PALI1 facilitates DNA and nucleosome binding by PRC2 and triggers an allosteric activation of catalysis

Qi Zhang1,3, Samuel C. Agius1,3, Sarena F. Flanigan1, Michael Uckelmann1, Vitalina Levina1, Brady M. Owen1 & Chen Davidovich1,2✉

The polycomb repressive complex 2 (PRC2) is a histone methyltransferase that maintains cell identities. JARID2 is the only accessory subunit of PRC2 that known to trigger an allosteric activation of methyltransferase. Yet, this mechanism cannot be generalised to all PRC2 variants as, in vertebrates, JARID2 is mutually exclusive with most of the accessory subunits of PRC2. Here we provide functional and structural evidence that the vertebrate-specific PRC2 accessory subunit PALI1 emerged through a convergent evolution to mimic JARID2 at the molecular level. Mechanistically, PRC2 methylates PALI1 K1241, which then binds to the PRC2-regulatory subunit EED to allosterically activate PRC2. PALI1 K1241 is methylated in mouse and human cell lines and is essential for PALI1-induced allosteric activation of PRC2. High-resolution crystal structures revealed that PALI1 mimics the regulatory interactions formed between JARID2 and EED. Independently, PALI1 also facilitates DNA and nucleosome binding by PRC2. In acute myelogenous leukemia cells, overexpression of PALI1 leads to cell differentiation, with the phenotype altered by a separation-of-function PALI1 mutation, defective in allosteric activation and active in DNA binding. Collectively, we show that PALI1 facilitates catalysis and substrate binding by PRC2 and provide evidence that subunit-induced allosteric activation is a general property of holo-PRC2 complexes.
over the course of evolution, gene families tend to expand1. Accordingly, the number of genes linked to the same function is commonly increased in vertebrates with respect to invertebrates, especially in cases of genes coding for transcriptional regulators3,4,5. For instance, the histone H3K4 methyltransferase MLL/COMPASS complex expanded from one in yeast to three sub-types in fly and six complexes in human, where additional subunits emerged through each expansion1. Similar expansion took place for the histone ubiquitin ligase polycomb repressive complex 1 (PRC1): from two complexes in the fly to at least six variants in vertebrates5,6. Vertebrate-specific subunits of histone modifiers provide the opportunity to identify molecular mechanisms that are fundamental to chromatin biology and, therefore, re-emerged through the course of evolution.

The polycomb repressive complex 2 (PRC2) is a histone methyltransferase complex that is required for the maintenance of cell identity in all multicellular organisms. At the molecular level, PRC2 maintains the repressed state of developmentally expressed genes through the tri-methylation of lysine 27 in histone H3 (H3K27me3), a hallmark of facultative heterochromatin6,7. The core PRC2 complex includes four subunits8–10, but it has a low histone methyltransferase activity and low affinity to DNA. Therefore, holo-PRC2 complexes include additional protein subunits—termed accessory subunits. Most of the accessory subunits of the vertebrate PRC2 emerged through gene duplication and some are vertebrate specific, with the latter poorly understood mechanistically.

The accessory subunits are collectively required for the recruitment of PRC2 to chromatin and for the regulation of its enzymatic activity11,12. Unbiased proteomic studies13–17 revealed two distinct holo-PRC2 complexes—PRC2.1 and PRC2.2—defined by their mutually exclusive accessory subunits16. The PRC2.2 complex is nearly identical in fly and human, and includes the accessory subunits AEBP2 and JARID2. Contrarily to PRC2.2, the PRC2.1 complex went through a massive expansion over evolution: from one accessory subunit in fly, to at least five in vertebrates13–17. The fly PRC2.1 accompanies a single accessory subunit: Pcl. The vertebrate PRC2.1 is far more complex: it contains one of the three polycomb-like (PCL) proteins16 (PHF1, MTF2 or PHF19) together with either EPOP13,16 or PALI116,18 (also annotated as LCOR-CRA_b, LCOR isoform 3 or C100RF12). Recent works indicate some non-redundant functions of the PRC2.1 and PRC2.2 complexes in mouse embryonic cells11,12, but the molecular basis is unknown.

The PRC2.2-specific subunit JARID2 has two activities that were implicated in nucleating H3K27me3: chromatin binding19,20 and allosteric stimulation of histone methyltransferase (HMTase)21. During JARID2-induced allosteric activation, PRC2 first di- or tri-methylates lysine 116 in JARID2 (JARID2-K116me2/3). Next, the di/tri-methyl-lysine binds to the regulatory subunit EED and triggers an allosteric activation of PRC221. This mechanism is thought as a “jump-start” to activate PRC22. After the nucleation of H3K27me3, histone tails carrying the H3K27me3 mark bind to the regulatory subunit EED22, to trigger further allosteric activation of PRC223. Yet, knockout of JARID2 in mouse ESC cells lacks major effect on H3K27me3 globally and locally11,21. This implies a parallel role taken by JARID2 in mouse ESC cells lacks major effect on H3K27me3, a hallmark of facultative heterochromatin6,7. The PRC2 maintains the repressed state of developmentally expressed genes through the tri-methylation of lysine 27 in histone H3 (H3K27me3), a hallmark of facultative heterochromatin6,7.

RESULTS

PALI1 K1241 is methylated in mouse and human cell lines. In a search for a PRC2.1 accessory subunit that could trigger an allosteric activation of PRC2, we first set out to map the PRC2 methylome in mouse and human cells. We reasoned that affinity purification of PRC2 followed by tandem mass spectrometry (AP-MS) will allow for the identification of methyl-lysines residing within assembled PRC2 complexes in vivo. Hence, we analysed multiple publicly available liquid chromatography with tandem mass spectrometry (LC-MS/MS) data originating from AP-MS experiments, where PRC2 subunits were used as baits15,26–29 (Fig. 1a and Supplementary Data 1). Although these studies15,26–29 were not focused on the methylation of PRC2 subunits, the high quality of the raw data allowed us to detect methyl-lysines in tryptic peptides (Supplementary Data 1). As expected, the PRC2 methylome contains the previously reported methylations in JARID2 K11621 and EZH2 K514 and K51530–32.

The two most frequently detected di- and tri-methyl lysines in the accessory subunits of PRC2 were in JARID2 K116 and PALI1 K1241 (Fig. 1a), with the former triggering an allosteric activation of PRC221. Specifically, PALI1 K1241 is methylated in five of the seven cell lines that were tested, including both human and mouse cell lines: HEK293T (human embryonic kidney), STS26T (human malignant peripheral nerve sheath tumour), LnCaP (human prostate cancer), U2OS (human osteosarcoma) and mouse embryonic stem cells (mESC). In four of these cell lines, PALI1 K1214 was identified either in its di- or tri-methyl form (i.e. PALI1 K1214me2/3).

In some of the cell lines, we also identified methylations in PALI1 K1214 and K1219, in agreement with a previous proteomic analysis in HCT116 cells30. The same study also identified methylations of EZH2 K510 and K515 that have more recently been shown to regulate PRC231,32. Yet, K1241 was not identified in that study30, which used antibodies against methyl-lysines for immunoaffinity purification ahead of the mass spectrometry30.

Hence, PALI1 and JARID2 were the only accessory subunits that were identified with di- or tri-methyl-lysine modifications that are evolutionarily conserved in mouse and human (Fig. 1a).
**Fig. 1 PALI1 is methylated in vitro and in vivo.**

**a** Schematic representation of the PRC2 methylome in vivo and in vitro, as identified from MS/MS data. Mouse and human icon represent the organism of origin and cell lines are indicated (see Methods section for references and accession numbers of the raw MS/MS data). Test tube icons representing methylations in the purified recombinant human PRC2-[MBP-PALI1]-long. PALI truncations used in this study are indicated in purple (upper right). Residues are indicated with asterisks where the position probability of the methylation is less than 0.95 (see Supplementary Data 1 for the values). PRC2.1 and PRC2.2 accessory subunits are in blue and red, respectively, and core subunits are in grey. Coomassie blue-stained SDS-PAGE of recombinant human PRC2-PALI1PIR complexes, as indicated. HMTase assay of the PRC2-[MBP-PALI1PIR]-long complex using mononucleosome substrates (Fig. 1c) confirmed that PALI1PIR is methylated. The MBP-cleaved and uncleaved PALI1PIR bands are indicated on the radiogram with asterisks. The HMTase assays were carried out three times with similar results and a representative gel is presented.

**Di- and tri-methylated JARID2 K116 (JARID2 K116me2/3) allosterically activates the PRC2.2 complex through direct interactions with the regulatory subunit EED.** Based on this observation, we set out to test the hypothesis that methyl-lysines in PALI1 allosterically activate the PRC2.1 complex.

**PALI1 K1241 is methylated by PRC2 in vitro.** We next set out to determine if PALI1 can be methylated by PRC2. We expressed and purified a recombinant human PRC2 (EZH2, EED, SUZ12 and RBBP4) in a complex with the PRC2-interacting region from PALI1 (PALI1PIR; Fig. 1b). In vitro HMTase assay using mononucleosome substrates (Fig. 1c) confirmed that the PRC2-PALI1PIR complex, comprising amino acids 1058–1250 from PALI1 (purple bar in Fig. 1a), is substantially more active than the core PRC2 complex. While this result is in agreement with previous reports, we also noted an additional band on the radiogram (Fig. 1c, marked with an asterisk). That band indicated a methylated protein that appeared only if PRC2 contained PALI1PIR and migrated with the apparent molecular weight of PALI1PIR. In order to confirm that the methylated protein is
indeed PALI1$_{P_{IR}}$, we purified a PRC2-PALI1$_{P_{IR}}$ complex with a 3C-cleavable MBP tag carried only by PALI1$_{P_{IR}}$. We then performed the HMTase assay in the presence or absence of human rhinovirus 3C protease. The 3C-specific cleavage of the MBP-tag on PALI1$_{P_{IR}}$ led to a large shift in the migration velocity of the methylated protein, confirming it is indeed PALI1$_{P_{IR}}$ (Fig. 1c, lane 2 versus lane 3).

