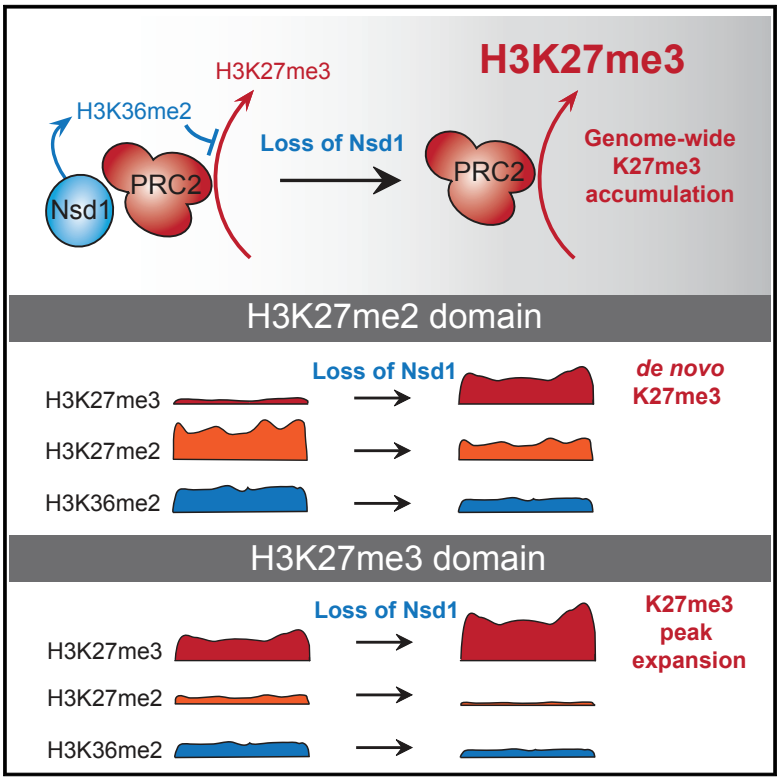
Short Article

Molecular Cell

The H3K36me2 Methyltransferase Nsd1 Demarcates PRC2-Mediated H3K27me2 and H3K27me3 Domains in Embryonic Stem Cells

Graphical Abstract



Authors

Gundula Streubel, Ariane Watson, Sri Ganesh Jammula, ..., Nevan J. Krogan, Adrian P. Bracken, Gerard Cagney

Correspondence

adrian.bracken@tcd.ie (A.P.B.), gerard.cagney@ucd.ie (G.C.)

In Brief

The Polycomb repressor complex 2 (PRC2) deposits H3K27me2 and H3K27me3 repressive histone modifications in spatially defined chromatin domains to maintain cellular identity. Streubel et al. identify the H3K36me2 methyltransferase Nsd1 as a key modulator of PRC2 to restrict H3K27me3 deposition and, thereby, to demarcate H3K27me3 from H3K27me2 domains in ESCs.

Highlights

- Endogenous interactome analysis identifies Nsd1 as a PRC2-associated protein in ESCs
- PRC2-mediated H3K27me2 and Nsd1-mediated H3K36me2 co-localize genome-wide in ESCs
- Nsd1 loss leads to a reduction of H3K36me2 and a gain of H3K27me3 across the ESC genome
- Nsd1 is required in ESCs to demarcate H3K27me3 and H3K27me2 chromatin domains

Streubel et al., 2018, Molecular Cell *70*, 371–379
April 19, 2018
https://doi.org/10.1016/j.molcel.2018.02.027

CellPress

Molecular Cell
Short Article

The H3K36me2 Methyltransferase Nsd1 Demarcates PRC2-Mediated H3K27me2 and H3K27me3 Domains in Embryonic Stem Cells

Gundula Streubel,[1,2,8] Ariane Watson,[2,8] Sri Ganesh Jammula,[3] Andrea Scelfo,[3] Darren J. Fitzpatrick,[1] Giorgio Oliviero,[2] Rachel McCole,[1] Eric Conway,[1] Eleanor Glancy,[1] Gian Luca Negri,[2] Eugene Dillon,[2] Kieran Wynne,[2] Diego Pasini,[3,4] Nevan J. Krogan,[5,6,7] Adrian P. Bracken,[1,9,10,*] and Gerard Cagney[2,9,*]

[1]Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland
[2]School of Biomolecular and Biomedical Science, Conway Institute, University College Dublin, Dublin 4, Ireland
[3]Department of Experimental Oncology, European Institute of Oncology, Via Adamello 16, 20139 Milan, Italy
[4]Department of Health Sciences, University of Milan, Via A. di Rudinì, 8, 20142 Milan, Italy
[5]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA
[6]Quantitative Biosciences Institute (QBI), University of California, San Francisco, San Francisco, CA 94148, USA
[7]Gladstone Institutes, San Francisco, CA 94158, USA
[8]These authors contributed equally
[9]These authors contributed equally
[10]Lead Contact
*Correspondence: adrian.bracken@tcd.ie (A.P.B.), gerard.cagney@ucd.ie (G.C.)
https://doi.org/10.1016/j.molcel.2018.02.027

SUMMARY

The Polycomb repressor complex 2 (PRC2) is composed of the core subunits Ezh1/2, Suz12, and Eed, and it mediates all di- and tri-methylation of histone H3 at lysine 27 in higher eukaryotes. However, little is known about how the catalytic activity of PRC2 is regulated to demarcate H3K27me2 and H3K27me3 domains across the genome. To address this, we mapped the endogenous interactomes of Ezh2 and Suz12 in embryonic stem cells (ESCs), and we combined this with a functional screen for H3K27 methylation marks. We found that Nsd1-mediated H3K36me2 co-locates with H3K27me2, and its loss leads to genome-wide expansion of H3K27me3. These increases in H3K27me3 occurred at PRC2/PRC1 target genes and as *de novo* accumulation within what were previously broad H3K27me2 domains. Our data support a model in which Nsd1 is a key modulator of PRC2 function required for regulating the demarcation of genome-wide H3K27me2 and H3K27me3 domains in ESCs.

INTRODUCTION

Polycomb group proteins are a family of evolutionarily conserved chromatin repressors with essential roles in maintaining cellular identity in stem, progenitor, and differentiated cells (Holoch and Margueron, 2017). Biochemically, Polycomb proteins form two main multiprotein Polycomb repressive complexes, PRC1 and PRC2 (Blackledge et al., 2015). The PRC2 complex consists of three core components, Eed, Suz12, and one of the two histone H3K27 methyltransferases Ezh1 or Ezh2, and it is responsible for all H3K27me2 and H3K27me3 (Ferrari et al., 2014; Margueron and Reinberg, 2011). PRC2 associates with several sub-stoichiometric components, including Polycomb-like proteins PCL1-3, JARID2, AEBP2, and EPOP, which are thought to confer distinct functional effects, including modifying the activity and association of the complex with chromatin (Holoch and Margueron, 2017). However, while methylations of H3K27 are understood to play key roles in cellular identity, little is known about how PRC2 is regulated to demarcate genome-wide domains of H3K27me2 and H3K27me3 (Conway et al., 2015).

The H3K27me2 and H3K27me3 modifications are located in different regions in the genome (Ferrari et al., 2014). The initial genome-wide mapping studies of PRC2 components established that they co-locate with H3K27me3 on the repressed promoters of 10%–15% of all genes with key roles in development and cell fate determination (Ringrose, 2007). The presence of the H3K27me3 mark leads to the recruitment of the canonical PRC1 (cPRC1). The genome-wide profiles of H3K27me3, PRC1, and PRC2 change dynamically during cellular differentiation, and they are required to maintain the repression of genes encoding drivers of alternative lineages (Bracken and Helin, 2009). However, most deregulated genes in PRC2-depleted cells are not direct H3K27me3 targets, suggesting that the H3K27me2 post-translational modification (PTM) might have an underappreciated role (Bracken et al., 2006; Ferrari et al., 2014). While all three H3K27 methylated states have been observed in H3K9me3-positive pericentric heterochromatin during early development (Puschendorf et al., 2008), in the euchromatin of embryonic stem cells (ESCs), H3K27me2 localizes at almost all regions except the H3K27me1-positive gene bodies of actively transcribed genes, the H3K27ac-positive regions at active enhancers, and the H3K27me3-positive regions at Polycomb repressed gene promoters (Ferrari et al., 2014; Lee et al., 2015; Wang et al., 2018). The ubiquitous genome-wide profile of H3K27me2

