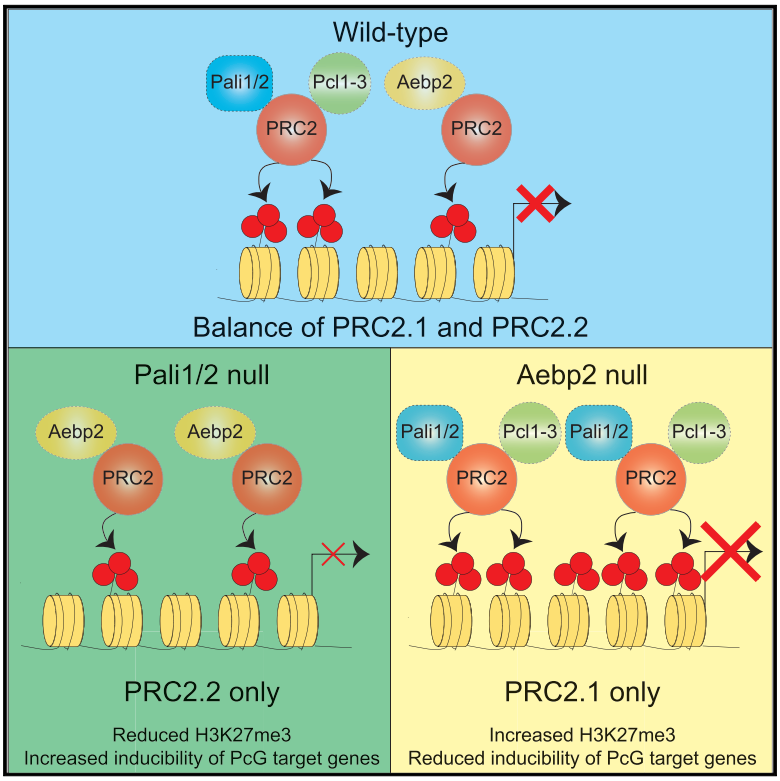
Article

Molecular Cell

# A Family of Vertebrate-Specific Polycombs Encoded by the *LCOR/LCORL* Genes Balance PRC2 Subtype Activities

## Graphical Abstract




## Highlights

- Pali1 and Pali2 form a new family of vertebrate-specific proteins that bind PRC2
- Pali1 promotes PRC2 methyltransferase activity *in vitro* and *in vivo*
- Pali1 defines a distinct PRC2.1 subtype essential for mouse development
- PRC2 subtype balance is essential for proper regulation of Polycomb target genes

## Authors

Eric Conway, Emilia Jerman, Evan Healy, ..., Raphaël Margueron, Haruhiko Koseki, Adrian P. Bracken

## Correspondence

adrian.bracken@tcd.ie

## In Brief

Pali1 and Pali2 form a new family of vertebrate-specific proteins that bind PRC2. Conway et al. show that Pali1 promotes PRC2 methyltransferase activity and defines a distinct PRC2.1 subtype essential for mouse development. They also establish that the balance of PRC2.1 and PRC2.2 activities is essential for proper regulation of polycomb target genes.

Conway et al., 2018, Molecular Cell *70*, 408–421
May 3, 2018 
https://doi.org/10.1016/j.molcel.2018.03.005

CellPress

CellPress

Molecular Cell
Article

# A Family of Vertebrate-Specific Polycombs Encoded by the *LCOR/LCORL* Genes Balance PRC2 Subtype Activities

Eric Conway,[1,10] Emilia Jerman,[1,10] Evan Healy,[1] Shinsuke Ito,[2] Daniel Holoch,[3] Giorgio Oliviero,[4] Orla Deevy,[1] Eleanor Glancy,[1] Darren J. Fitzpatrick,[1] Marlena Mucha,[1] Ariane Watson,[4] Alan M. Rice,[1] Paul Chammas,[5,6] Christine Huang,[7] Indigo Pratt-Kelly,[1] Yoko Koseki,[2] Manabu Nakayama,[8] Tomoyuki Ishikura,[2] Gundula Streubel,[1] Kieran Wynne,[5] Karsten Hokamp,[1] Aoife McLysaght,[1] Claudio Ciferri,[8] Luciano Di Croce,[5,6,9] Gerard Cagney,[4] Raphaël Margueron,[3] Haruhiko Koseki,[2] and Adrian P. Bracken[1,11,*]

[1]Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland
[2]Laboratory for Developmental Genetics, RIKEN Center for Integrative Medical Sciences (IMS), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[3]Institut Curie, Paris Sciences et Lettres (PSL) Research University, CNRS UMR 3215, INSERM U934, 75248 Paris Cedex 05, France
[4]School of Biomolecular and Biomedical Science, University College Dublin, Dublin 4, Ireland
[5]Gene Regulation, Stem Cells and Cancer Program, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr. Aiguader 88, 08003 Barcelona, Spain
[6]Universitat Pompeu Fabra (UPF), Dr. Aiguader 88, 08003 Barcelona, Spain
[7]Department of Structural Biology, Genentech, San Francisco, CA 94080, USA
[8]Chromosome Engineering Team, Department of Technology Development, Kazusa DNA Research Institute, Kisarazu 292-0818, Japan
[9]ICREA, Pg. Lluis Companys 23, 08010 Barcelona, Spain
[10]These authors contributed equally
[11]Lead Contact
*Correspondence: adrian.bracken@tcd.ie
https://doi.org/10.1016/j.molcel.2018.03.005

## SUMMARY

**The polycomb repressive complex 2 (PRC2) consists of core subunits SUZ12, EED, RBBP4/7, and EZH1/2 and is responsible for mono-, di-, and tri-methylation of lysine 27 on histone H3. Whereas two distinct forms exist, PRC2.1 (containing one polycomb-like protein) and PRC2.2 (containing AEBP2 and JARID2), little is known about their differential functions. Here, we report the discovery of a family of vertebrate-specific PRC2.1 proteins, "PRC2 associated LCOR isoform 1" (PALI1) and PALI2, encoded by the *LCOR* and *LCORL* gene loci, respectively. PALI1 promotes PRC2 methyltransferase activity *in vitro* and *in vivo* and is essential for mouse development. Pali1 and Aebp2 define mutually exclusive, antagonistic PRC2 subtypes that exhibit divergent H3K27-tri-methylation activities. The balance of these PRC2.1/PRC2.2 activities is required for the appropriate regulation of polycomb target genes during differentiation. PALI1/2 potentially link polycombs with transcriptional co-repressors in the regulation of cellular identity during development and in cancer.**

## INTRODUCTION

Polycomb group proteins are a family of chromatin regulators that are essential for maintaining cellular identity in higher eukaryotes (Margueron and Reinberg, 2011). They function primarily as two distinct multi-protein chromatin-associated complexes, the polycomb repressive complex 1 (PRC1) and PRC2 (Schuettengruber et al., 2017). PRC2 is composed of four core proteins, Suz12, Eed, Rbbp4/7, and one of the two histone H3K27 methyltransferases, Ezh1 or Ezh2 (Margueron and Reinberg, 2011). PRC2 mediates all H3K27 mono-, di-, and tri-methylation (H3K27me1/2/3) on chromatin (Conway et al., 2015; Højfeldt et al., 2018). PRC2 co-localizes with H3K27me3, which is required to recruit PRC1 through specific binding of its constituent CBX proteins via their chromodomain (Margueron and Reinberg, 2011). The recruited PRC1 complex is then thought to confer gene repression by chromatin compaction through CBX2 or its PHC1–3 subunits (Isono et al., 2013; Lau et al., 2017). As such, PRC2-mediated H3K27me3 contributes to maintaining cellular memory by directing PRC1-mediated repression of genes encoding key regulators of alternative lineages. Whereas less is known about the H3K27me2 mark, it is present on almost all intergenic euchromatin regions genome-wide (Conway et al., 2015; Ferrari et al., 2014; Lee et al., 2015). It has been proposed to contribute to cellular memory by forming a repressive "genomic blanket" to prevent the misfiring of cell-type-specific enhancers of alternative lineages.

An emerging paradigm is that there are two different subtypes of the PRC2 complex (Holoch and Margueron, 2017). In addition to the four core subunits, several auxiliary proteins associate with PRC2, including Jarid2, Aebp2, Pcl1–3 (Phf1, Mtf2, and Phf19), and Epop (Hauri et al., 2016; Smits et al., 2013). Much like the sub-functionalization within the canonical and non-canonical PRC1 complexes (Blackledge et al., 2015), recent comprehensive proteomic analyses of PRC2 have revealed

that it assembles primarily into two mutually exclusive combinations, termed PRC2.1 and PRC2.2 (Alekseyenko et al., 2014; Grijzenhout et al., 2016; Hauri et al., 2016). The PRC2.1 complex is defined as containing one of the three polycomb-like proteins (Pcl1–3), whereas PRC2.2 is defined as containing Aebp2 and Jarid2 (Holoch and Margueron, 2017). There is also additional variation within the PRC2.1 subtype, such that one Pcl1–3 protein is a constant and defining feature, whereas the presence of Epop is mutually exclusive of the presence of another uncharacterized protein, called C10ORF12 (Alekseyenko et al., 2014; Hauri et al., 2016).

Although little is known about why there are different forms of the PRC2 complex, the auxiliary PRC2 subunits characterized to date have roles in either promoting PRC2 activity or facilitating its interactions with chromatin or other proteins. For example, the polycomb-like proteins contribute to PRC2.1 association with chromatin via their Tudor chromatin reader and winged-helix domains and stimulate its methyltransferase activity *in vitro* (Brien et al., 2012, 2015; Choi et al., 2017; Li et al., 2017). The recently discovered Epop subunit bridges PRC2.1 with elongin B/C proteins, but unlike Pcl proteins, it appears not to promote PRC2.1 association with chromatin (Beringer et al., 2016; Liefke et al., 2016). The role of the most recently identified member of PRC2.1, C10ORF12, encoded by a putative isoform of the *LCOR* gene locus has not yet been elucidated. In terms of the PRC2.2 complex, its two specific subunits, Aebp2 and Jarid2, promote PRC2 methyltransferase activity *in vitro* (Cao and Zhang, 2004; Li et al., 2010; Son et al., 2013). Whereas the precise role of Aebp2 in this complex is unclear, Jarid2 may have a role in targeting PRC2.2; it was recently reported to bind to PRC1-mediated H2AK119ub1 via its ubiquitin interaction motif (Cooper et al., 2016). Consistent with this, the *in vitro* activity of PRC2.2 is enhanced in the presence of H2AK119ub1-modified nucleosomes (Kalb et al., 2014). However, it is still unclear why there are at least two PRC2 subtypes and whether they interplay with each other to regulate polycomb target genes during cell fate transitions.

Some intriguing clues as to the roles of PRC2.1 and PRC2.2 complexes have come from recent studies of Epop and Aebp2 in mouse embryonic stem cells (ESCs) (Beringer et al., 2016; Grijzenhout et al., 2016). Whereas depletion of Pcl and Jarid2 PRC2 subunits in ESCs correlate with partial reductions in the levels of both Ezh2 and H3K27me3 (Brien et al., 2012; Holoch and Margueron, 2017; Pasini et al., 2010), the loss of either Epop or Aebp2, members of PRC2.1 and PRC2.2, respectively, correlate with increased levels of H3K27me3 (Beringer et al., 2016; Grijzenhout et al., 2016). These paradoxical results are inconsistent with the fact that recombinant Epop and Aebp2 proteins promote core PRC2 histone methyltransferase activity *in vitro* (Cao and Zhang, 2004; Zhang et al., 2011). However, supporting the observations in Aebp2-null ESCs, *Aebp2*-null mice have a "Trithorax" phenotype, classically associated with decreased expression of *HOX* genes (Grijzenhout et al., 2016). These findings highlight how little is currently known about the respective roles of the non-core PRC2 components and also points to potential opposing activities within PRC2 complexes.

Here, we characterize a family of vertebrate-specific proteins, PALI1 and PALI2, which form part of the PRC2.1 complex. These proteins are encoded by isoforms of the *LCOR* and *LCORL* gene loci, respectively, with PALI1 encompassing the C10ORF12 protein, previously linked with PRC2.1. PALI1 and PALI2 are defined by their CTBP binding domains and a PALI interaction with PRC2 (PIP) domain. PALI1 also interacts with several other chromatin regulators, including the G9A complex, SET, and the deubiquitinases, USP11 and USP22. PALI1 is essential for mouse development and promotes PRC2 histone methyltransferase activity *in vivo* and *in vitro*. We characterize an antagonism between the PALI-PRC2.1 and AEBP2-PRC2.2 complexes in ESCs and during differentiation. Our data support a model in which the two PRC2 complexes promote H3K27me3 methylation to different extents and that loss of key members of either leads to an imbalance in the regulation of polycomb target genes. More broadly, our discovery of PALI1/2 links polycombs with transcriptional co-repressors in the regulation of cellular identity during development and in cancer.

## RESULTS

### PALI1 Associates with CTBP Co-repressor Proteins and PRC2.1 and Promotes HKMT Activity *In Vitro*

To identify PRC2.1-associated proteins, we performed mass spectrometric (MS) analysis of in-gel tryptic digestions from FLAG immunoprecipitations (IPs) of FLAG-PCL1/PHF1 stably expressed in HEK293T cells (Figure 1A). As expected, this identified peptides from the PCL1, EZH2, SUZ12, and EED proteins in the top 6 ranked proteins detected (Table S1). In addition, we also detected the p53 protein, which we previously reported to associate with PCL1 (Brien et al., 2015). Strikingly, peptides of the predicted C10ORF12 protein were the most abundant after SUZ12 and EZH2 (Table S1). Intriguingly, the peptides of both C10ORF12 and LCOR were detected at the top of the gel at a molecular weight corresponding to approximately 260 kDa, which is larger than the predicted ~50 kDa for LCOR and ~137 kDa for C10ORF12 (Figure 1A). LCOR, a widely expressed transcriptional co-repressor, is known to associate with CTBP proteins and nuclear receptors via conserved N-terminal domains (Fernandes et al., 2003; Palijan et al., 2009). To investigate the potential connection between LCOR, C10ORF12, and PCL1, we generated an antibody against C10ORF12 and performed western blots on an independent FLAG-PCL1 IP and found that both anti-C10ORF12 and anti-LCOR antibodies detected bands at about 260 kDa (Figure 1B). Supporting this result, a previous study reported that the *LCOR* gene locus encodes an isoform called *LCOR-Cra_b* (Hauri et al., 2016), which includes *C10ORF12* (Figure 1C). We decided to call the resulting 260 kDa protein "PALI1," for "PRC2-associated LCOR isoform 1".