A similar result was obtained when we performed the same experiment using a longer truncation of PALI1 that was designed based on the previous mapping of the PRC2-interacting region from PALI1 (PALI1 1058-1329; termed PALI1$_{P_{IR-long}}$ herein; long purple bar in Fig. 1a). PALI1$_{P_{IR-long}}$ co-purified with PRC2 as a soluble complex (Supplementary Fig. 1a, b), albeit in multiple truncated forms. In-gel digestion with mass spectrometry subsequently identified PALI1$_{P_{IR}}$ as a more stable truncation of PALI1$_{P_{IR-long}}$ while both constructs co-purified with PRC2 and enhanced HMTase (Fig. 1c and Supplementary Fig. 1c). These experiments confirmed that PALI1$_{P_{IR}}$ is sufficient to enhance the HMTase activity of PRC2 towards mononucleosome substrates and that PRC2 methylates PALI1$_{P_{IR}}$ in vitro. While the full-length PALI1 (PALI1FL) was not as stable as the PALI1$_{P_{IR-long}}$ construct (Supplementary Fig. 1d, e), it was methylated by PRC2 and enhanced methyltransferase (Supplementary Fig. 1f). In order to identify the methylated amino acids within PALI1, we performed LC-MS/MS analysis of the recombinant PRC2-PALI1$_{P_{IR-long}}$. As expected, we detected the previously reported methyl-lysines in EZH2 K514 and K515 (Supplementary Data 1). In PALI1$_{P_{IR-long}}$, we identified mono- and di-methylations in K1214, K1219 and K1241 (Fig. 1a, Supplementary Data 1 and Source Data file for the MS/MS spectra), in agreement with our proteomic analysis of in vivo AP-MS data (Fig. 1a). Additionally, K1214 and K1219 were detected also in tri-methyl forms. These methylations were identified either if the complex was pre-incubated with SAM or if not, indicating that a significant fraction of the complex was purified with these modifications (Supplementary Data 1). A similar observation was previously made for EZH2 automethylation, that occurs in the recombinant protein while co-expressed with other PRC2 subunits. These results confirm that three lysines within the PRC2-interacting region of PALI1 serve as a substrate for PRC2, including K1214, K1219 and K1241.

**PALI1 K1241 is required in order to enhance the HMTase activity of PRC2.** If methyl-lysines in PALI1 allosterically activate PRC2, the corresponding lysine residues are expected to be required for PALI1-mediated enhancement of HMTase. To determine if methyl-lysines in PALI1 are required or dispensable for PALI1-mediated enhancement of HMTase, we expressed and purified PRC2-PALI1$_{P_{IR}}$ and PRC2-PALI1$_{P_{IR-long}}$ mutant complexes, including all possible perturbations of the PALI1 mutations K1214A, K1219A or K1241A. Mutant complexes migrated on a gel filtration column similar to the wild type, excluding adverse effects on complex solubility (Supplementary Fig. 2b). To assess the ability of these lysine-to-alanine mutants to enhance HMTase, we carried out an in vitro HMTase assay using mononucleosomes as a substrate. The mutation K1241A in both PRC2-PALI1$_{P_{IR}}$ and PRC2-PALI1$_{P_{IR-long}}$ lead to ~50% reduction in HMTase activity, compared to the wild-type complexes (Fig. 2a and Supplementary Fig. 2c, d). The two other lysine-to-alanine mutants, K1214A and K1219A, did not affect the HMTase activity of the complexes (Fig. 2a and Supplementary Fig. 2c, d). The K1241A mutant significantly reduced PALI1-mediated enhancement of HMTase in all possible perturbations that we tested, while the other two lysines, K1214 and K1219, were dispensable for methyltransferase enhancement.

**PALI1-K1241me2/3 is sufficient in order to stimulate the HMTase activity of PRC2.** If the methylation of PALI1 K1241 is sufficient to trigger an allosteric activation of PRC2, we expected to mimic these regulatory interactions by using a short peptide, including a tri-methyl-lysine K1241 flanked by amino acids of the corresponding sequence from PALI1 (termed PALI1-K1241me3 peptide herein). Indeed, the PALI1-K1241me3 peptide significantly stimulated the HMTase activity of PRC2 towards mononucleosome substrates (Fig. 2b). Similar observations were made in the past for H3K27me3 and JARID2-K116me2/3 peptides, which allosterically activate PRC2. We also assayed a PALI1-K1219me3 peptide, which has a smaller positive effect on the HMTase activity of PRC2. The PALI1-K1241me3 peptide was ineffective in stimulating PRC2 (Fig. 2b). Another peptide, including tri-methylations on both K1219 and K1241, exhibited only a moderate stimulation of HMTase (Fig. 2b). As expected, no stimulation of HMTase observed by unmethylated wild-type or lysine-to-arginine mutant K1241 peptides that were used as negative controls (Fig. 2c).

The level of PALI1 K1241 methylation ranges from mono- to tri-methyl-lysine in different human and mouse cell lines (Fig. 1a and Supplementary Data 1). To determine how the methylation level of PALI1 K1241 affects its ability to stimulate PRC2, we produced a PRC2-PALI1$_{P_{IR-long}}$ complex with a 3C-cleavable MBP tag carried only by PALI1$_{P_{IR}}$, which we termed PALI1$_{P_{IR-long}}$ (Fig. 1a). PALI1$_{P_{IR-long}}$ co-purified with PRC2 as a soluble complex (Supplementary Fig. 1a, b), albeit in multiple truncated forms. In-gel digestion with mass spectrometry subsequently identified PALI1$_{P_{IR}}$ as a more stable truncation of PALI1$_{P_{IR-long}}$ while both constructs co-purified with PRC2 and enhanced HMTase (Fig. 1c and Supplementary Fig. 1c). These experiments confirmed that PALI1$_{P_{IR}}$ is sufficient to enhance the HMTase activity of PRC2 towards mononucleosome substrates and that PRC2 methylates PALI1$_{P_{IR}}$ in vitro. While the full-length PALI1 (PALI1FL) was not as stable as the PALI1$_{P_{IR-long}}$ construct (Supplementary Fig. 1d, e), it was methylated by PRC2 and enhanced methyltransferase (Supplementary Fig. 1f). In order to identify the methylated amino acids within PALI1, we performed LC-MS/MS analysis of the recombinant PRC2-PALI1$_{P_{IR-long}}$. As expected, we detected the previously reported methyl-lysines in EZH2 K514 and K515 (Supplementary Data 1). In PALI1$_{P_{IR-long}}$, we identified mono- and di-methylations in K1214, K1219 and K1241 (Fig. 1a, Supplementary Data 1 and Source Data file for the MS/MS spectra), in agreement with our proteomic analysis of in vivo AP-MS data (Fig. 1a). Additionally, K1214 and K1219 were detected also in tri-methyl forms. These methylations were identified either if the complex was pre-incubated with SAM or if not, indicating that a significant fraction of the complex was purified with these modifications (Supplementary Data 1). A similar observation was previously made for EZH2 automethylation, that occurs in the recombinant protein while co-expressed with other PRC2 subunits. These results confirm that three lysines within the PRC2-interacting region of PALI1 serve as a substrate for PRC2, including K1214, K1219 and K1241.

**PALI1-K1241me2/3 binds to the aromatic cage of the regulatory subunit EED to stimulate PRC2.** H3K27me3 and JARID2-K116me2/3 bind to the aromatic cage of the regulatory subunit EED to allosterically stimulate PRC2. We therefore wished to determine if there is a direct link between PALI1-K1241me2/3 and EED. We performed fluorescence anisotropy displacement titrations (Fig. 3a). As a positive control, we first quantified the affinity of an H3K27me3 peptide for EED, resulting with $K_d = 41.7 ± 2.8 \mu M$, in agreement with a previous study. The affinity of EED for PALI1-K1241me3 ($K_d = 7.49 ± 0.59 \mu M$) is similar to the affinity of EED for the JARID2-K116me3 peptide ($K_d = 8.07 ± 0.49 \mu M$; Fig. 3a and Supplementary Fig. 3a), consistent with previously published results. Then, we quantified the $K_d$ of unlabelled peptides, using fluorescence anisotropy displacement titrations (Fig. 3a). As a positive control, we first quantified the affinity of an H3K27me3 peptide for EED, resulting with $K_d = 41.7 ± 2.8 \mu M$, in agreement with a previous study. The affinity of EED for PALI1-K1241me3 ($K_d = 7.49 ± 0.59 \mu M$) is similar to the affinity of EED for the JARID2-K116me3 peptide ($K_d = 8.07 ± 0.49 \mu M$; Fig. 3a and Supplementary Fig. 3a). We observed weak interaction between mono-methylated K1241 peptide to EED ($K_d = 241 ± 38 \mu M$), while the di-methyl form increased the affinity for EED by ~10-fold ($K_d = 19.0 ± 2.3 \mu M$), almost to the level of the K1241.
**Fig. 2** PALI1 K1241 is required and PALI1-K1241me2/3 is sufficient to stimulate the HMTase activity of PRC2. 

**a** HMTase assays were carried out using 500 nM of wild-type or mutant recombinant complexes, as indicated, using 2 µM mononucleosomes substrate. The bar plot (right) represents mean HMTase activities, quantified using densitometry and normalised to the activity of the wild-type PRC2-PALI1PIR (upper line) and the core PRC2 (bottom line) complexes. 

**b** HMTase assay performed with 500 nM PRC2, 2 µM mononucleosomes and in the presence or absence of either 10 µM or 100 µM PALI1 peptide, as indicated. The bar plot (right) represents the relative HMTase activities of PRC2 in the presence of tri-methylated (blue) or K-to-A mutated (grey) PALI1 peptides, as indicated. 

**c** HMTase assays of PRC2 performed as above, in the presence or absence of PALI1-K1241 peptides, as indicated. 

**d** HMTase assays of PRC2 performed as above, in the presence or absence of PALI1-K1241 peptides with different methylation states, as indicated. The bar plots in **b–d** represents the relative HMTase activities, normalised to the HMTase activity of PRC2 in its basal state (dashed line). The numbers on the left side of the gels and radiograms in this figure represent the molecular weight marker in kDa. The bar plots in all panels represent the mean of the quantification performed using densitometry over three independent replicates. Error bars shown in this Figure represent standard deviation with the observed values plotted as dots. Uncropped gel images used to generate this figure are in Supplementary Fig. 2. Source data for this figure are provided as a Source Data file.
tri-methyl-lysine peptide ($K_d = 7.49 \pm 0.59 \mu M$) (Fig. 3a). In agreement with these binding assays, a titration experiment confirmed that the HMTase activity of the PRC2 core complex reached a saturation at a PALI1-K1241me3 peptide concentration of slightly above $K_d$ (Fig. 3b, c). Qualitatively, the results of these binding assays (Fig. 3a) are also in agreement with the HMTase assays done in the presence of the other peptides (Fig. 2d) and support a model where PALI1 K1241me2/3 binds to the aromatic cage in EED to stimulate the HMTase activity of PRC2.