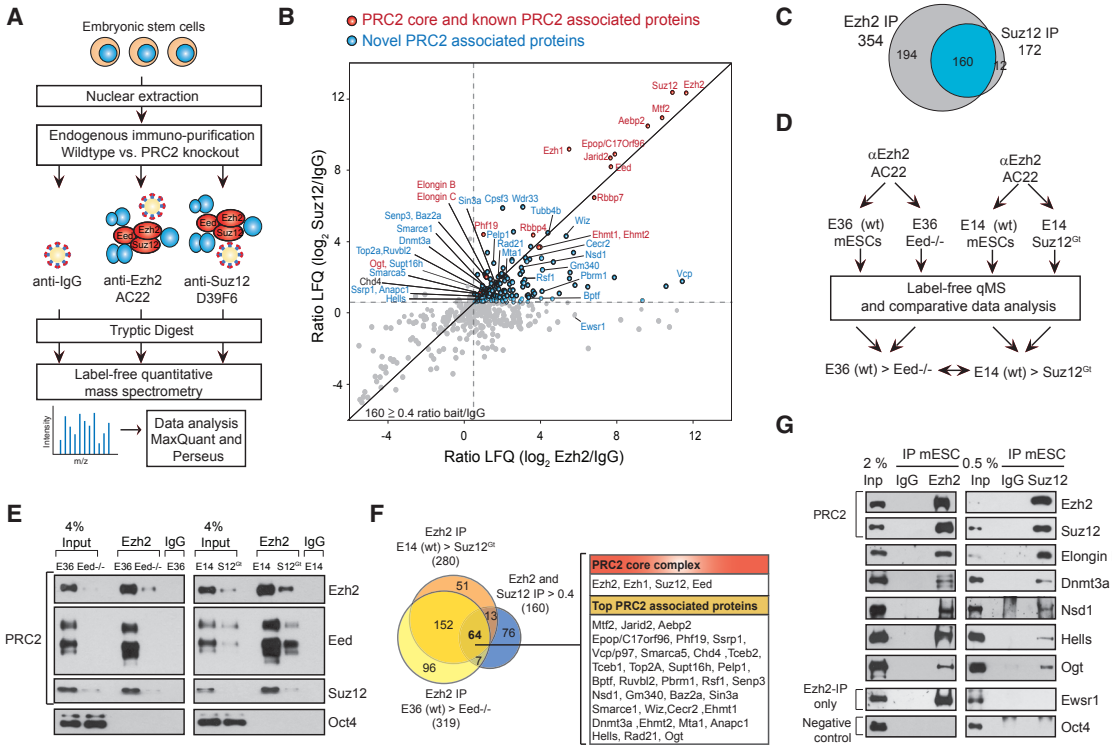


Figure 1. Ezh2- and Suz12-Associated Proteins Dependent on an Intact PRC2 Complex in Embryonic Stem Cells

(A) Experimental outline of the endogenous affinity immunoprecipitations of Ezh2 and Suz12 from mouse ESC nuclear lysates followed by label-free quantitative mass spectrometry. Immunoprecipitations were performed using monoclonal antibodies for Ezh2 (AC22), Suz12 (D39F), and IgG as a negative control. The samples were subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS) in triplicates, and data were analyzed by MaxQuant software.

(B) Scatterplot representation of the enrichment of proteins in Ezh2 and Suz12 purifications as log2 ratios of the label-free quantification (LFQ) intensities over IgG control from three independent replicates.

(C) Venn diagram showing the overlap between proteins enriched in Ezh2 and Suz12 immunoprecipitations (IPs) (log2 LFQ bait/IgG > 0.4).

(D) Experimental outline for label-free quantitative mass spectrometry of endogenous Ezh2 immunoprecipitations in wild-type compared to matched *Eed*$^{-/-}$ and Suz12 genetrap (GT) cells.

(E) Western blot analysis of Ezh2 immunoprecipitations in wild-type, *Eed*$^{-/-}$, and Suz12^{GT} ESCs with the antibodies as indicated. Oct4 is a negative control.

(F) High-confidence PRC2-associated proteins. Venn diagram analysis compares proteins identified in endogenous Ezh2 and Suz12 immunoprecipitations in embryonic day (E)14 ESCs from (B) and those identified in Ezh2 immunoprecipitations in wild-type compared to *Eed*$^{-/-}$ and Suz12^{GT} ESCs.

(G) Western blot analysis of immunoprecipitations of IgG, Ezh2, and Suz12 performed on nuclear lysates of wild-type mouse ESCs.

See also Figure S1 and Table S1.

suggests that its role is to act as a repressive blanket to prevent the misfiring of cell-type-specific enhancers required for alternative lineages (Conway et al., 2015). Consistent with this idea, the loss of PRC2 function in *Drosophila* cells leads to increases in the transcription of intragenic mRNAs at sites previously located within H3K27me2 domains (Lee et al., 2015).

Here we sought to identify factors that control the ability of PRC2 to mediate H3K27me2 and H3K27me3 in ESCs. Our observations describe a functional interplay between Nsd1 and PRC2 that provides important new insights into how the demarcation of H3K27me2 and H3K27me3 domains are regulated.

RESULTS

Identification of Ezh2- and Suz12-Associated Proteins that Depend on an Intact PRC2 Complex in ESCs

To generate a high-quality interaction network for the PRC2 complex in pluripotent cells, we performed mass spectrometry of endogenous immunoprecipitations of independent PRC2 subunits in matched wild-type, *Eed*$^{-/-}$, and Suz12^{Gt} mouse ESCs (Figures 1A and 1D). Initially, we performed mass spectrometry of endogenous immunoprecipitations of Ezh2 and Suz12 in wild-type ESCs (Figure 1B). The precipitated peptides were analyzed for three independent replicates of both the Ezh2 and Suz12 immunoprecipitations, then plotted against each other in order to highlight the proteins that co-purify with both PRC2 subunits (Figure 1B). This revealed that all core PRC2 proteins (Ezh2, Suz12, and Eed), as well as all previously characterized sub-stoichiometric subunits expressed in ESCs (Mtf2/Pcl2, Phf19/Pcl3, Aebp2, Jarid2, Rbbp4/7, and Epop), were immunoprecipitated by both antibodies (Figures 1B and S1A; Table S1). In addition, 148 candidate PRC2 interaction partners were also enriched in both the Ezh2 and Suz12 immunoprecipitations (Figure 1C; Table S1).

To determine which of the candidate proteins are dependent on an intact PRC2 complex for their association with Ezh2 and

Suz12, we performed endogenous Ezh2 immunoprecipitations coupled with mass spectrometry, this time on nuclear lysates of *Eed*$^{-/-}$ (Faust et al., 1998) and *Suz12*^{GT} (Pasini et al., 2004) in parallel with matched wild-type ESCs (Figure 1D). As expected, the lack of Eed led to decreased Suz12 in the Ezh2 immunoprecipitations (Figure 1E; Table S1), while the almost complete lack of Suz12 in the *Suz12*^{GT} cells led to decreased amounts of Eed in the Ezh2 immunoprecipitations (Figure 1E; Table S1). The Suz12^{GT} ESCs have a low level of Suz12 due to low-frequency skipping of the genetrap cassette, as described previously (Pasini et al., 2007). By combining these datasets, we identified a high-confidence set of 64 PRC2-interacting proteins (Figures 1F and S1B; Table S1). Among them were several proteins previously described to be both physically and functionally linked with PRC2, including Elongin B (Tceb2), Elongin C (Tceb1), Ehmt1, and Ogt (Beringer et al., 2016; Chu et al., 2014; Liefke et al., 2016; Mozzetta et al., 2014).

To evaluate our approach, we selected five candidate PRC2-interacting proteins, including Elongin B (Tceb1), Nsd1, Dnmt3a, Hells, and Ogt, and we performed independent co-immunoprecipitations of Ezh2 and Suz12 and confirmed their interaction (Figure 1G). Of note, while the Ewsr1 protein did immunoprecipitate with Ezh2 using the AC22 antibody (Bracken et al., 2006), it did not immunoprecipitate with a Suz12 antibody, demonstrating the advantage of using antibodies for two different PRC2 subunits (Figures 1B and 1G).