We next wanted to confirm that the PALI1 protein is encoded by the *LCOR-Cra_b* alternative splice form of *LCOR*. First, consistent with C10ORF12 being part of PALI1 and not an independent gene, we examined the chromatin IP sequencing (ChIP-seq) tracks of promoter-associated H3K4me3 from a number of different human cell lineages (Figure 1C). Whereas

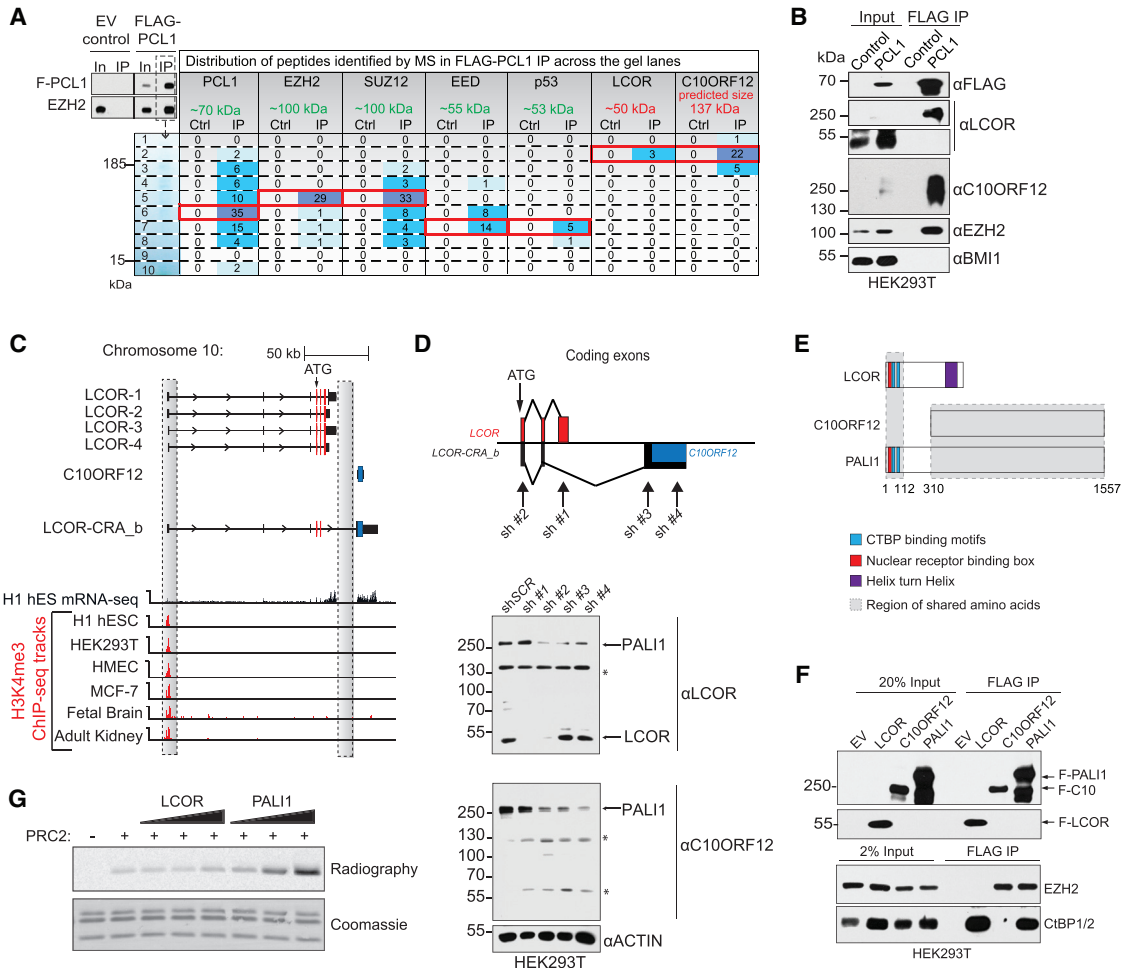



**Figure 1. PALI1 Associates with CTBP Co-repressor Proteins and PRC2.1 and Promotes HKMT Activity *In Vitro***

(A) (Left) Western blot and Coomassie staining of FLAG immunoprecipitations (IP) of nuclear lysates of HEK293T cells stably expressing FLAG-PCL1 or empty vector (EV). The Coomassie stained gel was subjected to in-gel tryptic digestion as 10 separate slices and subjected to liquid chromatography (LC)-mass spectrometry. (Right) The MaxQuant peptide counts of the indicated proteins are listed in the table adjacent to the respective gel band that they were identified in. The PCL1, EZH2, SUZ12, EED, and C10ORF12 proteins were the top 5 proteins IP (Table S1). The previously characterized PCL1 binding protein, p53, ranked 16th, whereas LCOR ranked 57th. Note that the peptides for LCOR and C10ORF12 do not match their predicted molecular weights on the gel.

(B) Western blot with the indicated antibodies of FLAG-IPs of nuclear lysates from control and FLAG-PCL1 stably expressing HEK293T cells.

(C) A genomic view of the *LCOR* gene locus showing five alternative splicing variants (LCOR1–4 and *LCOR-CRA_b*) and *C10ORF12*. RNA-seq of human embryonic stem cells and H3K4me3 ChIP-seq tracks in various cell lines, downloaded from ENCODE and Epigenetic Roadmap projects (Kundaje et al., 2015), are presented. Grey boxes highlight the predicted transcription start site of *LCOR* and *C10ORF12*.

(D) (Top) A schematic representation of the coding exons of *LCOR* (red), *C10ORF12* (blue), and *LCOR-Cra_b* (black). (Bottom) Western blot analysis on lysates of HEK293T cells stably expressing shRNAs targeting the indicated coding exons is shown. The asterisks represent non-specific background bands.

(E) A schematic of the LCOR and PALI1 proteins showing their shared N-terminal nuclear receptors box and CTBP1/2 binding motifs. The gray shaded box indicates regions of identical protein sequence.

(F) Western blot using the indicated antibodies of FLAG IPs of whole-cell lysates of HEK293T cells ectopically expressing FLAG-LCOR, FLAG-C10ORF12, or FLAG-PALI1 protein.

(G) Histone methyltransferase assay using recombinant PRC2 with a titrated addition of either LCOR or PALI1 protein, as indicated. Recombinant oligonucleosomes were used as a substrate. The PRC2 complex used contained FLAG-EED, SUZ12, EZH2, and RBBP4. Incorporation of 3H from radiolabeled SAM was used as a readout of methyltransferase activity. Coomassie staining shows equal loading of oligonucleosomes.

See also Figure S1 and Table S1.

H3K4me3 is enriched at the *LCOR* gene promoter, it is absent immediately upstream of the *C10ORF12* coding region, suggesting *C10ORF12* does not possess an independent promoter. We next cloned the 4,674-nucleotide PALI1 coding sequence by RT-PCR on cDNA from primary human mammary epithelial cells (Figure S1A). This confirmed that PALI1 is encoded by *LCOR-Cra_b*, a transcript comprising the first two coding exons of *LCOR* spliced to a third larger exon that contains *C10ORF12*

(Figures 1C and 1D). Accordingly, the resulting PALI1 protein sequence begins with the first 111 amino acids of the LCOR protein on its N terminus followed by 199 amino acids unique to PALI1 and, finally, 1,247 amino acids encoded by *C10ORF12* on its C terminus (Figures S1B and S1C). We confirmed that the peptides of LCOR and C10ORF12 proteins detected in the FLAG-PCL1 IP all mapped to PALI1 and not to the C-terminal part of the LCOR protein (Figures S1B and S1C). To characterize our anti-C10ORF12 (from here on referred to as anti-PALI1) and anti-LCOR antibodies and further substantiate the existence of the PALI1 protein, we designed four short hairpin RNA (shRNA) constructs to target specific coding exons that would either deplete LCOR on its own, LCOR and PALI1 together, or PALI1 on its own, in stably infected HEK293T cells (Figure 1D). The shRNA targeting the LCOR-specific exon (sh no. 1) reduced the mRNA and protein levels of LCOR, compared to control, scrambled shRNA, but did not affect levels of PALI1 (Figures 1D and S1D). The shRNA designed to target the predicted first common coding exon of the *LCOR* and *LCOR-Cra_b* mRNA transcripts (sh no. 2) reduced the mRNA and protein levels of both LCOR and PALI1. Finally, two shRNAs targeting *LCOR-Cra_b*, either in the unique region between *LCOR* and *C10ORF12* (sh no. 3) or within *C10ORF12* (sh no. 4), led to reduced levels of the mRNA and protein levels of PALI1 but did not affect LCOR (Figures 1D and S1D). Taken together, these data confirm the specificity of our antibodies to endogenous proteins and that C10ORF12 represents a portion of the ~260-kDa PALI1 protein rather than an independent isoform of the *LCOR* gene locus.

Next, to explore the potential common and unique interactions of PALI1 and LCOR with PRC2 and CTBP proteins (Figure 1E), we expressed FLAG/hemagglutinin (HA)-tagged PALI1, C10ORF12, or LCOR in HEK293T cells and performed FLAG-IPs (Figure 1F). This confirmed that, like LCOR, PALI1 IP CTBP1/2 co-repressor proteins, whereas only C10ORF12 and PALI1 IP EZH2. Because PRC2 is a histone lysine methyltransferase (HKMT), we analyzed whether PALI1 could modulate this activity *in vitro*. We incubated PRC2 with increasing amounts of recombinant PALI1 protein (Figures S1E and S1F) together with recombinant oligonucleosomes. This resulted in a dose-dependent increase in the methylation of histone H3, whereas the addition of LCOR did not promote HKMT activity, consistent with its inability to associate with the PRC2 complex (Figure 1G). Taken together, the data establish that PALI1 is an ~260-KDa isoform of LCOR capable of binding CTBP proteins and PRC2 and promoting PRC2 activity *in vitro*.

## PALI1 Is a Multi-domain Protein that Associates with Several Chromatin Regulators, Including the G9A and PRC2.1 Complexes

We next wished both to confirm that PALI1 specifically interacts with PRC2.1 and to further expand on its interactome. To do this, we performed mass spectrometric analysis of FLAG-PALI1, FLAG-C10ORF12, and FLAG-LCOR IPs (Figure 2A; Table S2). This revealed that both C10ORF12 and PALI1 proteins, but not LCOR, IP the components of the PRC2.1 complex, EZH2, EED, SUZ12, and PCL2, but not PRC2.2-specific components. Both LCOR and PALI1 proteins IP CTBP1/2 proteins, but C10ORF12 did not (Figure 2A). We also identified several additional proteins that associate with PALI1, most notably the G9A co-repressor complex, including the G9A, GLP, ZNF644, and WIZ subunits, as well as the SET protein and deubiquitinases USP11 and USP22 (Figures 2A and S2A). Whereas PRC2.2-specific members were previously reported to not interact with C10ORF12 (Alekseyenko et al., 2014; Hauri et al., 2016), we next wished to confirm that they do not interact with full-length PALI1 protein either. We ectopically expressed FLAG-tagged PALI1, PCL2, AEBP2, and JARID2 and performed FLAG IPs and western blots for endogenous PRC2 proteins (Figure 2B). This confirmed that PALI1 does not associate with AEBP2 or JARID2. Furthermore, whereas PALI1 is immunoprecipitated by both PCL2 and SUZ12, AEBP2 only IPs with SUZ12 (Figure S2B).

Next, to further delineate the potential interaction domains within the PALI1 protein for its various binding partners, we generated five truncations of the C10ORF12 region of PALI1, transiently expressed them as FLAG-HA-tagged proteins in HEK293T cells, and performed FLAG IPs (Figures 2C and 2D). This analysis revealed that fragment 1 and fragments 4/5 of C10ORF12 IP the G9A and PRC2.1 complexes, respectively (Figure 2D). This indicates that the G9A and PRC2.1 complexes bind to PALI1 via distinct regions. Consistent with these G9A and PRC2.1 interaction regions being potentially important, a disorder tendency analysis of the complete PALI1 protein predicted two "ordered," globular-like domains: one mapping to the G9A interaction region and the other to fragments 4/5, which we called PIP, for PALI interaction with PRC2 domain (Figure S2C).

These data establish that, in addition to CTBP binding motifs and a nuclear receptor binding box in its N terminus, the PALI1 protein also contains regions within its C terminus that allow it to associate with several different chromatin regulators, including the SET protein, deubiquitinases, and the G9A and PRC2.1 complexes.

## PALI1 Is a Member of a Vertebrate-Specific Family of PRC2.1 Proteins Defined by a Conserved PRC2 Interaction Domain

To define the minimal PIP domain within PALI1, we generated additional truncations of the extended PIP domain (Figures S3A and S3B). This identified a 51-amino-acid PRC2 interaction region (PIR2) sufficient to IP EZH2, but not BMI1 (Figure S3B). Remarkably, a protein homology search of this region identified only one other human protein: an orthologous sequence within a predicted splice form of the human *LCORL* gene (XM_011513822.2), the paralog of *LCOR*. We therefore named the predicted protein product of this isoform PALI2. Multiple sequence alignments of the mouse and human PALI1 and PALI2 proteins show high conservation of residues within this PIR2 region (Figure S3C). To validate the existence of the predicted PALI2 protein, we performed a detailed RT-PCR mapping analysis of all the *LCORL* exons on cDNA from HEK293T cells (Figure S3D). This confirmed that the *PALI2*-specific exon 8 (with homology to the extended PIP region of PALI1) exists within an alternative splice form of *LCORL*. This splice form contains exons 1–6 but lacks exon 7, which contains the nucleotide

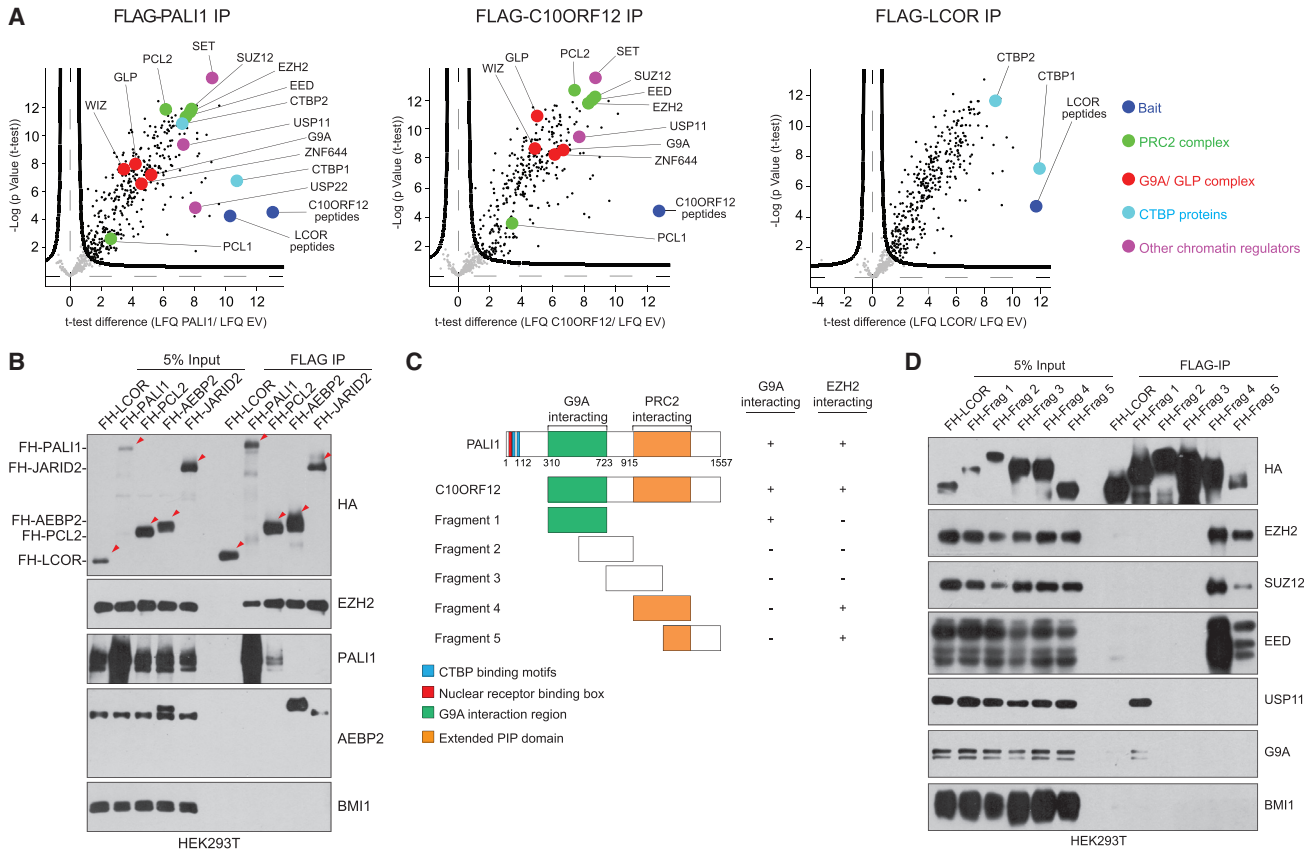



**Figure 2. PALI1 Is a Multi-domain Protein that Associates with Several Chromatin Regulators, Including the G9A and PRC2.1 Complexes**

(A) Identification of the interacting proteins of FLAG-tagged PALI1, C10ORF12, and LCOR, ectopically expressed in HEK293T cells. LC-MS analysis was performed using permutation-based false discovery rate (FDR)-corrected t test. The label free quantification (LFQ) intensity of the FLAG bait IP over the EV control is plotted against the −log10 (p value).

