



**HAL**  
open science

## Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping

V. Loubiere, G. L. Papadopoulos, Q. Szabo, A. M. Martinez, Giacomo Cavalli

### ► To cite this version:

V. Loubiere, G. L. Papadopoulos, Q. Szabo, A. M. Martinez, Giacomo Cavalli. Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping. *Science Advances*, 2020, 6 (2), pp.eaax4001. 10.1126/sciadv.aax4001 . hal-02896366

**HAL Id: hal-02896366**

**<https://hal.umontpellier.fr/hal-02896366>**

Submitted on 10 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## GENETICS

# Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping

V. Loubiere, G. L. Papadopoulos, Q. Szabo, A-M. Martinez\*, G. Cavalli\*

**Polycomb repressive complexes 1 and 2 have been historically described as transcriptional repressors, but recent reports suggest that PRC1 might also support activation, although the underlying mechanisms remain elusive. Here, we show that stage-specific PRC1 binding at a subset of active promoters and enhancers during *Drosophila* development coincides with the formation of three-dimensional (3D) loops, an increase in expression during development and repression in PRC1 mutants. Dissection of the *dachshund* locus indicates that PRC1-anchored loops are versatile architectural platforms that persist when surrounding genes are transcriptionally active and fine-tune their expression. The analysis of RING1B binding profiles and 3D contacts during neural differentiation in mice suggests that this role is conserved in mammals.**

## INTRODUCTION

Polycomb group proteins (PcG) assemble into two main epigenetic complexes called Polycomb repressive complexes 1 and 2 (PRC1 and PRC2), which are highly conserved across metazoans and collaborate at multiple levels to maintain their target genes in a repressed state (1, 2). PRC1 also binds a subset of active promoters and enhancers devoid of the PRC2-mediated H3K27me3 repressive mark in both *Drosophila* (3, 4) and mammals (5, 6). Loss-of-function experiments suggest that PRC1 might contribute to the transcriptional activation of a subset of genes (4–6). However, since PRC1 binds to a large number of sites, disentangling direct from indirect regulatory effects has been proven difficult, and the molecular mechanisms that might support transcriptional activation by PRC1 are obscure. Here, we tested whether PRC1 might mediate gene activation by forming enhancer-promoter loops, in addition to the repressive chromatin loops that were previously described (7, 8).

## RESULTS

### PRC1 binds a subset of active regulatory elements in eye discs

First, we characterized the three-dimensional (3D) epigenome of *Drosophila* eye-antennal imaginal discs (EDs) by producing chromatin immunoprecipitation–sequencing (ChIP-seq) maps for a comprehensive set of active (H3K4me2/3, H3K36me3, and H4K20me1), enhancer-related (H3K4me1), PRC2-mediated (H3K27me1/2), and PRC1-mediated (H2AK118Ub) posttranslational modifications (PTMs) that were analyzed in addition to our previously published ChIP-seq data for H3K27me3 and H3K27Ac (3). Furthermore, we generated a high-resolution Hi-C map from EDs (~1.25-kb resolution; fig. S1, see Materials and Methods). Genome-wide clustering of PTM ChIP-seq signals using self-organizing maps (SOMs) (Fig. 1A, fig. S2, and table S1) (9) identified five distinct chromatin types that exhibit hallmarks of active transcription start sites (aTSSs), active transcrip-

tion termination sites (aTTSs), enhancers, PcG-repressed, and the previously described null chromatin state (Fig. 1B and fig. S2) (10, 11). While aTSS and aTTS bins are essentially transcriptionally active, PcG-repressed bins are mostly silent, as expected (fig. S3). Moreover, aTSS and enhancer bins show high ATAC-seq (assay for transposase-accessible chromatin–sequencing) levels, as expected for these regulatory elements, which are known to be nucleosome depleted (fig. S3). Markedly, a subset of active enhancer bins found in the vicinity of transcriptionally active genes also show consistent enrichment for the PRC1-mediated H2AK118Ub mark without the PRC2-mediated H3K27me3 repressive mark (Fig. 1B and figs. S2 and S4). For instance, the *rainy head* gene is ubiquitously expressed in the ED (12) and retains detectable H2AK118Ub enrichment (fig. S4). Although previous studies using S2 cells suggest that the H2AK118Ub is virtually absent from active sites (11), our analysis suggests that catalytically active PRC1 complex binds to a subset of active enhancers in the developing ED. Further studies will be needed to address whether H2AK118Ub can genuinely mark active genes or whether the colocalization observed here might reflect the coexistence of multiple cell types in ED. Future investigations should also address the dynamics of PRC1 binding at these active sites, and whether it coincides with ongoing transcription.

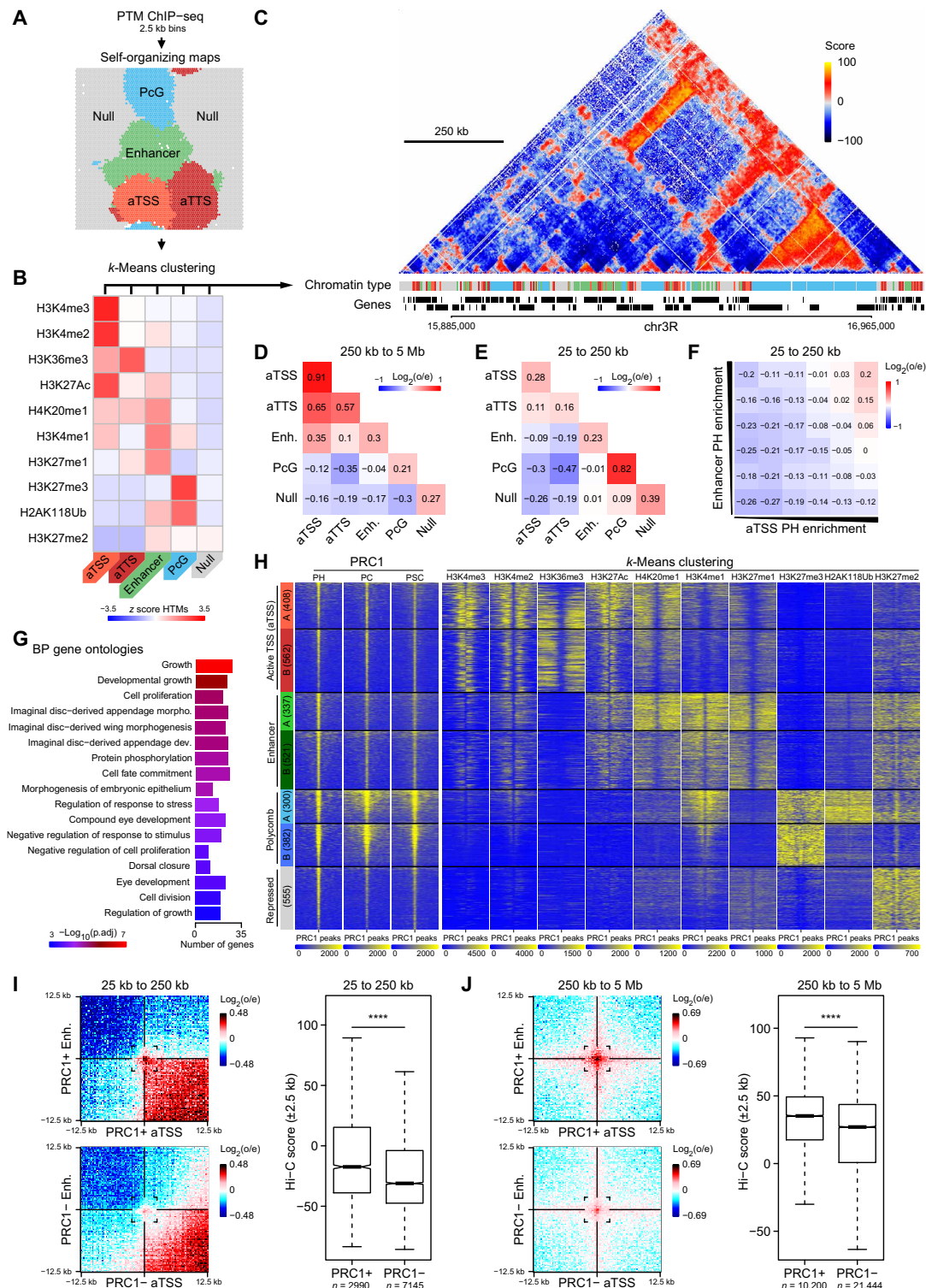
### PRC1 binding is associated with increased enhancer-promoter 3D contact frequencies in the ED

The five chromatin types (Fig. 1A) were used to segment the Hi-C map (Fig. 1C), and the results show that each type preferentially interacts with itself (Fig. 1, D and E) (13). In addition, aTSS and aTTS bins interact with each other, while they are segregated from PcG and null repressed chromatin at long-range distances (250 kb to 5 Mb) (Fig. 1D) (14). Enhancer bins behave like active chromatin at long-range distances, namely, they interact preferentially with themselves and with aTSS bins (Fig. 1D). The fact that they show no preferential short-range interactions with aTSS bins (25 to 250 kb) (Fig. 1E) is consistent with the possibility that enhancers only interact with a small subset of all surrounding promoters. To identify this subset, we complemented preexisting ChIP-seq data for the canonical PRC1 members PC (Polycomb) and PH (Polyhomeotic) (3) with a ChIP-seq for the PRC1 member PSC (Posterior Sex Combs), and ranked aTSS and enhancer bins based on PRC1 enrichment. We found

Copyright © 2020  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
NonCommercial  
License 4.0 (CC BY-NC).

Institute of Human Genetics, UMR 9002, CNRS–University of Montpellier, 34396 Montpellier, France.

\*Corresponding author. Email: anne-marie.martinez@igh.cnrs.fr (A-M.M.); giacomo.cavalli@igh.cnrs.fr (G.C.)