A Functional Screen for Modulators of PRC2-Mediated H3K27me2 and H3K27me3

Next, we explored if depletion of any of the PRC2-interacting proteins could modulate the levels or genomic balance of H3K27me2 and H3K27me3. To do this, we designed a high-content endoribonuclease-prepared small interfering RNA (esiRNA) screen of the 38 top-ranked PRC2-interacting proteins (Table S2), including the core subunits Suz12, Ezh2, and Eed as positive controls, and we used immunofluorescence for H3K27me2 and H3K27me3 as functional readouts (Figure 2A). We used the NTERA-2 pluripotent human embryonic carcinoma cell line because of its amenability to esiRNA transfection and immunofluorescence stainings. Efficient knockdowns were achieved for 32 of 38 targets (Figure S2A; Table S2), and immunofluorescence for total histone H3 was performed to normalize the H3K27me2 and H3K27me3 stainings (Figure 2A; Table S3). As expected, knockdown of EZH2, EED, and SUZ12 led to reduced levels of both H3K27me2 and H3K27me3 (Figures 2B–2D; Table S3). The knockdown of CHD4, OGT, and TOP2A (Figures 2B–2D) led to reductions in the global levels of both H3K27 methylated states (Figures 2B–2D; Table S3), consistent with previous reports of functional interplay with PRC2 (Gambetta et al., 2009; Sparmann et al., 2013; Thakurela et al., 2013). Western blot analysis of H3K27me2 and H3K27me3 levels in cells depleted of TOP2A and OGT validated these observations (Figure S2B).

Interestingly, knockdown of the H3K36me2 histone methyltransferase NSD1 had the strongest positive effect on H3K27me3 levels (Figures 2B, 2D, and S2B). Importantly, independent western blot analysis of cells depleted of NSD1 validated these observations (Figure S2B), while endogenous Ezh2 immunoprecipitations on nuclear lysates of *Eed*$^{-/-}$ in parallel with matched wild-type ESCs suggested that Nsd1 requires an intact PRC2 complex for its interaction with Ezh2 (Figure S2C). Finally, an immunoprecipitation of Nsd1 in mouse ESCs confirmed that it pulls down both Ezh2 and Suz12 proteins (Figure S2D). Importantly, while more than 40 peptides specific for mouse Nsd1 were identified in the Ezh2 and Suz12 mass spectrometry experiments (Figure S2E), the related Nsd2 or Nsd3 proteins were not detected (Table S1), consistent with their lower mRNA expression levels compared to *Nsd1* in mouse ESCs (Figure S2F). Moreover, we did not detect any other H3K36me2 methyltransferases in the Ezh2 and Suz12 immunoprecipitations in ESCs (Table S1).

Loss of Nsd1-Mediated H3K36me2 Leads to *De Novo* Accumulation of H3K27me3 Genome-wide

To characterize the requirement of Nsd1 for H3K27me3 in mouse ESCs, we depleted it using two independent small hairpin RNAs (shRNAs), and we observed global increases in H3K27me3 levels in both serum/leukemia inhibitory factor (LIF) (Figures 3A and S3A) and 2i/serum/LIF (Figures S3A and S3B) containing growth medium. As expected, the levels of H3K36me2 were reduced, while the bulk levels of H3K27me2 did not appear to change (Figures 3A and S3B). Of note, while the knockdown of the Setd2 methyltransferase led to a global loss of H3K36me3, it did not affect H3K27me3 levels (Figure S3B). Moreover, the changes in bulk H3K27me3 levels were not a consequence of phenotype changes in the ESCs, since the proliferation rate and expression of pluripotency markers were unaffected in the Nsd1-depleted ESCs (Figures S3C–S3E). Taken together, these results suggest that reductions in the global levels of H3K36me2, but not H3K36me3, can specifically lead to global increases in the levels of PRC2-mediated H3K27me3.

We next performed quantitative chromatin immunoprecipitation relative to a reference exogenous genome (ChIP-Rx) of H3K27me3 and H3K36me2, and we plotted the changes at intervals (or bins) of 50 kb between control and Nsd1-depleted ESCs (Figure 3B). This confirmed that the majority of genomic regions that lose H3K36me2 concomitantly gain H3K27me3 (Figure 3B; 99.1% for shNsd1.1 and 97.0% for shNsd1.2). As a control for the H3K27me3 antibody, we confirmed that it is capable of recognizing its epitope even in the presence of H3K36me2 on the same histone H3 tail (Figure S3F). Average plots of H3K36me2 and H3K27me3 across 40-kb genomic intervals up- and downstream of the center of all H3K27me3-positive regions, in both control and Nsd1-depleted ESCs, confirmed that the relative loss of H3K36me2 at these sites directly correlated with the degree of accumulation of H3K27me3 (Figure 3C). Furthermore, a heatmap representing all regions with a gain in H3K27me3 signal (n = 3,523) in the shNsd1.1 sample and the corresponding H3K36me2 signal intensity confirmed the relationship between both histone marks (Figure S3G). Moreover, we also found that weaker H3K27me3 peaks were more susceptible to an increase in H3K27me3 (Figure S3H).

We next quantified the changes of H3K27me3 and H3K36me2 at gene promoters, gene bodies, and intergenic regions, and we represented the distributions as boxplots (Figure 3D). This revealed that the gains of H3K27me3 and corresponding decreases in H3K36me2 occurred throughout the genome and were not restricted to any particular regions. Therefore, while

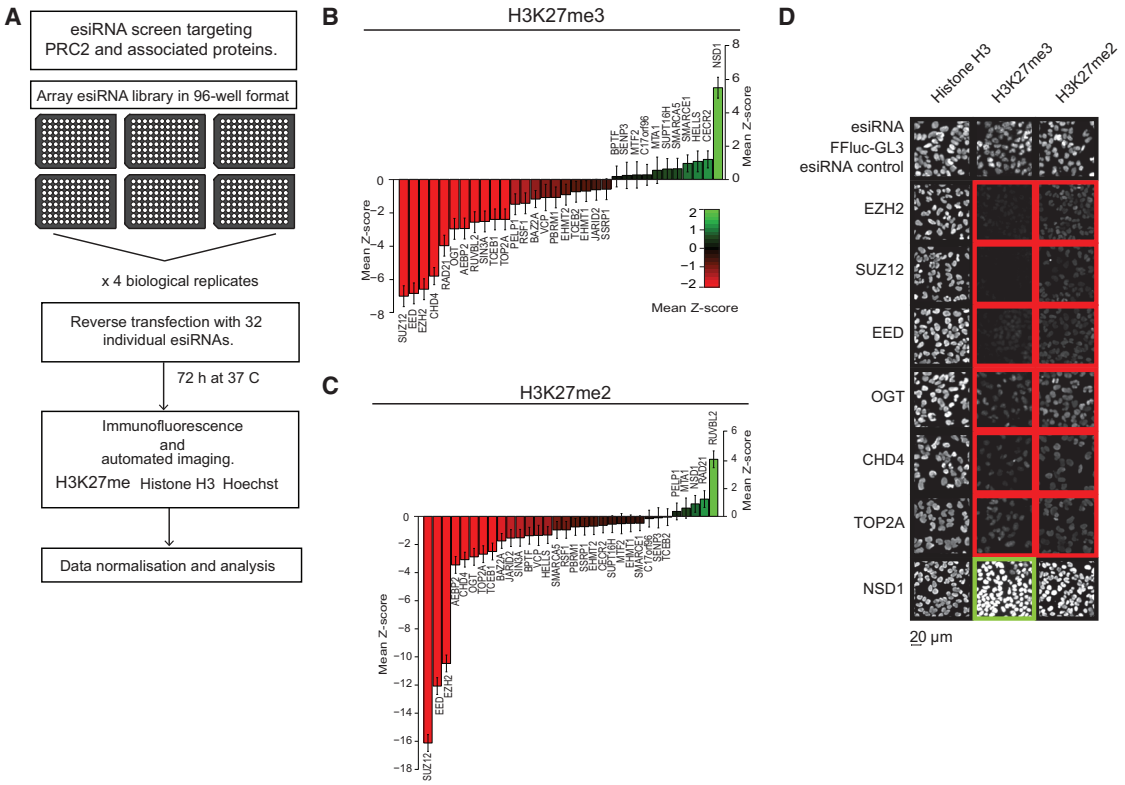


Figure 2. A Functional Screen for Modulators of PRC2-Mediated H3K27me2 and H3K27me3

(A) Experimental outline of functional screen of PRC2-associated proteins using H3K27me2 and H3K27me3 as functional readouts. The pluripotent embryonic carcinoma cell line NTERA-2 was reverse-lipofected with esiRNAs in a 96-well-plate format. Cells were stained for H3K27me2, H3K27me3, total histone H3, and Hoechst 72 hr post-transfection, and immunofluorescence signals were captured with CellInsight. The signals were normalized using histone H3 and Hoechst and then standardized using *Z* score. To adjust for multiple comparison, the p values were Bonferroni corrected.