(B) Western blots of FLAG-IP eluate of the indicated proteins following transient ectopic expression in HEK293T cells.

(C) (Left) A schematic representation of full-length PALI1 and the 5 fragments cloned and expressed as FLAG-tagged proteins in HEK293T cells. (Right) A summary table indicating the status of interaction of each fragment with EZH2 and G9A is shown.

(D) FLAG-IPs of the indicated proteins (LCOR or fragments of PALI1 as shown in C), transiently expressed in HEK293T cells, followed by western blotting analysis for the indicated proteins.

See also Figure S2 and Table S2.

sequence encoding the helix turn helix (HTH) domain of LCORL. Therefore, the organization of the PALI2 coding exons within the *LCORL* gene locus is similar to that of the organization of PALI1 within the *LCOR* gene locus.

To analyze the conservation and evolution of the LCOR/LCORL and PALI1/2 proteins in more depth, we aligned the amino acid sequences of all the possible coding exons of all four proteins across several vertebrate and invertebrate species (Figure 3A). This revealed that invertebrates have a single ortholog of LCOR, which shares the highly conserved HTH region with the vertebrate LCOR and LCORL proteins. However, invertebrate species lack the PIP domain and the large C terminus present in both the PALI1 and PALI2 proteins, suggesting these regions arose after the divergence of the vertebrate and invertebrate common ancestor. Importantly, the PIP domain of PALI1 and PALI2 is highly conserved within vertebrate species and overlaps with a predicted globular domain (Figure S2C), suggesting potential evolutionary importance.

To confirm that the PALI2 PIP domain is also capable of associating with PRC2, we cloned the paralogous sequence of the extended PALI1 PIP region from PALI2 and expressed it, together with the extended PIP region of PALI1, as FLAG-HA-tagged proteins in HEK293T cells (Figure 3B). FLAG-IPs then confirmed that it is also capable of immunoprecipitating EZH2, but not BMI1. In order to further dissect the PIP domain of PALI1 and PALI2, we closely analyzed the conservation and charges of its constituent amino acids (Figure 3C). Two highly conserved tryptophans (W1125 and W1186) were selected for mutation due to the potential importance of aromatic residues in protein-protein interactions and for structural stability within a modular domain (Mullin et al., 2017). Mutagenesis of these tryptophans to alanines individually did not disrupt the interaction with EZH2 (Figure 3D). However, the double mutation of

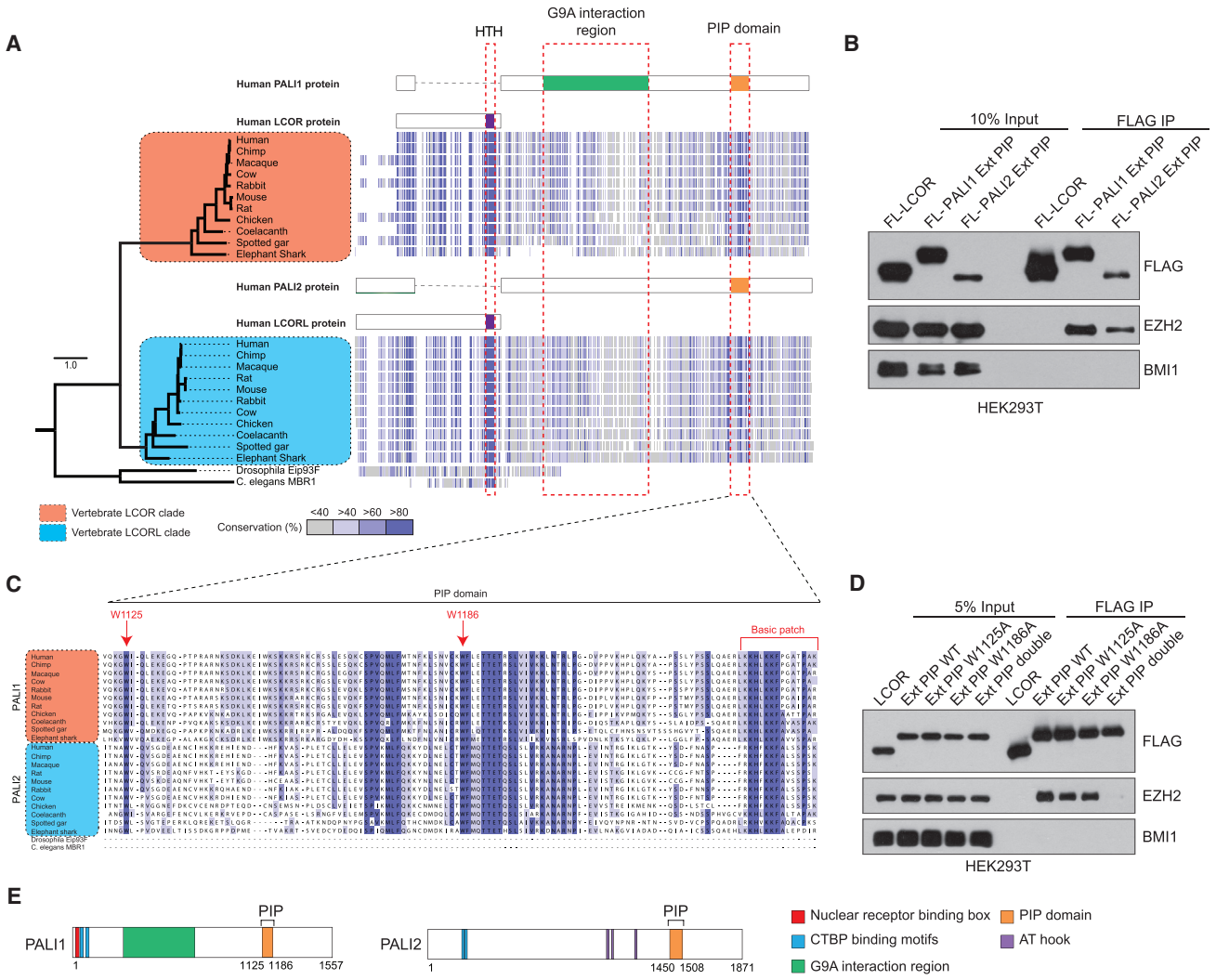



**Figure 3. PALI1 Is a Member of a Vertebrate-Specific Family of PRC2.1 Proteins Defined by a Conserved PRC2 Interaction Domain**

(A) A maximum-likelihood phylogenetic tree (left) inferred from aligned amino acid sequences of LCOR and LCORL exons (right). The percentage of the residues that agree with the consensus sequence is indicated by color. Positions corresponding to either the human LCOR/LCORL and PALI1/2 proteins are indicated above the alignments.

(B) FLAG-IPs of the indicated proteins (LCOR; extended PIP domain of PALI1 and the orthologous region of PALI2) transiently expressed in HEK293T cells, followed by western blotting using the indicated antibodies.

(C) Amino acid alignments of the PALI interaction with PRC2 (PIP) domain of PALI1 and PALI2 proteins from several vertebrate and invertebrate species. Two conserved tryptophans and a basic charged patch are indicated.

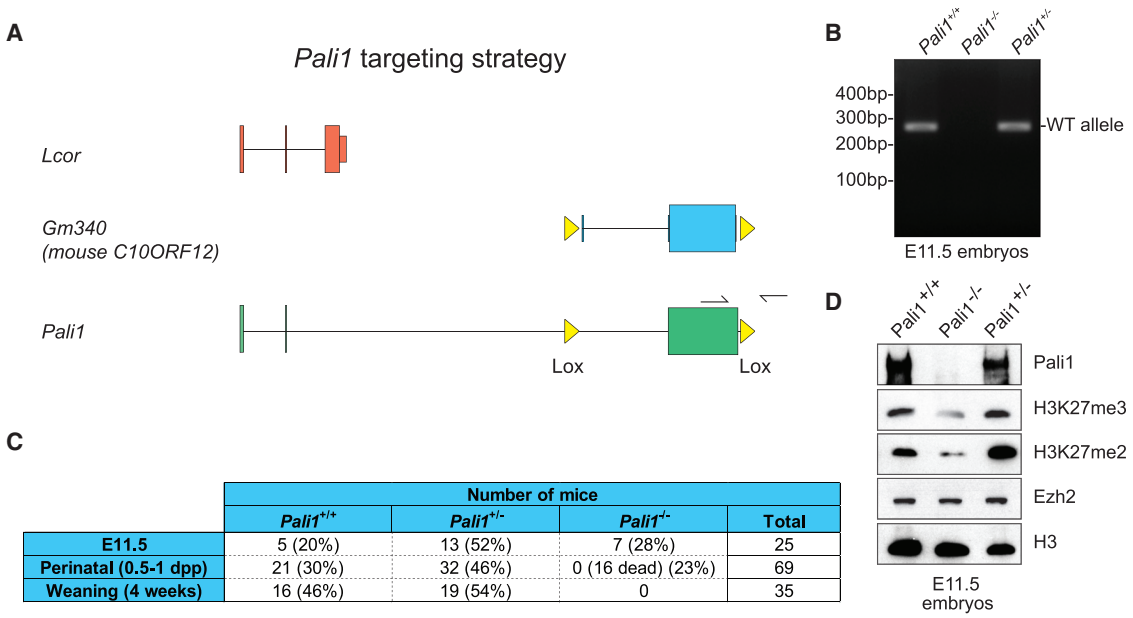
(D) FLAG-IPs of the indicated proteins (LCOR or PALI1 extended PIP wild-type [WT] or with the indicated mutations) transiently expressed in HEK293T cells, followed by western blot using the indicated antibodies.

(E) An illustration summarizing both the shared and unique domains and regions in the PALI1 and PALI2 proteins.

See also Figure S3.

W1125A and W1186A led to an almost complete loss in the ability of the PALI1 extended PIP region to IP EZH2, suggesting the two tryptophans are important for the interaction of the PIP domain with PRC2. Reciprocally, endogenous IP of SUZ12 revealed that the double mutation of W1125A and W1186A disrupts the association between the PRC2 core complex and the extended PIP domain (Figure S3E). The PALI1 and PALI2 PIP domains also contain a number of other highly conserved residues, including a basic patch (Figure 3C), which may also be involved in the interaction with PRC2. Despite the shared PIP domain and CTBP binding motifs of PALI1 and PALI2, the proteins have diverged somewhat in their protein organization. For example, PALI2 and LCORL do not share the nuclear receptor binding box present in PALI1 and LCOR, whereas PALI2 has three predicted adenine-thymidine rich (AT) hooks in its C terminus, which could confer the ability to bind to DNA (Figure 3E).

Taken together, we show that PALI1 is the founder member of a family of proteins that share CTBP binding domains in their




C

| | Number of mice | | | |
|---|---|---|---|---|
| | *Pali1*$^{+/+}$ | *Pali1*$^{+/-}$ | *Pali1*$^{-/-}$ | Total |
| E11.5 | 5 (20%) | 13 (52%) | 7 (28%) | 25 |
| Perinatal (0.5-1 dpp) | 21 (30%) | 32 (46%) | 0 (16 dead) (23%) | 69 |
| Weaning (4 weeks) | 16 (46%) | 19 (54%) | 0 | 35 |

**Figure 4. Pali1 Is Essential for Mouse Development and Promotes H3K27me3 *In Vivo***

(A) The strategy used to target the Pali1-specific exon (a.k.a. *Gm340*) for Cre-Lox-mediated knockout in mouse ESCs.

(B) Genotyping PCR, using the indicated primers in (A), shows loss of the *Pali1* allele in *Pali1*$^{-/-}$ mice embryos (E11.5).

(C) Table showing number of mice at various stages of mouse development following a cross of *Pali1*$^{+/-}$ mice.

(D) Western blots on whole-cell lysates harvested from E11.5 *Pali1*$^{+/+}$, *Pali1*$^{+/-}$, or *Pali*$^{-/-}$ embryos, with the indicated antibodies.

See also Figure S4.

N termini and bind PRC2.1 through a PIP domain in their C termini (Figure 3E).

## Pali1 Is Essential for Mouse Development and Promotes H3K27me3 *In Vivo*

To investigate the importance of Pali1 for mouse development, we generated mouse ESCs in which the Pali1-specific exon (a.k.a. *Gm340*; the mouse ortholog of *C10ORF12*) is flanked by LoxP sites (Figures 4A and S4A). Recombination between these LoxP sites, induced by the addition of TAT-Cre (Figure S4B), led to the loss of the Pali1-specific (*Gm340*) exon but did not affect the coding exons of *Lcor* (Figure 4A). The exclusive deletion of the Pali1-specific exon was confirmed by genotyping PCR and qRT-PCR (Figures S4C and S4D). Mice heterozygous for the Pali1-specific exon deletion (*Pali1*$^{+/-}$) were crossed, and the resulting embryos were genotyped and counted at various stages of development (Figures 4B and 4C). At embryonic day 11.5 (E11.5), *Pali1*$^{-/-}$ mice had the expected Mendelian ratio compared to matched wild-type and heterozygous mice. However, none of the *Pali1*$^{-/-}$ mice were viable at birth. We next analyzed E11.5 embryos from *Pali1*$^{-/-}$ mice by western blot to determine whether the lethality could be related to changes in the levels of PRC2 components or H3K27me3. Whereas levels of Ezh2, Lcor, H3K9me2, and H2AK119ub1 were unchanged in Pali1$^{-/-}$ E11.5 embryos (Figures 4D and S4E), a global reduction in the levels of H3K27me3 was observed compared to wild-type (Figure 4D), consistent with Pali1 being an important component of the PRC2.1 complex. However, no recurrent skeletal transformations were observed in *Pali1*$^{-/-}$ embryos at the E11.5 stage, potentially due to functional redundancy between the Pali1 and Pali2 proteins (Figure S4E). Nonetheless, our results establish Pali1 as an essential protein required for development and link reduced levels of PRC2-mediated H3K27me3 with its perinatal lethal phenotype.

## Pali1 and Aebp2 Define Mutually Exclusive Antagonistic PRC2 Subtypes

We next wished to dissect the role of Pali1 in mouse ESCs. To determine whether loss of core PRC2 would affect Pali1, we performed western blot analysis on total lysates of *Eed*$^{-/-}$ and matched wild-type ESCs and found that the levels of the Pali1 protein, like Pcl2, Epop, Aebp2, and Jarid2, was reduced in the absence of Eed (Figure 5A). This result, combined with the observation that Pali1 mRNA levels were unchanged in *Eed*$^{-/-}$ ESCs (Figure S5A), suggests that, consistent with being a PRC2.1-associated protein, Pali1 protein stability is reliant on the integrity of the complex. This destabilization of Pali1 in the absence of Eed, taken together with the high levels of conservation of the PIP domain and the finding that EZH2, SUZ12, and EED are among the highest scoring proteins in the FLAG-PALI1 IP-MS, suggest that PALI1 functions primarily as a part of the PRC2 complex.