**Fig. 1. PRC1 binding is associated with increased enhancer/promoter contacts in ED.** (A) Schematic view of the workflow used for the genome-wide stratification of PTMs using self-organizing maps. (B) Heatmap showing the average PTM z scores for each chromatin type (indicated at the bottom). (C) Hi-C screenshot of a ~1.4-Mb region on the chromosome 3R. Genome-wide stratification of the ED epigenome is reported at the bottom. (D and E)  $\text{Log}_2[\text{observed (o)}/\text{expected (e)}]$  average interaction frequencies between each chromatin type at long- (250 kb to 5 Mb) and short-range (25 to 250 kb) distances, respectively. (F)  $\text{Log}_2(\text{o/e})$  average interaction frequencies between aTSS and enhancer bins (25 to 250 kb) ranked by increasing PH enrichment. (G) Bar graph showing the ontologies of the top interacting aTSS bins from (F). (H) k-Means clustering of PTM ChIP-seq around a stringent set of PRC1 (PH and/or PC and/or PSC) binding sites ( $\pm 2.5$  kb). Clusters are shown on the left, together with PRC1 (PH, PC, and PSC) ChIP-seq tracks. (I and J)  $\text{Log}_2(\text{o/e})$  aggregate interactions within 25-kb windows centered on PRC1-bound (PRC1+) or PRC1-depleted (PRC1-) aTSS and enhancer (Enh.) sites between 25 to 250 kb (I) and 250 kb to 5 Mb (J). Corresponding quantifications of the Hi-C score ( $\pm 2.5$  kb) are shown on the right. Wilcoxon, \*\*\*\* $P < 1 \times 10^{-5}$ .

that high PRC1 levels are predictive of increased contact frequencies at both short- and long-range distances (Fig. 1F and figs. S5 and S6). Hence, the strength of enhancer-promoter contacts at genes involved in growth, imaginal disc morphogenesis, and eye development unexpectedly correlates with PRC1 binding to both elements (Fig. 1, F and G, and table S2).

To identify PRC1-bound TSSs and enhancers genome wide, we clustered PTM ChIP-seq signals around a stringent set of PRC1 peaks (see Materials and Methods and table S3). Notably, 60% of PRC1 sites lack detectable H3K27me3 in the ED and are decorated with the H3K27Ac active mark (Fig. 1H and fig. S7). In contrast with canonical H3K27me3-enriched Polycomb loci, these sites lie in the vicinity of transcriptionally active genes (fig. S8A and tables S4 and S5; see Materials and Methods) and correspond to either H3K4me3-enriched aTSSs (>76% at less than 500 bp; fig. S8E) (3) or H3K4me1-decorated enhancer-like binding sites (Fig. 1H and fig. S7). The 858 H3K27Ac-enriched putative enhancer sites lie in open chromatin regions (fig. S8B) (15, 16) that substantially overlap with previous mapping of developmental enhancers by STARR-seq (self-transcribing active regulatory region sequencing) in BG3 cells (fig. S8C) (17, 18), suggesting that they correspond to active enhancers.

PRC1-bound enhancers are located near critical genes regulating growth, imaginal disc, and eye development (fig. S8D and table S6). To test whether PRC1 binding mediates preferential interactions between active enhancer and promoter elements, we produced aggregate plots of the interaction frequencies centered on PRC1-bound aTSSs and enhancers. The results (Fig. 1, I and J) show that PRC1 binding induces stronger 3D contacts than control sets of aTSSs and enhancers characterized by similar pairwise distances, by similar ATAC-seq enrichments, and by the absence of PRC1 peaks (fig. S9; see Materials and Methods), both at short- and long-range distances. In contrast, PRC1-bound enhancers do not contact the TSSs of the corresponding genes. Therefore, PRC1 binding is associated with increased 3D contact frequencies between a subset of active regulatory elements in ED (Fig. 1, I and J and figs. S10 and S11).

### PRC1-anchored enhancer-promoter loops are developmentally dynamic

To test whether PRC1-anchored enhancer-promoter loops are developmentally dynamic, we analyzed the difference between late embryos (16 to 18 hours) and larval EDs. Sixty-three percent of PRC1 enhancers, which we named stage-specific enhancers (SSEs), show no detectable PRC1 peaks (Fig. 2, A to C, and fig. S12, A to C), lower H3K4me1 and H3K27Ac (fig. S12, D and E), and generally do not form enhancer-promoter loops in embryos (Fig. 2, D and E, and fig. S12, F and G). Unexpectedly, SSE target genes are transcriptionally up-regulated during development and down-regulated in EDs bearing PRC1 mutant clones (Fig. 2F and table S7). To eliminate the possibility that these transcriptional changes arise from a compensatory response by remaining heterozygous cells, we used an RNA interference (RNAi) system that allowed us to knock down *ph* in about 90% of the cells within the ED and reproduce the tumorigenic phenotype associated with *ph* mutation (fig. S13) (3). The *ph* RNAi transcriptome is skewed toward down-regulation at *ph*-bound enhancers, including SSEs (fig. S14). Consistent with the lack of H3K27me3 on these genes, mutations affecting the PRC2 methyltransferase *E(z)* or its PRC1 reader *Pc* have a much weaker transcriptional effect than *ph* and *Psc* mutations (Fig. 2F). These results suggest a previously unknown role for PRC1 in promoting SSE-promoter 3D contacts at

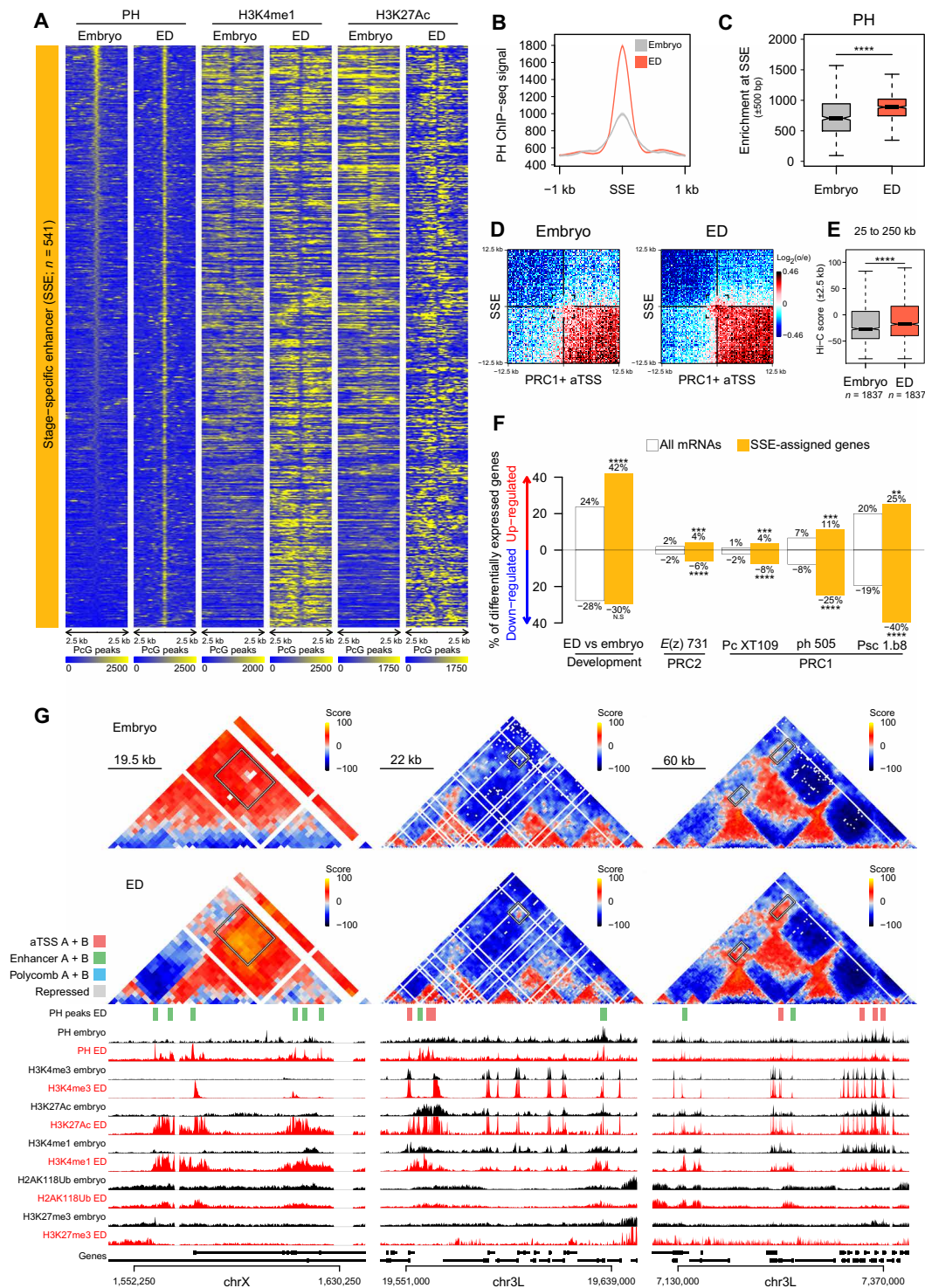
active genes enriched for H3K27Ac to favor their transcriptional induction during development (Fig. 2G).