(B) Graphs showing changes in H3K27me3 levels following esiRNA-mediated knockdowns, as indicated. The *Z* scores for each knockdown condition from four replicates are represented.

(C) Graphs showing changes in H3K27me2 levels following esiRNA-mediated knockdowns, as indicated. The *Z* scores for each knockdown condition from four replicates are represented.

(D) Representative immunofluorescence stainings of H3K27me2, H3K27me3, and histone H3 in cells transfected with control or esiRNAs targeting EZH2, SUZ12, and selected PRC2-associated proteins. The red boxes highlight global decreases in H3K27me2/3 levels, while the green box highlights a global increase in H3K27me3 levels in NSD1-depleted cells.

See also Figure S2 and Tables S2 and S3.

H3K27me3 is relatively high at promoter regions, it further increases in these regions upon depletion of Nsd1. In addition, while H3K27me3 is lower or absent on most gene bodies and intergenic regions, it accumulates in these regions in Nsd1-depleted cells. Consistent with this, we observed that H3K36me2 is detected throughout the genome and becomes depleted at promoters, gene bodies, and intergenic regions in Nsd1-depleted cells, independently of location.

Taken together, these observations led us to speculate that Nsd1-mediated H3K36me2 functions as a modulator of PRC2 activity to restrict the accumulation of H3K27me3 throughout the genome in ESCs.

Nsd1 Collaborates with the PRC2 Complex to Demarcate Regions of H3K27me2 and H3K27me3

To explore the degree of correlation or anti-correlation of the di- and tri-methylated forms of H3K27 and H3K36, we performed a principal component analysis (Figure 4A). This revealed a strong anti-correlation between H3K36me3 and H3K27me3, as reported previously (Ernst and Kellis, 2010). To our surprise, we observed a strong positive correlation between H3K27me2 and H3K36me2, despite previous demonstrations that both H3K36me2 and H3K36me3 peptides block PRC2-mediated H3K27me2/3 (Schmitges et al., 2011; Yuan et al., 2011). Supporting this observation, a positive correlation between H3K27me2 and H3K36me2 (Figure 4B, left panel), a moderate correlation between H3K27me3 and H3K36me2 (Figure 4B, middle panel), and a strong negative correlation between H3K27me3 and H3K36me3 (Figure 4B, right panel) were also observed in correlation plots.

The strong positive correlation between H3K27me2 and H3K36me2 prompted us to more closely monitor the H3K27me3 in the context of H3K27me3 and H3K27me2 domains (Figures 4C and S4A). This revealed that the increases in

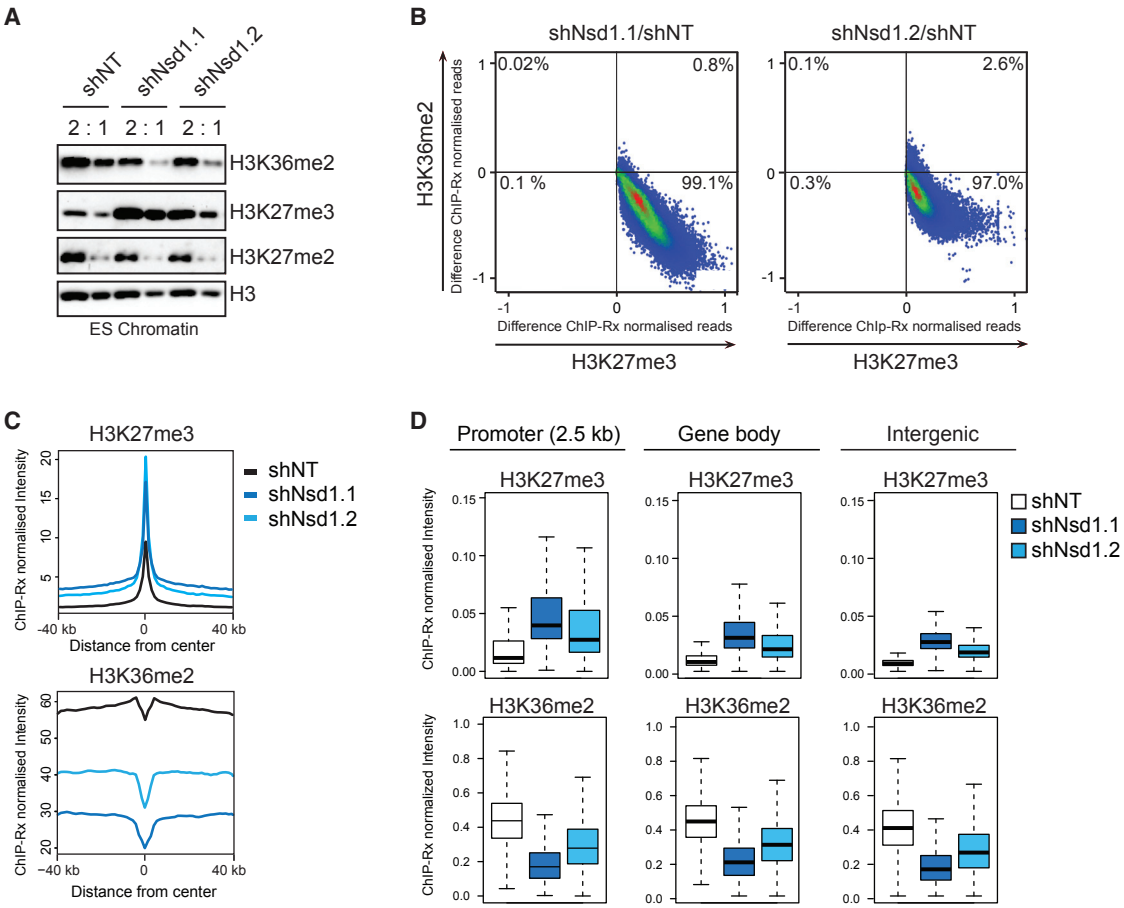


Figure 3. Loss of Nsd1-Mediated H3K36me2 Leads to Genome-wide Accumulation of H3K27me3

(A) Western blot analysis of chromatin lysates of ESCs transduced with control or two independent shRNAs targeting Nsd1 (shNsd1.1 and shNsd1.2).

(B) Scatterplot showing the relationship between the genome-wide changes in H3K27me3 and H3K36me2. The genome was segmented into small bins at a resolution of 50,000 bp, and the difference in the number of ChIP-Rx-normalized reads of H3K36me2 or H3K27me3 between control and Nsd1-depleted cells was plotted. The percentage of bins found in each quadrant is shown.

(C) Average ChIP-Rx signal profiles of H3K27me3 and H3K36me2 showing a depletion of H3K36me2 in regions enriched (false discovery rate [FDR] = 0.1) for H3K27me3 in wild-type and Nsd1-depleted cells. The enriched regions were extended ±40 kb from their midpoint.

(D) Distribution of H3K27me3 and H3K36me2 ChIP-Rx-normalized reads within gene promoters (±2.5 kb of transcription start site [TSS]), gene bodies, and intergenic regions in wild-type and Nsd1-depleted embryonic stem cells.

See also Figure S3.

H3K27me3 levels upon Nsd1 depletion occurred within H3K27me3 domains, but also *de novo* within H3K27me2 domains (Figures 4C and S4A). This *de novo* H3K27me3 accumulation within intergenic H3K27me2 domains correlated with reduced H3K27me2 (Figures 4D, 4E, and S4C), but without gene expression changes (Figure S4B). Interestingly, these changes correlated with increased Suz12 binding in both domains, consistent with PRC2's ability to bind its own mark (Figures S4D and S4E).

Altogether, our data demonstrate that, in ESCs, Nsd1 is an important regulator of H3K27me3 placement by PRC2, preventing an aberrant accumulation of the mark. We propose a mechanism by which Nsd-mediated H3K36me2 functions to restrict the deposition of H3K27me3 and, in doing so, contributes to the demarcation of PRC2-mediated H3K27me2 and H3K27me3 domains.

