplete H3K27me3 from genes.

G4 RNA, but neither control non-G4 PIM1 RNA nor poly(A) RNA, was also able to remove PRC2 from a pre-formed core-PRC2–substrate nucleosome complex (Fig. 3e,f).

To validate these findings, we measured the effect of RNA degradation on the binding of endogenous PRC2 in nuclear extract to wild-type mononucleosomes either lacking linker DNA (reconstituted with 147 bp DNA) or containing linker DNA (reconstituted with 183 bp DNA). We found that RNA depletion increased PRC2 binding to nucleosomes independently of linker DNA and independently of the DNA-binding accessory factors MTF2, AEBP2 and JARID2 (Fig. 3g and Supplementary Fig. 4e,f). Together, these data show that G4 RNA evicts PRC2 from the substrate core nucleosome particle via interactions with the PRC2 catalytic core, independently of accessory factors.

PRC2 transfers from chromatin to RNA upon gene activation. The transfer of PRC2 from chromatin to chromatin-associated G-tract RNA suggested that PRC2 also transfers from chromatin to nascent pre-mRNA upon gene activation (Fig. 5a). Cell transformation induced by oncogenic HRasV12 is accompanied by dynamic changes in PRC2 chromatin occupancy37,40,41. Notably, expression of HRasV12 leads to activation of Adcy7 and Sorcs2 and the subsequent loss of PRC2 from these genes (ref. 42). Loss of PRC2 from Sorcs2 is dependent on the Sorcs2 transcription start site43, suggesting for the nascent pre-mRNA. Consistent with this hypothesis, we found that activation of Adcy7 and Sorcs2 downstream of HRasV12 was accompanied by a change in PRC2 from binding chromatin to binding the pre-mRNA (Fig. 5b,c and Supplementary Fig. 6a–c). We next tested whether chromatin-associated G-tract RNA recapitulated the effect of gene activation on PRC2 binding at these genes. Tethering G-tract RNA, but not A-tract RNA, to Adcy7 reduced PRC2 binding and H3K27me3 at this gene but had no effect on PRC2 occupancy at Sorcs2 (Fig. 5d and Supplementary Fig. 6d). Reciprocally, tethering G-tract RNA to Sorcs2 reduced PRC2 occupancy and H3K27me3 at this gene but had no effect on Adcy7 (Fig. 5e and Supplementary Fig. 6e). Thus, PRC2 transfers from chromatin to nascent pre-mRNA upon gene activation, and this transfer can be recapitulated by tethering G-tract RNA to genes.

G-tract RNA reverses ectopic recruitment of PRC2 triggered by oncogenic HRas. Cell transformation mediated by oncogenes such as HRasV12 causes changes in PRC2 association with chromatin, including ectopic recruitment to specific genes. PRC2 activity can be inhibited in cancer cells with small molecules but this reactivates PRC2 target genes non-specifically. We postulated that G-tract RNA tethering would instead allow the specific reversal of deleterious PRC2 recruitment events (Fig. 6a). HRasV12-mediated recruitment of PRC2 to Smad6 is necessary for Ras-induced senescence44 and is dependent on transcriptional repression45. Consistent with this dependence on transcription repression reflecting loss of the competing nascent pre-mRNA, we found that expression of HRasV12 resulted in a switch in PRC2 binding from nascent pre-mRNA to chromatin at Smad6 (Fig. 6b). We next asked whether chromatin-associated G-tract RNA could reverse this recruitment of PRC2 to chromatin at Smad6 in HRasV12-expressing cells. We found that tethering G-tract RNA to Smad6 countered HRasV12-mediated PRC2 recruitment and reduced H3K27me3 at the gene (Fig. 6c and Supplementary Fig. 6f). As we had found for the other genes tested, G-tract RNA tethering and the resultant PRC2 loss was not sufficient to activate Smad6 transcription (Supplementary Fig. 6g). We conclude that G-tract RNA tethering allows the reversal of oncogene-mediated PRC2 recruitment events.