The PALI1-K1219me3 peptide binds to EED with high affinity ($K_d = 7.59 \pm 0.72 \mu M$; Fig. 3a), in agreement with its ability to stimulate PRC2 (Fig. 2b). Contrarily, the PALI1-K1214me3 peptides, which did not stimulate methyltransferase (Fig. 2b), binds to EED with a 4-fold lower affinity comparing PALI1-K1219me3 and PALI1-K1241me3 (Fig. 3a).

In order to directly link between PALI1-K1214me3 and the aromatic cage of EED within the context of PRC2, we reconstituted mutant PRC2 complex harbouring the defective cage mutation EED F97A. The PALI1-K1241me3 peptide did not lead to the activation of the cage-mutant PRC2 (Fig. 3d, e), in agreement with an EED-dependent allosteric activation of PRC2. On the same line of evidence, PALI1-K1241me3-induced activation of PRC2 was inhibited by an allosteric inhibitor of PRC2, A-395, but not the negative control A-395N (Supplementary Fig. 3). Collectively, our data support a mechanism where PALI1 K1241me2/3 binds to the aromatic cage in EED to trigger an allosteric activation of PRC2.

PALI1 and JARID2, but not H3, utilise the same interactions with the regulatory subunit EED. Given the functional identity between PALI1-K1214me3 and -K1219me3 to JARID2-K116me3, we wished to assess for structural resemblance. We, therefore, solved the crystal structures of EED76-441 co-crystallised with a PALI1-K1241me3 or PALI1-K1219me3 peptide (Fig. 4a and Table 1). We

Fig. 3 PALI1-K1241me3 binds to the aromatic cage of EED to stimulate PRC2 activity. a Fluorescence anisotropy displacement titrations, where unlabelled peptides competed a 5-FAM labelled JARID2-K116me3 peptide (40 nM) for binding to EED (10 μM). Each data points in the plot indicate the mean of the normalised anisotropy values and the error bars represent standard deviation over three independent replicates that were carried out on different days. Dissociation constants ($K_d$) and 95% confidence bounds on the coefficient are indicated in the table. The sequence of the peptides indicated in the table, with the methyl-lysine in red. See Supplementary Fig. 3a for the binding curve of the 5-FAM labelled JARID2-K116me3 to EED. b HMTase assay performed using PRC2 in the presence or increased concentrations of PALI1-K1241me3 peptides. c Data from b is presented, with the activity of the peptide-stimulated PRC2 is normalised to the activity of PRC2 in its basal state. Each data points indicate the mean of the relative HMTase activities from the three independent replicates, and the error bars represent standard deviations. The maximum activity of PRC2 is represented by a horizontal line. The vertical dashed line represents the $K_d$ of PALI1-K1241me3 peptide and EED, as quantified in the table, with the methyl-lysine in red. See Supplementary Fig. 3a for the binding curve of the 5-FAM labelled JARID2-K116me3 to EED.

| Key | Peptide | Sequence | $K_d$ (μM) |
|-----|---------|----------|-----------|
| PALI1-K1241me0 | JHHLKKFPGATY | n.d. |
| PALI1-K1214me1 | JHHLKK(me1)FPGATY | 241 ± 38 |
| PALI1-K1214me2 | JHHLKK(me2)FPGATY | 19.0 ± 2.3 |
| PALI1-K1214me3 | JHHLKK(me3)FPGATY | 7.49 ± 0.59 |
| PALI1-K1217me3 | DVPVK(me3)HPQK | 34.1 ± 2.6 |
| PALI1-K1219me3 | KHPLQK(me3)YAPSSY | 7.59 ± 0.72 |
| H3K27me3 | TKAAK(me3)SPATY | 41.7 ± 2.8 |
| 5-FAM labelled-JARID2-K116me3 | RLQAQRK(me3)FAQQS | 8.07 ± 0.49 |

6 NATURE COMMUNICATIONS | (2021) 12:4592 | https://doi.org/10.1038/s41467-021-24866-3 | www.nature.com/naturecommunications
compared the resulted structures with the crystal structures of the EED-H3K27me3 and EED-JARID2-K116me3 complexes. The structures indicate that the two tri-methyl-lysine PALI1 peptides (Fig. 4a, left two panels, and Supplementary Fig. 4a) bind to EED in a conformation resembling that seen for the JARID2-K116me3 peptide (Fig. 4a, the third panel). Specifically, in all three cases the tri-methyl-lysine and its adjacent aromatic residue, in the +1 position, adopting the same conformation when binding to EED (Fig. 4a, marked in dashed shapes). Contrarily, H3K27me3, does not have an aromatic residue at position +1, with respect to the tri-methyl-lysine, thus adopts a different binding mode to EED (Fig. 4a, right). The essential role of the phenylalanine at the +1 position has been previously demonstrated in the case of JARID2. Accordingly, a F1242A mutation in the PALI1-K1241me3 peptide prevents it from binding.
implies a biological signiﬁcant and the adjacent F1242, but not PALI1-K1214 or PALI1-K1219, are PRC2 across vertebrates. Of note PALI1-K1241 these data and structures indicate that PALI1 and JARID2 interact residue, despite no other sequence similarity and no common with EED using their tri-methyl-lysine and its adjacent aromatic 

Table 1 X-ray crystallography data collection and reﬁnement statistics.

|                        | PALI1-K1241me3 (PDB:6V3X) | PALI1-K1219me3 (PDB:6V3Y) |
|------------------------|---------------------------|---------------------------|
| Data collection        |                            |                           |
| Space group            | P 2,2,2,2                  | P 2,2,2,2                  |
| Cell dimensions        | a, b, c (Å)                | a, b, c (Å)                |
| Resolution (Å)         | 47.76-1.70                 | 47.76-1.70                 |
| Resolution (Å)         | 0.066 (0.631)              | 0.074 (0.643)              |
| Redundancy             | 16.5 (2.8)                 | 16.2 (3.2)                 |
| Completeness (%)       | 97.8 (96.5)                | 99.7 (97.9)                |
| Refinement             |                            |                           |
| Resolution (Å)         | 35.2-1.70                  | 48.8-1.63                  |
| No. reﬂections         | 47039 (4553)               | 56904 (5630)               |
| Rmerge/Reﬁnement      | 0.174/0.200                | 0.165/0.193                |
| Protein (chain A)      | 2850                      | 2898                      |
| Ligand (chain B)       | 39                        | 42                        |
| Water                  | 178                       | 267                       |
| B-factors              |                            |                           |
| Protein                | 22.9                      | 18.0                      |
| Ligand/ion             | 29.1                      | 24.6                      |
| Water                  | 30.1                      | 27.4                      |
| R.M.S. deviations      | Bond lengths (Å)           | Bond angles (°)           |
|                        | 0.006                     | 0.006                     |
|                        | 0.84                      | 0.89                      |
| Ramachandran plot      | Favoured regions (%)      | Disallowed regions (%)    |
|                        | 96.6                      | 3.4                       |
|                        | 96.4                      | 3.6                       |
|                        | Allowed regions (%)       | Disallowed regions (%)    |
|                        | 3.4                       | 0.0                       |
|                        | 3.6                       | 0.0                       |

Values in parentheses are for the highest-resolution shell.

to EED (Fig. 4c) and from stimulating PRC2 (Fig. 4d, e). Altogether, these data and structures indicate that PALI1 and JARID2 interact with EED using their tri-methyl-lysine and its adjacent aromatic residue, despite no other sequence similarity and no common ancestor (Fig. 4b and Supplementary Fig. 4b). Of note PALI1-K1241 and the adjacent F1242, but not PALI1-K1214 or PALI1-K1219, are fully conserved across different vertebrate species (Fig. 4b). This implies a biological signiﬁcance of PALI1-K1241me2/3 in regulating PRC2 across vertebrates.

PALI2 K1558 mimics PALI1 K1241 in the allosteric regulation of PRC2 in vitro. PALI1 is mutually exclusive with EPOP for PRC2 binding13,16 and share high sequence similarity with PALI18. We did not identify methyl lysines in EPOP (Fig. 1a) and have no factual basis to consider it as an allosteric regulator of PRC2, in agreement with the role of EPOP as a negative regulator of PRC2 in cells13,33. When it comes to PALI2, sequence alignment revealed that PALI2 K1558 and its adjacent phenylalanine are identical to the corresponding amino acids in PALI1: PALI1 K1241 and its adjacent phenylalanine (Fig. 5a). Accordingly, a PALI2-K1558me3 peptide, but not PALI1-K1558 (unmethylated), binds to EED with high afﬁnity (Kd = 8.81 ± 1.48 µM, Fig. 5b) and stimulates the enzymatic activity of PRC2 (Fig. 5c). Next, we reconstituted the recombinant PRC2-PALI21330-1641 complex, with the PALI2 construct designed based on its homology to PALI1156. Despite being less stable than PALI1156 (Supplementary Fig. 5b, c), the recombinant PRC2-PALI21330-1641 complex exhibited high enzymatic activity compared to PRC2 (Fig. 5d) and a small but notable level of PALI2 methylation observed (Supplementary Fig. 5d). Altogether, these results indicate that PALI2 may regulate PRC2 in a similar mechanism as identiﬁed above for PALI1. It is worth noting that PALI2 was not detected as an interacting partner of PRC2 in most of the AP-MS dataset that we analysed, with the only two exceptions where in human U2OS and mouse testis, where in both cases PALI2 detected with a poor coverage (0.6% and 6.3–12.8%, respectively). The poor coverage of PALI2 in these datasets was not allowed identifying its methylation. While it remains to be determined under what circumstances PALI2 regulates PRC2 in vivo, our data indicates that it has the capacity to bind PRC2 and to trigger an allosteric activation.