DISCUSSION

An outstanding question in Polycomb biology is how PRC2 is regulated to deposit H3K27me2 and H3K27me3 marks across the genomes of higher eukaryotic cells. Our data are consistent with a role for Nsd1-mediated H3K36me2 in modulating PRC2 activity across the genome, such that high H3K36me2 correlates with PRC2-mediated H3K27me2 deposition while low levels of H3K36me2 correlate with increased H3K27me3. More broadly, we propose that H3K36me2 modulates PRC2 activity by preventing random placement of H3K27me3 throughout the genome.

The co-location of H3K27me2 and H3K36me2 PTMs throughout the majority of the ESC genome suggests that PRC2 cooperates with Nsd1 at these locations. Moreover, while H3K36me2 is lower at H3K27me3-enriched regions, it is still

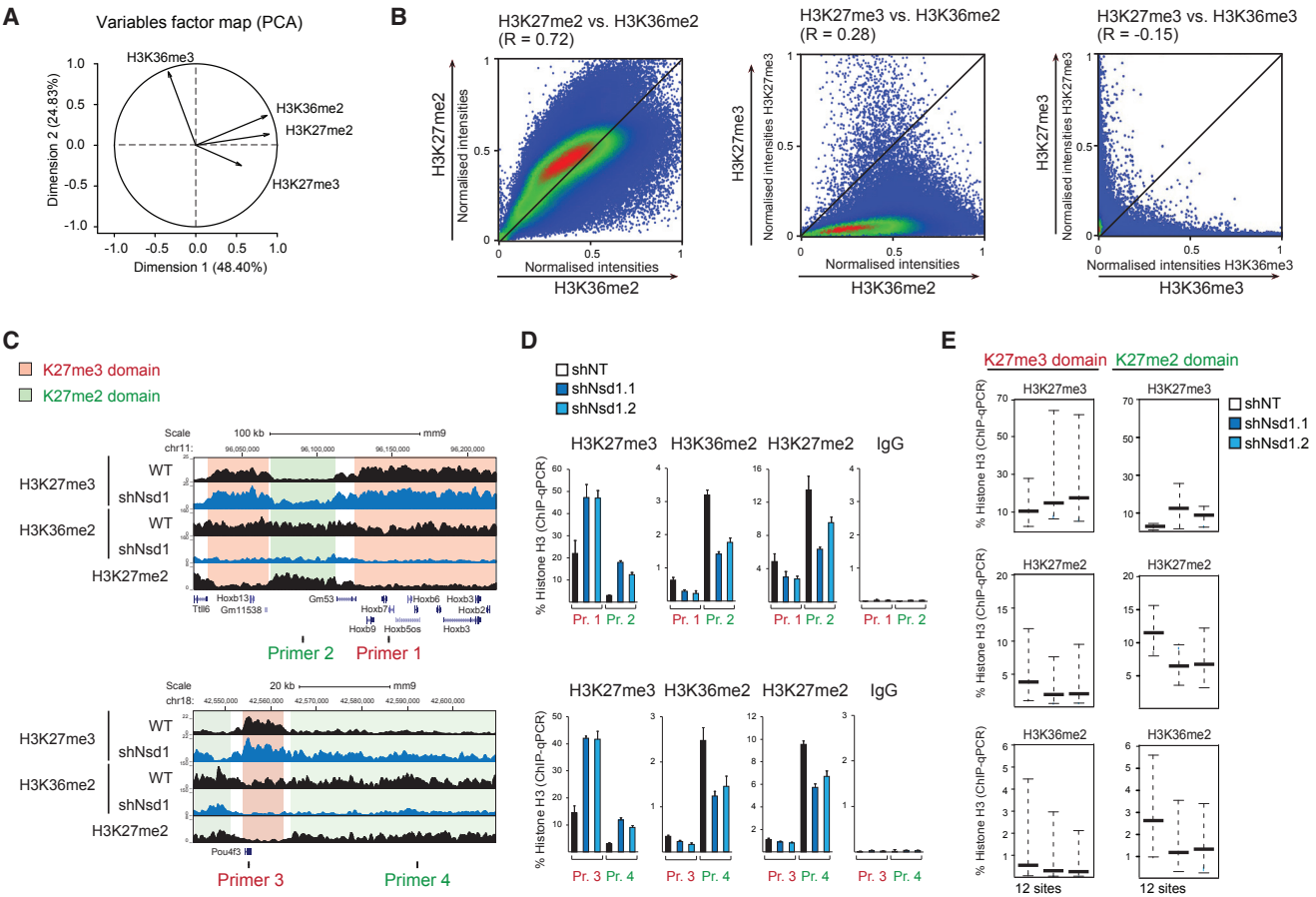


Figure 4. Nsd1-Mediated H3K36me2 Modulates PRC2 Activity to Demarcate Regions of H3K27me2 and H3K27me3

(A) Biplot representing the degree of correlation between the indicated H3K27 and H3K36 methylation states. Correlation coefficients with the first component are K36me2 = 0.8873187, K36me3 = −0.3305159, K27me2 = 0.8602497, and K27me3 = 0.5469931.

(B) Plots showing the genome-wide correlation between the H3K27me2 and H3K36me2 (left), H3K27me3 and H3K36me2 (middle), and H3K27me3 and H3K36me3 (right) modifications. The complete genome was fragmented into small bins of 10-kb resolution, and the total number of reads for each modification within each bin was computed and normalized to sequencing depth. Each modification type was then scaled to 0–1 and the R coefficient was determined with Pearson correlation. The p values for all three plots are <2.2e−16 (based on Pearson's product moment correlation coefficient, followed by t-distribution).

(C) Genome viewer representations of ChIP-Rx-normalized reads for H3K27 and H3K36 methylations at two loci, showing the increases of H3K27me3 in both pre-existing H3K27me2 (green) and H3K27me3 (red) domains in Nsd1-depleted ESCs. H3K27me2 was normalized to reads per genome content (RPGC).

(D) ChIP-qPCR analysis of H3K27me3, H3K36me2, and H3K27me2 changes at two H3K27me3 domains (red) and two H3K27me2 domains (green) using the primers indicated in (C). The means ± SD of technical replicates of a representative experiment are shown.

(E) Boxplot summary of ChIP-qPCR analyses of H3K27me3, H3K36me2, and H3K27me2 in shNT and shNsd1 cells at 12 independent genomic sites (primer pairs) representing H3K27me3 domains and 12 independent genomic sites representing H3K27me2 domains. The median (thick line) and the most outliers (thin lines) are indicated.

See also Figure S4.

present, as are low levels of H3K27me2. This co-localization of H3K27me2 and H3K36me2 is difficult to reconcile with the *in vitro* demonstration that synthetic H3K36me2- or H3K36me3-modified peptides inhibit PRC2-mediated di- and tri-methylation of H3K27 (Schmitges et al., 2011; Yuan et al., 2011). However, recent work clarified that, while symmetrically modified nucleosomes by H3K36me3 or H3K4me3 do indeed block PRC2 methyltransferase activity, asymmetrically modified nucleosomes do not (Voigt et al., 2012). Therefore, it is possible that the co-location of H3K36me2 and H3K27me2/3 observed by ChIP-Rx in this study might be due to asymmetric deposition of the PTMs. Nonetheless, in line with our findings, mass spectrometric quantifications of histone PTMs demonstrated that H3K36me2 and H3K27me2 PTMs co-occur on the same histone H3 tails (Jung et al., 2013; Yu et al., 2016). However, on the level of chromatin, we and others have not been able to chromatin-immunoprecipitate PRC2 or NSD1 at H3K27me2/H3K36me2-positive locations (data not shown; Ferrari et al., 2014), suggesting the interactions are transient and not stable. A potential explanation for this is that the ChIP method does not capture proteins that interact with chromatin with a short residence time (Schmiedeberg et al., 2009). Interestingly, the levels of

H3K36me2 accumulate within 2 hr after DNA replication, similar to PRC2-mediated H3K27me2 (Alabert et al., 2015). Therefore, it is possible that PRC2 and NSD1 cooperate to mediate their respective H3K27me2 and H3K36me2 marks by a mechanism that is coupled with DNA replication.