G-tract RNA tethering activates the tumor suppressor gene Cdkn2a in MRT cells. PRC2 silences tumor suppressor genes in a number of cancer types, including Cdkn2a (p16INK4a) in MRT and Dipy-V61L46,47. We therefore sought to determine the effect of tethering G-tract RNA to this gene in MRT cells. We found that recruitment of G-tract RNA, but not A-tract RNA, to Cdkn2a caused loss of PRC2 and H3K27me3 from Cdkn2a but had no effect on Evx2 (Fig. 7a and Supplementary Fig. 7a, b). Strikingly, this loss of PRC2...
Fig. 3 | G4 RNA inhibits interaction of the PRC2 catalytic core with the substrate core nucleosome particle. a, Fluorescence anisotropy measuring binding of the PRC2 catalytic core (EZH2−EED−SUZ12 VEFS domain) directly to fluorescein-labeled (G₄A₄)₄ RNA in either 100 mM K⁺ or Li⁺ buffer (mean and s.e.m., n = 3 independent experiments). b, Same as a, but with the G4-forming sequence within PIM1 RNA. c, Fluorescence intensity measuring binding of the PRC2 catalytic core directly to MDCC-labeled H3K27M obligate substrate core nucleosome particles (reconstituted with 147 bp DNA) in the presence of 500 nM (G₄A₄)₄ RNA or no RNA (mean and s.e.m., n = 3 independent experiments). d, Same as c but with 500 nM PIM1 G4 RNA or a control non-G4-forming 25-nt portion of PIM1 RNA. e, Titration of (G₄A₄)₄ and control A₄₀ RNA into a pre-formed complex of core PRC2 and MDCC-labeled substrate core nucleosome particle. The increase in fluorescence intensity with (G₄A₄)₄ RNA is interpreted as release of PRC2 from the nucleosome (mean and s.e.m., n = 3 independent experiments). f, Same as e, but with G4 and non-G4-forming PIM1 RNAs. g, Immunoblot for SUZ12, MTF2, HMGN1 and H3 after co-immunoprecipitation of PRC2 from mock or RNaseA-treated Mtf2GT/GT or Mtf2WT/WT ESC nuclear extract with nucleosomes containing biotin-tagged histone H2A (reconstituted with either 183 bp or 147 bp DNA). Representative of two independent experiments. Uncropped blot images are shown in Supplementary Data Set 1.
Fig. 4 | Chromatin-associated G-tract RNA removes PRC2 from specific genes in cells. a, Hypothesis: G-tract RNA, tethered to chromatin with dCas9, should compete with CGI chromatin for PRC2, reducing H3K27me3. The same-length RNA that is equally G-rich but lacking G tracts or RNA in which the G tracts are replaced with A tracts, both of which bind PRC2 only weakly, should both have no effect. b, Top, position of the Fgf11 sgRNA and primer pairs A and B. Bottom, change in HA-dCas9, SUZ12, H3K27me3 and total H3 occupancy at Fgf11 and Pax7 measured by ChIP-qPCR after dox-mediated induction of HA-dCas9 expression in cells containing the Fgf11 sgRNA, to which G-tract, G-rich or A-tract RNA is appended (mean and s.d., n = 3 independent dox inductions, no significant changes, Welch’s one-tailed t test). c, Change in H2A.K119ub, H3K27Ac and total H3 at Fgf11 and Pax7 before and after incubation with dox (mean and s.d., n = 3 independent dox inductions, no significant changes, Welch’s one-tailed t test). d, Change in HA-dCas9, SUZ12, H3K27me3 and total H3 occupancy at Fgf11 and Pax7 before and after dox treatment (day 6) and after subsequent dox washout (day 12) (mean and s.d., n = 3 independent dox inductions). Dose induction Fgf11-A G-tract RNA: SUZ12, P = 0.0018; H3K27me3, P = 0.14. Fgf11-B G-rich RNA: SUZ12, P = 0.0052; H3K27me3, P = 0.03. Fgf11-B G-rich RNA: SUZ12, P = 0.03, Welch’s one-tailed t test). e, Top, Fgf11 RNA sequence spanning the first exon–intron junction was appended to Fgf11 sgRNA. Bottom, same as b, but using Fgf11 sgRNA to which the Fgf11 RNA sequence has been appended (mean and s.d., n = 3 independent dox inductions. Fgf11 A: SUZ12, P = 2.9 × 10⁻⁴; H3K27me3, P = 0.0019. Fgf11 B: SUZ12, P = 0.045; H3K27me3, P = 0.026, Welch’s one-tailed t test).
was sufficient to activate CDKN2A and increase p16INK4a protein levels to a similar extent to the chemotherapeutic agent cisplatin and the EZH2 inhibitor E11 (Fig. 7b,c). Furthermore, CDNK2A upregulation was mirrored by an increase in the proportion of senescent cells (Fig. 7d and Supplementary Fig. 7c). We conclude that G-tract RNA tethering can be used to reverse Polycomb-mediated silencing of specific tumor suppressor genes in cancer cells.

Discussion

Current models of how PRC2 interacts with chromatin provide an explanation for the spatial distribution of PRC2 within the genome, but they do not account for the changes in PRC2 gene occupancy that occur during cell differentiation or in cancer. We have discovered that chromatin-associated G-tract RNA removes PRC2 from its target genes. We found that PRC2 directly binds G tracts within nascent RNAs, especially those at the first 5′ splice site predicted to form G4 structures. G4 RNA binds to the PRC2 catalytic core and antagonizes its interaction with the substrate core nucleosome particle. Consistent with these results, PRC2 is transferred from chromatin to nascent pre-mRNA during gene activation, and chromatin-associated G-tract RNA is sufficient to remove PRC2 from chromatin and deplete H3K27me3. These data support a model in which G tracts within nascent RNA remove PRC2 from chromatin during activation of Polycomb-repressed genes. We also demonstrate that this mechanism can be exploited to allow the targeted removal of PRC2 from tumor suppressor genes in cancer cells.

The mechanisms responsible for the removal of PRC2 from chromatin have been unclear. A number of recent studies have demonstrated that PRC2 recruitment is responsive to the activation state of the gene. Inhibition of RNA polymerase II or insertion of premature poly(A) signals triggers PRC2 recruitment to CGIs at active genes.31,41 Similarly, the removal of PRC2 from genes during transcriptional activation of Polycomb-repressed genes is dependent on their transcriptional activity.41 Furthermore, PRC2 is recruited to CGIs inserted into the genome, but not if they contain binding sites for transcriptional activators present in the cell or if they are positioned between an active promoter—enhancer pair.41 These results show that PRC2 is only able to stably associate with chromatin in
the absence of transcriptional activity. Based on experiments showing that global RNA degradation triggers PRC2 recruitment to transcribed genes and that RNA inhibits PRC2 nucleosome interaction and methyltransferase activity\textsuperscript{31,33,36–39}, we and others have suggested that one of the features of active genes that inhibits PRC2 function is the nascent pre-mRNA itself. Our results support this model, demonstrating that chromatin-associated RNA can prevent PRC2 recruitment to active genes and, in addition, that G-tract RNA removes stably associated PRC2 from genes.

Our data clarifies the nature of PRC2 RNA binding specificity and provides an explanation of its function. PRC2 RNA binding activity was first identified through its association with specific ncRNAs. Systematic measurement of direct RNA binding in cells later revealed that PRC2 binds the majority of nascent pre-RNAs and ncRNAs in a promiscuous manner\textsuperscript{11}, a conclusion also drawn from lower-stringency native RIP experiments\textsuperscript{54}. Although PRC2 was initially observed to bind a broad range of RNAs in vitro\textsuperscript{11}, later studies using more homogenous short oligonucleotides revealed specificity for repetitive G-tract sequences, especially when folded into G4 structures\textsuperscript{34,35}. Potentially consistent with this finding, it was also reported that G-tract sequences were enriched in RNA that co-precipitated with EZH2, but not with SUZ12, from formaldehyde-cross-linked HeLa cells\textsuperscript{36}. However, questions remained regarding the discordant results between EZH2 and SUZ12, whether the detected interactions were direct or indirect, the inability of the method to distinguish RNA strandedness, whether G-tract sequences were the most enriched sequences at PRC2 RNA binding sites and the locations of these sequences within RNA. Our measurements of PRC2 RNA cross-linking in cells at single-nucleotide resolution reveals that G-tract sequences with the potential to form G4 structures are the preferred RNA binding sites for PRC2 in cells and that these are predominantly localized just downstream of the first 5’ splice site. The concentration of these sequences at the 5’ end of nascent RNAs, in proximity to the site of PRC2 binding on chromatin, may aid the removal of PRC2 from genes. However, although PRC2 displays a preference for G-tract RNA, it can bind other RNA sequences in vitro and in cells; thus, other nascent RNA elements may also be able to remove PRC2 from chromatin, albeit less efficiently. Additional studies will be required to determine whether G-tract sequences located near the 5’ end of RNA are required for the removal of PRC2 from chromatin. Such experiments will need to avoid disrupting the functions of G-tract sequences in splicing\textsuperscript{55,56} and the function of the corresponding DNA sequences as transcription factor and PRC2 and PRC1 binding sites within CGI promoters.