PALI1 facilitates DNA binding by PRC2, with allosteric activation being dispensable for this function. The triple mutant complex, PRC2-PALI1PIR-long K1241A, K1219A and K1241A, is defective in allosterically stimulating HMTase but was still more active than the core PRC2 complex (Fig. 2a and supplementary Fig. 2c, d). This mutant complex cannot harbour any of the methylation that we identiﬁed in PALI1 (Fig. 1a). Hence, we suspected that the PRC2-interacting region within PALI1 regulates HMTase in an additional mechanism. In vitro HMTase assays previously demonstrated that the HMTase activity of PRC2 is enhanced by several of its DNA-binding accessory subunits, including MTF2, PHF19 and AEBP230. We therefore wished to determine if the PRC2-interacting region of PALI1 increases the afﬁnity of PRC2 to DNA.

To directly test if the PRC2-interacting region of PALI1 can facilitate DNA binding, we ﬁrst set out to quantify the afﬁnity of PRC2-PALI1PIR to DNA using ﬂuorescence anisotropy. We used a DNA probe designed to mimic 46 bases long dsDNA from a CpG island of the CDKN2B gene (termed CpG46 DNA, see Supplementary Table 1 for the DNA sequence). The afﬁnity of the PRC2-PALI1PIR to CpG46 DNA (Kd = 155 ± 26 nM) was >20-fold higher than the afﬁnity to the PRC2 core complex to the same DNA probe (Kd > 4 µM) (Fig. 6a), indicating that PALI1 facilitates DNA binding. Accordingly, Chromatin immunoprecipitation (ChIP) with qPCR conﬁrmed that ectopically expressed PALI1 binds to chromatin in cells (Supplementary Fig. 10), in agreement with a previous study that identiﬁed PALI1 in the chromatin-bound nuclear fraction18. In vitro, PALI1156 binds DNA even in the absence of PRC2, although with a reduced afﬁnity (Kd = 2.3 ± 0.3 µM; Supplementary Fig. 6c). The PALI1156 construct can enhance the HMTase activity of the core PRC2 even if they were not co-expressed together but rather combined just before the reaction (Supplementary Fig. 2h).

The DNA-binding activity of PALI1 was not speciﬁc to the CpG46 DNA probe: PRC2-PALI1PIR binds to DNA tightly even after the DNA probe was mutated to disrupt all the CpG sequences and to reduce the GC content from 79 to 21% (CpG46 DNA-probe, see Methods section for DNA sequence), with PALI1PIR did not signiﬁcantly increase the afﬁnity of PRC2 to a G-tract RNA (Supplementary Fig. 6a), which interacts with core PRC2 subunits36,37. Accordingly, this data indicates that the PRC2-interacting region of PALI1 facilitates high-afﬁnity interactions between PRC2 and DNA, not RNA, without an apparent DNA-sequence selectivity. Some level of target speciﬁcity might be achieved in cells by additional factors. For instance, PCL proteins were implicated in binding to CpG islands and can bind to PRC2 together with PALI118.
mononucleosomes are shifted (for quantification, see Supplementary Fig. 5). Source data for b, c are provided as a Source Data file.

Fig. 6 PALI1 facilitates DNA binding by PRC2. a Fluorescence anisotropy used to quantify the affinity of PRC2 complexes to fluorescein-labelled CpG46 or CpG46 mt DNA. Data represent the mean of three independent experiments that were carried out on different days and error bars represent standard deviation. Dissociation constants ($K_d$) and Hill coefficients are indicated in the table, including their standard error. Fluorescence anisotropy values are in arbitrary units (a.u.). b EMSA used to quantify the affinity of the indicated PRC2 complexes for a mixture of Cy5-labelled mononucleosomes and free DNA of the same sequence. Dashed boxes indicate the mononucleosome bands near the centred positions of nucleosomes, at either the centred or the off-centred positions, or between PRC2 and the free DNA (Fig. 6b).

In agreement with a previous work, the PRC2 core complex exhibited moderate affinity for mononucleosomes ($K_d = 330 \pm 30$ nM). Remarkably, the PRC2PIR-long construct increased the affinity of PRC2 to nucleosomes by >15-fold ($K_d = 19.0 \pm 0.6$ nM) compared to the PRC2 core complex (Supplementary Fig. 6b). To determine if PALI1 enhances substrate binding in a mechanism linked to allosteric activation, we quantified the affinity of the
concentrations were determined in the table. Days and error bars represent standard deviation. For the substrate quality are in Supplementary Fig. 7. Source data for this figure are provided as a Source Data file.

To directly examine the contribution of PALI1 to the high-affinity interactions between PRC2-PALI1PIR and nucleosome. Adding a DNA linker slightly increased the affinity of PRC2-PALI1PIR to the nucleosome constructs (Fig. 7b; $K_d = 17.1 \pm 0.6 \, \text{nM}$ for PRC2-PALI1PIR when using the a NCP182 probe), in agreement with our observation that PALI1 facilitates DNA binding (Fig. 6).

Given the high affinity of PALI1 to DNA (Fig. 6) and nucleosomes (Fig. 7), we anticipated that it would facilitate the chromatin-binding activity of PRC2 during histone methyltransferase. To test this, we next carried out quantitative histone methyltransferase assays (Fig. 7c) using a nucleosome array. In agreement with the function of PALI1 in allosterically activating PRC2 (Figs. 2–4), $k_{cat}$ of PRC2-PALI1PIR was over 7-fold higher than that of the core complex (Fig. 7c). Most importantly, $K_m$ values of the PRC2-PALI1PIR complex were almost insensitive to the salt concentration, with <2-fold increment in $K_m$ while comparing low salt to high salt reaction buffers ($K_m = 224 \pm 20.6 \, \text{nM}$ in 35 mM KCl and 392 ± 57.8 nM in 100 mM KCl; Fig. 7c).

Accordingly, the catalytic efficiency of PRC2-PALI1PIR was almost the same in high or low salt buffers (1.1-fold change,

**Fig. 7 PALI1 facilitates nucleosome binding by PRC2.**

**a** EMSA used to quantify the affinity of the indicated PRC2 complexes for H2A-Cy5-labelled nucleosome core particles substrate (top, NCP147) and mononucleosomes with 35 bp linker DNA substrate (bottom, NCP182). In all, 2-fold dilutions of the protein were carried out, starting from 2 μM PRC2-PALI1PIR and 4 μM PRC2. **b** Quantification of the EMSA from **a**. Data represent the mean of three independent experiments and the error bars represent standard deviations. Standard errors of dissociation constants ($K_d$) and Hill coefficients are indicated in the table. **c** Michaelis-Menten kinetic analysis of PRC2-PALI1PIR (15 nM) and PRC2 complex (50 nM) on a nucleosomal array substrate. Colour key for the plot and KCl concentrations are indicated in the table. Data represents the mean of three independent experiments that were carried out on different days and error bars represent standard deviation. $k_{cat}$, $K_m$, and the catalytic efficiency ($k_{cat}/K_m$) values are indicated with standard errors. Substrate concentrations were defined by the octamer concentrations in the arrays. Progress curves confirmed that the reaction is at the linear range and evidence for the substrate quality are in Supplementary Fig. 7. Source data for this figure are provided as a Source Data file.
see Fig. 7c for $k_{at}/K_M$ values). This suggests that the electrostatic potential is not the key driver for the interactions between PRC2-PALI1 in K562 chromatin during methyltransferase. Contrarily, the PRC2 core complex was very sensitive to the salt concentration, where increasing the KCl concentration from 35 mM to 100 mM significantly reduced its activity (Fig. 7). Collectively, our qualitative histone methyltransferase assays (Figs. 2–4), quantitative histone methyltransferase kinetic assays (Fig. 7c), the nucleosome binding assays (Fig. 7a, b) and the DNA binding assays (Fig. 6) support a mechanism where the PRC2-binding domain of PALI1 allosterically activates PRC2 while anchoring it on chromatin through interactions with DNA and nucleosomes to ensure a close proximity to the histone tail substrates.

Overexpression of PALI1 triggers cell differentiation in chronic myeloid leukaemia cells, with the phenotype altered by an allosteric-defective mutant. One paper reported a large increment of global H3K27me3 in HeLa cells after C00RF12 overexpression32. Yet, we did not detect a significant change of global H3K27me3 while overexpressing PALI1 in K562 (Supplementary Fig. 8a), HEK293T (Supplementary Fig. 8b, left) and HeLa (Supplementary Fig. 8b, right) cells. Little is known about the cellular function of PALI1, but our data thus far indicates a resemblance to JARID2 at the molecular level. JARID2 is frequently deleted in the leukaemic transformation of chronic myeloid malignancies40. Accordingly, the overexpression of JARID2 leads to reduced proliferation in leukaemia cell lines and it has been proposed to serve as a tumour suppressor in leukaemia41. Deletion of the LCOR locus has been reported in B-cell acute lymphoblastic leukaemia42. Given the functional resemblance between PALI1 to JARID2 and their potential role as tumour suppressors in hematopoietic malignancies, we wished to determine if PALI1 has a negative effect on the proliferation of myelogenous leukaemia cells. Indeed, competitive proliferation assays indicated that the overexpression of PALI1, but not the negative control LacZ, in a human chronic myeloid leukaemia cell line (K562) leads to a strong reduction in cell proliferation (Fig. 8a, b).

In addition to reduced proliferation, we noticed that the overexpression of PALI1, but not the negative control LacZ, in K562 cells led to the pelleted cells becoming red in colour (Fig. 8c). This observation suggested differentiation along the erythroid lineage43, in accord with the reduced cell proliferation (Fig. 8a, b). We therefore set out to detect the erythroid differentiation marker CD235a, which increases during erythropoiesis44, and the erythroid precursor marker CD44 that resembles between PALI1 to JARID2: (i) Both PALI1 and JARID2 have no common ancestor: PALI1 is a 193 amino acids in PALI1 (1058–1250) are sufficient to bind PRC2, bind DNA and stimulate methyltransferase, with the other >1500 amino acids in PALI1 likely available to engage in other tasks.