We next determined whether the stability of PRC2 complex proteins is affected in *Pali1*$^{-/-}$ single knockout (Pali1 sKO) and *Pali1*$^{-/-}$, *Pali2*$^{-/-}$ double knockout (Pali1/2 dKO) ESCs. To do this, we generated Pali1/2 dKO ESCs using CRISPR/Cas9 gene editing to introduce out-of-frame deletions in the Pali2-specific exon, upstream of the PIP domain in the Pali1 sKO

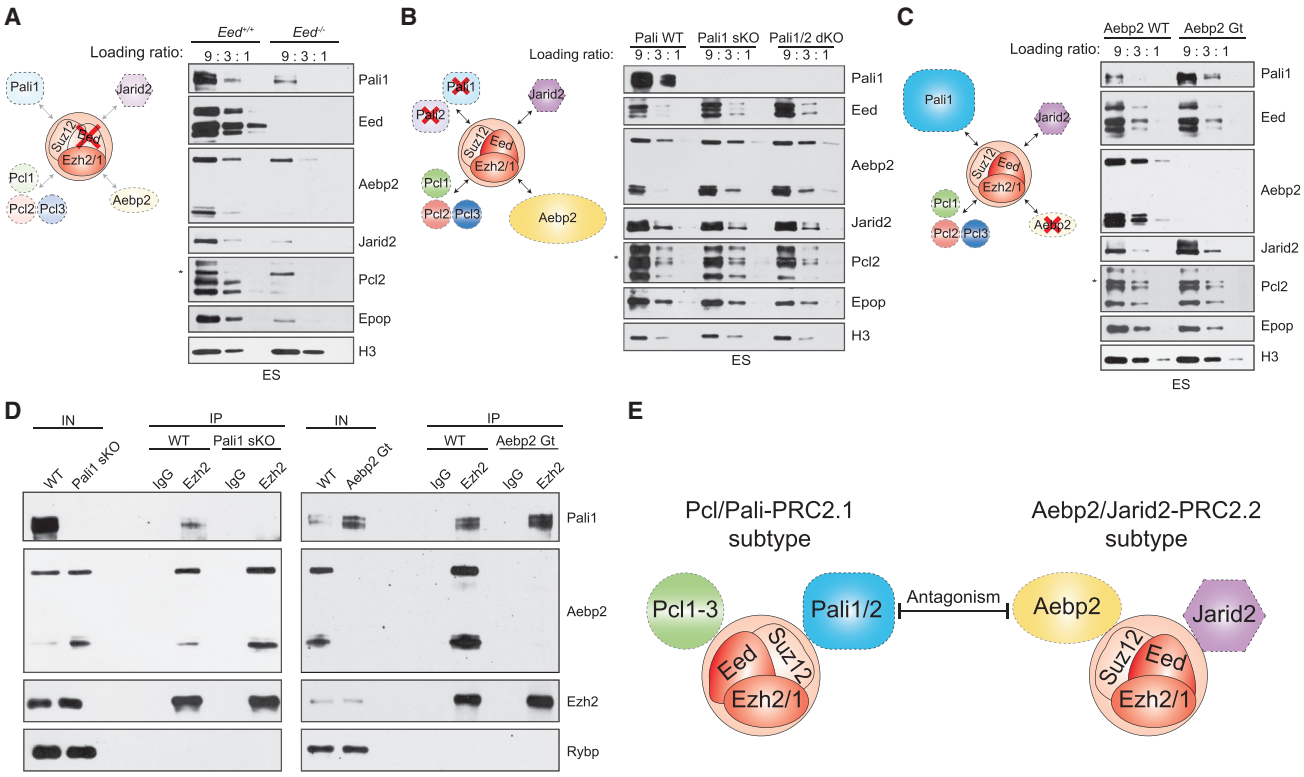



**Figure 5. Pali1 and Aebp2 Define Mutually Exclusive Antagonistic PRC2 Subtypes**

(A–C) Western blot on a dilution series of whole-cell lysates of (A) *Eed*$^{-/-}$, (B) *Pali1*$^{-/-}$ (Pali1 sKO), *Pali1*$^{-/-}$, *Pali2*$^{-/-}$ (Pali1/2 dKO), (C) *Aebp2*$^{gt/gt}$ (Aebp2 Gt), and matched wild-type ESCs.

(D) Endogenous IPs of Ezh2 in Pali1 sKO or Aebp2 Gt ESCs, compared to matched wild-type controls, followed by western blot with the indicated antibodies.

(E) Model representing the antagonism between Pcl/Pali-PRC2.1 and Aebp2/Jarid2-PRC2.2 complexes.

See also Figure S5.

ESCs (Figure S5B). Western analysis of total lysates from the Pali1 sKO and Pali1/2 dKO ESCs showed no changes in the levels of Eed, Pcl2, Epop, or Jarid2 proteins (Figure 5B). However, we observed a 2- to 3-fold increase in the amount of the short isoform of Aebp2 (Figure 5B), which correlated with a similar increase in the mRNA levels of the corresponding short, but not long, mRNA transcript of *Aebp2* (Figure S5C). As a control, we investigated the consequences of loss of the other PRC2.2 component, Jarid2. Western analyses on total lysates from *Jarid2*$^{-/-}$ ESCs showed that its loss does not affect Pali1 protein levels but does lead to a destabilization of Aebp2 protein, consistent with Aebp2 and Jarid2 being part of the same PRC2.2 subtype (Figure S5D).

Intriguingly, Pali1 protein levels were ~3-fold increased in *Aebp2*$^{gt/gt}$ (Aebp2 Gt) ESCs compared to matched wild-type control cells (Figure 5C), whereas its mRNA levels were unchanged (Figure S5E). Taken together with the fact that Pali1 is destabilized in *Eed*$^{-/-}$ ESCs (Figure 5A), this suggests that Pali1 is stabilized by increased association with the core PRC2 complex in the absence of Aebp2. This is consistent with a possible competition between the Aebp2 and Pali1 proteins for association with PRC2. To test this, we performed endogenous IPs of Ezh2 in Pali1 sKO, Aebp2 Gt, and matching wild-type ESCs (Figure 5D). This confirmed our hypothesis, revealing that loss of Pali1 leads to an increase of the short isoform of Aebp2 in the PRC2 complex, and reciprocally, loss of Aebp2 led to an increase in the amount of Pali1 in PRC2. These results uncover an antagonistic interplay between two protein subunits of the different PRC2.1 and PRC2.2 subtypes (Figure 5E).

## Pali1-PRC2.1 and Aebp2-PRC2.2 Exhibit Divergent H3K27-Methylation Activities *In Vitro* and *In Vivo*

We next wished to explore how the antagonism between Pali1 and Aebp2 might affect PRC2 activity on polycomb target genes. We performed ChIPs in Pali1 sKO, Pali1/2 dKO, and matched wild-type ESCs and observed a reduction in H3K27me3 at the promoters of several PRC2 target genes in the absence of the Pali proteins (Figure 6A). ChIPs of Eed KO and wild-type ESCs were performed for comparison and revealed a complete loss of H3K27me3 at polycomb target genes (Figure S6A) consistent with the complete disruption of the PRC2 complex integrity in these cells (Figure 5A). In contrast, Aebp2 Gt ESCs have increased enrichments of H3K27me3 on polycomb target genes compared to wild-type ESCs (Figure 6B), as reported previously (Grijzenhout et al., 2016). To further explore this, we performed quantitative ChIP-seq of H3K27me3 with a spike in of an exogenous *Drosophila* reference genome (ChIP-Rx) in the Pali1 sKO, Pali1/2 dKO, and wild-type ESCs (Figure 6C). The average plots

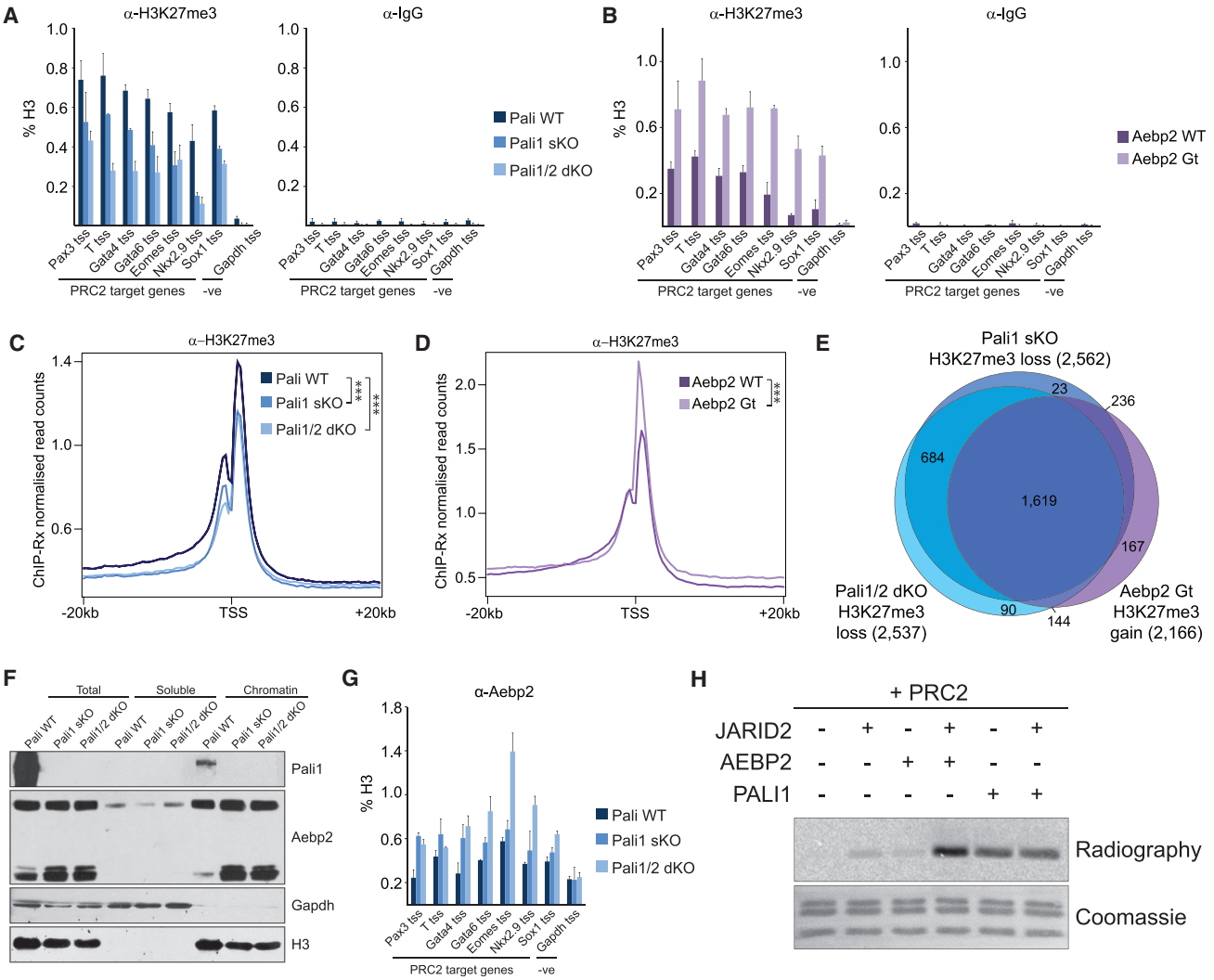



**Figure 6. Pali1-PRC2.1 and Aebp2-PRC2.2 Exhibit Divergent H3K27-Methylation Activity *In Vitro* and *In Vivo***

(A) ChIP-qPCRs at polycomb target genes using the indicated antibodies in Pali1 sKO and Pali1/2 dKO and matched wild-type ESCs. The Gapdh promoter is included as a negative control.

(B) ChIP-qPCRs at polycomb target genes using the indicated antibodies in Aebp2 Gt and matched wild-type ESCs. The Gapdh promoter is included as a negative control.

(C) Average ChIP-Rx signal profiles for H3K27me3 at the TSS (±20 kb) of all genes in Pali1 sKO and Pali1/2 dKO and matched wild-type ESCs. H3K27me3 is significantly less in both Pali1 sKO and Pali1/2 dKO ESCs compared to wild-type ($p < 2.2 \times 10^{-16}$).

(D) Average ChIP-Rx signal profiles for H3K27me3 at the TSS (±20 kb) of all genes in Aebp2 Gt and matched wild-type ESCs. H3K27me3 is significantly greater in the Aebp2 Gt ESCs compared to wild-type ($p = 4.26 \times 10^{-12}$).

(E) Venn diagram analysis of the changes in level of H3K27me3 at the promoters of the top 3,000 H3K27me3 genes in Pali1 sKO, Pali1/2 dKO, and Aebp2 Gt ESCs.

(F) Western blot analysis, using the indicated antibodies, of soluble and chromatin fractions of Pali1 sKO and Pali1/2 dKO and matched wild-type ESCs.

(G) ChIP-qPCRs at polycomb target genes using the indicated antibodies in Pali1 sKO, Pali1/2 dKO, and matched wild-type ESCs. The Gapdh promoter is included as a negative control.

(H) Histone methyltransferase assay using recombinant PRC2 with addition of AEBP2, JARID2, and/or PALI1 protein, as indicated. Recombinant oligonucleosomes were used as a substrate. The PRC2 complex used contained FLAG-EED, SUZ12, EZH2, and RBBP4. Incorporation of 3H from radiolabeled SAM was used as a readout of methyltransferase activity. Coomassie staining shows equal loading of oligonucleosomes.

(A, B, and G) All experiments were performed at least three times and a representative experiment is shown. Error bars are representative of technical triplicates. See also Figure S6.

of H3K27me3 at all transcription start sites confirmed a reduction in H3K27me3 levels in Pali1 sKO ESCs and a slightly greater reduction in H3K27me3 levels in Pali1/2 dKO ESCs (Figure 6C). A ChIP-Rx of H3K27me3 in the Aebp2 Gt and wild-type ESCs confirmed an increase in H3K27me3 levels at transcription start sites (TSSs) in the absence of Aebp2 (Figure 6D). Furthermore, 72% (2,166/3,000) of all PRC2 target genes gain H3K27me3 in the Aebp2 Gt ESCs, whereas 84.5% (2,537/3,000) lose

H3K27me3 in the Pali1/2 dKO ESCs (Figure 6E). Taken together, these data further support our hypothesis that Pali1/2 and Aebp2 regulate the same set of polycomb target genes.

Consistent with the idea of an antagonistic interplay between Pali1 and Aebp2, Pali1 protein levels, which are completely absent from the soluble fraction, increase on chromatin in Aebp2 Gt ESCs (Figure S6B), mirroring the increase in total abundance of Pali1. The levels of Jarid2 and Pcl2 also increase on chromatin in Aebp2 Gt ESCs (Figure S6B). Furthermore, chromatin westerns and ChIPs of Aebp2 in Pali1 sKO and Pali1/2 dKO, compared to wild-type ESCs, established that the short isoform of Aebp2 increases on chromatin and on polycomb target genes (Figures 6F and 6G). This is consistent with a general increase in the Aebp2 short isoform protein levels. Meanwhile, the chromatin bound levels of Jarid2, Pcl2, and Epop appear to be unchanged (Figure S6C). We speculate that the reason for the moderate 2-fold increase of Aebp2 levels in the ChIP analysis of the Pali1/2 KO ESCs compared to the perhaps 5- to 10-fold increase in short Aebp2 protein levels on chromatin in these cells is because the Aebp2 antibody recognizes both the short and the long forms of Aebp2 in the ChIP experiment. Interestingly, we observe a varying ratio of long to short Aebp2 protein between different wild-type ESC lines. However, despite this, we only observe consequences on the levels of the short form of Aebp2 in the Pali1/2 KO ESCs.

Taken together, these data support the idea that the Pali1 and Aebp2 proteins are central to an antagonism between the PRC2.1 and PRC2.2 complexes, and their absence disrupts the balance of PRC2-mediated H3K27me3 on polycomb target genes.

To explore the potential mechanism behind the changes in PRC2-mediated H3K27me3 upon loss of Aebp2 or Pali1/2, we speculated that PALI1-PRC2.1 has a stronger histone methyltransferase activity compared to AEBP2-PRC2.2. To test this, we performed a histone methyltransferase assay with the core PRC2 complex together with either recombinant human AEBP2, JARID2, or PALI1 proteins (Figure S6D). This established that PALI1 individually is the strongest at promoting PRC2 methyltransferase activity *in vitro*, whereas JARID2 and AEBP2 perform best synergistically (Figures 6H and S6E; Son et al., 2013). Furthermore, the addition of JARID2 to methyltransferase assays together with PALI1 did not increase activity any further compared to PALI1 alone, consistent with their mutual exclusivity within the PRC2.2 and PRC2.1 complexes, respectively. These data are consistent with the idea that the increases and decreases in H3K27me3 levels in the absence of either Aebp2 or Pali1, respectively, can be explained by consequent imbalances in the proportions of the Pali-PRC2.1 and Aebp2-PRC2.2 complexes, which have different degrees of histone methyltransferase activity individually.