PRC2, augmented by the accessory subunits PHF1, MTF2 or PHF19 (in PRC2.1) or JARID2 and AEBP2 (in PRC2.2), forms multivalient interactions with the nucleosome core, modified histone tails and DNA\textsuperscript{1}. Recent structural analysis showed the details of the interaction between the catalytic EZH2 SET domain and the substrate nucleosome and between EED and the K27-methylated nucleosome\textsuperscript{3}; the non-catalytic lobe of PRC2 has recently been shown to cooperate with AEBP2 and JARID2 to form a further nucleosome interaction surface\textsuperscript{5,6}. Consistent with previous results\textsuperscript{1,39}, we found that the minimal catalytically active PRC2 core (EZH2, EED and the SUZ12 VEFS domain) binds RNA and does so preferentially in conditions favoring G4 formation. We therefore focused on potential antagonism between G4 RNA and the substrate core nucleosome particle for binding to the PRC2 catalytic core. Using a well-defined system consisting of the PRC2 catalytic core and an obligate substrate nucleosome reconstituted with 147-bp DNA, we found that G4 RNA blocks the binding of the PRC2 catalytic core to the substrate core nucleosome particle. Notably, titration of G4 RNA disrupted a pre-formed complex of the PRC2 catalytic core and the substrate nucleosome, which is consistent with our finding that chromatin-associated G-tract RNA evicts PRC2 from chromatin in cells. The competitive effect of RNA on PRC2 nucleosome binding in nuclear extracts was also unaffected by the absence of PRC2 accessory factors. These experiments demonstrate that G4 RNA blocks a fundamental aspect of PRC2 function that is common to both PRC2.1 and PRC2.2. Other PRC2 RNA binding surfaces have been identified in JARID2 (refs. \textsuperscript{34,37,39}) and AEBP2.
nascent pre-mRNA is not merely a passive intermediary; it plays a role in positively acting regulators with chromatin, including the transcription factor YY1 (ref. 52) and the histone methyltransferases Set1 and Set2 (ref. 53). Other studies have also demonstrated a role for nascent pre-mRNA in countering the function of negative-acting chromatin modification. For example, nascent pre-mRNA interacts with DNMT1, and RNA blocks DNMT1 activity (ref. 60). On the other hand, nascent pre-mRNA promotes the interaction of positively acting regulators with chromatin, including the transcription factor YY1 (ref. 61) and the histone methyltransferases Set1 and Set2 (ref. 53). Unspliced, chromatin-associated RNA has also been found to promote HNRNPU oligomerization and chromatin decompaction (ref. 62). Together with the data shown here, these studies argue that nascent pre-mRNA is not merely a passive intermediary; it plays a direct role in altering chromatin state to promote its own production. It will be important to determine the contribution of G-tract sequences and G4 structures to these other regulatory roles of nascent RNA on chromatin.

By showing that tethered G-tract RNA removes PRC2 from chromatin, we have discovered a means to selectively remove PRC2 from chromatin and deplete repressive chromatin regulators. For example, nascent pre-mRNA interacts with DNMT1, and RNA blocks DNMT1 activity (ref. 60). On the other hand, nascent pre-mRNA promotes the interaction of positively acting regulators with chromatin, including the transcription factor YY1 (ref. 61) and the histone methyltransferases Set1 and Set2 (ref. 53). Unspliced, chromatin-associated RNA has also been found to promote HNRNPU oligomerization and chromatin decompaction (ref. 62). Together with the data shown here, these studies argue that nascent pre-mRNA is not merely a passive intermediary; it plays a direct role in altering chromatin state to promote its own production. It will be important to determine the contribution of G-tract sequences and G4 structures to these other regulatory roles of nascent RNA on chromatin.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41594-019-0293-z.

Received: 30 November 2018; Accepted: 5 August 2019; Published online: 23 September 2019