Discussion

Our data indicate that the PRC2-interacting domain of PALI1 is sufficient to enhance the HMTase activity of PRC2 by two independent mechanisms: (i) allosteric activation of catalysis (Fig. 2) and (ii) DNA binding (Figs. 6 and 7). Hence, as little as 193 amino acids in PALI1 (1058–1250) are sufficient to bind PRC2, bind DNA and stimulate methyltransferase, with the other >1500 amino acids in PALI1 likely available to engage in other tasks.

A convergent evolution between PALI1 and JARID2. The nucleation of H3K27me3 de novo takes place when transcription programmes are changed during cell differentiation and newly repressed genes acquire the H3K27me3 mark. In the context of the PRC2.2, JARID2 facilitates the nucleation of H3K27me324, aided by its chromatin-binding activity19,20,46 and its ability to allosterically stimulate PRC221. In the context of PRC2.1, MTF2 functions in the nucleation of H3K27me324, with a proposed contribution to its DNA-binding activity47,48. Although MTF2 does not allosterically stimulate PRC2, it coexists in the PRC2.1 complex with PALI116,18,25. Through the discovery that PALI1 allosterically activates PRC2 (Figs. 2 and 3) and facilitates substrate binding (Figs. 6 and 7), we reveal a striking functional resemblance between PALI1 to JARID2: (i) Both PALI1 and JARID2 bind to nucleosomes: the former likely with the aid of its DNA binding activity (Fig. 6) and the latter through interactions with H2AK119-ubiquitinated chromatin19,20,46. (ii) JARID2 comprises the PRC2.2 complex together with the chromatin-binding subunit AEBP2 while PALI1 binds to the PRC2.1 complex together with a polycomb-like DNA binding subunit (PHF1, MTF2 or PHF19)16,18,25, (iii) PRC2 methylates PALI1 and JARID2, with the di- or tri-lysine then binds to EED for allosteric activation of PRC2 (Figs. 2 and 3) and Sanulli et al.21). (iv) For their interactions with EED, both PALI1 and JARID2 using an aromatic residue, located at the +1 position with respect to the methylated lysine (Fig. 4a). The importance of the +1 adjacent aromatic residue supported by the JARID2 F117A mutation that prevents both EED binding and the stimulation of PRC225 and data herein using the PALI1 F1242A mutant peptide (Fig. 4c–e).

Strictly, despite these mechanistic and structural similarities, PALI1 and JARID2 have no common ancestor: PALI1 is a vertebrate-specific protein18 and JARID2 is conserved in fly and human (Fig. 4b). Therefore, we propose that PALI1 has emerged in vertebrates as the result of convergent evolution, under a selection pressure to mimic some of the molecular functions of JARID2 within the context of the PRC2.1 complex.
PALI1 provides PRC2 with means to gauge its own enzymatic activity before adding a stimulus. EZH2 automethylation is proposed to modulate the HMTase activity of PRC2 in response to molecular cues, including the presence of histone H3 tails and SAM concentration. A similar principle might apply for PALI1. This mechanism seems to affect mainly the methylation of histone substrates, not the automethylation of EZH2. Specifically, a similar level of EZH2 methylation observed in the presence of the methyl-defective mutants PALI1 K1214A, K1219A and K1241A or the presence of PALI1-methyl peptides (Supplementary Fig. 2c, e). Future studies are still needed in order to identify whether PALI1 (and JARID2) are methylated in cis or by another PRC2 complex and if the degree of methylation changes during normal development or in cancer and other pathologies. Yet, our analysis indicates that PRC2 has the capacity to methylate PALI1 (Fig. 1b, c) and that the degree of PALI1 K1241 methylation varies between cell lines (Fig. 1a). Our data also indicates that PRC2-induced methylation of PALI1 in vitro is more efficient in the presence of nucleosome substrates (Supplementary Fig. 2h).

Given these observations, it is plausible that PALI1 provides PRC2 with means to gauge its own enzymatic activity before applying an additional stimulus.

PALI1 as a potential bridge between the H3K27me3 and H3K9me2 repressive marks. Amino acids in regions of PALI1 dispensable for the regulation of PRC2 bind to the H3K9-methyltransferase G9a. PRC2 and G9a share a portion of their genomic targets and are physically associated. In ES cells, G9a contributes to H3K27 methylation in vivo, with the global H3K27me1—but not H3K27me2/3—reduced upon knockout of G9a. Affinity purification mass spectrometry (AP-MS) with either C10ORF12 or PALI1 used as baits detected the sub-units of the G9a-GLP H3K9me1/2 methyltransferase complex, including G9a, GLP and WIZ. Importantly, the PRC2-interacting domain of PALI1 is distinct to its G9a-interacting region. Our data indicate that the PRC2-binding region of PALI1 is sufficient (i) to bind to PRC2 (Fig. 1), (ii) to promote DNA and nucleosome
Proteins binding to the N-terminal domain of PALI1 are likely to be attributed to interactions with the G9a complex and the CTBP1, which share most of their target genes\(^1\), with their accessory subunits and other PRC2 core subunits in orange. Dashed red arrows represent methylation, with the red stars represent methyl-lysines. The blue arrow represents H3K9me2 de novo during cell differentiation, it does have the required for the stability of PALI1 in cells (Supplementary Table S1). Small variation between the activity of the full-length sequences encoding for human EZH2, SUZ12, RBBP4, EED and AEBP2 (UniProtKB: Q15910-2, Q15022-1, Q99028-1, Q75530-1 and Q6ZN18-2, respectively) into a pFastBac1 expression vectors, modified to include either a PreScission-cleavable N-terminal histidine-MBP tag or TEV-cleavable N-terminal hexahistidine tag, were previously described\(^{16,25,38}\). Full-length PALI1 was subcloned and subcloned into the pFB1.HMBP.A3.Prs.ybbR vector using Gibson Assembly (primers as indicated in Table S1) with gene synthesized N-terminal fragment (amino acids 1-310) of PALI1 (GeneScript) and commercially available C1000bp-PCR3 primer design encoding PALI1 amino acids 311-1557 (Müllingen Science #MHS6278-202756878) as templates (as seen in Supplementary Table 1 for the full-length PALI1 ORF sequence). The PIR (amino acids 1058-1250) and PIR-long (amino acids 1058-1329) fragments of PALI1 were subcloned into the pFB1.HMBP.A3.Prs.ybbR vector digested with Xnal and Xhal (NEB), under a PreScission-cleavable N-terminal histidine-MBP tag, using Gibson Assembly (NEB #E2611L) with primers as indicated in Supplementary Table 1.

Mutations were introduced to plasmids coding for PALI1 and its truncations using Takara PrimeSTAR HS DNA Polymerase (Clontech #R045A), with primers indicated in Supplementary Table 1. Baculovirus production, titration, infection, and cell harvesting and the purification of PCR2, PCR2-PALI1tag, PCR2-PALI1long and their mutants performed as previously described\(^6\). The expression and purification of PCR2 in complexes with MBP-fused PALI1 truncations, PCR2-[MBP-PALI1tag] and PCR2-[MBP-PALI1long] performed as above, with the exception that PCR2 core subunits were expressed under TEV-cleavable histidine-tag and PALI1 truncations under PreScission-cleavable N-terminal histidine-MBP tag. During the purification of these constructs, only TEV was used to cleave tags selectively from the PCR2 core subunits, with the MBP tag on the PALI1 construct remaining intact. All the complexes were snap-frozen in liquid nitrogen and stored at −80 °C as single-use aliquots.

For the structure-function study of EED, two fragments of human EED (amino acids 40-441 and 76-441) were subcloned into a pgEX-MLH expression vector with a TEV-cleavable N-terminal GST-tag (a gift from the lab of Asst. Prof. Yufeng Tong, University of Windsor) using primers as in Supplementary Table 1. The recombinant proteins were overexpressed in E.coli BL21 (DE3) at 17 °C overnight and then purified by Glutathione-agarose (Sigma #G68410). Briefly, harvested cells were resuspended in an ice-cold lysis buffer (20 mM Tris-HCl pH 7.5 at 25 °C, 250 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1 mM DTT) and lysed using sonication. The cleared lysate was batch-bound to Glutathione-agarose and washed using ice-cold 10 column volumes (v/c) of lysis buffer before proteins were eluted using ice-cold elution buffer (20 mM Tris-HCl pH 7.5 at 25 °C, 150 mM NaCl, 10 mM reduced glutathione). Tag cleaved using TEV, overnight at 4 °C. The protein was subsequently purified by heparin HP column (GE #170-40701), using a buffer containing 20 mM Tris-HCl pH 7.5 at 4 °C and a 150–200 mM NaCl gradient. Gel filtration purification carried out using HiLoad 16/600 Superdex 200 size exclusion column (GE #F28-9993-35), using a buffer containing 20 mM HEPES pH 7.5 and 150 mM NaCl. The peak fractions were pooled, concentrated to a buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine (TCEP) and snap-frozen as single-use aliquots.

For measuring the DNA binding of PALI1tag, The PIR (amino acids 1058-1250) fragment of PALI1 were subcloned into the pMAL-mlh vector (a gift from the lab of Dr. Yufeng Tong, University of Windsor) digested with BsrR (NEB), under a TEV-cleavable N-terminal MBP tag, using Gibson Assembly (NEB #E2611L) with primers as indicated in Supplementary Table 1. The recombinant protein was overexpressed in E.coli BL21 (DE3) at 17 °C overnight and then purified by Amylose resin (NEB #E8021). The tag was removed and further purified using the same procedure and buffer from EED purification with the exception that the gel filtration purification was carried out using Superdex 75 Increase 10/300 GL column. The peak fractions were pooled, concentrated to a buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM TCEP and snap-frozen as single-use aliquots.

For purifying PALI1tag from the pFB1.HMBP.A3.Prs.ybbR vector, Baculovirus production, titration, infection and cell harvesting was performed as previously described\(^6\). The purification was performed as previously described\(^6\), with the exception that the N-terminal hexahistidine-MBP tag was retained and the gel filtration purification was carried out using Superdex 75 Increase 10/300 GL column.