### The Balance of PRC2.1 and PRC2.2 Activities Is Essential for the Proper Regulation of Polycomb Target Genes during ESC Differentiation

Next, we evaluated the consequences of the deregulated H3K27me3 levels in the Pali1 sKO, Pali1/2 dKO, and Aebp2 Gt ESCs on the regulation of polycomb target genes during induction to differentiate to embryoid bodies (EBs). RNA-seq analysis of knockout and matched wild-type ESCs, both before and after induction to differentiate to EBs, revealed that loss of Aebp2 led to an inability to properly activate a cohort of 525 polycomb target genes with functions associated with differentiation and germ layer development, including *Wnt3*, *Fgf8*, *Olig3*, *Msx3*, *Pou3f2*, *Otx1*, *FoxA2*, *Nova2*, *T*, and *Gata6* (Figure S7A). This failure to correctly activate polycomb target genes correlated well with those genes that have increased H3K27me3 levels in Aebp2 Gt ESCs (Figure 7A). The opposite phenotype was observed in the Pali1 sKO and Pali1/2 dKO ESCs, which had significantly higher expression levels of these genes compared to matched wild-type ESCs, 8 days post-induction to differentiate (Figure S7A). This result was largely coincident with those polycomb target genes that lose H3K27me3 in Pali1 sKO and Pali1/2 dKO ESCs (Figure 7A). The majority (84%) of the polycomb target genes that became more activated during differentiation in Pali1 sKO and Pali1/2 dKO were also impaired in their activation in Aebp2 Gt ESCs (Figure 7B). To validate these observations, we monitored the mRNA expression of two representative polycomb target genes, *Gata6* and *T*, by RT-PCR in an independent experiment (Figures 7C and 7D). These results are consistent with the changes in the levels of H3K27me3 on the promoters of these polycomb target genes in the knockout ESCs. Whereas the Aebp2 Gt ESCs have increased H3K27me3 levels, consistent with their impaired ability to activate polycomb target genes appropriately during differentiation, the Pali1 sKO and Pali1/2 dKO ESCs have decreased H3K27me3, consistent with a greater propensity to activate these genes during differentiation. Example genome viewer tracks exhibiting this trend are shown for the *Nova2* and *Msx3* genes (Figure S7B). Taken together, these results are consistent with a model in which the Pali1/2 and Aebp2 proteins are central to balancing the respective activities of PRC2.1 and PRC2.2 and that the loss of either leads to the dysregulation of polycomb target genes during differentiation.

## DISCUSSION

Here, we report the discovery of a family of vertebrate-specific proteins, PALI1 and PALI2, encoded by the *LCOR* and *LCORL* gene loci, respectively. These multi-domain proteins are defined by a PIP domain, which confers the ability to bind to the PRC2 complex. We anticipate that this discovery will open up new avenues of research linking polycomb function with CTBP corepressors and nuclear receptors in the regulation of cellular identity during development and in cancer.

The discovery of PALI1 and PALI2 raises interesting questions about how the PRC2 complex might be targeted to new target genes during cell fate decisions in mammals, the mechanisms for which remain poorly defined. As well as their PIP domain, the PALI1 and PALI2 proteins also contain several additional motifs and domains, thereby notably expanding the potential mechanisms by which the PRC2 complex could be targeted to specific chromatin sites. For instance, PALI1 shares its N terminus with the LCOR protein, which contains both CTBP binding motifs and a nuclear receptor binding box that has been reported to interact with the estrogen, progesterone, and thyroid hormone receptors (Fernandes et al., 2003; Palijan et al., 2009; Song et al.,

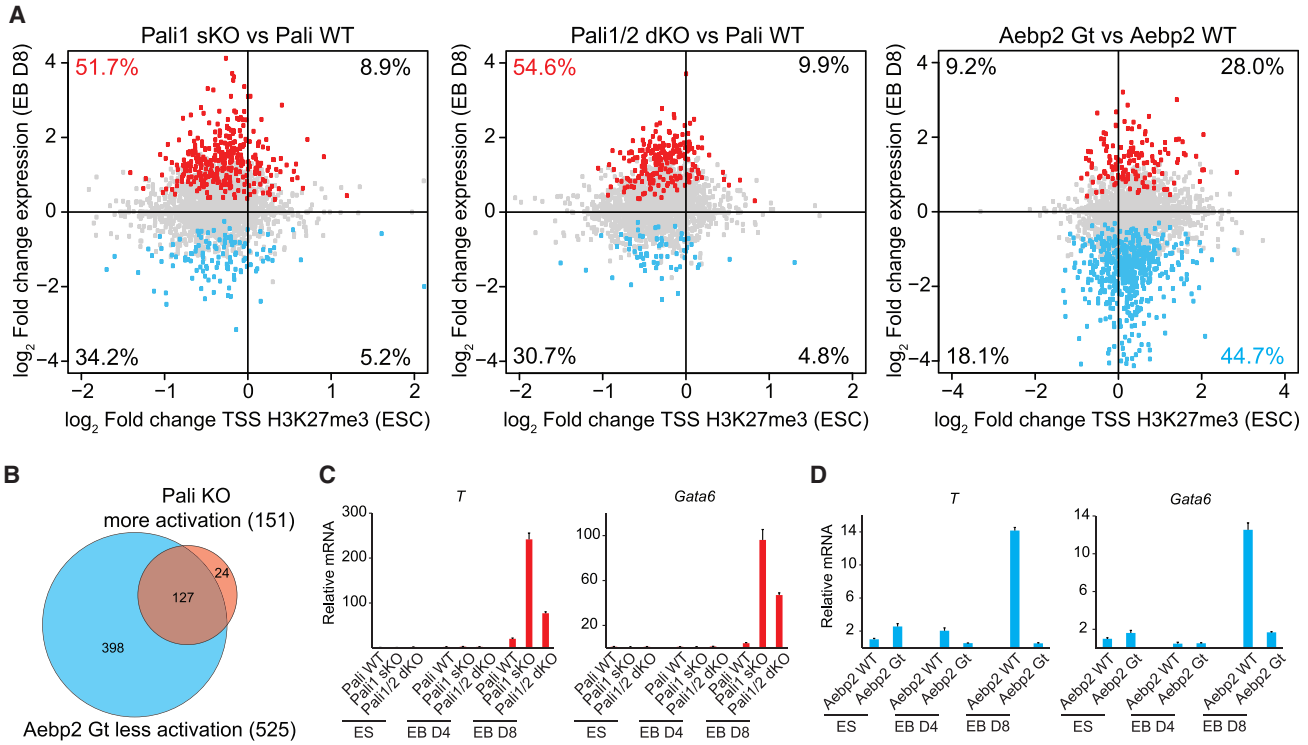



**Figure 7. The Balance of PRC2.1 and PRC2.2 Activities Is Essential for the Proper Regulation of Polycomb Target Genes during ESC Differentiation**

(A) The relationship between changes in H3K27me3 levels in Pali1 sKO, Pali1/2 dKO, and Aebp2 Gt ESCs and gene expression in EBs (day 8) for the top 3,000 H3K27me3-positive genes. Genes showing significant differential expression are shown in red (increased) and blue (decreased). The percentage of all genes in each quadrant is indicated.

(B) Venn diagram showing the overlap of H3K27me3-positive genes that are increased in both Pali1 sKO and Pali1/2 dKO and decreased in Aebp2 Gt EBs (D8).

(C) qRT-PCR analyses of Pali1 sKO, Pali1/2 dKO, and matched wild-type ESCs during induction to differentiate to embryoid bodies (EBs). The results are normalized to the mRNA levels of the Gapdh housekeeping gene.

(D) qRT-PCR analyses of Aebp2 Gt and matched wild-type ESCs during induction to differentiate to EBs. The results are normalized to the mRNA levels of the Gapdh housekeeping gene.

(C and D) All experiments were performed at least three times and a representative experiment is shown. Error bars are representative of technical triplicates. See also Figure S7.

2012). In addition to possessing two CTBP binding motifs, the PALI2 protein also contains three AT hooks, with the potential for direct interaction with DNA. Therefore, the multi-domain PALI1 and PALI2 proteins provide potential mechanisms by which CTBP1/2 co-repressors and nuclear receptors could contribute to targeting or regulating PRC2 activity. However, despite several efforts to determine whether Pali1 localizes at PRC2 target genes, we were unsuccessful by both endogenous and epitope-tagged ChIP approaches (data not shown). It will be important to determine whether these are just technical issues or instead reflective of a transient interaction of Pali1 with chromatin. Furthermore, it will be important to delineate the respective functions of the several PALI1 and PALI2 domains and their requirement for their interplay with co-repressors, nuclear receptors, and other chromatin regulators during differentiation and development.

Whereas the knockout of core PRC2 members, such as *Eed*, *Suz12*, and *Ezh2*, all result in early embryonic lethality (~E7.5–E8.5) and defects in gastrulation (Laugesen and Helin, 2014), the *Pali1*$^{-/-}$ mice generated here resulted in lethality between E11.5 and birth. This perinatal lethal phenotype is consistent with recent results from an independent group who knocked out both Lcor and Pali1 at a shared exon (Shalom-Barak et al., 2018). Whereas *Eed* and *Suz12* knockout mice lack H3K27me3, the *Pali1* knockout mice have reduced, but not a total loss of, H3K27me3 at E11.5. These data are consistent with the fact that there are two subtypes of the PRC2 complex with potential redundancy. Supporting this, *Pcl2* knockout mice have viable offspring whereas *Jarid2*- and *Aebp2*-null mice survive until mid- to late-embryonic stages (Grijzenhout et al., 2016; Laugesen and Helin, 2014). These data all support the idea that, upon the impairment of the activity of one subtype of the PRC2 complex, embryos can remain viable until later in development due to the activity of the other PRC2 subtype. However, it is clear that, for proper development in the mid- to late-stage embryo, the balance of both PRC2.1 and PRC2.2 is essential.

The discovery of the PALI1 and PALI2 proteins provides additional mechanisms by which the activity of the PRC2 complex could be deregulated in cancer and potential new opportunities to treat it. PRC2 function is frequently deregulated in many

cancer types due to genetic perturbation (Conway et al., 2015). These include recurrent heterozygous "change-of-function" mutations in *EZH2*, which occur in 22% of B cell lymphomas; loss-of-function mutations in *EZH2*, *SUZ12*, *EED*, and *AEBP2* in T cell leukemias and malignant peripheral nerve sheath tumors; and H3K27M mutations in ~78% of pediatric gliomas (Comet et al., 2016; Conway et al., 2015). More broadly, whereas we have not yet characterized the potential interplay between PALI1 and nuclear receptors, it will be exciting to investigate this link, considering their role as key drivers in hormone-dependent cancers, including prostate and breast cancer (Deroo and Korach, 2006; Heinlein and Chang, 2004).

In summary, this study adds to the growing evidence of an expanded and sub-functionalized polycomb system in vertebrates. We believe this work will also contribute to a better mechanistic understanding of the interplay between polycombs, transcriptional co-repressors, and nuclear receptors during cell fate decisions and carcinogenesis in higher eukaryotic cells.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - Cell Culture
  - FLAG IPs for Mass Spectrometric Analyses
  - Mass Spectrometry Analysis
  - Cloning and Plasmid Generation
  - Generation of Knockdown and Overexpression Cell Lines
  - Preparation of Whole-Cell Protein Lysates
  - RNA Preparation and RT-PCR Analysis
  - Baculovirus Production and Purification PRC2, PALI1, and LCOR Recombinant Proteins
  - Histone Methyltransferase Assay
  - FLAG IPs
  - Evolutionary Analysis
  - Endogenous IPs
  - Cellular Fractionations
  - Chromatin IPs and Quantitative ChIP-Rx
  - RNA-Seq and Bioinformatics Analysis
  - ESC and Mouse Targeting Strategy
  - Generation of Pali1/2 Double Knockout Cells
- DATA AND SOFTWARE AVAILABILITY

## SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and three tables and can be found with this article online at https://doi.org/10.1016/j.molcel.2018.03.005.

## ACKNOWLEDGMENTS

We thank members of the Bracken lab for helpful discussions and critical reading of the manuscript. We also thank A. Fisher, A. Wutz, and N. Brockdorff for *Jarid2*, *Eed*, and *Aebp2* knockout ESCs, respectively. We thank R. Liefke, Y. Shi, and A. Leitner for sharing data and A. Flaus for advice on the characterization of the PIP domain. We also acknowledge support of the Spanish Ministry of Economy and Competitiveness (BFU2016-75008-P) to LDC. Work in the Bracken Lab is supported by Science Foundation Ireland under the SFI Investigators Programme (SFI/16/IA/4562) and the BBSRC-SFI (SFI/17/BBSRC/3415), the Irish Research Council, St. Vincent's Foundation and the Irish Cancer Society Collaborative Cancer Research Centre, and BREAST-PREDICT grant CCRC13GAL (http://www.breastpredict.com). E.G. is supported by a research grant from Science Foundation Ireland (SFI) under grant number 15/IA/3104. D.H. was supported by a postdoctoral fellowship from the Fondation pour la Recherche Médicale (SPF20150934266).

## AUTHOR CONTRIBUTIONS

E.C., E.J., and A.P.B. conceived the project. E.J. identified the PALI1 protein in mass spectrometric analysis of FLAG-PCL1/PHF1 and performed the initial characterization of its association with PRC2. E.J., G.O., E.C., A.W., K.W., and G.C. processed mass spectrometry samples and helped interpret the results. E.C., E.J., O.D., M.M., and I.P.-K. contributed to the biochemical characterization of the PIP domain in PALI1 and PALI2. A.M.R. and A.M. performed the evolutionary analyses of the proteins encoded by the *LCOR* and *LCORL* gene loci. E.G. performed the fractionation experiments. E.H., C.H., and C.C. purified recombinant PALI1 and LCOR for histone methyltransferase assays, which D.H. and R.M. designed and D.H. performed. M.N., G.S., T.I., Y.K., S.I., and H.K. generated *Pali1* knockout ESCs and characterized the *Pali1* knockout mouse. P.C., L.D.C., D.J.F., and K.H. performed the RNA-seq and bioinformatics analyses. E.C. characterized the function of PALI proteins in mouse ESCs. E.C., E.J., and A.P.B. co-wrote the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interest.

Received: July 6, 2017
Revised: January 11, 2018
Accepted: March 2, 2018
Published: April 5, 2018

## REFERENCES

Akasaka, T., Kanno, M., Balling, R., Mieza, M.A., Taniguchi, M., and Koseki, H. (1996). A role for mel-18, a Polycomb group-related vertebrate gene, during theanteroposterior specification of the axial skeleton. Development *122*, 1513–1522.

Alekseyenko, A.A., Gorchakov, A.A., Kharchenko, P.V., and Kuroda, M.I. (2014). Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP-XL cross-linking and affinity purification. Proc. Natl. Acad. Sci. USA *111*, 2488–2493.

Beringer, M., Pisano, P., Di Carlo, V., Blanco, E., Chammas, P., Vizán, P., Gutiérrez, A., Aranda, S., Payer, B., Wierer, M., and Di Croce, L. (2016). EPOP functionally links elongin and polycomb in pluripotent stem cells. Mol. Cell *64*, 645–658.

Blackledge, N.P., Rose, N.R., and Klose, R.J. (2015). Targeting polycomb systems to regulate gene expression: modifications to a complex story. Nat. Rev. Mol. Cell Biol. *16*, 643–649.

Bracken, A.P., Pasini, D., Capra, M., Prosperini, E., Colli, E., and Helin, K. (2003). EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. EMBO J. *22*, 5323–5335.

Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev. *20*, 1123–1136.

Bracken, A.P., Kleine-Kohlbrecher, D., Dietrich, N., Pasini, D., Gargiulo, G., Beekman, C., Theilgaard-Mönch, K., Minucci, S., Porse, B.T., Marine, J.C., et al. (2007). The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells. Genes Dev. *21*, 525–530.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

Brien, G.L., Gambero, G., O'Connell, D.J., Jerman, E., Turner, S.A., Egan, C.M., Dunne, E.J., Jurgens, M.C., Wynne, K., Piao, L., et al. (2012). Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. Nat. Struct. Mol. Biol. *19*, 1273–1281.

Brien, G.L., Healy, E., Jerman, E., Conway, E., Fadda, E., O'Donovan, D., Krivtsov, A.V., Rice, A.M., Kearney, C.J., Flaus, A., et al. (2015). A chromatin-independent role of Polycomb-like 1 to stabilize p53 and promote cellular quiescence. Genes Dev. *29*, 2231–2243.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.

Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol. Cell *15*, 57–67.

Choi, J., Bachmann, A.L., Tauscher, K., Benda, C., Fierz, B., and Müller, J. (2017). DNA binding by PHF1 prolongs PRC2 residence time on chromatin and thereby promotes H3K27 methylation. Nat. Struct. Mol. Biol. *24*, 1039–1047.

Comet, I., Riising, E.M., Leblanc, B., and Helin, K. (2016). Maintaining cell identity: PRC2-mediated regulation of transcription and cancer. Nat. Rev. Cancer *16*, 803–810.

Conway, E., Healy, E., and Bracken, A.P. (2015). PRC2 mediated H3K27 methylations in cellular identity and cancer. Curr. Opin. Cell Biol. *37*, 42–48.

Cooper, S., Grijzenhout, A., Underwood, E., Ancelin, K., Zhang, T., Nesterova, T.B., Anil-Kirmizitas, B., Bassett, A., Kooistra, S.M., Agger, K., et al. (2016). Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. Nat. Commun. *7*, 13661.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Deroo, B.J., and Korach, K.S. (2006). Estrogen receptors and human disease. J. Clin. Invest. *116*, 561–570.

Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics *21*, 3433–3434.

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nat. Protoc. *7*, 1728–1740.

Fernandes, I., Bastien, Y., Wai, T., Nygard, K., Lin, R., Cormier, O., Lee, H.S., Eng, F., Bertos, N.R., Pelletier, N., et al. (2003). Ligand-dependent nuclear receptor corepressor LCoR functions by histone deacetylase-dependent and -independent mechanisms. Mol. Cell *11*, 139–150.

Ferrari, K.J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., Fischle, W., Bonaldi, T., and Pasini, D. (2014). Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. Mol. Cell *53*, 49–62.

Finn, R.D., Clements, J., Arndt, W., Miller, B.L., Wheeler, T.J., Schreiber, F., Bateman, A., and Eddy, S.R. (2015). HMMER web server: 2015 update. Nucleic Acids Res. *43* (W1), W30–W38.

Grijzenhout, A., Godwin, J., Koseki, H., Gdula, M.R., Szumska, D., McGouran, J.F., Bhattacharya, S., Kessler, B.M., Brockdorff, N., and Cooper, S. (2016). Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. Development *143*, 2716–2723.

Hauri, S., Comoglio, F., Seimiya, M., Gerstung, M., Glatter, T., Hansen, K., Aebersold, R., Paro, R., Gstaiger, M., and Beisel, C. (2016). A high-density map for navigating the human polycomb complexome. Cell Rep. *17*, 583–595.

Heinlein, C.A., and Chang, C. (2004). Androgen receptor in prostate cancer. Endocr. Rev. *25*, 276–308.

Højfeldt, J.W., Laugesen, A., Willumsen, B.M., Damhofer, H., Hedehus, L., Tvardovskiy, A., Mohammad, F., Jensen, O.N., and Helin, K. (2018). Accurate H3K27 methylation can be established de novo by SUZ12-directed PRC2. Nat. Struct. Mol. Biol. *25*, 225–232.

Holoch, D., and Margueron, R. (2017). Mechanisms regulating PRC2 recruitment and enzymatic activity. Trends Biochem. Sci. *42*, 531–542.

Isono, K., Endo, T.A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B.E., and Koseki, H. (2013). SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. Dev. Cell *26*, 565–577.

Kalb, R., Latwiel, S., Baymaz, H.I., Jansen, P.W., Müller, C.W., Vermeulen, M., and Müller, J. (2014). Histone H2A monoubiquitination promotes histone H3 methylation in polycomb repression. Nat. Struct. Mol. Biol. *21*, 569–571.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., et al.; Roadmap Epigenomics Consortium (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

Landeira, D., Bagci, H., Malinowski, A.R., Brown, K.E., Soza-Ried, J., Feytout, A., Webster, Z., Ndjetehe, E., Cantone, I., Asenjo, H.G., et al. (2015). Jarid2 coordinates Nanog expression and PCP/Wnt signaling required for efficient ESC differentiation and early embryo development. Cell Rep. *12*, 573–586.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lau, M.S., Schwartz, M.G., Kundu, S., Savol, A.J., Wang, P.I., Marr, S.K., Grau, D.J., Schorderet, P., Sadreyev, R.I., Tabin, C.J., and Kingston, R.E. (2017). Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. Science *355*, 1081–1084.

Laugesen, A., and Helin, K. (2014). Chromatin repressive complexes in stem cells, development, and cancer. Cell Stem Cell *14*, 735–751.

Lee, H.G., Kahn, T.G., Simcox, A., Schwartz, Y.B., and Pirrotta, V. (2015). Genome-wide activities of Polycomb complexes control pervasive transcription. Genome Res. *25*, 1170–1181.

Li, G., Margueron, R., Ku, M., Chambon, P., Bernstein, B.E., and Reinberg, D. (2010). Jarid2 and PRC2, partners in regulating gene expression. Genes Dev. *24*, 368–380.

Li, H., Liefke, R., Jiang, J., Kurland, J.V., Tian, W., Deng, P., Zhang, W., He, Q., Patel, D.J., Bulyk, M.L., et al. (2017). Polycomb-like proteins link the PRC2 complex to CpG islands. Nature *549*, 287–291.

Liefke, R., Karwacki-Neisius, V., and Shi, Y. (2016). EPOP interacts with elongin BC and USP7 to modulate the chromatin landscape. Mol. Cell *64*, 659–672.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Margueron, R., and Reinberg, D. (2011). The polycomb complex PRC2 and its mark in life. Nature *469*, 343–349.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D., and Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol. Cell *32*, 503–518.

Mullin, N.P., Gagliardi, A., Khoa, L.T.P., Colby, D., Hall-Ponsele, E., Rowe, A.J., and Chambers, I. (2017). Distinct contributions of tryptophan residues within the dimerization domain to Nanog function. J. Mol. Biol. *429*, 1544–1553.

Orlando, D.A., Chen, M.W., Brown, V.E., Solanki, S., Choi, Y.J., Olson, E.R., Fritz, C.C., Bradner, J.E., and Guenther, M.G. (2014). Quantitative ChIP-seq normalization reveals global modulation of the epigenome. Cell Rep. *9*, 1163–1170.

Palijan, A., Fernandes, I., Verway, M., Kourelis, M., Bastien, Y., Tavera-Mendoza, L.E., Sacheli, A., Bourdeau, V., Mader, S., and White, J.H. (2009). Ligand-dependent corepressor LCoR is an attenuator of progesterone-regulated gene expression. J. Biol. Chem. *284*, 30275–30287.

Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. *23*, 4061–4071.

Pasini, D., Cloos, P.A., Walfridsson, J., Olsson, L., Bukowski, J.P., Johansen, J.V., Bak, M., Tommerup, N., Rappsilber, J., and Helin, K. (2010). JARID2 regulates binding of the polycomb repressive complex 2 to target genes in ES cells. Nature *464*, 306–310.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat. Protoc. *11*, 1650–1667.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44* (W1), W160–W165.

Sanulli, S., Justin, N., Teissandier, A., Ancelin, K., Portoso, M., Caron, M., Michaud, A., Lombard, B., da Rocha, S.T., Offer, J., et al. (2015). Jarid2 methylation via the PRC2 complex regulates H3K27me3 deposition during cell differentiation. Mol. Cell *57*, 769–783.

Schoeftner, S., Sengupta, A.K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., and Wutz, A. (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. EMBO J. *25*, 3110–3122.

Schuettengruber, B., Bourbon, H.M., Di Croce, L., and Cavalli, G. (2017). Genome regulation by polycomb and Trithorax: 70 years and counting. Cell *171*, 34–57.

Shalom-Barak, T., Liersemann, J., Memari, B., Flechner, L., Devor, C., Bernardo, T.M., Kim, S., Matsumoto, N., Friedman, S.L., Evans, R.M., et al. (2018). Ligand-dependent corepressor (LCoR) is a rexinoid-inhibited PPARγ-RXRα coactivator. Mol. Cell. Biol., MCB.00107-17.

Shen, L., Shao, N., Liu, X., and Nestler, E. (2014). ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics *15*, 284.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Smits, A.H., Jansen, P.W., Poser, I., Hyman, A.A., and Vermeulen, M. (2013). Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Res. *41*, e28.

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

Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. *4*, 1521.

Song, Y., Shan, S., Zhang, Y., Liu, W., Ding, W., Ren, W., Xia, H., Li, X., Zhang, Q., Zhao, L., et al. (2012). Ligand-dependent corepressor acts as a novel corepressor of thyroid hormone receptor and represses hepatic lipogenesis in mice. J. Hepatol. *56*, 248–254.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.

Streubel, G., Fitzpatrick, D.J., Oliviero, G., Scelfo, A., Moran, B., Das, S., Munawar, N., Watson, A., Wynne, K., Negri, G.L., et al. (2017). Fam60a defines a variant Sin3a-Hdac complex in embryonic stem cells required for self-renewal. EMBO J. *36*, 2216–2232.

Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., et al. (2016). Ensembl 2016. Nucleic Acids Res. *44* (D1), D710–D716.

Zhang, Z., Jones, A., Sun, C.W., Li, C., Chang, C.W., Joo, H.Y., Dai, Q., Mysliwiec, M.R., Wu, L.C., Guo, Y., et al. (2011). PRC2 complexes with JARID2, MTF2, and esPRC2p48 in ES cells to modulate ES cell pluripotency and somatic cell reprogramming. Stem Cells *29*, 229–240.

Zhu, S., Xiang, J.F., Chen, T., Chen, L.L., and Yang, L. (2013). Prediction of constitutive A-to-I editing sites from human transcriptomes in the absence of genomic sequences. BMC Genomics *14*, 206.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Anti-FLAG | Sigma-Aldrich | F3165; RRID: AB_259529 |
| Anti-EZH2 (Western) | Prof. Kristian Helin; Pasini et al., 2004 | BD43 |
| Anti-EZH2 (IP) | Prof. Kristian Helin; Bracken et al., 2006 | AC22 |
| Anti-LCOR | Santa Cruz Biotechnology | sc-163009; RRID: AB_10844786 |
| Anti-C10ORF12/ PALI1 | Merck | ABE1367 |
| Anti-BMI1 | Prof. Kristian Helin; Bracken et al., 2007 | DC9 |
| Anti-CTBP | Santa Cruz Biotechnology | sc-11390; RRID: AB_2086634 |
| Anti-ACTIN | MP Biomedicals | #08691001; RRID: AB_2335127 |
| Anti-PCL1 | Proteintech | 15663-1-AP |
| Anti-G9A | Cell Signaling Technology | #3306 (C6H3); RRID: AB_2097647 |
| Anti-H2Aub | Cell Signaling Technology | #8240 (D27C4); RRID: AB_10891618 |
| Anti-H3K27me3 (Western) | Active Motif | #61017; RRID: AB_2614987 |
| Anti-H3K27me3 (ChIP-qPCR) | Merck | #07-449 |
| Anti-H3 | Abcam | ab1791; RRID: AB_302613 |
| Anti-EED | Prof. Kristian Helin; Bracken et al., 2003 | #AA19.30 |
| Anti-AEBP2 (Western) | Cell Signaling Technology | #14129 (D7C6X) |
| Anti-AEBP2 | Prof. Neil Brockdorff; Grijzenhout et al., 2016 | N/A |
| Anti-JARID2 | Abcam | ab48137; RRID: AB_881370 |
| Normal mouse IgG | Merck | #12371 |
| Anti-FLAG M2 affinity gel | Sigma-Aldrich | A2220; RRID: AB_10063035 |
| Anti-SUZ12 | Cell Signaling Technology | #3737 (D39F6) |
| Anti-HA | Cell Signaling Technology | #3724 (C29F4); RRID: AB_1549585 |
| Anti-RYBP | Sigma-Aldrich | PRS2227; RRID: AB_1847589 |
| Anti-GAPDH | Santa Cruz Biotechnology | sc-25778; RRID: AB_10167668 |
| Anti-H3K9me2 | Abcam | 1220; RRID: AB_449854 |
| HRP-linked rabbit secondary antibody | Sigma-Aldrich | A0545; RRID: AB_257896 |
| HRP-linked mouse secondary antibody | Merck | 401253 |
| HRP-linked goat secondary antibody | Santa Cruz Biotechnology | sc-2020; RRID: AB_631728 |
| Anti-SET | Santa Cruz Biotechnology | sc-25564; RRID: AB_2185634 |
| Anti-USP11 | Cambridge Biosciences | A301-613A |
| Anti-USP22 | Abcam | ab4812; RRID: AB_2212745 |
| Anti-H3K27me2 | Cell Signaling Technology | #9728 (D18C8) |
| Anti-PCL2/MTF2 | Proteintech | 16208-1-AP; RRID: AB_2147370 |
| Anti-EPOP | Prof. Luciano Di Croce; Beringer et al., 2016 | N/A |
| Anti-H3K27me3 (ChIP-RX) | Cell Signaling Technology | #9733 (C36B11); RRID: AB_2616029 |
| Bacterial and Virus Strains | | |
| *E. coli* (DH5α)-One Shot Top10 chemically competent | Thermo Fisher Scientific | C404003 |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Lentivirus (pLKO-Puro) | This study | N/A |
| Baculovirus (pVL) | This study | N/A |
| **Biological Samples** | | |
| E11.5 mouse embryos | This study | N/A |
| **Chemicals, Peptides, and Recombinant Proteins** | | |
| FLAG/His-PALI1 | This study | N/A |
| FLAG/His-LCOR | This study | N/A |
| Recombinant PRC2 | This study | N/A |
| TAT-Cre | Prof. Diego Pasini | N/A |
| LIF | This study | N/A |
| 3X FLAG Peptide | Sigma-Aldrich | F4799 |
| FLAG/His-AEBP2 | This study | N/A |
| FLAG/His-JARID2 | This study | N/A |
| **Critical Commercial Assays** | | |
| Lipofectamine 2000 | Thermo Fisher Scientific | #11668019 |
| Protein-A-Sepharose 4B | Sigma-Aldrich | P9424 |
| FastStart Universal SYBR Green Master | Roche | #4913914001 |
| LR Clonase II plus enzyme | Thermo Fisher Scientific | #12538-120 |
| pCR8/GW/TOPO Cloning Kit | Thermo Fisher Scientific | K2500-20 |
| High Capacity cDNA Reverse Transcription Kit | Thermo Fisher Scientific | #4368813 |
| Rneasy purification kit | QIAGEN | #74106 |
| Cellfectin II Reagent | Thermo Fisher Scientific | #10362100 |
| Benzonase nuclease | Sigma-Aldrich | E1014 |
| Geneart Site-Directed Mutagenesis System | Thermo Fisher Scientific | A13282 |
| NEBNext ultra DNA library preparation kit | NEB | E7370 |
| Kapa library quantification for Illumina | Roche | #KK4835 |
| Truseq stranded total RNA library prep kit ribo-zero human/ mouse/ rat | Illumina | 20020612 |
| **Deposited Data** | | |
| RNA-seq data | This paper | GEO: GSE100679 |
| H3K27me3 ChIP-RX data | This paper | GEO: GSE108302 |
| H3K27me3 ChIP-seq data | Streubel et al., 2017 | GEO: GSE81081 |
| Deposited raw imaging data | This paper | https://doi.org/10.17632/mhynrhktv4.1 |
| **Experimental Models: Cell lines** | | |
| HEK293T | This lab | Brien et al., 2012 |
| HMEC | This lab | Brien et al., 2012 |
| sf9 | Prof. Diego Pasini | N/A |
| High Five | Prof. Diego Pasini | N/A |
| Eed WT and null ESCs | Prof. Anton Wutz | Schoeftner et al., 2006 |
| Aebp2 WT and gt ESCs | Prof Neil Brockdorff | Grijzenhout et al., 2016 |
| Jarid2 WT and null ESCs | Prof Amanda Fisher | Landeira et al., 2015 |
| Pali1/2 WT and null ESCs | This study | N/A |
| **Oligonucleotides** | | |
| Primers for cloning, gRNA, shRNA, RT-qPCR and ChIP qPCR, see Table S3 | This study | N/A |
| **Recombinant DNA** | | |
| pX458 pSpCas9(BB)-2A-GFP | Addgene | #48138 |
| pPAX8 | Brien et al., 2012 | N/A |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| pVSVG | Brien et al., 2012 | N/A |
| pLKO.1 TRC Puro | Addgene | #10878 |
| pMINKIO FLAG/HA GW Dest | Brien et al., 2015 | N/A |
| pLENTI EFIA FLAG/HA GW Dest | Bracken Laboratory | N/A |
| pVL no tag GW Dest | Prof. Diego Pasini | N/A |
| pVL FLAG/His GW Dest | Prof. Diego Pasini | N/A |
| pCR8-LCOR | This study | N/A |
| pCR8-C10ORF12 | This study | N/A |
| pCR8-PALI1 full length | This study | N/A |
| pCR8-PALI1 Fragment 1 | This study | N/A |
| pCR8-PALI1 Fragment 2 | This study | N/A |
| pCR8-PALI1 Fragment 3 | This study | N/A |
| pCR8-PALI1 Fragment 4 | This study | N/A |
| pCR8-PALI1 Fragment 5 | This study | N/A |
| pCR8-PALI1 PIR1 | This study | N/A |
| pCR8-PALI1 PIR2 | This study | N/A |
| pCR8-PALI1 PIR3 | This study | N/A |
| pCR8-PALI1 PIR4 | This study | N/A |
| pCR8-PALI1 PIR5 | This study | N/A |
| pCR8-PALI1 PIR6 | This study | N/A |
| pCR8-PALI1 Fragment 4 W1125A | This study | N/A |
| pCR8-PALI1 Fragment 4 W1186A | This study | N/A |
| pCR8-PALI1 Fragment 4 W1125A, W1186A | This study | N/A |
| pCR8-PALI2 Fragment 4 | This study | N/A |
| pCR8-AEBP2 | Prof. Kristian Helin | N/A |
| pCR8-JARID2 | This study | N/A |
| pCR8-PCL2 | Brien et al., 2015 | N/A |
| pCR8-PCL1 | Brien et al., 2015 | N/A |
| pCR8-EZH2 | Bracken Laboratory | N/A |
| pCR8-EED | Bracken Laboratory | N/A |
| pCR8-SUZ12 | Bracken Laboratory | N/A |
| Software and Algorithms | | |
| Maxquant 1.3.0.5 and 1.1.36 | Cox and Mann, 2008 | http://www.coxdocs.org/doku.php?id=maxquant:start |
| ProteinProspector MS-Product | | http://prospector.ucsf.edu/prospector/mshome.htm |
| RaxML | Stamatakis, 2014 | https://sco.h-its.org/exelixis/web/software/raxml/index.html |
| Clustal Omega | Sievers et al., 2011 | https://www.ebi.ac.uk/Tools/msa/clustalo |
| TBLASTN | Camacho et al., 2009 | https://blast.ncbi.nlm.nih.gov/Blast.cgi |
| HMMER | Finn et al., 2015 | http://www.hmmer.org |
| Ensembl | Yates et al., 2016 | http://www.ensembl.org//uswest.ensembl.org/index.html?redirectsrc=//www.ensembl.org%2Findex.html |
| IUPRED | Dosztányi et al., 2005 | http://iupred.enzim.hu |
| Bowtie v.2.1.0 | Langmead and Salzberg, 2012 | http://bowtie-bio.sourceforge.net/index.shtml |
| deepTools suite | Ramírez et al., 2016 | http://deeptools.readthedocs.io/en/latest |
| Kallisto v.0.43.1 | Bray et al., 2016 | http://pachterlab.github.io/kallisto//releases/2017/03/20/v0.43.1 |
| tximport | Soneson et al., 2015 | https://bioconductor.org/packages/release/bioc/html/tximport.html |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| DESeq2 | Love et al., 2014 | https://bioconductor.org/packages/release/bioc/html/DESeq2.html |
| macs2 v.2.1.1 | Feng et al., 2012 | https://github.com/taoliu/MACS |
| HISAT v2.0.5 | Pertea et al., 2016 | https://ccb.jhu.edu/software/hisat/index.shtml |
| bamTobw.sh utility | Zhu et al., 2013 | N/A |