### PREs in the *dac/CG5888* locus are required for faithful gene activation

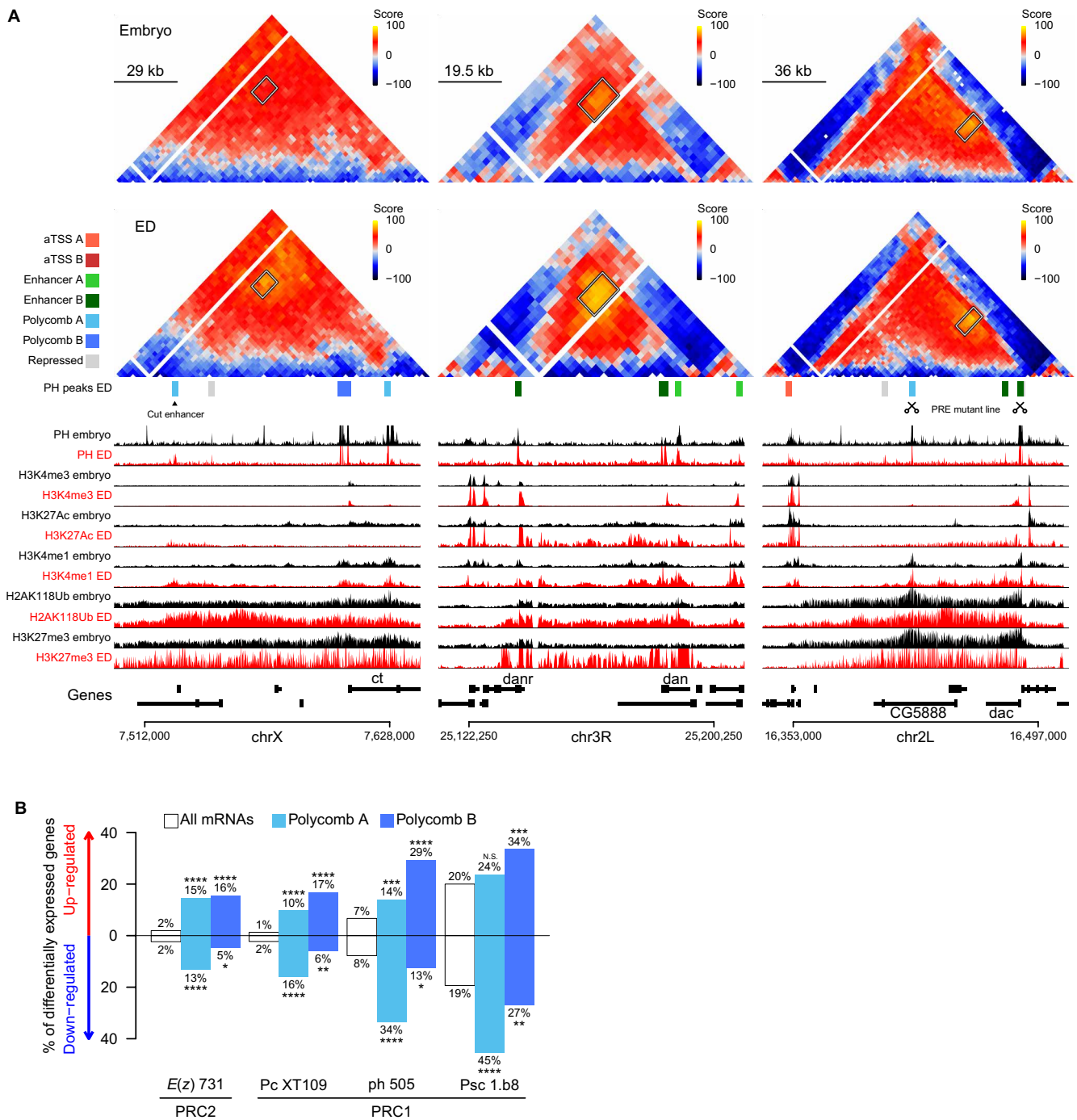
Although the finding of PRC1-anchored repressive loops at H3K27me3-enriched loci is not unprecedented (7, 8), the presence of PRC1 loops at fully active sites (Fig. 2G) was unknown. It was recently suggested that a subset of Polycomb response elements (PREs) might be marked by H3K27me3 to maintain silencing in some cells, while they may switch to function as developmental enhancers in other cells (19), but the chromatin conformation of these elements in the active or the repressed state was not directly analyzed. We therefore tested the alternative possibility that a subset of canonical PREs contains enhancers that may activate transcription via PRC1-mediated looping. Clustering analysis shows that the Polycomb cluster can be subdivided into two subclusters, characterized by differential transcription, chromatin marking, and 3D looping frequencies. Polycomb cluster A (Fig. 1H) is characterized by higher transcription (fig. S8A) and H3K4me1 levels (fig. S7) than the Polycomb cluster B (Fig. 1H and fig. S7), suggesting that developmental enhancers in the Polycomb A cluster are in a poised state (20). We found that Polycomb A sites form stronger loops between them and are marked by higher H2AK118Ub levels than Polycomb B sites (fig. S15). Some of these sites are located at key developmental gene loci that show increased H3K4me1 and 3D contacts in EDs compared with embryos (Fig. 3A). One of these loci includes the *dan* and *danr* promoters (21), which exhibit stage-specific H3K4me3 peaks and are up-regulated in the ED ( $\times 5.0$  and  $\times 7.6$ , respectively; Fig. 3A). Furthermore, a significant fraction of genes from the H2AK118Ub-enriched Polycomb cluster A is down-regulated in *ph* and *Psc* mutants (Fig. 3B) and in *ph* RNAi EDs (fig. S14). Down-regulated genes include the core members of the Retinal Determination Gene Network *twin of eyeless*, *eyegone*, *sine oculis*, *eyes absent*, and *tiptop* (table S7 and fig. S14), which are major determinants of eye development (22). These data suggest that PRC1 might sustain transcriptional activation in the expression domains at a subset of targets, which are otherwise marked by the canonical PRC2-mediated H3K27me3 mark in their silencing domains.

To directly test whether PRC1-anchored sites, which form repressive chromatin loops in the cells where genes are silent, might also play an activating role in the normal gene expression domain, we focused on the PRC1 loop located near the uncharacterized *CG5888* gene and the eye selector gene *dachshund* (*dac*) (8). We found that this loop is maintained in the ED, although the spreading of the H3K4me1 and H3K27Ac marks coincides with increased *dac* transcription (by 4.5-fold) (Fig. 3A). To overcome the problem of cell heterogeneity in the ED, we took a microscopy-based approach focusing on the *dac* gene in its active expression domain (a broad stripe of cells including the morphogenetic furrow (MF), as well as its anterior and posterior neighbors) and in its repressed state (anterior to the MF) (fig. S16, A and B). CRISPR-mediated deletion of the two PRC1-binding sites located in the *dac/CG5888* locus (Fig. 4A and fig. S16C) (8) leads to reduced *dac* expression (Fig. 4, A and B, and fig. S16, A and B). Because the *dac* TSS-proximal binding site is close to the *dac* promoter (fig. S16C) and might therefore dampen its expression via close cis-regulatory effects, we tested the impact of the sole removal of the 3' *dac* TSS-distal PRC1 binding site. No transcriptional impact on *dac* transcripts was found in this mutant condition (Fig. 4, A and B). Furthermore, in a line where the





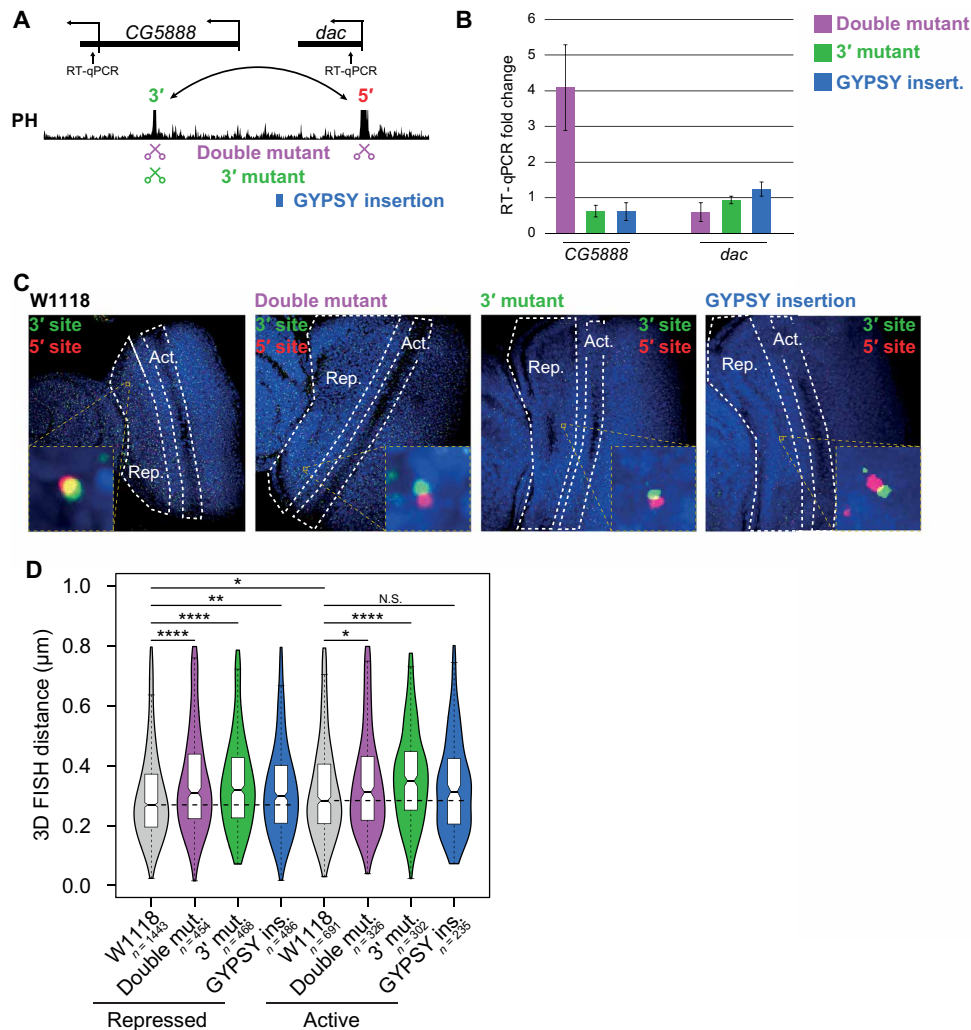
**Fig. 2. Reinforced PRC1 binding is associated with increased interactions and transcriptional activation at the larval stage.** (A) PH, H3K4me1, and H3K27Ac ChIP-seq tracks centered on stage-specific enhancer (SSE) sites that only overlap with PRC1 peaks at the larval stage ( $\pm 2.5$  kb). (B and C) PH ChIP-seq average tracks in a 2-kb window centered on SSEs at embryonic (in gray) and larval stages (in orange). Corresponding quantifications ( $\pm 500$  bp) are shown in (C). Wilcoxon, \*\*\*\* $P < 1 \times 10^{-5}$ . (D and E)  $\text{Log}_2(o/e)$  aggregate interactions within 25-kb windows centered on PRC1-bound aTSSs and SSEs between 25 and 250 kb in embryos (left) and EDs (right). Corresponding quantifications ( $\pm 2.5$  kb) are shown in (E). Wilcoxon, \*\*\*\* $P < 1 \times 10^{-5}$ . (F) Percentage of significantly up- (top) and down-regulated (bottom) SSE-assigned genes during development (on the left) or between control and PRC2 [ $E(z)^{731}$ ] or PRC1 ( $Pc^{XT109}$ ,  $ph^{505}$ , and  $Psc^{1.B8}$ ) mutant EDs (on the right, padjust cutoff = 0.001). Hypergeometric test, \*\*\*\* $P < 1 \times 10^{-5}$ , \*\*\* $P < 1 \times 10^{-3}$ , \*\* $P < 1 \times 10^{-2}$ , N.S., not significant. (G) Screenshots of Hi-C maps centered on developmentally regulated enhancer-promoter loops anchored at active PRC1 binding sites in embryo (top panels) and ED (middle panels). PRC1 peaks from ED and their assigned clusters are shown below Hi-C screenshots (legend on the upper left) and projected on the maps as white squares. At the bottom, PTM ChIP-seq tracks are shown for embryos (black) and EDs (red).