**Methods**

**Protein expression and purification.** The cloning of constructs for the expression of the full-length sequences encoding for human EZH2, SUZ12, RBBP4, EED and AEBP2 (UniProtKB: Q15910-2, Q15022-1, Q99028-1, Q75530-1 and Q6ZN18-2, respectively) into a pFastBac1 expression vectors, modified to include either a PreScission-cleavable N-terminal histidine-MBP tag or TEV-cleavable N-terminal hexahistidine tag, were previously described\(^{16,25,38}\). Full-length PALI1 was assembled and subcloned into the pFB1.HMBP.A3.Prs.ybbR vector using Gibson Assembly (primers as indicated in Table S1) with gene synthesized N-terminal fragment (amino acids 1-310) of PALI1 (GeneScript) and commercially available C1000bp-PCR3 primer design encoding PALI1 amino acids 311-1557 (Müllingen Science #MHS6278-202756878) as templates (as seen in Supplementary Table 1 for the full-length PALI1 ORF sequence). The PIR (amino acids 1058-1250) and PIR-long (amino acids 1058-1329) fragments of PALI1 were subcloned into the pFB1.HMBP.A3.Prs.ybbR vector digested with Xnal and Xhal (NEB), under a PreScission-cleavable N-terminal histidine-MBP tag, using Gibson Assembly (NEB #E2611L) with primers as indicated in Supplementary Table 1.

Recent studies showed that PRC2.1 and PRC2.2 synergise and share most of their target genes\(^11\), with their accessory subunits collectively required\(^12\). While a previous attempt to identify the binding sites of PALI1 on chromatin using ChIP was reported as unsuccessful\(^{16}\), the PRC2.1 complex is localised at a minority of unique genomic sites, independently of PRC2.2\(^11\). We measured small variation between the affinity of the PRC2-PALI1PIR complex for the CpG-reach DNA (CpG46 \(K_d = 155 ± 26\) nM) comparing a size-matched DNA lacking CpG sequences (CpG46 mt; \(K_d = 73.7 ± 10\) nM) (Fig. 6a). It is, therefore, possible that PALI1 could allow for some degree of target specificity, utilising some variations in affinities for DNA combined with context-specific chromatin binding. Such context-specific binding could be attributed to interactions with the G9a complex and the CTBP proteins binding to the N-terminal domain of PALI1\(^{16}\). This model is in agreement with the view that a combination of factors and interactions are responsible for the recruitment of PRC2 to its target genes\(^1\).

Collectively, our data indicate that the PRC2-binding domain of PALI1 enhances H3K27-methyltransferase by two independent mechanisms (Fig. 9): (i) DNA- and nucleosome-binding and (ii) allosteric stimulation. The remarkable mechanistic resemblance between PALI1 and JARID2 indicates convergent evolution of the regulation of the PRC2.1 and PRC2.2 complexes, respectively. More broadly, it implies that subunit-induced allosteric activation is an indispensable property of a homo-PRC2 complex in vertebrates.
reactants were then diluted away by repeat concentrating the solution 25× times using 4 mL Amicon Ultra 10,000 MWCO, while topping up the solution by labelling Buffer A (Promega) in a final volume of 4 mL (Promega P8978). The final DNA was then labelled using PCR as described above, except that a5'-Cy5 linked forward primer was used. The template for amplifying 147-base-pair DNA was obtained from Addgene (plasmid #26656) and the template for the amplification of the 182-base-pair DNA was obtained from gene synthesis, with both sequences and the sequences of the primers used for their PCR amplification are indicated in Supplementary Table 1. Purification was performed via HiTrap Q HP column (GE #17-1154-01) with 0.1 M gradient starting with buffer A (20 mM Tris-HCl, pH 7.5 at 25 °C, 150 mM NaCl) into 500 mM buffer B (20 mM Tris-HCl, pH 7.5 at 25 °C, 2 M NaCl). The peak fractions were pooled, DNA was purified by ethanol precipitation and was then dissolved in TE buffer.

Nucleosomes were reconstituted using the salt dialysis method. Reconstituted mononucleosomes were dialysed against a buffer consisting of 20 mM Tris-HCl pH 7.5 (at 25 °C), 25 mM KCl, 1 mM EDTA and 1 mM DTT and concentrated with Amicon Ultra-0.5 ml centrifugal filter (Merck #UCS03906). Nucleosomes were stored at 4 °C and the quality of nucleosomes was assessed by 5-6% TBE gel.

For the generation of nucleosomal arrays, a genomic region (GRCh37/hg19, chr9:47478129-97517591) including the ATOH1 gene was cloned into pUC18 vector linearised with restriction enzyme Smal (NEB #R0141). ATOH1 DNA was amplified in a large scale, typically of 10–15 ml, with homemade Pfu DNA polymerase with primers as indicated in Supplementary Table 1. The amplified DNA was purified by ion exchange chromatography using Q resin (GE Healthcare). Purified DNA was concentrated by precipitation using isopropanol and then dissolved in TE buffer.

Arrays were assembled using the gradient dialysis approach, starting with 0.7 mg/ml ATOH DNA and in the presence of 20-fold molar excess of histone octamers in a dialysis tubing (MWCO: 12–14 kDa, Spectra/Por 4 Dialysis Membrane). The starting buffer was 20 mM Tris-HCl pH 7.5, 2 mM NaCl, 1 mM EDTA, 1 mM DTT and the buffer was gradually exchanged into 250 mM KCl, 20 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT over 27 hours at 1 ml/min using peristaltic pump. The final step dialysis was performed into the 20 mM Tris pH 7.5, 2.5 mM KCl, 1 mM EDTA, 1 mM DTT. All dialysis were done at 4 °C. The concentration of the nucleosome core particles in the arrays was determined by measuring the concentration of the octamers using BCA assay (ThermoFisher #23252). Arrays were stored at 4 °C and the quality was assessed by 0.8% agarose TBE gel.

In vitro HMTase activity assays using radiolabelled S-adenosyl-l-methionine

HMTase activity assays were performed as previously described, with some modifications. In brief, each 10 µL HTMase reaction contained 500 nM PRC2, 2 µM mononucleosomes and 5.0 µM S-[(methyl-14C)]-adenosyl-l-methionine (PerkinElmer, #NEC363050UC) in the presence or absence of stimulatory or control peptide or 100 mM KCl as indicated in the text. Reactions were incubated at 30 °C, 25 µM SAM, 0.5 mM MgCl2, 0.1% Tween-20, 5 mM DTT and either 35 mM or 100 mM KCl as indicated in the text. Reactions were incubated at 30 °C for 2 h and quenched using 1 µl of 2.4% v/v TFA. For each enzyme and condition, an aliquot of the JARID2-K116me3 peptide for EED, various concentrations of EED (amino acids 76–441) were incubated with 40 nM of 5-FAM labelled JARID2-K116me3 peptide as internal standard. For each enzyme and condition, an aliquot of the MTase-Glo reaction mixture was added and luminescence was developed using reagents supplied with the MTase-Glo assay system. Fluorescence anisotropy data were collected using a FluoroMax-4 fluorescence spectrophotometer (Jobin Yvon) and the ratio of F/F0 was measured. The fluorescence anisotropy data were collected using MaxQuant by searching against a database containing the protein subunits of the PRC2-PALI1P21/peripheral complex. DNA binding assays using fluorescence anisotropy

3′-fluorescein-labelled CpG64 and CpG64 mt DNAs were synthesised by Integrated DNA Technologies, Inc. The CpG64 DNA probe was designed to form a hairpin including a 46 bases long double-stranded DNA with a nucleotide sequence originating from a CpG island of the human CDK6282 PRC2-target gene, with this sequence mutated in CpG64 mt DNA to remove the CpG sequences and reduce the GC content. For DNA probes sequence see Supplementary Table 1). The RNA-binding activities of proteins were quantified using fluorescence anisotropy. Experiments were carried out as we previously described. Brieﬂy, a 3′-fluorescein-labelled G4 24 RNA was hybridised with the sequence (UUAGGG)4 was used and the reaction took place as above, with the exception that the initial incubation at 95 °C was limited to 1 min and the binding buffer was 50 mM Tris-HCl pH 7.5 at 25 °C, 100 mM KCl, 2 mM 2-mercaptoethanol, 0.1 mg/ml BSA, 0.05% NP-40 and 0.1 mg/ml fragmented yeast tRNA (Sigma, #R836). Protein binding assays using fluorescence anisotropy

For assaying the affinity of the JARID2-K116me3 peptide for EED, various concentrations of EED (amino acids 40–441) were incubated with 40 nM of 5-FAM labelled JARID2-K116me3 peptide and quenched using 1 µl of 2.4% v/v TFA. For each enzyme and condition, an aliquot of the MTase-Glo reaction mixture was added and luminescence was developed using reagents supplied with the MTase-Glo assay system.
peptide in binding buffer (50 mM Tris-HCl pH 7.5 at 25 °C, 100 mM KCl, 2 mM 2-mercaptoethanol, 0.1 mg/mL BSA, 0.05% NP-40, 2.5% glycerol) at 30 °C for 30 min before a fluorescence anisotropy measurement took place using a PHERAScan plate reader. Data processing was carried out as previously described166, with some modifications. Specifically, with changing the concentration of EED protein (P), we recorded ΔRobs: the observable anisotropy of the mixtures after the subtraction of the observable anisotropy of the 5-FAM labelled JARID2-K116me3 peptide ligand. Data were fitted to equation (1), below, by using non-linear least-squares fit (Matlab, MathWorks) to estimate the anisotropy difference Δr and the dissociation constant between the protein EED to the ligand JARID2-K116me3 peptide, Kc:

$$\Delta R_{obs} = \frac{(K_c + [P]_0 + [L]_0) - \sqrt{(K_c + [P]_0 + [L]_0)^2 - 4([P]_0 - [L])}}{2}$$

$$[L] = ([L]_0 - [PL])$$

where the concentrations of the protein-ligand complex [PL] and the free ligand [L] are calculated from Eqs. (2) and (3) below, respectively, with $P_0$ and $L_0$ indicating the total concentration of the protein and the ligand, respectively.

$$[PL] = \frac{([L]_0 - [PL])}{K_c + [P]}$$

Fluorescence anisotropy displacement titrations were used to assay the dissociation constants of the unlabelled peptides (N) and the protein (P). Assays were carried out as described above, with the exception that 2-fold serial dilutions of unlabelled peptides were combined with EED at a final concentration of 10 µM and 5-FAM labelled JARID2-K116me3 peptide supplemented with 10% FBS (Cellzera AU-BFS/FS) and kept on ice. The data were recorded signals from the Cy5 dye. The fractions of bound Cy5-DNA were calculated based on the unbound nucleosomes band and the total concentration of the protein and the ligand, respectively. The reaction mixtures were incubated at 4 °C for 30 min and then subjected to centrifugation at 4 °C at 1000 g for 10 min on ice before adding NP-40 to a concentration of 1% (v/v) and mixing. Samples were centrifuged at 4 °C and 5000 g for 5 min and the supernatant was kept as the cytoplasmic fraction. The nuclear fraction was washed in cytoplasmic extraction buffer with 1% (v/v) NP-40 twice, by centrifugation at 4 °C at 5000 g for 5 min, and then resuspended in Laemmli buffer.