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Adrian Bracken (adrian.bracken@tcd.ie).

## METHOD DETAILS

### Cell Culture

HEK293T were cultured in DMEM supplemented with 10% FBS (GIBCO), 100 U/mL penicillin (GIBCO) and 100 U/mL streptomycin (GIBCO). Embryonic stem cells (ESCs) were grown on gelatinised culture dishes in GMEM (sigma) supplemented with 10% ESC qualified FBS (Millipore), 100 U/mL penicillin (GIBCO), 100 U/mL streptomycin (GIBCO), 50μM β-mercaptoethanol (sigma), 1:100 Glutamax (GIBCO), 1:100 non-essential amino acids (GIBCO), 1mM sodium pyruvate (GIBCO) and 1:500 homemade leukemia inhibitory factor (LIF). For Embryoid body differentiation, 2-3x$10^6$ million ESCs were washed three times with PBS (Lonza) and seeded to non-adherent Petri-dishes in ESC media without LIF. The media was changed every 2 days. Sf9 and High Five cells were cultured in Hinks TNMFH (sigma) supplemented with 10% FBS (GIBCO) and 0.5% Gentamycin (Sigma). S2 cells were cultured in Schneider media (Sigma S0146) supplemented with 10% heat inactivated FBS (GIBCO).

Primary Human Mammary Epithelial Cells (HMECs) were passaged in the M87A medium: 50% MM4 medium [DMEM/F12 (GIBCO), 10μg/ml insulin (Sigma), 10nM tri-iodothyronine (Sigma), 1nM β-estradiol (Sigma), 0.1μg/ml hydrocortisone (Sigma), 0.5% FBS (26140, GIBCO), 5ng/ml epidermal growth factor (Peprotech), 2mM glutamine (Lonza), 1ng/ml cholera toxin (Sigma)] and 50% MCDB170 medium [MEGM media (Lonza) supplemented with 5μg/ml transferrin (Lozna), $10^{-5}$ isoproterenol (Sigma), and 2mM glutamine (Lonza). M87A media was further supplemented with 0.1nM oxytocin (Bachem) and 0.1% Albu-Max 1 (Invitrogen). For passaging, ~80% confluent cells were trypsinised by STV (5.37mM KCl, 6.9mM $NaHCO_3$, 136.9mM NaCl, 5.55mM D-Glucose, 0.54mM EDTA, 500mg/L Trypsin (1:250)) buffer and split at a 1 to 5 ratio.

### FLAG IPs for Mass Spectrometric Analyses

Nuclear or whole cell pellets were re-suspended in 4-5 packed cell volumes of IPH buffer (50mM Tris-HCl pH 8.0, 150mM NaCl, 0.5mM EDTA, 0.5% NP-40 detergent, 0.5 μM DTT, 2 μg/mL aprotinin, 1 μg/mL leupeptin, 1mM PMSF) and incubated for 30 min at 4°C with rotation. $MgCl_2$ was then added to a final concentration of 7mM before the digestion of chromatin with 250U/mL benzonase-nuclease overnight at 4°C with rotation. The IPH whole cell or nuclear lysates were pre-cleared to remove non-specifically binding proteins by incubation with mouse IgG agarose beads (Sigma). FLAG-tagged proteins were IP from these lysates using anti-FLAG-M2 affinity gel agarose (Sigma; A2220) for 4-16 hr at 4°C with rotation. The agarose beads were washed five times with IPH buffer and IP material was obtained in solution by competitive elution with 250μg/mL 3xFLAG peptide (Sigma) while shaking at 25°C.

### Mass Spectrometry Analysis

For in gel digests, after IP of FLAG-PCL1, the eluates were denatured in SDS-loading buffer and then separated on 4%–12% gradient Nu-PAGE gels (Novex) and stained with GelCode Blue solution (Thermo Scientific). The control and FLAG-PCL1 lanes were excised and cut into 10 fragments from the top to the bottom of the gel. Each fragment was diced into small (~1 $mm^3$) pieces and washed three times with 25mM $NH_4HCO_3$/50% acetone, dehydrated in acetonitrile/$NH_4HCO_3$ (3:2) and rehydrated with 50mM $NH_4HCO_3$ twice, followed by dehydration with undiluted acetonitrile and reduction of proteins in 10mM DTT in 25mM $NH_4HCO_3$ for 1 hr at 56°C and a subsequent incubation in 55mM iodoacetamide for 1 h. The gel pieces were washed with 25mM $NH_4HCO_3$, then 25mM $NH_4HCO_3$/50% acetonitrile and dehydrated. The peptides were then digested with one half of Trypsin Singles reaction (Sigma, T7575) in 25mM $NH_4HCO_3$ overnight at 37°C. Finally, peptides were extracted from the gel pieces with 50% acetone/ 5% formic acid (Sigma, 94318) and dried by vacuum concentration. The final peptide sample was resuspended in 20 μL 0.1% formic acid.

Material IP with FLAG-LCOR, FLAG-C10ORF12 and FLAG-PALI1 proteins, as well as negative control IP, were treated with trypsin while still bound to FLAG-agarose. The digested peptides of all FLAG IPs were analyzed on a Thermo Scientific LTQ Orbitrap. Each sample was injected onto a nanoACQUITY Symmetry C18 trap (5 μm particle size, 180 μm x 20mm) in buffer A (0.1% formic acid in water) at a flow rate of 4 μl/min and then separated over a nanoACQUITY BEH C18 analytical column (1.7 μm particle size, 100 μm x 100mm) over 1 h with a gradient from 2% to 25% buffer B (99.9% acetone/0.1% formic acid) at a flow rate of 0.4 μl/min.

The mass spectrometer continuously collected data in a data-dependent manner, collecting a survey scan in the Orbitrap mass analyzer at 60,000 resolution with an automatic gain control (AGC) target of 1 × 10E6 followed by collision-induced dissociation (CID) MS/MS scans of the 10 most abundant ions in the survey scan in the ion trap with an AGC target of 5,000, a signal threshold of 1,000, a 2.0 Da isolation width, and 30ms activation time at 35% normalized collision energy. Charge state screening was employed to reject unassigned or 1+ charge states. Dynamic exclusion was enabled to ignore masses for 30 s that had been previously selected for fragmentation. Raw files were processed using version 1.1.36 of MaxQuant. For protein identification the human UniProt database (release 2013_12; 67,911 entries) was combined with the reversed sequences and sequences of widespread contaminants, such as human keratins. Carbamidomethylation was set as fixed modification. Variable modifications were oxidation (M) and N-acetyl (protein). Initial peptide mass tolerance was set to 6 ppm and fragment mass tolerance was set to 20 ppm. The peptide and protein false discovery rates (FDR) were set to 0.01.

### Cloning and Plasmid Generation

Full-length ORFs of *PCL1, PCL2, LCOR, C10ORF12* and PALI1, as well as all protein fragments used in this study, were PCR amplified (primers available upon request) from cDNA generated from either HMECs or HEK293T human cell lines. The human *JARID2* ORF was PCR amplified from Open Biosystems image clone 4520786. All PCR products were inserted into the pCR8/GW/TOPO Gateway cloning entry vector (Invitrogen). The pCR8-*AEBP2* plasmid was a kind gift from Kristian Helin. All ORFs were subsequently sub-cloned into Gateway compatible expression vectors by recombination using LR-Clonase enzyme (Invitrogen). Single and double point mutants of PALI1 extended PIP were generated using the GeneArt site-directed mutagenesis system (Invitrogen), in accordance with manufacturer's instructions. For generation of stable shRNA knockdowns in HEK293T cells, the complimentary DNA oligonucleotides, containing indicated sequences, were annealed and ligated into TRC2 pLKO lentiviral expression vector (Sigma). The shSCR (CAACAAGATGAAGAGCACC) (Sigma, SHC202) was purchased to use as a negative control. Oligonucleotide sequences are available in Table S3.

### Generation of Knockdown and Overexpression Cell Lines

Lentiviral particles were produced by calcium phosphate transfection of HEK293T cells, which were seeded at 5 million cells per T75 flask the day prior to transfection with 10 μg of pLKO shRNA expression vector, 10 μg of viral packaging vector (pPAX8) and 8 μg of viral envelope vector (pVSVG) DNA. The media containing viral particles was collected at 48 and 72 hr post transfection, filtered through a 0.45 μm filter, and used directly to infect target cells for 8 hr in the presence of 5μg/ml polybrene. 24-48 hr later selection with puromycin (Simga) (0.5μg/ml).

Transient or stable ectopic expression of proteins were achieved using the calcium phosphate transfection of HEK293T with either pMINKIO FLAG-HA or pLENTI FLAG-HA expressions vectors. 1-10μg of DNA was transfected into HEK293T for 24-48 hr. For generation of stable cell lines the cells were selected with puromycin (1μg/mL) 48 hr post transfection.

### Preparation of Whole-Cell Protein Lysates

Cells were scraped down to collect them, washed three times in PBS and resuspended in ice cold High Salt buffer (50mM Tris-HCl, pH 7.2, 300mM NaCl, 0.5% (v/v) NP-40, 1mM EDTA pH7.4, 2 μg/mL Aprotonin, 1 μg/mL Leupeptin, 1mM PMSF). Cells were then sonicated and incubated for 20 min at 4°C while rotating to ensure sufficient lysis. The lysates were then clarified at 14,000RPM at 4°C for 20 mins.

### RNA Preparation and RT-PCR Analysis

Total RNA was extracted using the RNeasy kit (QIAGEN) and cDNA was generated by reverse transcription PCR using the TaqMan Reverse Transcription kit (Applied Biosystems). Relative mRNA expression levels were determined by the SYBR Green I detection chemistry (Applied Biosystems) on the ABI Prism 7500 Fast Real–Time PCR System. The levels of *RPLPO* or *GAPDH* were used as normalizers. The error bars indicate standard deviation of triplicate qPCR data.

### Baculovirus Production and Purification PRC2, PALI1, and LCOR Recombinant Proteins

Recombinant baculoviruses for Flag/His-PALI1 and Flag/His-LCOR (N-terminal tag) were generated in *Spodoptera frugiperda,* Sf9 cells. Briefly, a baculovirus transfer vector (pVL1392) containing either PALI1 or LCOR coding ORFs and a linearized baculovirus DNA construct (Allele Biotech #ABP-BVD-10001) were co-transfected into Sf9 cells. The supernatants were collected from transfected cells 10 days after transfection (passage 1 [P1]) and passaged to P3, which were used for subsequent recombinant protein purifications. Flag/His-PALI1 and Flag/His-LCOR were then expressed in *Trichoplusia ni*, High Five, cells. The cells were incubated at 28°C and harvested 44-48hr post infection, washed twice in PBS and lysed in BPL2B buffer (20mM Tris, 500mM NaCl, 20% Glycerol, 4mM MgCl2, 3mM β-mercaptoethanol, 0.05% NP40, 2 μg/mL Aprotonin, 1 μg/mL Leupeptin, 1mM PMSF). Lysates were sonicated and cleared by centrifugation for 30min at 20,000 g. The supernatant was loaded onto Ni-NTA His resin (Novagen #70666) and incubated for 3hrs rotating at 4°C. Ni-NTA His resin was then washed a number of times in BPL2B buffer to remove non-specific interactions. The His-tagged proteins were subsequently eluted in Elution buffer (350mM Imidazole, 50mM Tris, 5mM MgCl2, 150mM NaCl, 0.05% NP40). Eluted fractions were then snap frozen in liquid nitrogen.