References
1. Shalaska, L., Beltran-Nebot, M., Ule, J. & Jenner, R. G. Regulatory feedback from nascent RNA to chromatin and transcription. Nat. Rev. Mol. Cell Biol. 18, 331–337 (2017).
2. Holoch, D. & Margueron, R. Mechanisms regulating PRC2 recruitment and enzymatic activity. *Trends Biochem. Sci.* 42, 531–542 (2017).
3. Schuettengruber, B., Bourdon, H. M., Di Croce, L. & Cavalli, G. Genome regulation by polycomb and trithorax: 70 years and counting. *Cell* 171, 34–57 (2017).
4. Comet, I., Rüising, E. M., Leblanc, B. & Helin, K. Maintaining cell identity: PRC2-mediated regulation of transcription and cancer. *Nat. Rev. Cancer* 16, 803–810 (2016).
5. Qi, W. et al. Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. *Proc. Natl Acad. Sci. USA* 109, 21360–21365 (2012).
6. Knutson, S. K. et al. Durable tumor regression in genetically altered malignant rhahdoid tumors by inhibition of methyltransferase Ezh2. *Proc. Natl Acad. Sci. USA* 110, 7922–7927 (2013).
7. Caganova, M. et al. Germinal center dysregulation by histone methyltransferase Ezh2 promotes lymphomagenesis. *J. Clin. Invest.* 123, 5009–5022 (2013).
8. Beguelin, W. et al. Ezh2 is required for germinal center formation and somatic Ezh2 mutations promote lymphoid transformation. *Cancer Cell* 23, 677–692 (2013).
9. McCabe, M. T. et al. Ezh2 inhibition as a therapeutic strategy for lymphoma with Ezh2-activating mutations. *Nature* 492, 108–112 (2012).
10. Mohd., F. et al. Ezh2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. *Nat. Med.* 23, 483–492 (2017).
11. Boyer, L. A. et al. Polycomb complexes repress developmental regulators in mammalian embryonic stem cells. *Nature* 441, 345–353 (2006).
12. Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H. & Helin, K. Genome-wide maps of chromatin state in pluripotent mammalian ES cells. *Genome Res.* 23, 2838–2849 (2013).
13. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent mammalian ES cells. *Science* 328, 92–96 (2010).
14. Mohn, F. et al. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Cell Reports* 5, 755–766 (2013).
15. Agger, K. et al. The H3K27me3 demethylase JMJD3 contributes to the activation of the INK4A-ARF locus in response to oncogene- and stress-induced senescence. *Genes Dev.* 23, 1171–1176 (2009).
16. Barradas, M. et al. Histone demethylase JMJD3 contributes to epigenetic control of INK4a/ARF by oncogenic Ras. *Genes Dev.* 23, 1177–1182 (2009).
17. Berg, T. et al. A transgenic mouse model demonstrating the oncogenic role of mutations in the polycomb-group gene Ezh2 in lymphomagenesis. *Blood* 123, 3914–3924 (2014).
18. Hosogane, M., Funayama, R., Nishida, Y., Nagashima, T. & Nakayama, K. Lack of transcriptional repression due to Ezh2 mutations is also associated with mesenchymal transition of INK4b-ARFINK4a locus. *Mol. Cell Biol.* 20, 1341–1350 (2013).
19. Kwon, C. K., Marsico, G., Sahakyan, A. B., Chambers, V. S. & Balasubramanian, S. rG4-seq reveals widespread formation of G-quadruplex telomere DNA and TERRA-binding protein TLS/FLJ. *Cell. Biol.* 20, 340–350 (2013).
20. Agger, K. et al. Activation of Bmp2-Smad1 signal and its regulation by histone methylation. *Mol. Cell Biol.* 33, 628–640 (2013).
21. Sourovialas, G. P. et al. An oncogenic Ezh2 mutation induces tumors through global redistribution of histone 3 lysine 27 trimethylation. *Blood* 123, 3914–3924 (2014).
22. de Vries, N. A. et al. Prolonged Ezh2 depletion in glioblastoma causes a robust switch in cell fate resulting in tumor progression. *Cell Reports* 10, 383–395 (2017).
23. Jermann, P., Hoerner, L., Burger, L. & Schubeler, D. Short sequences can efficiently recruit histone h3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. *Proc. Natl Acad. Sci. USA* 111, 13915–13921 (2014).
24. Lynch, M. D. et al. An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate polycomb complex recruitment. *EMBO J.* 31, 317–329 (2011).
25. Mendenhall, E. M. et al. GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet.* 6, e1001239 (2010).
26. Li, H. et al. Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature* 549, 287–291 (2017).
27. Perino, M. et al. MTF2 recruits polycomb repressive complex 2 by helical shape-selective DNA binding. *Nat. Genet.* 50, 1002–1010 (2018).
28. Cooper, S. et al. Jarid2 binds mono-ubiquitylated H2A lysine 11 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. *Nat. Commun.* 7, 13661 (2016).
29. Kalb, R. et al. Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat. Struct. Mol. Biol.* 21, 569–571 (2014).
30. Beltran, M. et al. The interaction of PRC2 with RNA or chromatin is mutually antagonistic. *Genome Res.* 26, 896–907 (2016).
63. Dominguez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. Nat. Rev. Mol. Cell Biol. 17, 5–15 (2016).

Acknowledgements
We thank the UCL Cancer Institute Genomics Core Facility and Bill Lyons Informatics Centre, both supported by the Cancer Research UK-UCL Centre (award C416/A18088). We thank A. Bracken (Trinity College Dublin), N. Brockdorff (University of Oxford), A. Fisher (London Institute for Medical Sciences) and B. Vanhaesebroeck (UCL) for cell lines. We also thank I. Ruiz de los Mozos and J. Ule for assistance with iCount and feedback on the manuscript and to M. Vila de Mucha for assistance with flow cytometry. The research was funded by grants from the European Research Council (ERC, 311704), Worldwide Cancer Research (13-0256) and Bloodwise (18008) to R.G.J., CoNaCyT (411064) to M.T., ERC (309952) and the Helmholtz Society to T.B., and Cancer Research UK (FC001078), Medical Research Council (FC001078) and Wellcome Trust (FC001078) grants to the Francis Crick Institute (funding N.J., S.K., S.J.G. and J.R.W.).

Author contributions
M.B. co-designed and performed all experiments, except where noted below. M.T. performed the nucleosome IPs with different linker DNA lengths. N.J., assisted by S.K., measured competition between G4 RNA and the substrate core nucleosome particle for the PRC2 catalytic core in experiments co-designed by J.R.W. G.K. performed bioinformatics analysis, assisted by J.A. and R.G.J. K.B.W. helped with qRT–PCR experiments. B.M.F. and A.T. produced nucleosomes. J.H., T.B., S.J.G., J.R.W. and R.G.J. supervised the research. R.G.J. co-designed experiments and wrote the paper with help from all authors.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41594-019-0293-z.

Correspondence and requests for materials should be addressed to R.G.J.

Peer review information Anke Sparmann was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019
RNA tethering. Tethered sequences were placed at the 3′ end of the sgRNA sequence (taken from plko.1-puro U6 sgRNA), separated by a spacer. The sequences were synthesized as gBlocks (Integrated DNA Technologies (IDT)) that also comprised BsuAI-stuffer (taken from plko.1-puro U6 sgRNA) and a Pol III T terminator (sequences in Supplementary Table 2). The gBlocks were digested with AgeI and EcoRI (New England Biolabs) and ligated into plko.1-puro U6 sgRNA BsuAI stuffer (Addgene plasmid #50920, a gift from R. Maehr and S. Wolfe). The sgRNA targeting sequence for Sorcs2 was previously described. 5 Other sgRNAs were designed using CHOPCHOPv2 (ref. 5) (sequences in Supplementary Table 2), synthesized as oligonucleotides, annealed, and inserted into the vector using the BsuAI site.

A G-rich sequence with the same G content as the tethered G-tract sequence, appended to the Fgf11 sgRNA and a 5′ spacer sequence, was ordered as a gBlock (sequence in Supplementary Table 2), digested with NdeI and EcoRI and cloned into plko.1-puro U6 sgRNA BsuAI stuffer.

Input iCLIP. iCLIP data for PRC2 (antibody to SUZ12), FUS and HNRNPC CLIPs as a washout experiment, fresh media was added on day 6, changed every 2 days, until day 12. When indicated, cells were treated with cisplatin (Sigma) at 3.3 μM for 24 h or with E11 (Geron) 10 μM for 6 d.