**Fig. 3. PRC1-anchored loops are not restricted to transcriptionally silent PcG targets in EDs. (A)** Screenshots of Hi-C maps centered on developmentally regulated loops anchored at H3K27me3-enriched PRC1 binding sites in embryos (top panels) and EDs (middle panels). PRC1 peaks from EDs and their assigned clusters are shown below Hi-C screenshots (legend on the upper left) and projected on the maps as white squares. At the bottom, PTM ChIP-seq tracks are shown for embryos (black) and EDs (red). **(B)** Percentage of significantly up- (top) and down-regulated (bottom) genes from the Polycomb A (in cyan) and Polycomb B (in blue) clusters between control and PRC2 [*E(z)<sup>731</sup>*] or PRC1 (*Pc<sup>XT109</sup>*, *ph<sup>505</sup>*, and *Psc<sup>1.B8</sup>*) mutant EDs (padjust cutoff = 0.001). Hypergeometric test, \*\*\*\* $P < 1 \times 10^{-5}$ ; \*\*\* $P < 1 \times 10^{-3}$ ; \*\* $P < 1 \times 10^{-2}$ ; \* $P < 5 \times 10^{-2}$ ; N.S., not significant.

insertion of a gypsy insulator sequence (8) blocks the interaction between the two PRC1-binding sites (Fig. 4, A and C), *dac* expression is not affected (Fig. 4B). Three-dimensional fluorescence in situ hybridization (3D FISH) measurements using DNA probes located

next to the PRC1 anchor sites (Fig. 4, C and D, and fig. S16C) indicate that the loop is normally present not only in the silent cells but also in the cells where *dac* is transcribed, and it is disrupted in both regions upon CRISPR-mediated deletion of both PRC1 binding



**Fig. 4. The PRC1-anchored loop at the *dac* locus fine-tunes the transcription of neighboring genes.** (A) Schematic view of the three CRISPR mutant lines in which (i) the two PRC1 anchors located at the *dac* locus are excised (double mutant, purple), (ii) only the 3' anchor is removed (3' mutant, green), and (iii) a GYPSY insulator element is inserted between the two anchors (in blue). (B) RT-qPCR of the *CG5888* and *dac* transcripts in the double mutant (purple), the 3' mutant (green), and the GYPSY insertion line (blue) versus WT (*w1118*) EDs. (C) Pictures of the 3D FISH of 5' (red) and 3' (green) PRC1 anchors in WT (*w1118*, left), double mutant (middle left), 3' mutant (middle right), and GYPSY insertion EDs (right). An example of the segmentation of the active ("Act.") and the repressed ("Rep.") parts of the tissue is shown (white dashed lines); insets show higher magnification. (D) Violin plot of 3D FISH distances between anchor PRC1 sites in the repressed (left) and the active (right) parts of the tissue in WT (*w1118*, gray), double mutant (purple), 3' mutant (green), and GYPSY insertion lines (in blue). Wilcoxon, \*\*\*\* $P < 1 \times 10^{-5}$ ; \*\* $P < 1 \times 10^{-2}$ ; \* $P < 5 \times 10^{-2}$ ; N.S., not significant.

sites or the *dac* TSS-distal 3' site only (Fig. 4D). These data suggest that PRC1 does induce a loop in the *dac/CG5888* locus both in the silent and in the active part of the tissue, but disruption of the loop does not play a major transcriptional role for the *dac* gene in this stage.

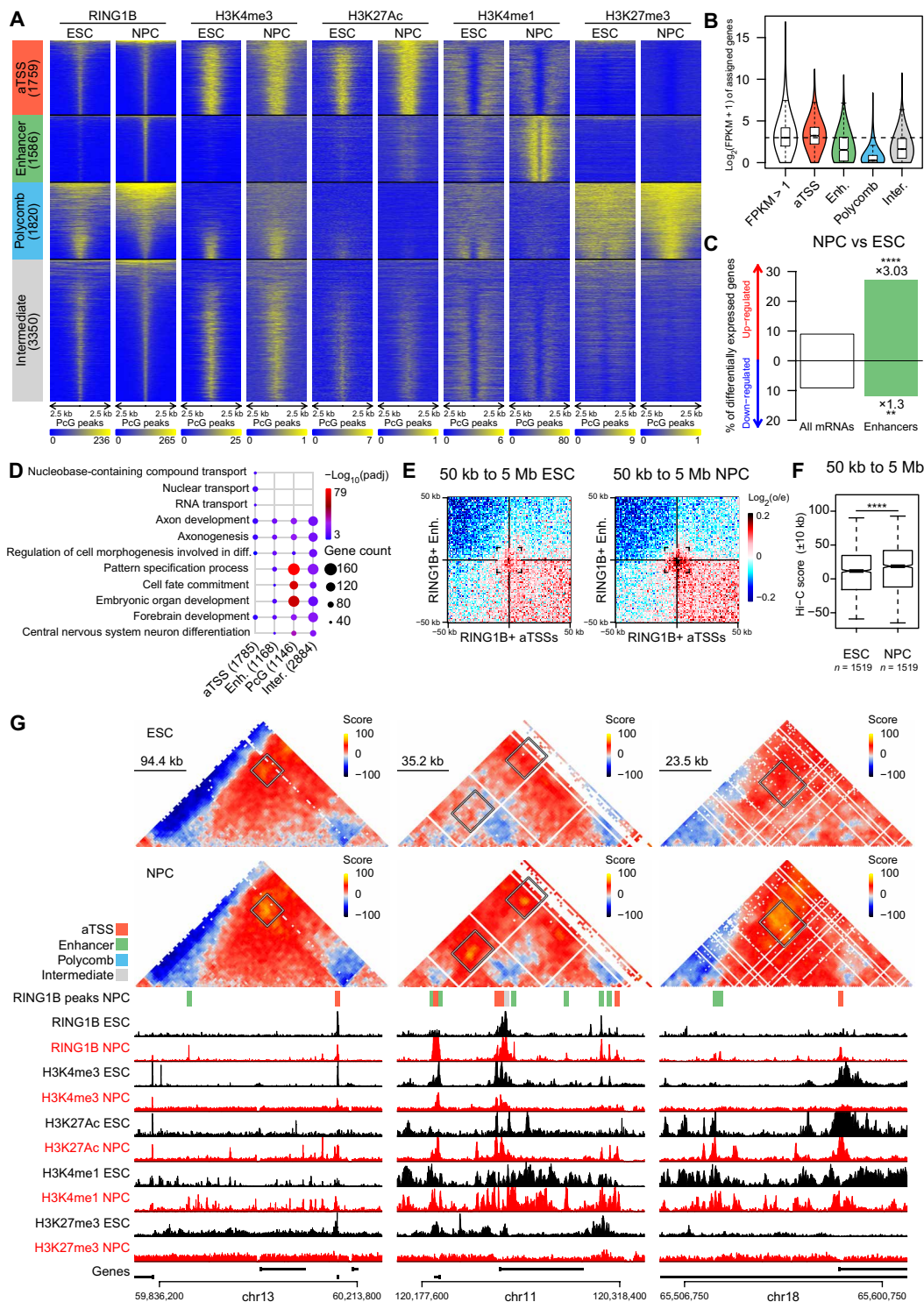
We therefore tested whether the *CG5888* gene might be affected. In contrast to *dac*, deletion of the two PRC1-binding sites induced *CG5888* expression, suggesting that loss of silencing in complete absence of Polycomb proteins (8) might be the dominant effect for this gene. Markedly, however, *CG5888* expression is reduced upon deletion of the 3' PRC1 binding site, as well as upon insertion of a gypsy insulator sequence (Fig. 4, A and B). These data suggest that the loop in the *dac/CG5888* plays an instructive role for expression of *CG5888*, whereas it is not required for transcription of *dac* in normal conditions. In the case of silencing, the PRC1-dependent loop was shown to play an instructive role for the *dac* gene in sensitized

genetic conditions. Whether this is true also for activation remains to be studied.