It is likely that other H3K36me2 methyltransferases function to demarcate H3K27me2 and H3K27me3 in other cells types. Supporting this, Lu and colleagues showed that knockdown of NSD2 and NSD3 in mesenchymal cells leads to global increases in H3K27me3 (Lu et al., 2016). In *Drosophila*, the Ash1 Trithorax protein, which mediates H3K36me2, counteracts PRC2-mediated H3K27me3 deposition on Polytene chromosomes (Dorighi and Tamkun, 2013). Furthermore, the mammalian Ash1 ortholog Ash1L counteracts Polycomb-mediated repression by promoting H3K36me2 at gene bodies of Hox genes (Miyazaki et al., 2013). Although our data suggest that Nsd1-mediated H3K36me2 has a general regulatory impact on H3K27me3 genome-wide, it is also possible that the Nsd1 protein itself or an H3K36me2-binding protein, rather than the modification, may inhibit PRC2 activity, e.g., by sterical competition. Finally, in terms of a potential role in directing PRC2 activity to CpG islands, it is interesting to note that H3K36me2 levels are lower at promoters with CpG islands compared to non-CpG island promoters in ESCs (Blackledge et al., 2010).

Our results have implications for our understanding of the potential mechanisms by which the NSD1–3 proteins contribute to development and disease. Consistent with an important role for Nsd1 during mouse development, its loss leads to early lethality and the inability to activate Hox genes (Rayasam et al., 2003). As depletion of Polycomb group proteins causes embryonic developmental defects and a deregulated Hox gene (Bracken and Helin, 2009), one could speculate that the Nsd1 knockout phenotype is related to aberrantly high levels of PRC2-mediated H3K27me3. Our results suggest that the frequent deregulation of the NSD1–3 proteins in cancer and the oncogenic H3K36M mutation (Brien et al., 2016; Lu et al., 2016) are associated with deregulated demarcation of PRC2-mediated H3K27me2 and H3K27me3 domains. It will be important to determine the consequences on altered H3K27me2 and H3K27me3 profiles in these cancers. In the broader context, one could speculate that the loss of function of Nsd proteins and H3K36M mutations would phenocopy EZH2 change-of-function mutations and, as such, could be targeted by PRC2 inhibition strategies (Conway et al., 2015). Another potential implication of the functional interplay between NSD1 and EZH2 described here might provide new insight into the pathogenic defects in Sotus and Weaver overgrowth syndromes. Constitutional *NSD1* loss-of-function and *EZH2* change-of-function mutations cause Sotos and Weaver syndromes, respectively, yet both share considerable phenotypic overlap (Tatton-Brown and Rahman, 2013). Our data suggest that the NSD1 and EZH2 mutations in these syndromes would lead to increased levels and redistribution of H3K27me3 coupled with decreased levels of H3K27me2. These potentially shared epigenomic defects could be explored in order to understand the common phenotypic features of these syndromes.

We propose that H3K36me2 functions to restrict the accumulation of PRC2-mediated H3K27me3 across the genome in ESCs. Our study adds to the growing evidence of interplay between methylations at the H3K27 and H3K36 residues, and it provides new insights into the regulation of PRC2 activities 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
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mammalian cell culture
  - Stable cell lines
- METHOD DETAILS
  - Antibodies
  - Production of esiRNAs and transfections
  - Functional esiRNA screen for H3K27me2 and H3K27me3
  - Western blotting
  - Peptide dot blots
  - Immunoprecipitations and proteomic analysis
  - Mass spectrometry
  - ChIPs, ChIP-Rx and library preparation
  - Next generation sequencing analysis
  - Quantitative real-time PCR
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and four tables and can be found with this article online at https://doi.org/10.1016/j.molcel.2018.02.027.

ACKNOWLEDGMENTS

We thank members of the Bracken and Cagney laboratories for helpful discussions and critical reading of the manuscript. We also thank D. Rossi and L. Cedrone for technical support with ChIP-Rx and K. Hokamp and C. O'Brien for support with bioinformatic analysis. We thank A. Wutz for providing Eed knockout ESCs. We acknowledge the help of the UCD Conway Proteomics Facility. Work in the Bracken Lab is supported by Science Foundation Ireland under the SFI Investigators Programme (SFI/16/IA/4562) and the BBSRC-SFI joint funding award (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). G.S. was partially funded by an EMBO short-term fellowship ASTF 341-2015. E.G. is supported by a research grant from Science Foundation Ireland (SFI) under grant 15/IA/3104. Work in the Cagney laboratory is supported by Science Foundation Ireland (SFI 10/1N.1/B3.19) and an Irish Research Council award to A.W.

AUTHOR CONTRIBUTIONS

G.C. and A.P.B. planned the research. G.S., A.P.B., and G.C. designed the experiments. G.S. performed the immunoprecipitations, most western blots, and ChIP and RT-PCR experiments. A.W. designed and performed the esiRNA screen and analysis. N.J.K. supervised the esiRNA screen and provided materials. S.G.J. and D.J.F. performed genomic bioinformatic analyses. G.S. and A.S. performed ChIP-Rx experiments. R.M., E.C., and E.G. contributed to ChIP analysis and gene expression analysis. R.M. performed the dotblot analyses. G.O., A.W., and K.W. processed mass spectrometry samples and G.O. and E.D. performed proteomic bioinformatic analyses. G.L.N. set up the data

analysis pipeline for mass spectrometry data. D.P. supervised the sequencing of ChIP experiments and provided expertise and reagents for ChIP-Rx. A.P.B. and G.S. generated the figures. G.S., A.P.B., and G.C. co-wrote the manuscript.

DECLARATIONS OF INTERESTS

The authors declare no competing interests.

Received: June 5, 2017
Revised: December 22, 2017
Accepted: February 23, 2018
Published: March 29, 2018

REFERENCES

Alabert, C., Barth, T.K., Reverón-Gómez, N., Sidoli, S., Schmidt, A., Jensen, O.N., Imhof, A., and Groth, A. (2015). Two distinct modes for propagation of histone PTMs across the cell cycle. Genes Dev. *29*, 585–590.

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., Zhou, J.C., Tolstorukov, M.Y., Farcas, A.M., Park, P.J., and Klose, R.J. (2010). CpG islands recruit a histone H3 lysine 36 demethylase. Mol. Cell *38*, 179–190.

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., and Helin, K. (2009). Polycomb group proteins: navigators of lineage pathways led astray in cancer. Nat. Rev. Cancer *9*, 773–784.

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.

Brien, G.L., Valerio, D.G., and Armstrong, S.A. (2016). Exploiting the epigenome to control cancer-promoting gene-expression programs. Cancer Cell *29*, 464–476.

Chu, C.S., Lo, P.W., Yeh, Y.H., Hsu, P.H., Peng, S.H., Teng, Y.C., Kang, M.L., Wong, C.H., and Juan, L.J. (2014). O-GlcNAcylation regulates EZH2 protein stability and function. Proc. Natl. Acad. Sci. USA *111*, 1355–1360.

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.

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.

Dorighi, K.M., and Tamkun, J.W. (2013). The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in Drosophila. Development *140*, 4182–4192.

Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. Nat. Biotechnol. *28*, 817–825.

Faust, C., Lawson, K.A., Schork, N.J., Thiel, B., and Magnuson, T. (1998). The Polycomb-group gene eed is required for normal morphogenetic movements during gastrulation in the mouse embryo. Development *125*, 4495–4506.

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.

Gambetta, M.C., Oktaba, K., and Müller, J. (2009). Essential role of the glycosyltransferase sxc/Ogt in polycomb repression. Science *325*, 93–96.

Heninger, A.K., and Buchholz, F. (2007). Production of endoribonuclease-prepared short interfering RNAs (esiRNAs) for specific and effective gene silencing in mammalian cells. CSH Protoc. *2007*, pdb.prot4824.

Henschel, A., Buchholz, F., and Habermann, B. (2004). DEQOR: a web-based tool for the design and quality control of siRNAs. Nucleic Acids Res. *32*, W113–W120.

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

Jung, H.R., Sidoli, S., Haldbo, S., Sprenger, R.R., Schwämmle, V., Pasini, D., Helin, K., and Jensen, O.N. (2013). Precision mapping of coexisting modifications in histone H3 tails from embryonic stem cells by ETD-MS/MS. Anal. Chem. *85*, 8232–8239.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res. *12*, 996–1006.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. J. Stat. Softw. *25*, 1–18.

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.

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.

Lu, C., Jain, S.U., Hoelper, D., Bechet, D., Molden, R.C., Ran, L., Murphy, D., Venneti, S., Hameed, M., Pawel, B.R., et al. (2016). Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. Science *352*, 844–849.

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

Miyazaki, H., Higashimoto, K., Yada, Y., Endo, T.A., Sharif, J., Komori, T., Matsuda, M., Koseki, Y., Nakayama, M., Soejima, H., et al. (2013). Ash1l methylates Lys36 of histone H3 independently of transcriptional elongation to counteract polycomb silencing. PLoS Genet. *9*, e1003897.