**Electrophoretic mobility shift assay.** Cy5-DNA or Cy5-H2A labelled nucleosomes were diluted using binding buffer (50 mM Tris-HCl pH 7.5 at 25 °C, 100 mM KCl, 2 mM 2-mercaptoethanol, 0.05% v/v NP-40, 0.1 mg/mL BSA, 5% glycerol) with additional 5% v/v glycerol for measuring binding for Cy5-H2A labelled nucleosomes). In total, 2-fold serial dilutions of protein in binding buffer were combined with nucleosomes probes, to a final probe concentration of 5 nM. The reaction mixtures were incubated at 4 °C for 30 min and then subjected to non-denaturing gel electrophoresis at 6.6 V/cm over a 0.7% agarose gel buffered with 1× TBE at 4 °C for 1.5 h. Gels were imaged using Typhoon 5 Imager (GE Healthcare) to record signals from the Cy5 dye. The fractions of bound Cy5-DNA labelled nucleosomes were calculated based on the unbound nucleosomes band and the fractions of bound Cy5-H2A labelled nucleosomes were calculated based on the bound nucleosomes band, with the densitometry analysis carried out using ImageJ. Data were fitted with GraphPad Prism software using non-linear regression for specific binding with Hill slope function. All experiments were performed in triplicate.

**Cell culture.** K562 cells were cultured in RPMI-1640 (Merck #R8758) growth medium and HEPES and Hela cells were cultured in DMEM growth medium. In all cell culture media were supplemented with 10% FBS (Cellzera AU-BFS/FS) and 1% (v/v) penicillin-streptomycin (Thermo Scientific #15140122) and incubated at 37 °C with 5% CO2. K562 and HeLa cells were acquired from ATCC and were tested periodically for mycoplasma contamination.

**Plasmid transfection, generation of lentiviruses and lentiviral transduction.** Flag-PALI1 WT and K1241A mutant, and Flag-LacZ (ORF originated from Addgene #25893), were subcloned into Sma1 (NEB #R01411) linearised pBluescript-EGFP (Addgene #21373) or pDHIV-dTomato (Addgene #21374) vectors using Gibson Assembly (see Table S1 for primers) and NEB stable Competent Ecoli (NEB #C3040). For plasmid transfection, 10-5 HEK293T or HeLa cells were seeded in a six-well plate. The following day, the medium was replaced with 2 ml of antibiotic-free DMEM. The transfection mixture contained 9 µl Lipofectamine® LTX Reagent with 3 µl PLUS™ Reagent (Thermo Scientific #15338100) and 3 µg of a plasmid to 500 µl Opti-MEM™ (Thermo Scientific #31980062). The transfection mixture was incubated at room temperature for 25 min and then added to the cells before returning them to the incubator. The growth medium was replaced after 24 h. For cell harvesting after immunoblotting, 48 h after transfection the media was removed and replaced with 100–400 µl of Laemmli buffer (1% (v/v) SDS, 10% (v/v) glycerol, 35 mM Tris-HCl pH 7.5 at 25 °C, 0.01% (w/v) bromophenol blue, 5 mM MgCl2, 1% (v/v) 2-mercaptoethanol) and 25 µM/µl Benzonase (Merck #70746). For the generation of lentiviruses, HEK293T cells were transfected as above, with 0.5 µg pMD2-G plasmid (Addgene #12259), 1 µg psPAX2 plasmid (Addgene #12600) and 1.5 µg of transfer plasmid. After 48 and 72 h, the culture supernatant containing the lentivirus was collected and stored at −80 °C. For lentiviral transduction, 250 µl of lentiviral supernatant was added to 3 x 104 K562 cells to a final volume of 550 µl, with polybrene at a final concentration of 8 µg/ml.

**Immunoblotting.** For nuclear fractionation of K562 cells, the cells were washed twice with phosphate-buffered saline (PBS) by centrifugation at 500 x g for 5 min then resuspended in cytoplasmic extraction buffer (20 mM Tris pH 7.5 at 25 °C, 0.1 mM EDTA, 2 mM MgCl2, 20 mM BME and protease inhibitor cocktail (Sigma #4639312001)) to a density of 2 x 106 cells/ml. The cells were incubated for 2 min at room temperature then 10 min on ice before adding NP-40 to a concentration of 1% (v/v) and mixing. Samples were centrifuged at 4 °C and 5000 g for 5 min and the supernatant was kept as the cytoplasmic fraction. The nuclear fraction was washed in cytoplasmic extraction buffer with 1% (v/v) NP-40 twice, by centrifugation at 4 °C at 5000 g for 5 min, and then resuspended in Laemmli buffer.

**Samples containing 50 µg total protein were loaded on a 10%, 16.5% or 8-16% acrylamide gel (Biorad #4561103) for SDS-PAGE and then transferred to a nitrocellulose membrane (GE Life Sciences #10600002). Membranes were incubated in blocking buffer (Thermo Scientific #37539) for 1 hour at room temperature before applying antibodies. Signal was detected using SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Scientific #34580) and images were taken on a ChemiDoc™ imager. All experiments were performed in triplicates.**

**The antibodies used for immunoblotting include: anti-Flag HRP-conjugated (Sigma #A8592, 1:1000), anti-LCOR (Merck #ABE1367, 1:250), anti-active EZH2 (Active Motif #39875, 1:1000), anti-EED (Abclonal #A12773, 1:5000), anti-H3 (Abcam #Ab1791, 1:100000), anti-H3K27me3 (Merck #07-449, 1:5000 or Cell Signalling #9733S, 1:4000), mouse anti-mouse HRP-conjugated (Santa Cruz Biotechnology #sc-2357, 1:5000).**

**Competitive cell proliferation assay.** Cells were transduced with lentiviruses carrying the specified gene constructs and were then cultured, each sample separately, for 7 days. At this point, an equal number of eGFP or dTomato positive cells were sorted by flow cytometry as described below, and a mix of the same collection tube and placed in the same well for the competition experiment. Cells of the two competing treatments were cultured together in the same well for 7 days. Next, the number of eGFP- and dTomato-expressing cells counted using flow cytometry with the B530-A and YG586-A detectors, respectively. Three independent biological replicates, starting from lentivirus transduction, were initiated on three different days and were carried out as described above.

**Detection of eGFP and dTomato using Flow cytometry.** Before sorting or analysis by flow cytometry, cells were centrifuged at 500 x g for 5 min, and the supernatant removed. The cells were then resuspended in flow cytometry buffer (PBS supplemented with 10% FBS and 615 µM EDTA) to a density of 106 cells/mL, and resuspended into a cell strainer (Falcon #352355) and incubation of GFP or Dtomato, the cells were sorted by flow cytometry on a BD Influx™ cell sorter using the 488 nm or the 561 nm lasers, respectively. For the selection of
transduced cells, and unless stated otherwise, gates for sorting were set to include the top 10 or 20% of the GFP or d’Tomato positive cells in the samples transduced with Palling or wild type, based on the first replicate. For the analysis of transduced cells, ~0.5–1.0 × 10^5 intact single cells were analysed on a BD LSRII Fortessa® X-20 analyser, where the threshold for GFP or d’Tomato positive cells was defined based on the intensity observed at the top 0.1% of untransduced K562 cells. Gating strategies used for cell sorting are shown in Supplementary Fig. 11. Data were analysed using BD FACSDiva™ and GraphPad Prism.

Detection of CD235a or CD44 using Flow Cytometry. Cells were transduced with lentiviruses carrying the specified gene constructs and were then cultured for 7 days. Cells were sorted for high GFP or d’Tomato expression, as described above, and then cultured for 7 additional days before the growth medium was removed by spinning the cells at 500 × g for 5 min. Cells at a density of 2 × 10^6 cells/mL were incubated on ice for 15 min in flow cytometry buffer with 2.5 µL of Pacific Blue™- conjugated anti-CD235a antibody (BioLegend #439108) per 100 µL for detection of CD235a, or 2.5 µL of APC or PE conjugated anti-CD44 antibody (Biologend #130301, Biolegend #130007) per 100 µL for the detection of CD44. The cells were centrifuged again at 500 × g for 5 minutes and the supernatant removed, then washed with antibody-free flow cytometry buffer. The cells were then analysed by flow cytometry for the quantification of CD235a or CD44 in the GFP or d’Tomato positive cells using the V450-A and R670-A or YG586-A detectors, respectively. Three independent biological replicates, starting from lentivirus transduction, were initiated on three different days and were carried out as described above. The data was analysed using FlowJo and GraphPad Prism.

Doxycycline-induced CRISPR/Cas9 knockout of EED. K562 cells were first transduced with lentiviruses generated using a vector for expression of Cas9 with bacterialin selection (Addgene #52962), and then selected using 10 µg/mL of bacterialin, and were continuously grown in the presence of that antibiotic. For the doxycycline-induced expression of the gRNA, DNA oligos with the sequence of the target site in EED (Supplementary Table 1) were annealed together and cloned as a NotI, KpnI fragment into the pSpCas9(2A)PGK-puro plasmid (Addgene #12324) (Andrews et al., 2013). The plasmid was packed into lentiviruses, as described above, which were then used to transduce K562 cells that were next selected using flow cytometry, based on EGFP expression. Next, the gRNA was used to induce 3 µg/mL of doxycycline (Merck #63131).