RNA quantification. RNA was purified using TRIzol (Bioline), treated with Turbo DNase (Thermo Fisher Scientific) for 30 min at 37 °C and reverse transcribed using SuperScript III (Life Technologies) and random hexamer primers. Specific RNAs were quantified via qPCR (Applied Biosystems) using QuantScript SYBR Green PCR Kit (Qing) with the primers shown in Supplementary Table 2.

NATuRe STRuCTuRAl & MoleCulAR BioloGy | www.nature.com/nsmb

Methods
Cell culture. E14, ES99, A2058, and A2058#7 (kind gifts from N. Brockdorff) and Jar50E75 (gift from A. Fisher) mouse ES cells were maintained on 0.1% gelatin in KDMEM, 10% FCS, 5% knockout serum replacement, non-essential amino acids, L-glutamine, 2-mercaptoethanol, penicillin-streptomycin and 1000 U/ml leukemia inhibitory factor (03-0011-100, Stemgent). Mfg2ΔGT and Mfg2ΔTUTW (gifts from A. Bracken) were maintained in GMEM with the same supplements, except with no serum replacement and replacing L-glutamine with GlutaMAX. G-401 cells were acquired from Sigma with certification from the European Collection of Authenticated Cell Cultures (ECACC) and maintained in McCoy 5A media supplemented with 10% FBS, 2 mM L-glutamine and penicillin-streptomycin. NIH-3T3 cells (gift from B. Vanhaesebroeck) were cultured in DMEM, 10% FBS and penicillin-streptomycin. All cell lines were tested negative for mycoplasma. A NIH-3T3 cell line expressing H-RasV12 was generated by transfecting pWZL hygro H-Ras V12 (gift from S. Lowe, Addgene plasmid #1874 (ref. 61)) with Eugene HD (Promega) and selection in hygromycin (2 μg/ml). 5′ DNA strands for RNA tethering, cells were transfected with pHAGE TRE dCas9 (Addgene plasmid #50915, a gift from R. Maehr and S. Wolfe) and selected with 2 μg/ml G418. The dCas9 cell lines was then transfected as before with plko.1-puro U6 sgRNA constructs and selected with puromycin at 1 μg/ml dCas9 expression was induced using doxycycline (2 μg/ml) for 6 days (with media changed every 2 days). For RNA tethering experiments, fresh media was added on day 6, changed every 2 days, until day 12. When indicated, cells were treated with cisplatin (Sigma) at 3.3 μM for 24 h or with E11 (Geron) 10 μM for 6 d.

Chromatin immunoprecipitation. Cells were trypsinized from the plates, washed with PBS and cross-linked with 1% formaldehyde for 20 min. ChIP was performed as described previously30 except cells were sonicated for five cycles for NIH-3T3 cells or four cycles for G-401 cells (30 s on, 30 s off) using a Diagnen Picoruptor. ChIP was performed using antibodies to SUZ12 (Cell Signaling #3737), HA-ducR9 (3F10, Roche 11867423001), H3K27me3 (Abcam ab6002 or ab192985), H3K27ac (ab4729), H2AK119ub (Cell Signaling #8240s), total H3 (Abcam #ab71913) or specific IgG control (Abcam ab4540). Enrichment of specific genomic regions was measured relative to input DNA by qPCR (Applied Biosystems) using QuantScript SYBR Green PCR Kit (Qing) with the primers shown in Supplementary Table 2.

Histone methyltransferase assays. 30 nM PRC2 (EZH2–SUZ12–EED–RBBP4 or RBBP7; Active Motif 31387) was incubated in the presence of 0.8 μM nucleosomes in 20 mM HEPES, pH 7.9, 150 mM NaCl/KCl/LiCl, 20% glycerol 0.05% IGEPAL CA-630, 0.25 mM EDTA 1 mM DTT, 320 μM SAM and Cm protease inhibitor for 30 min at 25 °C.

RNAs for binding experiments. G4-forming PIM1 and control AG4 sequences were taken from ref. 61. Two additional control RNAs, one for which the Gs within the G4-forming sequence were mutated to non-Gs (G-to-H) and a second for which the Gs within the G4-forming sequence were mutated to non-Gs and an equal number of non-G nucleotides outside of the G4-forming sequence were replaced to Gs (G-rich), were synthesized as gBlocks (IDT, sequences in Supplementary Table 2) and cloned into pCNDNA3.1. Linearized vectors were transfected using the MAXscript T7 Transcription Kit (Thermo Fisher Scientific), and RNA was treated with Turbo DNase (Thermo Fisher Scientific). Biotin-14-CTP (159916 Life Technologies) was added in a 0.4:1 ratio relative to cTP RNA integrity was verified by polyacrylamide gel electrophoresis. G4 structure formation was confirmed using a reverse transcription assay66. 5′-biotinylated-(Gρ4rA4) 40-mer RNA oligonucleotides were obtained from IDT. Native gel electrophoresis to measure formation of secondary structure was performed as described previously14. RNA was folded either as described previously79 or in pulldown buffer to confirm maintenance of RNA structure during PCR2 pulldown assays. Radiolabeled RNA was visualized using a Typhoon phosphorimager (GE) and ImageQuantTL (GE).

RNA pulldowns. Biotinylated RNA was incubated in pulldown buffer containing 10 mM HEPES, pH 7.9, 150 mM KCl/LiCl or 0.25 mM EDTA, pH 8.0, 1 mM DTT, 5% glycerol, 0.05% IGEPAL CA-630, 35 ng/ml BSA, RNaseOUT (Invitrogen) and Cm protease inhibitor and G4 formation promoted by heating to 95 °C before cooling on ice and incubation at 37 °C for 30 min. 50, 50 or 5 ng/μl folded biotinylated-RNA was bound to MyOne Streptavidin T1 Dynabeads (Thermo Fisher Scientific) for 1 h at 4 °C, washed and then incubated with 1.5 ng/ml of recombinant PRC2 (Active Motif 31387) for 3 h at 4 °C. Beads were washed 3× with binding buffer and then resuspended in NuPAGE loading buffer. In vitro–transcribed biotinylated PIM1 or PIM1Δ-G4 RNA were folded in pulldown buffer containing NaCl and G4 formation was promoted as described above. RNA was bound to MyOne Streptavidin T1 Dynabeads, added to 0.1 μg/ml ESC nuclear extract, prepared as described previously81, and the pulldown was allowed to proceed as described above.