Mozzetta, C., Pontis, J., Fritsch, L., Robin, P., Portoso, M., Proux, C., Margueron, R., and Ait-Si-Ali, S. (2014). The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. Mol. Cell *53*, 277–289.

Oliviero, G., Brien, G.L., Watson, A., Streubel, G., Jerman, E., Andrews, D., Doyle, B., Munawar, N., Wynne, K., Crean, J., et al. (2016). Dynamic protein interactions of the Polycomb Repressive Complex 2 during differentiation of pluripotent cells. Mol. Cell. Proteomics *15*, 3450–3460.

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.

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., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol. Cell. Biol. *27*, 3769–3779.

Puschendorf, M., Terranova, R., Boutsma, E., Mao, X., Isono, K., Brykczynska, U., Kolb, C., Otte, A.P., Koseki, H., Orkin, S.H., et al. (2008). PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. Nat. Genet. *40*, 411–420.

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*, W160–W165.

Rayasam, G.V., Wendling, O., Angrand, P.O., Mark, M., Niederreither, K., Song, L., Lerouge, T., Hager, G.L., Chambon, P., and Losson, R. (2003). NSD1 is essential for early post-implantation development and has a catalytically active SET domain. EMBO J. *22*, 3153–3163.

Ringrose, L. (2007). Polycomb comes of age: genome-wide profiling of target sites. Curr. Opin. Cell Biol. *19*, 290–297.

Schmiedeberg, L., Skene, P., Deaton, A., and Bird, A. (2009). A temporal threshold for formaldehyde crosslinking and fixation. PLoS ONE *4*, e4636.

Schmitges, F.W., Prusty, A.B., Faty, M., Stützer, A., Lingaraju, G.M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. Mol. Cell *42*, 330–341.

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.

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.

Sparmann, A., Xie, Y., Verhoeven, E., Vermeulen, M., Lancini, C., Gargiulo, G., Hulsman, D., Mann, M., Knoblich, J.A., and van Lohuizen, M. (2013). The chromodomain helicase Chd4 is required for Polycomb-mediated inhibition of astroglial differentiation. EMBO J. *32*, 1598–1612.

Tatton-Brown, K., and Rahman, N. (2013). The NSD1 and EZH2 overgrowth genes, similarities and differences. Am. J. Med. Genet. C. Semin. Med. Genet. *163C*, 86–91.

Thakurela, S., Garding, A., Jung, J., Schübeler, D., Burger, L., and Tiwari, V.K. (2013). Gene regulation and priming by topoisomerase IIα in embryonic stem cells. Nat. Commun. *4*, 2478.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740.

Voigt, P., LeRoy, G., Drury, W.J., 3rd, Zee, B.M., Son, J., Beck, D.B., Young, N.L., Garcia, B.A., and Reinberg, D. (2012). Asymmetrically modified nucleosomes. Cell *151*, 181–193.

Wang, L., Joshi, P., Miller, E.L., Higgins, L., Slattery, M., and Simon, J.A. (2018). A role for monomethylation of histone H3-K27 in gene activity in Drosophila. Genetics *208*, 1023–1036.

Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nat. Methods *6*, 359–362.

Yu, Y., Chen, J., Gao, Y., Gao, J., Liao, R., Wang, Y., Oyang, C., Li, E., Zeng, C., Zhou, S., et al. (2016). Quantitative profiling of combinational K27/K36 modifications on histone H3 variants in mouse organs. J. Proteome Res. *15*, 1070–1079.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J. Biol. Chem. *286*, 7983–7989.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. *9*, R137.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-Ezh2-BD43, WB	Pasini et al., 2004, Supernatant	Ezh2 clone BD43
Mouse monoclonal anti-Ezh2-AC22, WB	Pasini et al., 2004, Supernatant	Ezh2 clone AC22; RRID: AB_2056149
Mouse monoclonal anti-Eed-AA19, WB	Bracken et al., 2003, Supernatant	Eed-AA19
Rabbit polyclonal anti-Suz12 D39F6, WB, IP	Cell Signaling	Cat#3737; RRID: AB_2196850
Rabbit polyclonal anti-Oct4, WB	Abcam	Cat#Ab19857; RRID: AB_445175
Rabbit polyclonal anti-ElonginB/Tceb2 FL-118, WB	Santa Cruz	Cat#sc-11447; RRID: AB_641309
Rabbit polyclonal anti-Nsd1, WB	Abbexa	Cat#abx135901
Mouse monoclonal anti-Dnmt3a [64B1446], WB	Abcam	Cat#ab13888; RRID: AB_300714
Rabbit polyclonal anti-Hells, WB	Bethyl Laboratories	Cat#A300-226A; RRID: AB_2263996
Goat polyclonal anti-Ewsr1 C-19, WB	Santa Cruz	Cat#sc-6533; RRID: AB_2262518
Rabbit polyclonal anti-Ogt H-300, WB	Santa Cruz	Cat#sc-32921; RRID: AB_2156938
Rabbit monoclonal anti-H3K27me3 C36B11, WB, ChIP, IF, Custom made	Cell Signaling	Cat #9733; RRID: AB_2616029
Rabbit monoclonal anti-H3K27me2 D18C8, WB, ChIP, IF	Cell signaling	Cat#9728; RRID: AB_1281338
Mouse monoclonal anti-H3K27me1 MABI 0321, WB, IF	Active Motif	Cat#61015; RRID: AB_2715573
Rabbit polyclonal anti-Histone H3, WB, ChIP	Abcam	Cat#ab1791; RRID: AB_302613
Rabbit monoclonal anti-H3K36me2 C75H12, WB	Cell Signaling	Cat#2901; RRID: AB_1030983
Rabbit monoclonal anti-H3K36me3 D5A7, WB, ChIP	Cell Signaling	Cat# 4909S; RRID: AB_1950414
Rabbit polyclonal H3K36me2, ChIP	Abcam	Cat# ab9049; RRID: AB_1280939
Rabbit polyclonal H3K36me2, ChIP-Rx, Dot blot	Abcam	Cat#ab176921
Rabbit polyclonal anti-Top2a H-231, WB	Santa Cruz	Cat#sc-13058; RRID: AB_2240766
Rabbit polyclonal anti-Ogt H-300, WB	Santa Cruz	Cat#sc-32921; RRID: AB_2156938
Rabbit polyclonal anti-β-Tubulin H-235, WB	Santa Cruz	Cat#sc-9104; RRID: AB_2241191
Goat polyclonal HRP-linked secondary anti-rabbit, WB	Sigma-Aldrich	Cat#A0545; RRID: AB_257896
Mouse monoclonal HRP-linked secondary anti-light chain rabbit, WB	Abcam	Cat#ab99697; RRID: AB_10673897
Goat polyclonal HRP-linked secondary anti-mouse, WB	Merck	Cat#401253; RRID: AB_10696557
Mouse monoclonal HRP-linked secondary anti-goat, WB	Santa Cruz	Cat#sc-2020; RRID: AB_631728
Mouse monoclonal Histone H3 MABI 0301, IF	Active Motif	Cat#39763; RRID: AB_2650522
Goat anti-rabbit Alexa Fluor-conjugated secondary antibody	Thermo Fisher	A-11012; RRID: AB_2534079
Goat anti-mouse IgG2b Alexa Fluor-conjugated secondary antibody	Thermo Fisher	A-21242; RRID: AB_2535811
Goat anti-mouse IgG2a Alexa Fluor-conjugated secondary antibody	Thermo Fisher	A-21131l; RRID: AB_2535771
Bacterial and Virus Strains		
STBL3	Thermo Fisher Scientific	*C737303*
Chemicals, Peptides, and Recombinant Proteins		
LATKAARKSAPATGGVKKPHR (unmodified H3 aa20-40)	Peptide 2.0	Sequence_ID: 165451. Batch_No.: 121506-004
LATKAARKSAPATGGVK(me2)KPHR (H3K36me2)	Peptide 2.0	Sequence_ID: 165449. Batch_No.: 121506-002
LATKAARKSAPATGGVKKPHR (H3K27me3-K36me2)	Peptide 2.0	Sequence_ID: 165450. Batch_No.: 121506-003
LATKAARK(me3)SAPATGGVKKPHR (H3K27me3)	Peptide 2.0	Sequence_ID: 165448. Batch_No.: 121506-001