### Dynamic PRC1-Anchored enhancer-promoter loops form during in vitro differentiation of mouse neural progenitor stem cells

We note that the results described above differ from the case of the *Meis2* locus in mice, in which the activation of the gene involves the disruption of a RING1B-dependent tripartite interaction between the promoter of the gene, a distal repressive site, and a tissue-specific enhancer (23). Therefore, we analyzed high-resolution Hi-C maps from synchronized populations of mouse embryonic stem cells (ESCs) and neural progenitor stem cells (NPCs) obtained from an in vitro differentiation system (24) to address whether PRC1 also mediates enhancer-promoter contacts in mammalian cells. NPCs are characterized by a





**Fig. 5. PRC1 redeployment at developmental enhancers correlates with transcriptional activation and increased enhancer-promoter loops during NPC specification in mouse cells.** (A) *k*-Means clustering of RING1B, H3K4me3, H3K27Ac, H3K4me1, and H3K27me3 ChIP-seq signals from NPCs centered on RING1 peaks from NPCs ( $\pm 2.5$  kb). Clusters are depicted on the left. Corresponding ChIP-seq tracks from ESCs are shown for comparison. (B) Violin plot of the  $\log_2(\text{FPKM} + 1)$  of the genes assigned to each cluster in NPCs. (C) Bar graph of the percentage of significantly up- (top) and down-regulated (bottom) genes between ESCs and NPCs (padjust cutoff = 0.001). Hypergeometric test, \*\*\*\* $P < 1 \times 10^{-5}$ ; \*\* $P < 1 \times 10^{-2}$ . (D) Comparative Gene Ontology (GO) analysis of the genes assigned to each cluster. (E and F)  $\log_2(\text{o/e})$  aggregate interactions within 100-kb windows centered on RING1B-bound aTSS and enhancer sites between 50 kb and 5 Mb in ESCs (left) and NPCs (right). Corresponding quantifications are shown in (F) ( $\pm 10$  kb). Wilcoxon, \*\*\*\* $P < 1 \times 10^{-5}$ . (G) Screenshots of Hi-C maps centered on developmentally regulated enhancer-promoter loops anchored at active RING1B binding sites in ESCs (top panels) and NPCs (middle panels). RING1B peaks from NPCs, and their assigned clusters are shown below Hi-C screenshots (see legend on the upper left) and are projected on the maps as white squares. On the bottom panels, PTM ChIP-seq tracks are shown for ESCs (black) and NPCs (red).



set of peaks bound by the RING1B component of PRC1 complexes, which are absent in ESCs, enriched for H3K4me1, devoid of H3K27me3, and located near genes implicated in neural function (25). Thus, we clustered RING1B peaks based on RING1B, H3K4me3/1, H3K27Ac, and H3K27me3 enrichments in NPCs (24, 25) and identified a set of aTSS (H3K4me3) and enhancer (H3K4me1) sites (Fig. 5A). RING1B and H3K4me1 are sharply increased at enhancer sites in NPCs compared with ESCs and exhibit weak but significant H3K27Ac enrichments compared with H3K27me3-repressed sites (Fig. 5A and fig. S17). Accordingly, genes flanking these sites show higher transcription (Fig. 5B) and up-regulation in NPCs (Fig. 5C) and are enriched for neural-related ontologies (Fig. 5D). Furthermore, RING1B enhancer sites form stage-specific loops with RING1B-bound promoters in NPCs (Fig. 5, E to G). These observations confirm the data from *Drosophila* and suggest that PRC1-anchored active enhancer-promoter contacts are an evolutionary conserved feature in metazoans.

## DISCUSSION

PRC1 plays important roles during normal physiology and in cancer (1–3), but how it might play a dual silencing and an activating role is a matter of great interest. The data provided here suggest that, rather than behaving as transcriptional repressors, PRC1-mediated loops establish versatile architectural platforms that can induce repression and activation. In the absence of transcription factors, PRC1 might cooperate with PRC2 to establish repressive loops and to form silent Polycomb domains, whereas the binding of developmental transcription factors might exploit PRC1-dependent enhancer-promoter contacts to coordinate the timely induction of cognate genes during development. The net effect of looping appears to be gene specific since expression of the *CG5888* gene is more sensitive to disruption of a PRC1-dependent loop than the neighboring *dac* gene. At *dac*, one of the PRC1 binding sites is within a few hundred base pairs (bp) from the TSS, whereas the PRC1-binding site closest to the two alternative *CG5888* alternative promoters is located, respectively, around 16 and 25 kb away. One possibility is therefore that the relative location of regulatory elements might modulate the effects of PRC1-dependent loops. In addition to the role in 3D looping, recent work has shown that PRC1 might assist transcription by modulating occupancy and phosphorylation of RNA polymerase II, as well as association of the pausing-elongation factor Spt5 to enhancers and promoters (4). In future studies, it will be important to analyze the relative role and the interplay of these mechanisms in individual cells at the onset of silencing, as well as during transcriptional activation to understand the chain of molecular events that triggers this dual function for Polycomb-bound regions.

## MATERIALS AND METHODS

### Fly stocks and handling

Flies were raised on standard corn meal yeast extract medium at 25°C. The Oregon-R *w1118* line, referred to as wild type (WT), was obtained from R. Paro [Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH), University of Heidelberg, Germany]. For the transcriptome of *Pc*<sup>XT109</sup> null mutant, mosaic imaginal discs bearing control or mutant clones were generated using the glass multimer reporter (GMR)–hid genetic system described in (26). The FRT2A GMR–hid stock is available at the Bloomington Drosophila Stock Center (no. 6343). The genotypes of control and *Pc* mutant lines are the following:

*UAS-mCD8-GFP*, *UAS-nls lacZ/ey-Gal4 UAS-FLP*; *FRT2A/FRT2A GMR-hid l(3)CL-L<sup>1</sup>* (control line); *UAS-mCD8-GFP*, *UAS-nls lacZ/ey-Gal4 UAS-FLP*; *FRT2A Pc<sup>XT109</sup>/FRT2A GMR-hid l(3)CL-L<sup>1</sup>* (mutant line). *w* and *ph* RNAi clones were generated by crossing an eye-specific Flip-out system (BL no. 64095), with *white* UAS-RNAi (BL no. 33623) or *polyhomeotic* UAS-RNAi (Vienna Drosophila Resource Center no. 50028) lines. The respective genotypes of resulting flies are the following:  $y^1 w^* P\{ry[+t7.2] = ey-FLP.N\}2 P\{w[+mC] = GAL4-Act5C(FRT.CD2).P\}D$ ;  $M\{w[+mC] = UAS-NowGFP-NLS\}ZH-86Fb/UAS-white$  RNAi (control line);  $y^1 w^* P\{ry[+t7.2] = ey-FLP.N\}2 P\{w[+mC] = GAL4-Act5C(FRT.CD2).P\}D$ ;  $M\{w[+mC] = UAS-NowGFP-NLS\}ZH-86Fb / UAS-polyhomeotic$  RNAi (*ph* knockdown line). The CRISPR lines in which the two PRC1-bound PREs and the *dac*-TSS distal PRE in the *dac* locus were deleted are described in (8). This line was derived from the Oregon-R *w1118* line, which was therefore used as a control.

### ChIP-seq from EDs or whole *Drosophila* embryos

ChIP from ED or whole 14- to 16-hour embryos were performed in duplicate as previously described (27). Briefly, ~200 EDs from wandering larvae were used per replicate for PTM ChIP-seq and ~500 for PSC ChIP-seq [PSC antibody is described in (28)]. For each replicate of PTM ChIP-seq from 14- to 16-hour embryos, 2-hour collections were performed at 25°C with about 500 flies. When necessary, several batches of dissection/collection were snap frozen in liquid nitrogen and stored at –80°C to collect enough material. Samples were then transferred into a tight Tenbrock and cross-linked for 15 min at room temperature (RT) using 1.8% formaldehyde and continuous homogenization before being lysed on ice for 2 hours with 1% SDS lysis buffer 2 [see (27)]. Chromatin was sonicated using a Bioruptor (Diagenode) for 18 cycles (settings, 30 s on and 30 s off; high power), and the quality of the sonication was checked on a 1.5% agarose gel (~300-bp fragments). Then, samples were precleared overnight (O/N) using 15 µl of protein A Dynabeads (Thermo Fisher Scientific, catalog no. 10001D) at 4°C, and an aliquot was kept apart to constitute the INPUT. Antibodies were added (table S8) 4 hours before the addition of 30 µl of protein A Dynabeads and O/N incubation at 4°C. Last, the beads were washed, and the precipitated chromatin was eluted before O/N de-cross-linking at 65°C, proteinase K treatment, phenol/chloroform purification, and O/N ethanol precipitation. Libraries were prepared using TruSeq ChIP Sample Preparation kit from Illumina following the manufacturer's instructions. PSC ChIP-seq libraries were subjected to 50-bp single-end sequencing on a HiSeq 2500, PTM ChIP-seq libraries from ED were subjected to 50-bp paired-end sequencing on a HiSeq 2000, and PTM ChIP-seq from 14- to 16-hour embryos were subjected to 75-bp paired-end sequencing on a HiSeq 2000 (~40 to 50 million reads per replicate).