Nucleosome pulldowns. Recombinant human histones were expressed in Escherichia coli and purified as described previously66. Nucleosomes were assembled by salt deposition dialysis using a biotinylated 601 sequence—containing a 185 bp DNA fragment, as described previously14. 50 nM of nucleosomes was incubated with 1.5 ng/ml recombinant PRC2 (Active Motif 31387), 10 μM MyOne Streptavidin T1 Dynabeads (Thermo Scientific) and 200, 20 or 2 ng/μl of pre-folded RNA, in pulldown buffer (10 mM HEPES, pH 7.9, 150 mM LiCl or KCl, 0.25 mM EDTA, pH 8.0, 1 mM DTT, 5% glycerol, 0.05%, IGEPAL CA-630, 320 μM SAM, and Cm protease inhibitor) for 3 h at 4 °C. Beads were washed three times with 4X in LiCl or KCl pulldown buffer supplemented with 1 M urea. For nucleosomes pulldown using nuclear extract, 50 nM nucleosomes was incubated with 0.2 μg/ml ESC nuclear extract, in nucleosome pulldown buffer containing NaCl instead of KCl or LiCl.

When measuring the effect of linker DNA, in order to ensure pulldown of PRC2 binding to intact nucleosomes and not to any potential free DNA, we used nucleosomes pulldown buffer containing NaCl in 200 μl of buffer (100 mM Tris-Cl, pH 7.4, 50 mM NaCl, 10 mM EDTA) with 10 μM proteinase K (Roche 03115828001) for 20 min at 1.100 rpm. and 37 °C. An equal volume of PK buffer-containing 7 M urea was added, and a second incubation was performed. Supernatant was collected, and RNA was purified via phenol–chloroform extraction.
Immunoblotting. Immunoblotting was performed for SUZ12 (Santa Cruz sc-462634), EZH2 (CST 3147), JARID2 (CST 13594), AEBP2 (CST 14112), MTF2 (Proteintech 16208-1-AP), p16INK4a (Santa Cruz sc-63530), ACTB (CST 4967), HMGN1 (Bethyl Laboratories A307-3623A), HRAS G12V (DH21L, CST 13412), H3K27me1 (Abcam 61015) and H3 (Abcam ab73197). Proteins were visualized using Amersham ECL Western Blotting Detection Reagent (GE) and detected using an ImageQuant LAS 4000 imager and ImageQuant TL (GE). Contrast and brightness were altered in a linear fashion equally across the whole image. The main figures present cropped images; uncropped images are presented in Supplementary Data Set 1.

Recombinant protein production. PRC2 core complex ([EZH2 E2D SUZ12 V5FS domain]) was purified as described previously30. Yeast histone octamer containing wild-type H3 or H3 with the K27M mutation was expressed in E. coli and purified using a two-step method described previously30. The octamer, the mutation K18C was introduced to histone H3, and the fluorophore (7-Diethylamino-3-(((2-Maleimidyl)ethyl)amino)carbonyl)coumarin (MDCC) was attached by mixing 40 μM octamer with 200 μM MDCC under non-reducing conditions for 30 min in the dark, after which, labeled nucleosomes were purified using a 30-nt sliding window as described above. The uniqueness of the labeling reaction was verified via MS. Nucleosomes were reconstituted with 147-bp DNA containing the Widom 601 sequence using standard procedure2.

Fluorescence binding experiments. Direct binding between RNA and PRC2 was analyzed via fluorescence anisotropy using a fluorescein-labeled (G4)10 RNA (AAAAAGGGAAAAAGGGAAAAAGGGAAAAAGGGAAAA) or a 28-nt G4-forming portion of PIM1 RNA (ATGCCGGGGGGGGGGGGGUGGGGUGG). RNA was heated to 95 °C in 100 mM KCl or LiCl, cooled on ice and incubated at 37 °C. All binding experiments were performed at 20 °C, and fluorescence was measured on a Jaco FP-8500 spectrophotometer, with excitation at 495 nm and emission at 525 nm. PRC2 was titrated into 20 nM labeled RNA in assay buffer (50 mM Tris-HCl, pH 7.5, 0.01 % Brij-35, 400 μM SAM) with either 100 mM KCl or 100 mM LiCl. Fluorescence anisotropy data were analyzed using GraphPad Prism (GraphPad Software, USA) and DYNAtBio (BioKin Ltd).

The binding of PRC2 to nucleosomes was performed using fluorescence intensity measurements utilizing MDCC-labeled nucleosomes (excitation 430 nm, emission 476 nm). PRC2 was titrated into 10 nM labeled nucleosomes in 25 mM Tris-HCl, pH 7.5, 40 mM KCl, 0.01 % Brij-35, 10 μM BSA, and 400 μM SAM. Binding was indicated by a decrease in fluorescence intensity. For the competition experiment, PRC2 was titrated into labeled nucleosomes in the presence of 500 nM unlabeled (G4)10 RNA. PIM1 G4 RNA (GGGGGGUGGGGGGGGGUGGGGUGG) or a control non-G4-forming portion of PIM1 RNA (GAGUUCUGCUG AAUGCGGGAGAAGAGAU) using the buffer conditions detailed above. For the PRC2 eviction experiment, the PRC2–nucleosome complex was pre-formed by mixing 50 nM PRC2 and 10 nM MDCC-labeled nucleosomes, and then titration was performed with either (G4)10, (G4)5, PIM1 G4 or PIM1 non-G4-forming control RNA. Binding affinities were determined in DynaFit (BioKin Ltd) by applying a simple 1:1 binding model.

Measuring cell senescence. The proportion of senescent cells was measured using the Senescence Assay Kit (Abcam ab228562). Cells were treated as described above, stained with 2H and 2+H2Et according to the manufacturer's instructions. Cells were harvested via trypsinization, and signal was measured using the FL-1 channel on a Fortessa X20 flow cytometer and quantified with FlowJo (BD Biosciences).

G4 structure prediction. G4 scores were calculated across the mm9 genome using G4Hunter21, using a 25-nt sliding window. Sequences with a G4 score above a threshold of 1.2 were selected and overlapped with splice sites defined by iCount as described previously31. The unique molecular identifiers were classified as non-G4 forming if no G4 structures were predicted by G4Hunter using a 30-nt sliding window as described above. To normalize the cross-link density for G content, the G frequency at each position was calculated for both groups, and the cross-link density for the non-G4 group divided by the non-G4/G4 G-frequency ratio. The number of cross-link sites per 5-nt window was displayed using the heatmap2 function from the ggplot2 package in R.