Chromatin immunoprecipitation. HEK293T cells were transfected in one 16 cm dish per two IP samples, with the transfection carried out by scaling the transfection process described above (in “Plasmid transfection, generation of lentiviruses and lentiviral transduction”). The next day, the cells were passed into 15 cm dishes. In all, 48 h after transfection, the cells were washed with room temperature PBS, and then crosslinked with 1% (v/v) formaldehyde in PBS for 10 min at room temperature. The formaldehyde was quenched by adding glycine to a final concentration of 1.0 M. After 5 min, cells were washed twice in diluted PBS, and then crosslinked with 1% (v/v) formaldehyde in PBS for 10 min at room temperature, and the two eluates were combined. The protein was collected and resuspended in PBS at a concentration of 1 mg/mL. The DNA was fragmented using sonication at 4 °C in a Bioruptor set to high, with 20 cycles of 30 s on and 30 s off. The samples were then centrifuged to remove large DNA at 10,000 × g for 5 min, and washed once using PBS. After the PBS was removed, the cell pellets were stored at −80 °C.

DNA was quantified using qPCR in a 10 µL reaction volume in a Biorad CFX384™ Real-Time C1000 Touch™ Thermal Cycler. Each reaction contained 2.5 µL of template DNA that was diluted 1:10, 1.25 µL of 2.8 µM forward primer, 1.25 µL of 2.8 µM reverse primer (0.7 µM final concentration per primer), 5 µL QuantNova SYBR Green 2X master mix (Quant No. #208054) and was prepared using Qiagen QIAbility. For each biological replicate, Cq values from two technical replicates were averaged from both the input and IP samples, and the following formula was used to calculate the fraction of DNA in the IP: 2^(-(Cq[IP]−Cq[input])) × (input volume)/(IP volume). Three independent biological replicates, starting from transfection, were initiated on three different days and were carried out as described above. The data was analysed using FlowJo and GraphPad Prism.

The antibodies used for ChIP include: anti-LCOR (Merck #ABE1367, 3 µg), anti- EZH2 (Active Motif #59875, 6 µg), anti-H3K27me3 (Cell Signalling #9733 S, 10 µL) and anti-Flag (Merck #F1804, 3 µg).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 6V3X and 6Y3T. The mass spectrometry data have been deposited to Monash University research repository Figshare, with doi.org/10.26180/14752/599. Source data are provided with this paper.

Received: 20 April 2020; Accepted: 7 July 2021; Published online: 28 July 2021

References

1. Holland, L. Z. et al. The amphibious genome illuminates vertebrate origins and ecephalochordate biology. Genome Res. 18, 1100–1111 (2008).
2. Prachumwat, A. & Li, W. H. Gene number expansion and contraction in vertebrate genomes with respect to invertebrate genomes. Genome Res. 18, 221–232 (2008).
3. Larroux, C. et al. Genesis and evolution of metazoan transcription factor gene classes. Mol. Biol. Evol. 25, 966–978 (2008).
4. Shilatifard, A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annu. Rev. Biochem. 81, 65–95 (2012).
5. Schreurnann, J. C., Gutierrez, L. & Muller, J. Histone H2A monoubiquitination and Polycomb repression: the missing pieces of the puzzle. Fly 6, 162–168 (2012).
6. Schuetzengruber, B., Bourbon, H. M., Di Croce, L. & Cavalli, G. Genome regulation by polycomb and trithorax: 70 years and counting. Cell 171, 34–57 (2017).
7. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. Nature 469, 343–349 (2011).
8. Yu, J. B., Lee, C. H., Oksuz, O., Stafford, J. M. & Reinberg, D. PRC2 is high maintenance. Genes Dev. 33, 903–935 (2019).
9. Laugesen, A., Hoffeldt, J. W. & Helin, K. Molecular mechanisms directing PRC2 recruitment and H3K27 methylation. Mol. Cell. 74, 8–18 (2019).
10. Devey, O. & Bracken, A.P. PRC2 functions in development and congenital disorders. Development 146, dev181354 (2019).
11. Healy, E. et al. PRC2.1 and PRC2.2 synergize to coordinate H3K27 trimethylation. Mol. Cell. 76, 437–452.e6 (2019).
12. Hoffeldt, J. W. et al. Non-core subunits of the PRC2 complex are collectively required for its target-site specificity. Mol. Cell 76, 423–436.e3 (2019).
13. Brcisunic, M. et al. EPOF links elongin and polycomb in pluripotent stem cells. Mol. Cell. 64, 645–658 (2016).
14. Zhang, Z. et al. PRC2 complexes with JAKID2, MTIF2, and eSPRC2p48 in ES cells to modulate ES cell pluripotency and somatic cell reprogramming. Stem Cells 29, 229–240 (2011).
15. Kloet, S. L. et al. The dynamic interacome and genomic targets of Polycomb complexes during stem-cell differentiation. Nat. Struct. Mol. Biol. 23, 682–690 (2016).
16. Hauri, S. et al. High-density map for navigating the human polycomb puzzle. Rev. Biochem. 81, 95 (2012).
17. Smits, A. H., Jansen, P. W., Poser, I., Hyman, A. A. & Vermeulen, M. Genome-wide mapping of mono- and biubiquitination of histone H3 by immunoprecipitation and mass spectrometry. Nucleic Acids Res. 343–349 (2011).
18. Conway, E. et al. A family of vertebrate-specific polycomb group genes balances PRC2 subtypes during embryogenesis and cell fate specification. Cell Rep. 17, 583–595 (2016).
19. Smits, A. H., Jansen, P. W., Poser, I., Hyman, A. A. & Vermeulen, M. Structure-function relationship of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Res. 41, e28 (2013).
20. Conway, E. et al. A family of vertebrate-specific polycombs encoded by the LCOR/LCORL genes balance PRC2 subtypes and their activities. Mol. Cell. 70, 408–421. e8 (2018).
20. Kalb, R. et al. Histone H2A monoubiquitination promotes histone H3 methylation in polycomb repression. Nat. Struct. Mol. Biol. 21, 569–571 (2014).

21. Margueron, R. et al. Role of the polycomb protein EED in the propagation of PRC2-mediated repressive domain formation. Mol. Cell 70, 1149–1162.e5 (2018).

22. Shi, Y. et al. C10ORF12 modulates PRC2 histone methyltransferase activity during cell differentiation. Mol. Cell 57, 769–783 (2015).

23. Margueron, R. et al. Histone H2A monoubiquitination promotes histone H3 methylation in polycomb repression. Nat. Struct. Mol. Biol. 21, 569–571 (2014).

24. Oksuz, O. et al. Capturing the onset of PRC2-mediated repressive domain formation. Mol. Cell 70, 1149–1162.e5 (2018).

25. Alekseyenko, A. A., Gorchakov, A. A., Kharchenko, P. V. & Kuroda, M. I. Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP–XL-linking and affinity purification. Proc. Natl Acad. Sci. USA 111, 2480–2491 (2014).

26. Ragazini, R. et al. EZH2 inhibits Polycomb Repressive Complex 2 activity in germ cells. Nat. Commun. 10, 3858 (2019).

27. Olivierio, G. et al. Dynamic protein interactions of the polycomb repressive complex 2 during differentiation of pluripotent cells. Mol. Cell Proteomics 15, 3450–3460 (2016).

28. Vassel, M. et al. EZH2/1/2 function mostly within canonical PRC2 and exhibit proliferation-dependent redundancy that shapes mutational signatures in cancer. Proc. Natl Acad. Sci. USA 116, 6075–6080 (2019).

29. Jain, S. U. et al. PFA epigenomically-associated protein EZH2 inhibits PRC2 activity through a H3 K27M-like mechanism. Nat. Commun. 10, 2146 (2019).

30. Guo, A. et al. Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. Mol. Cell Proteomics 13, 372–387 (2014).

31. Wang, X. et al. Regulation of histone methylation by automethylation of PRC2. Genes Dev. 33, 1416–1427 (2019).

32. Lee, C. H. et al. Automethylation of PRC2 promotes H3K27 methylation and H3K27me3 deposition during cell differentiation. Mol. Cell 64, 695–672 (2016).

33. Zhang, Q. et al. RNA exploits an exposed regulatory site to inhibit the enzymatic activity of PRC2. Nat. Struct. Mol. Biol. 26, 237–247 (2019).

34. Long, Y. et al. Conserved RNA-binding specificity of polycomb repressive complex 2 is achieved by dispersed amino acid patches in EZH2. Elife 6, e31558 (2017).

35. Jager, K., Rechsteiner, T. J. & Richmond, T. J. Expression and purification of recombinant histones and nucleosome reconstitution. Methods Mol. Biol. 119, 1–16 (1999).

36. Wang, X. et al. Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. Nat. Struct. Mol. Biol. 24, 1028–1038 (2017).

37. Puda, A. et al. Frequent deletions of JARID2 in leukemia transformation of chronic myeloid malignancies. Am. J. Hematol. 87, 245–250 (2012).

38. Su, C. L., Deng, T. R., Shang, Z. & Xiao, Y. JARID2 inhibits leukemia cell proliferation by regulating CCND1 expression. Int. J. Hematol. 102, 76–85 (2015).

39. Sinclair, P. B. et al. Dynamic clonal progression in xenografts of acute lymphoblastic leukemia with intrachromosomal amplification of chromosome 21. Haematologica 103, 634–644 (2018).

40. Andersson, L. C., Jokinen, M. & Gahmberg, C. G. Induction of erythroid differentiation in the human leukemia cell line K562. Nature 278, 364–365 (1979).

41. Andersson, L. C., von Willebrand, E., Jokinen, M., Karhi, K. K. & Gahmberg, C. G. Glycoporphin A as an erythroid marker in normal and malignant hematopoiesis. Haematologica Transfus. 26, 338–344 (1981).

42. Chen, K. et al. Resolving the distinct stages in erythroid differentiation based on dynamic changes in membrane protein expression during erythropoiesis. Proc. Natl Acad. Sci. USA 106, 17413–17418 (2009).

43. Son, J., Shen, S. S., Margueron, R. & Reinberg, D. Nucleosome-binding activities within JARID2 and EZH1 regulate the function of PRC2 on chromatin. Genes Dev. 27, 2663–2677 (2013).

44. Li, H. et al. Polycomb-like proteins link the PRC2 complex to CpG islands. Nature 549, 287–291 (2017).
Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-24866-3.

Correspondence and requests for materials should be addressed to C.D.

Peer review information Nature Communications thanks the anonymous reviewers for their contributions to the peer review of this work. Peer review reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

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