### Hi-C from EDs or whole *Drosophila* embryos

Hi-C experiments from ED and 16- to 18-hour embryos were conducted as previously described by Ogiyama *et al.* (8). For larval Hi-C triplicates, ~200 EDs were quickly dissected (<1 hour) from wandering larvae at RT in Schneider's insect medium before being directly processed. Briefly, formaldehyde was added to reach a 2% final concentration, and the samples were homogenized for 3 min using a tight tenbrock at RT and transferred on a rotating wheel (total time for fixation was 10 min). Then, cell membranes were lysed on ice, and nuclei were subjected to Dpn II treatment O/N at 37°C. The next day, restriction sites were end repaired and biotinylated

using Klenow [New England BioLabs (NEB), catalog no. M0210] and biotin-14-dATP (Life Technologies, catalog no. 19524-016) before being religated in situ using T4 DNA ligase (NEB, catalog no. M0202). Then, proteins were degraded with proteinase K before O/N de-cross-linking at 68°C and subsequent DNA purification using phenol/chloroform and ethanol precipitation. Purified DNA was sheared with an LE220 Covaris sonicator (volume, 130  $\mu$ l; fill level, 10; duty cycle, 15; PIP (Peak Incident Power), 500; cycles/burst, 200; time, 58 s). Sonication was checked on an agarose gel, and 300- to 500-bp DNA fragments were size selected using AMPure XP beads (Agencourt, catalog no. A63881) before biotin pull-down with MyOne Streptavidin T1 beads (Life Technologies, catalog no. 65602). Biotin was removed from unligated ends, and pulled-down DNA fragments were end repaired and A-tailed before NEXTFLEX (Bioo Scientific, 514101) ligation of the adaptors. Last, libraries were polymerase chain reaction (PCR) amplified (10 to 12 cycles) using NEBNext Ultra II Q5 Master Mix (NEB, catalog no. M0544 L), and DNA was double size selected using AMPure XP beads (Agencourt, catalog no. A63881) to isolate 300- to 800-bp fragments. Last, libraries were subjected to 50-bp paired-end sequencing on an Illumina HiSeq2000 (~900 million reads per replicate).

### RNA-seq from *Drosophila* EDs

About 150 EDs were dissected from wandering larvae on ice, and total RNA was extracted using TRIzol reagent. Total RNA was purified using the RNA clean and concentrator kit from Zymo Research (catalog no. R1015). Then, poly-A RNAs were selected, and libraries were prepared using the TruSeq Stranded mRNA Sample Preparation kit from Illumina (*Pc<sup>XT109</sup>*) or the NEBNext Ultra RNA Library Prep Kit for Illumina (*ph RNAi*) and subjected to either 50-bp single-end sequencing (*Pc<sup>XT109</sup>* mutant EDs, ~50 million reads per replicate, two replicates) or 150-bp paired-end sequencing (*white UAS-RNAi* and *polyhomoeotic UAS-RNAi*, ~40 million reads per replicate, four replicates) on a HiSeq 2000.

### Three-dimensional fluorescence in situ hybridization

To generate the DNA probes for 3D FISH in the *dac* locus, genomic DNA of *Drosophila* was PCR amplified twice using the PCR primers that were designed in (8) and are available in table S9. PCR products were purified using the QIAquick PCR purification kit (QIAGEN, catalog no. 28104), and the probes were labeled with Alexa Fluor 555 and 647 dyes using the FISH Tag DNA Multicolor Kit (Invitrogen Life Technologies, catalog no. MP32951). Then, FISH experiments on ED were conducted as previously described by Bantignies and Cavalli (29). Briefly, ~20 EDs were quickly dissected (<20 min) from wandering larvae and fixed for 20 min at RT using a 4% formaldehyde dilution in PBT [phosphate-buffered saline (PBS) + 0.1% Tween 20]. EDs were treated with RNase A (200  $\mu$ g/ml) for 2 hours and permeabilized in PBTr (PBS + 0.3% Triton X-100) for 1 hour at RT. Then, EDs were progressively transferred into a prehybridization mixture [50% formamide, 4 $\times$  Saline-Sodium Citrate (SSC) buffer, 100 mM NaH<sub>2</sub>PO<sub>4</sub> (pH 7.0), and 0.1% Tween 20], and DNA was denatured for 15 min at 80°C. EDs were subsequently transferred into FISH buffer [50% formamide, 10% dextran sulfate, 2 $\times$  SSC, and Salmon sperm DNA (0.5 mg/ml)] containing ~30 ng of each DNA FISH fluorescent probe and hybridized for 14 to 20 hours at 37°C while protected from light. The following day, DNA was counterstained with 4',6-diamidino-2-phenylindole (DAPI), and EDs were washed before being mounted on microscopy slides using ProLong Gold antifade reagent (Invitrogen, catalog no. P36930). Images were

acquired on a Leica SP8 confocal microscope (63 $\times$  objective, 59-nm pixel size, 300-nm z-steps), and microscopy pictures were mounted using the FIJI software (30).

### Microscopy image analysis

3D FISH analysis was conducted using the “Image Processing Toolbox” in MATLAB (version R2018a). Briefly, the channels were smoothed using 3D Gaussian filters ( $\Sigma = 3$  and 1 for DAPI and FISH channels, respectively) and segmented in 3D; FISH objects that were smaller than 50 voxels or located outside of the segmented DAPI were discarded. Last, 3D distances between the centroids of FISH objects were calculated, and only the mutual nearest neighbors that were closer than 0.8  $\mu$ m were considered.

For measuring the area of DAPI- and green fluorescent protein (GFP)-positive cells in eye discs, images of individual discs were smoothed using Gaussian filter ( $\Sigma = 6$ ) and segmented using Otsu’s method. Holes within segmented discs were filled using the imfill function before area measurement. Last, a DAPI mask was applied to the GFP segmentation before calculating the GFP/DAPI overlapping fraction.

### Reverse transcription quantitative PCR

Wandering larvae were collected and briefly washed with 70% ethanol. Then, ~30 EDs were quickly dissected (<30 min) on ice in Schneider’s medium, and total RNA was extracted using TRIzol reagent. Total RNA was purified using the RNA clean and concentrator kit from Zymo Research (catalog no. R1015) following the manufacturer’s instructions and using the DNase I from AMBION (catalog no. AM2222). Purified RNA was quantified on a NanoDrop, and 250 ng of RNA was used for the reverse transcription using the SuperScript III First-Strand Synthesis System kit from Thermo Fisher Scientific (catalog no. 18080051) following the manufacturer’s recommendations. Last, reverse transcription quantitative PCR (RT-qPCR) was performed on a LightCycler480 (Roche). PCR primers are listed in table S9.

### Immunostainings

For DAC immunostainings in WT (*W1118*) and pre-mutant EDs, experiments were systematically conducted in parallel. Wandering larvae were collected, and EDs were quickly dissected at RT in 1 $\times$  PBS before being fixed for 20 min in 4% Formaldehyde. EDs were permeabilized for 1 hour in 1 $\times$  PBS + 0.5% Triton X-100 and blocked for 1 hour in 3% bovine serum albumin (BSA) PBTr (1 $\times$  PBS + 0.025% Triton X-100). Then, primary antibodies were added (table S8) in 1% BSA PBTr and incubated O/N at 4°C on a rotating wheel. The following day, discs were washed in PBTr, and secondary antibodies or rhodamine phalloidin was added (table S8) and incubated for 2 hours at RT on a rotating wheel. Then, nuclei were counterstained with DAPI, and EDs were washed in PBTr before being mounted on microscope slides using ProLong Gold antifade reagent (Invitrogen, catalog no. P36930). Last, images were acquired on a Zeiss Axiomager Z2 Apotome Leica SP8 confocal microscope using the same settings for both genotypes, and images were mounted using the Fiji software (30).

### Alignment, peak calling, and visualization of ChIP-seq, ATAC-seq, and FAIRE-seq from *Drosophila*

After initial quality checks of newly generated data using fastqc, paired-end and single-end .fastq files were aligned to the dm6 version of the *Drosophila* genome using bowtie 2 with default parameters

[v2.1.0; (31)]. Newly generated and publicly available data used in this study are listed in table S10. Reads with low mapping quality (mapq, <30) were discarded using samtools (32). As a control, ChIP-seq duplicates generated for this study were clustered together on the basis of their Pearson correlation coefficients (fig. 18). Peak calling was performed on each replicate separately and on the merged replicates using MACS2 (v2.1.1.20160309) (33) with the following parameters: -g dm --nomodel --keep-dup 1 --extsize 200. For the visualization and the quantification of ChIP-seq tracks, reads per kilo base per million mapped reads (RPKM)-normalized bigWig binary files were generated using the bamCoverage function from Deeptools2 (v2.5.1) (34) with the following parameters: -of = bigwig --samFlagExclude 128 --ignoreDuplications -e 200 --normalizeUsingRPKM. Last, replicates were merged using the bigWigMerge and bedGraphToBigWig tools from the University of California, Santa Cruz (35), with default parameters. For the visualization of the tracks, the Integrative Genomics Viewer (36) and the “seqplots” R package (v1.20.0) (37) were used.

### k-Means clustering of PTM ChIP-seq signals around PRC1 peaks in ED

Highly confident PRC1 peaks in ED were defined as the union of nonoverlapping PH, PC, and PSC peaks that (i) were detected in the two biological replicates with a  $q$  value of at least 2 and (ii) had a  $q$  value of at least 25 when using merged replicates (see the previous section for the details regarding the peak calling with MACS2). Average ChIP-seq signals were computed around PRC1 peak summits ( $\pm 2.5$  kb) for H3K4me3, H3K4me2, H3K36me3, H3K27Ac, H4K20me1, H3K4me1, H3K27me1, H3K27me3, H2AK118Ub, and H3K27me2 in ED using bwtools (38). Then, ChIP-seq signals were scaled and stratified into seven clusters using the “kmeans” function in R (R version 3.5.1, www.R-project.org). Last, PRC1 enhancer sites from ED that did not overlap with the PRC1 peaks from late embryos ( $\pm 1$  kb) were defined as SSE sites. Heatmaps and average tracks were plotted using the “seqplots” R package (v1.20.0) (37).