Characteristics of G4-forming sequences. The number of G tracts in each sequence, the number of Gs within each G tract, the number of nucleotides in the loops, the base composition within the loops, and the position of the cross-linked G tracts were calculated. Sequences were filtered using custom scripts and plotted in R. The expression level of the genes (RPKM) in each group was obtained from total RNA-seq data and log2 transformed.

Alternative splicing. RNA-seq data from ref. 39 (WT ESC: GSM1399452, GSM1399453, GSM1399454 and SUZ12 + Ras: ESC: GSM1399458, GSM1399459, GSM1399460) were filtered and processed using custom scripts and plotted in R. The significance of the genes (RPKM) in each group was obtained from total RNA-seq data and log2 transformed.

Statistical analysis. The significance of the increase in the cross-link density across the set of G4-forming first exon–intron junctions (n=942) versus the set of non-G4-forming first exon–intron junctions (n=760) was estimated using a Wilcoxon rank-sum test. The significance of the decrease in the number of G tracts per cross-linked, predicted G4 versus non-cross-linked, predicted G4 was estimated using a Wilcoxon rank-sum test. Measurements of PRC2 RNA or nucleosome binding were performed in triplicate, and data plotted in GraphPad Prism (GraphPad Software, USA), with error bars representing the standard error of the mean. The significance of changes in HA-dCas9, SUZ12, H3K27me3 and H2A.K19199 occupancy after addition of dox relative to untreated cells was estimated using an unpaired t-test (Student’s t test) with Welch’s correction factor 2. For comparison, the number of alternative splicing events that occur during differentiation of ESC to neural precursor cells was calculated using MISO with the same thresholds using data from ref. 39 (ESC: GSM1180294 and GSM1180295; NPC day 3: GSM1184609 and GSM1184610).

Data availability. Input iCLIP sequencing data have been deposited in the Gene Expression Omnibus (GEO) with accession code GSE120696. Previously published ICLIP sequencing data and RNA-seq data are available in GEO under accession code GSE66829. The positions of predicted G-quadruplex RNA structures and the positions of PRC2 cross-link sites around first 5’ splice sites are provided in Supplementary Table 1.
Supplementary Data Set 2 contains t statistics, confidence intervals, effect sizes and degrees of freedom for all significance tests. Raw quantitative PCR data and all other data are available upon reasonable request. Requests for data and materials should be addressed to R.G.

References
64. Kearns, N. A. et al. Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. Development 141, 219–223 (2014).
65. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res. 44, W272–W276 (2016).
66. Huppertz, I. et al. iCLIP: protein-RNA interactions at nucleotide resolution. Methods 65, 274–287 (2014).
67. Bouwman, R. D. et al. Human immunodeficiency virus Tat associates with a specific set of cellular RNAs. Retrovirology 11, 53 (2014).
68. Kwok, C. K. & Balasubramanian, S. Targeted detection of G-quadruplexes in cellular RNAs. Angew. Chem. Int Ed. Engl. 54, 6751–6754 (2015).
69. Bartke, T. et al. Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell 143, 470–484 (2010).
70. Saravanan, M. et al. Interactions between the nucleosome histone core and Arp8 in the INO80 chromatin remodeling complex. Proc. Natl Acad. Sci. USA 109, 20883–20888 (2012).
71. Dyer, P. N. et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzym. 375, 23–44 (2004).
72. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28, 511–515 (2010).
73. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36 (2013).
74. Kata, Y., Wang, E. T., Airoldi, E. M. & Burge, C. B. Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat. Methods 7, 1099–1105 (2010).
75. Hon, G. C. et al. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. Mol. Cell 56, 286–297 (2014).
Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- n/a  Confirmed
- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
  - *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  - *Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated
- Clearly defined error bars
  - *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on *statistics for biologists* may be useful.

Software and code

Policy information about *availability of computer code*

Data collection

- ImageQuant LAS 4000 software (GE Healthcare Life Sciences); https://www.gelifesciences.com/en/gb/shop/molecular-biology/nucleic-acid-electrophoresis--blotting--and-detection/molecular-imaging-for-nucleic-acids/imagequant-las-4000-series-p-00039
- 7500 Software v2.0.6; https://www.thermofisher.com/uk/en/home/life-science/pcr/real-time-pcr/real-time-pcr-instruments/7500-fast-real-time-pcr-system.html
- QuantStudio 5 Real-Time PCR System software (ThermoFisher); https://www.thermofisher.com/uk/en/home/life-science/pcr/real-time-PCR/real-time-PCR-instruments/quantstudio-3-5-real-time-PCR-system/quantstudio-5-5.html
- GE Healthcare Typhoon scanner control v50
- Jasco FP-8500 spectrofluorometer software; https://jasocon/products/spectroscopy/fluorometer/spectrofluorometer-models/fp-8500-spectrofluorometer/
- Illumina HiSeq Software Suite; http://emea.support.illumina.com/sequencing/sequencing_instruments/hiseq_2500/downloads.html?langsel=/en/
  - Fortessa X20 Model 657675R1

Data analysis

- iCount; Curk et al. (2016) iCount: protein RNA interaction iCLIP data analysis (in preparation). https://github.com/omazcz/iCount/tree/master/iCount
- Bowtie; http://bowtie-bio.sourceforge.net/index.shtml
- CHOPCHOPv2; (Labun et al., 2016) https://gtcr.cn.org/chopchop-v2-improved-online-crisprcas9-target-prediction-tool
- Cufflinks; (Trapnell et al., 2010) http://cole-trapnell-lab.github.io/cufflinks/
- Cuffmerge; (Trapnell et al., 2010) http://cole-trapnell-lab.github.io/cufflinks/
For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Input iCLIP sequencing data have been deposited in the Gene Expression Omnibus (GEO) with accession code GSE120696. Previously published iCLIP sequencing data and RNA-seq data are available in GEO under accession code GSE66829. The positions of predicted G-quadruplex RNA structures and the positions of PRC2 crosslink sites around first 5’ splice sites are provided in Supplementary Table 1. Synthetic DNA and primer sequences are provided in Supplementary Table 2. Supplementary Dataset 1 contains uncropped images of all blots. Supplementary Dataset 2 contains t-statistics, confidence intervals, effect sizes and degrees of freedom for all significance tests. Raw quantitative PCR data is available upon request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | No sample size calculation was performed. |
|-------------|-------------------------------------------|
| Data exclusions | One replicate of the UV-RIP was discarded because of lack of detection of RNA in the IP samples. No other data were excluded from any of the other analyses. |
| Replication | Unless stated, all experiments were performed in triplicate. One representative replicate is shown for biotin-RNA pull-down, biotin-nucleosome pull-down, p16INK4a immunoblotting, RT stalling, and native acrylamide gel electrophoresis experiments. Further replicates of the data in Fig. 2a and 2c are shown in Supplementary Fig. 3d and e, respectively. Each input iCLIPs were performed once but each of the three inputs showed similar RNA crosslink distributions as expected for background RNA crosslinking. |
| Randomization | Randomisation was not used. |
| Blinding | Blinding was used for the first replicate of the senescence experiment but not for any other experiments. |