Baculovirus expression plasmids for Flag-His-tagged Ezh2, His-tagged Eed, His-tagged Suz12, untagged RbAp48, Flag-His-tagged Jarid2 (amino acid 1-530) and Flag-His-tagged Aebp2 were used as previously described (Margueron et al., 2008; Sanulli et al., 2015; Son et al., 2013). Recombinant proteins were produced in Sf9 cells grown in Insect-XPRESS medium (Lonza). After 72h of infection, Sf9 cells were resuspended in BC300 (50 mM Tris-HCl pH 7.9, 300 mM KCl, 2 mM EDTA, 10% glycerol, supplemented with 0.5 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 1 μM pepstatin A, 10 μg/mL aprotinin and 1 μg/mL leupeptin) and lysed by sonication using a Branson Sonifier at 10% amplitude for 10 cycles of 10 s each. The lysate was incubated overnight with Flag-M2 agarose beads (Sigma). After washing with BC300, proteins were eluted with 0.2 mg/mL Flag peptide in BC300 (Jarid2, Aebp2) or BC150 (PRC2, 150 mM KCl instead of 300 mM). PRC2 was further purified by anion-exchange chromatography on Mono-Q resin in 50 mM Tris-HCl pH 7.9, 2 mM EDTA, 1 mM phenylmethylsulfonyl fluoride with a gradient of KCl increasing from 100 mM to 600 mM.

### Histone Methyltransferase Assay

Preparation of substrates: *Xenopus* histones were expressed in *E. coli*, purified from inclusion bodies, solubilized in 7 M guanidine hydrochloride or 7 M urea, 20mM Tris-HCl pH 8, 10mM DTT, combined in equimolar ratios and dialyzed against refolding buffer (10mM Tris-HCl pH 7.5, 1mM EDTA, 2 M NaCl, 5mM β-mercaptoethanol). Octamers were isolated from the dialyzed mixture by size-exclusion chromatography using a Superdex 200 column. Before use in HMT assays, octamers were dialyzed through a gradient to 10mM Tris-HCl pH 7.9, 1mM EDTA, 0.4M NaCl, 5mM β-mercaptoethanol and then by single-step dialysis to 10mM Tris-HCl pH 7.9, 1mM EDTA, 1mM β-mercaptoethanol. Oligonucleosomes were prepared by combining plasmid DNA containing 50 repeats of a 177-bp nucleosome positioning sequence with octamers in equal microgram amounts in refolding buffer, followed by gradient dialysis to 10mM Tris-HCl pH 7.9, 1mM EDTA, 0.4 M NaCl, 1mM β-mercaptoethanol and then by single-step dialysis to 10mM Tris-HCl pH 7.9, 1mM EDTA, 1mM β-mercaptoethanol. Enzyme, protein cofactors and substrate (2μg oligonucleosomes/reaction) were combined in HMT buffer (50mM Tris-HCl pH 8.5, 2.5mM MgCl2, 4mM DTT) containing 30kBq 3H-S-adenosyl-methionine per 25μL reaction. Reactions were incubated 30 min at 30°C, stopped by addition of sample buffer, separated by 12.5% SDS-PAGE, and transferred to PVDF membranes. Equal loading of substrates was confirmed by staining the membrane with Coomassie Brilliant Blue and incorporation of radiolabelled methyl groups is visualized by radiography.

### FLAG IPs

Cells were resuspended in 650μL of High Salt buffer (50mM Tris-HCl, pH 7.2, 300mM NaCl, 0.5% (v/v) NP-40, 1mM EDTA pH7.4, 2 μg/mL Aprotonin, 1 μg/mL Leupeptin, 1mM PMSF) and sonicated once for 10 s. The lysates rotated at 4°C for 20 min before diluting the lysates with 650μL of No Salt buffer (50mM Tris-HCl, pH 7.2, 0.5% (v/v) NP-40, 1mM EDTA pH7.4, 2 μg/mL Aprotonin, 1 μg/mL Leupeptin, 1mM PMSF). The lysates were then clarified at 20,817 g at 4°C. The lysates were pre-cleared with equilibrated Mouse IgG coupled beads (Sigma) for 1 hr. The lysates were then incubated with 20μL of FLAG beads (Sigma) overnight in the presence of 250U/mL Benzonase nuclease. Beads were washed 5 times with Wash buffer (1:1 dilution of High Salt: No Salt buffer). FLAG tagged complexes were eluted by competitive elution using 250μg/mL 3xFLAG peptide (Sigma) while shaking at 25°C.

### Evolutionary Analysis

A maximum likelihood phylogenetic tree was inferred using RaxML (Stamatakis, 2014) from a Clustal Omega (Sievers et al., 2011) alignment of LCOR and LCORL exonic ortholog sequences identified using TBLASTN (Camacho et al., 2009) and HMMER (Finn et al., 2015) in representative genomes (Yates et al., 2016).

### Endogenous IPs

Embryonic stem cells were resuspended in Buffer C (20mM HEPES pH 7.9, 0.2mM EDTA, 1.5mM $MgCl_2$, 20% glycerol, 420mM NaCl, 2 μg/mL Aprotonin, 1 μg/mL Leupeptin, 1mM PMSF), sonicated 3x 15 s and dounced 20 times with a tight pestle. Lysates were incubated for 20 min rotating at 4°C and clarified by centrifugation at 20,817 g at 4°C for 20 min. Lysates were dialysed for 5 hr at 4°C against 50 volumes of Buffer C100 (20mM HEPES pH 7.9, 0.2mM EDTA, 1.5mM MgCl2, 20% glycerol, 125mM KCl). Lysates were again clarified by centrifugation at 20,817 g at 4°C for 20 min. 5 μg antibody was coupled to 20 μL packed Protein A beads (Sigma) by incubation in 1mL PBS (0.1% Tween-20) at 4°C rotating overnight. Beads were collected by centrifugation at 5,440 g at room temperature and washed twice in 1mL 0.2M Sodium Borate pH 9.0. Antibodies were then crosslinked to beads by incubation in 1mL 0.2M Sodium Borate pH 9.0 (containing 20mM dimethyl pimelimidate dihydrochloride) at room temperature rotating for 30 min. Reaction was quenched by washing beads once in 1mL 0.2M Ethanolamine pH 8.0 and incubating for 2 hr at room temperature rotating in 1mL 0.2M Ethanolamine pH 8.0. Beads were washed once in Buffer C100 and blocked for 60 min at 4°C rotating in Buffer C100 (0.1mg/mL Insulin (Sigma), 0.2mg/mL Chicken egg albumin (Sigma), 0.1% (v/v) fish skin gelatin (Sigma). Antibody-crosslinked beads were incubated with protein lysates, in the presence of 250U/mL Benzonase nuclease, at 4°C rotating for 3 hr and washed 5 times in Buffer C100 (+0.02% NP-40). After the final wash beads were resuspended in 100μL of SDS-PAGE sample buffer. IP material was eluted by boiling for 5 min with shaking before centrifuging the beads at 20,817 g for 5 min and keeping the resulting supernatant.

### Cellular Fractionations

Embryonic stem cells were lysed in 400μL pre-extraction buffer (20mM HEPES pH 7.2, 0.5% Triton X-100, 50mM NaCl, 3mM MgCl2, 300mM Sucrose, 2 μg/mL Aprotinin, 1 μg/mL Leupeptin, 1mM PMSF). Incubate at 4°C for 30 min. 200μL of suspension was removed and labeled "Total extract," 200 μL of 2X SDS-PAGE sample buffer was added to the "Total extract." Clarify the remaining lysate suspension at 20,817 g in a 4°C centrifuge for 10 min. Supernatant was kept and labeled "Soluble." 200 μL of 2X SDS-PAGE sample buffer was added to this "Soluble" fraction. The insoluble (chromatin) pellet was washed once in 1mL of pre-extraction buffer before resuspension in 200 μL of pre-extraction buffer and 200μL of 2X SDS-PAGE sample buffer. All samples were boiled at 99°C for 5 min before sonicating 3 times for 10 s at 60% amplitude.

### Chromatin IPs and Quantitative ChIP-Rx

Embryonic stem cells were washed once with PBS before crosslinking for 10 min with PBS containing 1% formaldehyde (Sigma). Crosslinking was quenched with 0.125M Glycine for 5 min before two PBS washes. The crosslinked cells were lysed in 6mL of SDS-Lysis buffer (100mM NaCl, 50mM Tris pH8.1, 5mM EDTA pH 8.0, 0.02% $NaN_3$, 0.5% SDS, 2 μg/mL Aprotonin, 1 μg/mL Leupeptin, 1mM PMSF). Chromatin was pelleted by centrifugation at 1200RPM for 5 min at room temperature. The supernatant was then discarded and the chromatin was resuspended in 3mL of ChIP buffer (2:1 dilution of SDS-Lysis buffer: Triton dilution buffer [100mM Tris pH 8.6, 100mM NaCl, 5mM EDTA pH 8.0, 0.02% $NaN_3$, 5% Triton X-100, 2 μg/mL Aprotonin, 1 μg/mL Leupeptin, 1mM PMSF]). Chromatin was sheared to approximately 100bp-1000bp fragments by successive 30 s rounds of sonication at 5%–8% amplitude. Sonicated chromatin was pre-cleared for 30 min using equilibrated protein A or G beads (Sigma) that had been blocked in TE (10mM Tris pH8.1, 1mM EDTA pH 8.0) containing 0.5mg/mL BSA and 0.2mg/mL Herring Sperm DNA. 10-100μg (DNA) of chromatin was incubated overnight with antibody while rotating at 4°C. Following clarification of insoluble precipitates the chromatin was incubated for 3 hr with 50μL of blocked protein A or G beads. After incubation, the beads were washed three times in Mixed Micelle Buffer (150mM NaCl, 20mM Tris pH 8.1, 5mM EDTA pH 8.0, 5.2% Sucrose, 0.02% $NaN_3$, 1% Triton X-100, 0.2% SDS), twice with Buffer 500 (0.1% Sodium Deoxycholate, 1mM EDTA pH 8.0, 50mM HEPES pH7.5, 1% Triton X-100, 0.02% $NaN_3$), twice with LiC detergent wash (0.5% Sodium Deoxycholate, 1mM EDTA pH 8.0, 250mM LiCl, 0.5% NP-40, 10mM Tris pH 8.0, 0.02% $NaN_3$) and finally one wash with TE. IP material was eluted from the beads with Elution buffer (0.1M $NaHCO_3$, 1% SDS) while shaking for 1 hr at 65°C. The supernatant of the elution was retained and incubated overnight at 65°C while shaking to reverse the crosslinks. The eluted complexes were then subject to RNase (Thermo Fisher) and Proteinase K (Sigma) treatment prior to DNA clean up through Phenol Chloroform clean up and Ethanol precipitation. ChIP enrichment was analyzed by qPCR using the SYBR Green I detection chemistry (Applied Biosystems) on the ABI Prism 7500 Fast Real–Time PCR System. We also performed quantitative chromatin IP relative to a reference exogenous genome (ChIP-Rx) coupled with massively parallel DNA sequencing for the genome-wide mapping of histone modifications, as described previously (Orlando et al., 2014). For ChIP-Rx 250 μg of chromatin (protein) was used per ChIP with a 1.67% spike in of *Drosophila* S2 chromatin and 5μg of antibody was used per ChIP. For ChIP-Rx we used the NEBNext Ultra DNA library preparation kit (NEB) for library prep and the Kapa library quantification for Illumina (Roche #KK4835). Sequencing was performed with Illumina HiSeq v4 chemistry on a HiSeq 2500. Sequencing reads were aligned to the mouse reference genome (mm10) using Bowtie2 v2.1.0 (Langmead and Salzberg, 2012). Only unique alignments were retained for downstream analyses. Sequencing reads were also aligned to the *Drosophila* genome (dm6) and normalization factors calculated (Orlando et al., 2014). Ambiguous reads, i.e., reads that aligned to both mouse and *Drosophila* genomes were removed from all downstream analyses. Bigwig files were generated at a resolution of 10bp using the bamCoverage utility from the deepTools suite (Ramírez et al., 2016) and data were subsequently visualized as ChIP-Rx normalized tracks using the UCSC genome browser. ChIP-Rx normalized average profiles were constructed for transcription start site regions (TSS ± 20kb) using ngsplot (Shen et al., 2014). Knockout and gene trap profiles were compared to their respective wild-type conditions using a one-tailed paired Wilcoxon test. To determine the set of genes affected by both Pali1 and Aebp2, we took the top three thousand H3K27me3 positive genes and computed the intersection of those genes that lose H3K27me3 in the Pali1/2 double knockout and gain H3K27me3 in the Aebp2 gene trap. Data for H3K27me3 was taken from GSE81081 and processed as previously described (Streubel et al., 2017). In brief, reads were aligned to the mouse reference genome (mm10) using Bowtie2 v2.1.0 (Langmead and Salzberg, 2012) and peaks determined using macs2 v2.1.1 (Feng et al., 2012). H3K27me3 positive genes were defined as the set of genes with an H3K27me3 peak (FDR = 0.05) overlapping the promoter regions (TSS ± 2kb) of mm10 RefSeq genes. Genes were ranked based on the proportion of the promoter region covered by an H3K27me3 peak.

### RNA-Seq and Bioinformatics Analysis

For RNA seq the libraries were prepared using the Truseq Stranded Total RNA library prep kit Ribo-Zero Human/ Mouse/ Rat (Illumina) according to manufacturer's instructions. Sequencing was performed on an Illumina HiSeq 2500 sequencer using HiSeq v4 chemistry following the manufacturer's protocols. Raw sequencing reads were aligned to the mouse reference genome (mm10) using kallisto v0.43.1 (Bray et al., 2016). Sequence reads were aggregated into a count for each gene using tximport (Soneson et al., 2015). Differentially expressed genes were identified using DESeq2 (Love et al., 2014). Sequence tracks were generated by aligning reads to the mm10 reference using HISAT v2.0.5 (Pertea et al., 2016). Resulting bedGraph files were converted to bigwig format and scaled using hits per billion with the bamTobw.sh utility (Zhu et al., 2013) for visualization on the UCSC Genome Browser. H3K27me3 positive genes were determined in the same manner as for the ChIP-seq analysis above.

To visualize the relationship between changes in H3K27me3, in ESCs, and gene expression, in EB at day 8, for the Aebp2 Gt, Pali1 sKO and Pali1/2 dKO cells, the log2 fold change of H3K27me3 was computed for the promoter regions (TSS ± 2.5kb) of the top 3,000 H3K27me3 positive genes and plotted against the log2 fold change in gene expression. The percentage of genes in each quadrant was subsequently calculated.

### ESC and Mouse Targeting Strategy

Targeting vectors to knock out Pali1 (Gm340) were introduced into R1 or M1 ESCs. The morula aggregation method was used to generate chimeras. Conditional mutants were crossed with Rosa26::CreERT2 (ERT2-Cre) transgenic mice (Artemis Pharmaceuticals, Germany). Skeletal analyses were performed on E11.5 mice and cleared skeletons were analyzed under a stereomicroscope as described previously (Akasaka et al., 1996). All animal experiments were performed according to the in-house guidelines for the care and use of laboratory animals of the RIKEN Center for Integrative Medical Sciences, Yokohama, Japan [Approval number: Kei-27-001(7)]. All gene recombination experiments were carried out according to the in-house guidelines for the genetic recombination experiments of the RIKEN Center for Integrative Medical Sciences (Approval number: *Kei-29-011*).

### Generation of Pali1/2 Double Knockout Cells

Pali1 and Pali2 double knockout mouse ESCs were generated by transfecting Pali1 knockout ESCs with pSpCas9 (BB)-2A-eGFP vector (Addgene, #48138) containing a guide RNA targeting the 5′ end of the Pali2 specific exon, using Lipofectamine 2000. The same vector without a gRNA sequence was used as the negative control. 48 hr after transfection the GFP high population of cells was collected by FACs and individual cells were seeded to each well of a 96 well plate. Individual clones were expanded and genotyped by amplifying the region surround the PAM site and Sanger sequencing to identify indels.

## DATA AND SOFTWARE AVAILABILITY

The accession number for the RNA-seq data reported in this paper is GEO: GSE100679.

The accession number for the ChIP-Rx data reported in this paper is GEO: GSE108302.

The accession number for the publically available H3K27me3 data reported in this paper is GEO: GSE81081.

Original images have been deposited to Mendeley Data and are available at: https://doi.org/10.17632/mhynrhktv4.1