### k-Means clustering of PTM ChIP-seq signals around ATAC-seq peaks in ED

For the clustering of ATAC-seq peaks in ED, only the ATAC-seq peaks with a  $-\log_{10} P$  value of at least 25 were considered. Then, clustering was performed with the same approach as for the PRC1 peak summits (see previous paragraph). Last, the ATAC-seq peak summits from the “aTSS” and “enhancer” clusters that do not overlap with highly confident PRC1 peak summits (defined in the previous section)  $\pm 2.5$  kb were used as PRC1-depleted control sets.

### Genome-wide clustering of PTM ChIP-seq signals from ED using SOMs

SOMs have been previously shown to be a powerful method to identify complex relationships between genomic data (9). This algorithm relies on a machine learning-based process that trains for a given number of iterations to build a model, which is used to produce intuitively readable maps. Average ChIP-seq enrichments were computed genome wide (2.5-kb bins) for H3K4me3, H3K4me2, H3K36me3, H3K27Ac, H4K20me1, H3K4me1, H3K27me1, H3K27me3, H2AK118Ub, and H3K27me2 in ED using bwtools (38). The resulting matrix was scaled and used to train an  $80 \times 80$  toroidal SOM for 100 iterations using the “kohonen” R package (v3.0.7) (39, 40). Then, we overimposed a  $k$ -means clustering directly on the SOM nodes using the “kmeans” function in R (R version 3.5.1,

www.R-project.org; fig. S2). Last, the maps were plotted using an in-house R script.

### Hi-C analysis from ED or whole *Drosophila* embryos

Raw Hi-C sequencing data were processed using the “scHiC2” pipeline (41). Construction of expected models, Hi-C contact scoring, and Hi-C aggregate plots were performed using the “shaman” R package [https://bitbucket.org/tanaylab/shaman, (42)]. Hi-C replicates generated for this study cluster together and show high Pearson correlation coefficients (fig. S18). Moreover, our newly generated Hi-C replicates from whole 16- to 18-hour embryos cluster with the two replicates generated in (8) and show high Pearson correlation coefficients (fig. S18). Therefore, these three replicates were merged to increase the resolution of the maps. Hi-C map visualization, segmented interaction frequencies, and downstream analyses were conducted using in-house R scripts using the “misha” package (https://github.com/msauria/misha-package).

Moreover, to facilitate the visualization of newly generated Hi-C maps, raw sequencing data were processed using the Hi-C-Pro pipeline (version 2.11.1, REF: 26619908) with default parameters and normalized at 2-, 5-, and 10-kb resolutions using the ice\_norm flag. Last, the hicpro2juicebox utility was used to convert the allValidPairs output of the pipeline into Juicebox .hic format at fragment resolution. These processed files are available in the GEO database repository.

### Alignment and differential expression analysis of *Drosophila* RNA-seq

After initial quality checks of newly generated data using fastqc, paired-end and single-end .fastq files were aligned to the dm6 version (dmel\_r6.15 annotation) of the *Drosophila* genome using STAR with default parameters (v2.4.0i) (43). Then, the reads were counted using htseq-count (44), and the differential expression analysis was performed using the “DESeq2” R package (v1.22.1) (45).

### Collection and alignment of ESC and NPC datasets

Raw ChIP-seq data for RING1B from ESCs and NPCs, H3K4me1, H3K4me3, H3K27Ac, and H3K27me3 from NPCs were collected (table S10) and aligned to the mm10 version of the mouse genome using bowtie 2 with default parameters [v2.1.0; (31)]. For the visualization and the quantification of ChIP-seq signals, RPKM-normalized bigWig binary files were generated with the same method used for *Drosophila* ChIP-seq (see upper sections). H3K4me1, H3K4me3, H3K27Ac, and H3K27me3 ChIP-seq datasets from ESCs, RNA-seq, and Hi-C data from ESCs and NPCs were taken from (24) (table S10).

### k-Means clustering of RING1B peaks from NPCs

RING1B PRC1 peak coordinates reported in (25) were lifted over to the mm10 version of the mouse genome, and the peaks that were closer than 500 bp have been merged. Average ChIP-seq enrichments for RING1 B, H3K4me3, H3K27Ac, H3K4me1, and H3K27me3 from NPCs were computed around RING1B peaks from NPCs ( $\pm 2.5$  kb), mean normalized, and stratified using the kmeans function in R ( $k = 4$ ).

### Hi-C analysis from ESCs and NPCs

Hi-C quantification, aggregate plots, and Hi-C maps visualization were generated with the same methods as for *Drosophila* Hi-C (see upper sections).



## Gene assignment strategy and Gene Ontology analysis

For each cluster, PRC1 peaks were assigned to overlapping genes. In the case where a peak does not overlap with any gene, it was assigned to the gene corresponding to the closest annotated TSS. To identify and visualize enriched Gene Ontologies [DAVID database, “Biological Process” (BP) ontologies (46)], the “clusterProfiler” R package (47) has been used. For mammalian data, the same methods were used as for *Drosophila*.

## Manipulation of genomic coordinate files

Computations on genomic coordinate files (.bed) were conducted using the “BEDTools” suite (v2.26.0) (48).

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/2/eaax4001/DC1>

Fig. S1. Resolution of Hi-C maps.

Fig. S2. SOMs of 10 PTM ChIP-seq from ED.

Fig. S3. aTSS and aTSS SOM clusters are transcriptionally active.

Fig. S4. A subset of active enhancers shows substantial H2AK118Ub enrichment in the ED.

Fig. S5. High levels of PRC1 and H2AK118Ub are associated with increased enhancer-promoter interaction frequencies at short-range distances (25 to 250 kb).

Fig. S6. High levels of PRC1 are associated with increased enhancer-promoter interaction frequencies at long-range distances (250 kb to 5 Mb).

Fig. S7. Quantification of PTM ChIP-seq signal for each cluster of PRC1 peaks.

Fig. S8. PRC1 clusters have distinct genomic features.

Fig. S9. k-Means clustering of ED ATAC-seq peaks provides control sets of aTSSs and enhancers.

Fig. S10. PRC1-bound aTSSs form loops with the TTS of the corresponding genes, whereas PRC1-bound enhancers do not.

Fig. S11. PRC1 binding is associated with increased aTSS-aTSS and enhancer-enhancer contact frequencies in ED.

Fig. S12. PRC1 redeployment at a subset of SSEs coincides with increased enhancer marks and enhancer-promoter contacts at the larval stage.

Fig. S13. UAS-*ph* RNAi clones phenocopy *ph* mutant clones.

Fig. S14. The transcriptional response to *ph* depletion varies upon chromatin types.

Fig. S15. Canonical Polycomb sites enriched for H2AK118Ub show higher H3K4me1 levels and higher contact frequencies in ED.

Fig. S16. Removal of the two PRC1 anchors at the *dac* locus dampens its expression, while its expression pattern remains unchanged.

Fig. S17. RING1B is redeployed at active enhancer sites in NPCs.

Fig. S18. Clustering of ChIP-seq and Hi-C replicates from ED and whole *Drosophila* embryos.

Table S1. Chromatin types from *Drosophila* ED.

Table S2. Gene ontologies of the top interacting aTSS bins identified on the basis of PH ChIP-seq enrichment.

Table S3. Stringent set of PRC1 peak summits used for the clustering in ED.

Table S4. List of the genes assigned to each cluster of PRC1 peaks in ED.

Table S5. FPKM of *Drosophila* genes for the different conditions used in this study.

Table S6. Comparative Gene Ontology analysis of the genes assigned to each cluster of PRC1 peaks in ED.

Table S7. Differential gene expression analyses during development and in polycomb mutant Eds.

Table S8. List of the antibodies used in the study.

Table S9. PCR primers used in the study.

Table S10. NGS datasets used in the study.

[View/request a protocol for this paper from Bio-protocol.](#)