Reporting for specific materials, systems and methods
Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ☐   | Unique biological materials |
| ☒   | Antibodies |
| ☒   | Eukaryotic cell lines |
| ☒   | Palaeontology |
| ☐   | Animals and other organisms |
| ☒   | Human research participants |

Methods

| n/a | Involved in the study |
|-----|-----------------------|
| ☒   | ChIP-seq |
| ☐   | Flow cytometry |
| ☒   | MRI-based neuroimaging |

Unique biological materials

Policy information about availability of materials

Obtaining unique materials All unique materials are available from the authors upon request.

Antibodies

**Antibodies used**

- Rabbit monoclonal RAS G12V (D2H12) Cell Signaling Technology (CST) Cat#14412; AB_2714031, Lot 1
- Rabbit monoclonal SUZ12 [D39H6] CST Cat#3732; AB_10828723; Lot 6
- Goat polyclonal SUZ12 [P-15] Santa Cruz Biotech Cat#sc-62664; AB_2196857; Lot D0615;
- Mouse monoclonal EZH2 [AC27] Cell Signaling Technology (CST) Cat#3147; AB_10694383; Lot 4
- Rabbit mAb antibody JARID2 (D6M9X) Cell Signaling Technology (CST) Cat#13594; AB_2798269
- Rabbit mAb AEFP2 [D7C6X] Cell Signaling Technology (CST) Cat#14129; AB_2798398 Lot 1
- Rabbit polyclonal PCL2 Proteintechn 16208-1-AP AB_2147370
- Rabbit polyclonal HMG11 Bethyl Laboratories A302-363A; AB_1907346 Lot 1
- Mouse monoclonal p16INK4a Santa Cruz Biotech Cat#sc-56330; AB_785018
- Rabbit polyclonal to Histone H3 (acetyl K27) - ChIP Grade Abcam ab4729; AB_2118291 lot GR3231937-1
- Rabbit monoclonal H2AK119ub Cell Signaling Technology (CST) Cat#4820; AB_10891618 Lot 6
- Rabbit polyclonal ACTB CST Cat#4967; AB_330288; Lot 7
- Non-specific Rabbit Anti-Mouse IgG H&L Abcam Cat#ab64640; AB_2614925
- Rat monoclonal HA [3F10] Roche Cat#11867423001; AB_390918; Lot 27573500
- Mouse monoclonal Histone H3 (tri-methyl K27) Abcam Cat#ab6002; AB_305237; Lot GR2759011-8
- Rabbit monoclonal Histone H3 (tri-methyl K27) Abcam Cat#ab192985; AB_2605559; Lot GR3202355-5
- Rabbit polyclonal Histone H3 (mono-methyl K27) Abcam Cat#ab175037
- Rabbit polyclonal Histone H3 Abcam Cat#ab1791; AB_302613; Lot GR3300978-2

Validation

RAS G12V antibody was tested by the manufacturers by western blot analysis of extracts from various human cell lines with our without the specific Ras mutation. The antibody only gave a band in 3T3 cells expressing HRasV12. (Supplementary Fig. 6a, SUZ12 and EZH2 antibodies were validated in the lab by western blot in wild-type and knock-out mouse ESC extracts, as reported in Beltran et al., 2016. AEBP2, JARID2 and PCL2 antibodies were validated in the lab by western blot in wild-type and knock-out mouse ESC extracts.

Antibodies against histone H3 and methylated K27 were validated by the manufacturer using several strategies as noted on their website. The mouse monoclonal histone H3 [tri-methyl K27] antibody is blocked by tri-methyl K27 peptide and slightly by di-methyl K27 peptide (there is <1% cross reactivity with di-methyl K27 as determined by ELISA). It is not blocked by mono-methyl K4, di-methyl K4, tri-methyl K4, mono-methyl K9, di-methyl K9, tri-methyl K9, mono-methyl K27 or unmodified K27 peptides.

Rabbit monoclonal histone H3 [tri-methyl K27] antibody is tested by the manufacturers using various strategies, including ELISA, peptide array. The antibody is blocked by tri-methyl K27 peptide and slightly by di-methyl K27 peptide (there is 14% cross-reactivity with di-methyl K27 as determined by ELISA).

Antibodies against histone H3 K7 acetylated was validated by the manufacturer using chromatin immunoprecipitation and western blot.

Ubiquityl-Histone H2A antibody recognizes endogenous levels of histone H2A protein only when ubiquitinated at Lys119. The manufacturer tested cross-reaction with other ubiquitinated proteins or free ubiquitin. That the HA antibody specifically recognized HA-dCas9 was verified by western blotting (Supplementary Fig. 5a).

ACTB and non-specific IgG antibodies were not validated but are extensively used by the community.

Eukaryotic cell lines

**Cell line source(s)**

NIH/3T3: From ATCC via Bert Vanhaesebroeck’s laboratory.

mESC.Eh3/Er (Kanhere et al., 2010).

ES-E14 and and IARID2/GT/GT: From Amanda Fisher’s laboratory

Aebp2/W1/W1 and Aebp2/G1/G1: from Neil Brockdorff’s laboratory

Pdl2GT/GT and Pdl2WT/WT: Adrian Bracken’s laboratory

G401 from Sigma-Aldrich

**Authentication**

The G-401 cell line has been certified by the European Collection of Authenticated Cell Cultures (ECACC).
**Flow Cytometry**

**Plots**
- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

**Methodology**

| Sample preparation       | Samples were incubated with dye for two hours, washed and trypsinized |
|--------------------------|------------------------------------------------------------------------|
| Instrument               | Fortressa X20 Model 657675R1, SN:H657675R1011                           |
| Software                 | Flowjo (BD Biosciences)                                                |
| Cell population abundance| NA                                                                      |
| Gating strategy          | Cells were selected for single-cell events and senescence was measured using BS30/30 channel. |

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.