## REFERENCES AND NOTES

- B. Schuettengruber, H. M. Bourbon, L. Di Croce, G. Cavalli, Genome regulation by polycomb and trithorax: 70 years and counting. *Cell* **171**, 34–57 (2017).
- V. Loubiere, A. M. Martinez, G. Cavalli, Cell fate and developmental regulation dynamics by polycomb proteins and 3D genome architecture. *BioEssays* **41**, e1800222 (2019).
- V. Loubiere, A. Delest, A. Thomas, B. Bonev, B. Schuettengruber, S. Sati, A.-M. Martinez, G. Cavalli, Coordinate redeployment of PRC1 proteins suppresses tumor formation during *Drosophila* development. *Nat. Genet.* **48**, 1436–1442 (2016).
- M. Pherson, Z. Misulovin, M. Gause, K. Mihindukulasuriya, A. Swain, D. Dorsett, Polycomb repressive complex 1 modifies transcription of active genes. *Sci. Adv.* **3**, e1700944 (2017).
- L. Morey, A. Santanach, E. Blanco, L. Aloja, E. P. Nora, B. G. Bruneau, L. D. Croce, Polycomb regulates mesoderm cell fate-specification in embryonic stem cells through activation and repression mechanisms. *Cell Stem Cell* **17**, 300–315 (2015).
- I. Cohen, D. Zhao, C. Bar, V. J. Valdes, K. L. Dauber-Decker, M. B. Nguyen, M. Nakayama, M. Rendl, W. A. Bickmore, H. Koseki, D. Zheng, E. Ezhkova, PRC1 fine-tunes gene repression and activation to safeguard skin development and stem cell specification. *Cell Stem Cell* **22**, 726–739.e7 (2018).
- K. P. Eagen, E. L. Aiden, R. D. Kornberg, Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 8764–8769 (2017).
- Y. Ogiyama, B. Schuettengruber, G. L. Papadopoulos, J.-M. Chang, G. Cavalli, Polycomb-dependent chromatin looping contributes to gene silencing during *Drosophila* development. *Mol. Cell* **71**, 73–88 (2018).
- A. Mortazavi, S. Pepke, C. Jansen, G. K. Marinov, J. Ernst, M. Kellis, R. C. Hardison, R. M. Myers, B. J. Wold, Integrating and mining the chromatin landscape of cell-type specificity using self-organizing maps. *Genome Res.* **23**, 2136–2148 (2013).
- G. J. Filion, J. G. van Bommel, U. Braunschweig, W. Talhout, J. Kind, L. D. Ward, W. Brugman, I. J. de Castro, R. M. Kerkhoven, H. J. Bussemaker, B. van Steensel, Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* **143**, 212–224 (2010).
- H.-G. Lee, T. G. Kahn, A. Simcox, Y. B. Schwartz, V. Pirrotta, Genome-wide activities of Polycomb complexes control pervasive transcription. *Genome Res.* **25**, 1170–1181 (2015).
- J. Jacobs, M. Atkins, K. Davie, H. Imrichova, L. Romanelli, V. Christiaens, G. Hulsemans, D. Potier, J. Wouters, I. I. Taskiran, G. Paciello, C. B. González-Blas, D. Koldere, S. Aibar, G. Halder, S. Aerts, The transcription factor Grainy head primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. *Nat. Genet.* **50**, 1011–1020 (2018).
- T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, H. Parrinello, A. Tanay, G. Cavalli, Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
- E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, J. Dekker, Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
- K. Davie, J. Jacobs, M. Atkins, D. Potier, V. Christiaens, G. Halder, S. Aerts, Discovery of transcription factors and regulatory regions driving in vivo tumor development by ATAC-seq and FAIRE-seq open chromatin profiling. *PLoS Genet.* **11**, e1004994 (2015).
- D. J. McKay, J. D. Lieb, A common set of DNA regulatory elements shapes *Drosophila* appendages. *Dev. Cell* **27**, 306–318 (2013).
- J. O. Yanez-Cuna, C. D. Arnold, G. Stampfel, L. M. Boryi, D. Gerlach, M. Rath, A. Stark, Dissection of thousands of cell type-specific enhancers identifies dinucleotide repeat motifs as general enhancer features. *Genome Res.* **24**, 1147–1156 (2014).
- C. D. Arnold, D. Gerlach, C. Stelzer, L. M. Boryi, M. Rath, A. Stark, Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* **339**, 1074–1077 (2013).
- J. Erceg, T. Pakozdi, R. Marco-Ferreres, Y. Ghavi-Helm, C. Girardot, A. P. Bracken, E. E. M. Furlong, Dual functionality of cis-regulatory elements as developmental enhancers and Polycomb response elements. *Genes Dev.* **31**, 590–602 (2017).
- A. Rada-Iglesias, R. Bajpai, T. Swigut, S. A. Brugmann, R. A. Flynn, J. Wysocka, A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
- J. Curtiss, M. Burnett, M. Mlodzik, distal antenna and distal antenna-related function in the retinal determination network during eye development in *Drosophila*. *Dev. Biol.* **306**, 685–702 (2007).
- T. L. Davis, I. Rebay, Master regulators in development: Views from the *Drosophila* retinal determination and mammalian pluripotency gene networks. *Dev. Biol.* **421**, 93–107 (2017).
- T. Kondo, K. Isono, K. Kondo, T. A. Endo, S. Itoharu, M. Vidal, H. Koseki, Polycomb potentiates *meis2* activation in midbrain by mediating interaction of the promoter with a tissue-specific enhancer. *Dev. Cell* **28**, 94–101 (2014).
- B. Bonev, N. M. Cohen, Q. Szabo, L. Fritsch, G. L. Papadopoulos, Y. Lubling, X. Xu, X. Lv, J.-P. Hugnot, A. Tanay, G. Cavalli, Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**, 557–572.e24 (2016).
- S. L. Kloet, M. M. Makowski, H. I. Baymaz, L. van Voorthuysen, I. D. Karemaker, A. Santanach, P. W. T. C. Jansen, L. D. Croce, M. Vermeulen, The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation. *Nat. Struct. Mol. Biol.* **23**, 682–690 (2016).
- R. S. Stowers, T. L. Schwarz, A genetic method for generating *Drosophila* eyes composed exclusively of mitotic clones of a single genotype. *Genetics* **152**, 1631–1639 (1999).
- V. Loubiere, A. Delest, B. Schuettengruber, A.-M. Martinez, G. Cavalli, Chromatin immunoprecipitation experiments from whole *drosophila* embryos or larval imaginal discs. *Bio Protoc.* **7**, e2327 (2017).

28. H. Strutt, R. Paro, The polycomb group protein complex of *Drosophila melanogaster* has different compositions at different target genes. *Mol. Cell. Biol.* **17**, 6773–6783 (1997).
29. F. Bantignies, G. Cavalli, Topological organization of *Drosophila* Hox genes using DNA fluorescent in situ hybridization. *Methods Mol. Biol.* **1196**, 103–120 (2014).
30. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
31. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
32. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
33. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoutte, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
34. F. Ramirez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
35. W. J. Kent, A. S. Zweig, G. Barber, A. S. Hinrichs, D. Karolchik, BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204–2207 (2010).
36. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
37. P. Stempor, J. Ahringer, SeqPlots - Interactive software for exploratory data analyses, pattern discovery and visualization in genomics. *Wellcome Open Res.* **1**, 14 (2016).
38. A. Pohl, M. Beato, bwtool: A tool for bigWig files. *Bioinformatics* **30**, 1618–1619 (2014).
39. T. Kohonen, Self-organizing neural projections. *Neural Netw.* **19**, 723–733 (2006).
40. R. Wehrens, L. M. C. Buydens, Self- and Super-organizing Maps in R: The kohonen Package. *2007* **21**, 19%J Journal of Statistical Software (2007).
41. T. Nagano, Y. Lubling, C. Várnai, C. Dudley, W. Leung, Y. Baran, N. M. Cohen, S. Wingett, P. Fraser, A. Tanay, Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017).
42. N. M. Cohen, P. Olivares-Chauvet, Y. Lubling, Y. Baran, A. Lifshitz, M. Hoichman, A. Tanay, SHAMAN: bin-free randomization, normalization and screening of Hi-C matrices. *bioRxiv* **187203**; doi: <https://doi.org/10.1101/187203>.
43. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
44. S. Anders, P. T. Pyl, W. Huber, HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
45. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
46. D. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
47. G. Yu, L.-G. Wang, Y. Han, Q.-Y. He, clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
48. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

**Acknowledgments:** We thank L. Fritsch and M. Fitz-James for helping in the preparation of FISH probes and for reading and correcting this manuscript, respectively. We thank the Montpellier GenomiX (Montpellier, France) sequencing facility, the BioCampus *Drosophila* facility (Montpellier, France) and the MRI (Montpellier, France) imaging facility for technical support. **Funding:** Research in the laboratory of G.C. was supported by the CNRS, the University of Montpellier, and by grants from the European Research Council (ERC-2008-AdG No 232947 and 2017-AdG No 788972), the European Union's Horizon 2020 research and innovation program under grant agreement no. 676556 (MuG), the Agence Nationale de la Recherche (N. ANR-15-CE12-006-01), the Fondation pour la Recherche Médicale (grant no. DEI20151234396), the French National Cancer Institute (INCA, project no. PLBIO16-222), and the Laboratory of Excellence EpiGenMed (no. ANR-10-LABX-12). A.-M.M. is supported by a grant of the Fondation ARC pour la Recherche sur le Cancer (contract no. PJA 20171206301). V.L. is supported by a doctoral fellowship from the Laboratory of Excellence EpiGenMed and the Fondation ARC pour la Recherche sur le Cancer. Q.S. is supported by a doctoral fellowship from the French Ministry of Higher Education and Research and la Ligue nationale contre le cancer. **Author contributions:** V.L. performed molecular biology, genomics, microscopy, and genetics experiments. V.L. analyzed the ChIP-seq and RNA-seq data. G.L.P. performed the analyses of Hi-C data. Q.S. analyzed 3D FISH experiments. A.-M.M. and G.C. jointly supervised the work. V.L., A.-M.M., and G.C. wrote the manuscript, and all the authors commented on and corrected the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials, and with regard to the genomic data generated for this study, they have been deposited in GEO under the accession number GSE126985. Additional data related to this paper may be requested from the authors.

Submitted 20 March 2019  
Accepted 6 November 2019  
Published 10 January 2020  
10.1126/sciadv.aax4001

**Citation:** V. Loubiere, G. L. Papadopoulos, Q. Szabo, A.-M. Martinez, G. Cavalli, Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping. *Sci. Adv.* **6**, eaax4001 (2020).

## Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping

V. Loubiere, G. L. Papadopoulos, Q. Szabo, A-M. Martinez and G. Cavalli

*Sci Adv* 6 (2), eaax4001.  
DOI: 10.1126/sciadv.aax4001

ARTICLE TOOLS	<a href="http://advances.sciencemag.org/content/6/2/eaax4001">http://advances.sciencemag.org/content/6/2/eaax4001</a>
SUPPLEMENTARY MATERIALS	<a href="http://advances.sciencemag.org/content/suppl/2020/01/06/6.2.eaax4001.DC1">http://advances.sciencemag.org/content/suppl/2020/01/06/6.2.eaax4001.DC1</a>
REFERENCES	This article cites 46 articles, 10 of which you can access for free <a href="http://advances.sciencemag.org/content/6/2/eaax4001#BIBL">http://advances.sciencemag.org/content/6/2/eaax4001#BIBL</a>
PERMISSIONS	<a href="http://www.sciencemag.org/help/reprints-and-permissions">http://www.sciencemag.org/help/reprints-and-permissions</a>

Use of this article is subject to the [Terms of Service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).