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
H3K27me3 ChIP-Rx	This paper; GEO	GEO: GSE107773
H3K36me2 ChIP-Rx	This paper; GEO	GEO: GSE107773
H3K27me2 ChIP-Seq	Ferrari et al., 2014	GEO: GSM1234538
H3K36me3 ChIP-Seq	Ferrari et al., 2014	GEO: GSM1234540
RNA-Seq WT ES cells	Ferrari et al., 2014	GEO: GSM1249334
RNA-Seq WT embryoid bodies	Ferrari et al., 2014	GEO: GSM1249336
Proteome data	This paper; PRIDE; Tables S1 and S2	PRIDE: PXD005381
Deposited original data	This paper	https://doi.org/10.17632/4pzh3gjkm9.1
Experimental Models: Cell Lines		
Human: HEK293T	ATCC	ATCC CRL-3216
Human: NTERA-2 cl.D1	ATCC	RRID: CVCL_3407; ATCC CRL-1973
Mouse: ES cell line E14	Pasini laboratory	N/A Strain of origin 129P2/Ola
Mouse: EED KO ES cells	Schoeftner et al., 2006	N/A background strain: 129/Ola
Mouse: E36 WT cells	Schoeftner et al., 2006	N/A background strain: 129/Ola
Mouse: Suz12^{GT} ES cells	Pasini et al., 2004	N/A Strain of origin 129P2/Ola
Oligonucleotides		
qPCR Rrimers	This paper; Table S4	N/A
esiRNA sequences	This paper: Table S4	N/A
Recombinant DNA		
pLKO.TRC1.shmNsd1.1, puro	Sigma-Aldrich	TRCN0000123381
pLKO.TRC1.shmNsd1.2, puro	Sigma-Aldrich	TRCN0000123379
pLKO.1-puro Luciferase shRNA Control, puro	Sigma-Aldrich	SHC007
pLKO.TRC1 shmSetd2, puro_1	Ferrari et al., 2014; Sigma-Aldrich	TRCN0000238533
pLKO.TRC1 shmSetd2, puro_4	Ferrari et al., 2014; Sigma-Aldrich	TRCN0000238536
Software and Algorithms		
Perseus software, version 1.4	Tyanova et al., 2016	N/A
MaxQuant software, version 1.3.0.5	Cox and Mann, 2008	N/A
Bowtie version 1.2.1.1	Langmead et al., 2009	http://bowtie-bio.sourceforge.net/index.shtml
macs2 version 2.1.1.2	Zhang et al., 2008	https://github.com/taoliu/MACS
deepTools version 2.5.3	Ramírez et al., 2016	https://deeptools.readthedocs.io/en/latest/
ngsplot version 2.61	Shen et al., 2014	https://github.com/shenlab-sinai/ngsplot

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).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mammalian cell culture

Eed KO ESCs and parental cell lines were provided by Dr. A Wutz (Schoeftner et al., 2006), while E14 ES and the Suz12^{GT} cell line were provided by Dr. D. Pasini (Pasini et al., 2004), (originated from 129P2/Ola, male). Mouse embryonic stem cells were grown on gelatinized culture dishes in GMEM media (Sigma-Aldrich) supplemented with 15% FBS (GIBCO) (v/v), 1000 U/ml leukemia inhibitory factor (EMD Millipore), 100 U/ml penicillin, 100 U/ml streptomycin (GIBCO), 2 mM GlutaMAX (GIBCO), 1:100 MEM non-essential amino acids (GIBCO), 1 mM sodium pyruvate (GIBCO) and 50 μM β-mercaptoethanol (Sigma-Aldrich). For growth in 2i growth conditions, the GSK3 inhibitor CHIRON99021 (Millipore) and MEK inhibitor PD0325901 (Millipore) were added at a final concentrations of

3 μM and 1 μM, respectively. For embryoid bodies were generated with the hanging drop method as previously described (Pasini et al., 2007). The NTERA-2 cells (pluripotent embryonal carcinoma, male, ATCC CRL-1973) and HEK293T cells (human embryonic kidney, contains the SV40 T-antigen, female, ATCC CRL-3216) were grown in DMEM (Sigma-Aldrich) supplemented with 10% (v/v) FBS (GIBCO), 100 U/ml penicillin and 100 U/ml streptomycin (GIBCO). All cell lines were propagated at 37°C with 5% CO_2.

Stable cell lines

To obtain stable knockdowns in mouse ESCs, pLKO.1 lentiviral shRNA expressing vectors directed against mouse Nsd1 (pLKO.TRC1.shmNsd1.1; GCTCGTTAAGACACCAGGAAA; TRCN0000123381; Sigma-Aldrich) and pLKO.TRC1.shmNsd1.2; CCACTGTTCTTCTTTCCTTAT; TRCN0000123379; Sigma-Aldrich; CCAGATTTCTTTCTCCGCCTT) and Setd2, (TRCN0000238533; TRCN0000238536), (Ferrari et al., 2014) were used. A control shRNA, shLuciferase (Sigma, SHC007) was used. Lentiviral particles were generated in HEK293T transfected with pLKO.1 and plasmids encoding the necessary packaging (pPax8) and envelope proteins (pVSVG). Lentiviral particles were added to E14 ES cells together with 10 μg/ml Polybrene (Sigma-Aldrich) for 12-16 hours and then selected with 2 μg/μl puromycin.

METHOD DETAILS

Antibodies

For western blotting, the antibodies used were Ezh2-BD43 (Pasini et al., 2004), Eed-AA19 (Bracken et al., 2003), Suz12 (Cell Signaling, D39F6), Oct4 (Abcam, Ab19857), ElonginB (also known as Tceb2) (Santa Cruz, sc-11447), Nsd1 (Abbexa, abx135901), Dnmt3a (Abcam, ab13888), Hells (Bethyl, A300-226A), Ogt (Santa Cruz, sc-32921), Ewsr1 (Santa Cruz, sc-6533), H3K27me3 (Cell Signaling, C36B11), H3K27me2 (Cell signaling, D18C8), H3K27me1 (Active Motif, MABI 0321), Histone H3 (Abcam, ab1791), H3K36me2 (Cell Signaling, C75H12), H3K36me3 (Cell Signaling, D5A7), Top2a (Santa Cruz, sc-13058), Ogt (Santa Cruz, sc-32921), βTubulin (Santa Cruz, sc-9104). HRP-linked secondary antibodies, anti-rabbit (Sigma, A0545), anti-light chain rabbit (Abcam, ab99697), anti-mouse (Merck, 401253) and anti-goat (Santa Cruz, sc-2020) were used. For immunoprecipitations, the antibodies used were Ezh2-AC22), (Bracken et al., 2006), Suz12 (Cell Signaling, D39F6), rabbit IgG (EMD Millipore, 12-370) and mouse IgG (EMD Millipore, 12-371). For Chromatin-Immunoprecipitations, the antibodies used were H3K27me3 (Cell signaling, C36B11), H3K27me2 (Cell signaling, D18C8), H3K36me2 (Abcam, ab9049, ab176921), Histone H3 (Abcam, ab1791). For Immunofluorescence experiments, the antibodies H3K27me3 (Cell Signaling, C36B11), H3K27me2 (Cell Signaling, D18C8), H3K27me1 (Active Motif, MABI 0321) and Histone H3 (Active Motif, MABI 0301) and Alexa Fluor-conjugated secondary antibodies (A-11012, A-21242, A-21131) were used.

Production of esiRNAs and transfections

Endonribonuclease-prepared siRNAs (esiRNAs) were synthesized in-house as described in Heninger and Buchholz (2007). Briefly, target regions for each gene of interest were designed using DEQOR software (Henschel et al., 2004). Following *in vitro* transcription and purification of these target regions, double-stranded RNAs (dsRNAs) were digested into ~21 nt fragments using *E. coli* purified RNase III mutant E38A. The resulting esiRNAs consisted of a heterozygous population of ~21 nt siRNA molecules for each gene of interest. NTERA-2 and mouse ESCs were transfected with esiRNAs using Lipofectamine RNAiMAX, according to the manufacturer's protocol. For transfections in ESCs, 300 ng of esiRNAs were used together with 3 μL RNAiMax in 600 μL Optimem (GIBCO) for reverse transfection of $1.8x10^5$ cells per 6-well in Penicillin and Streptomycin free ES growth medium. Following incubation for 16 hours, the medium was changed and the cells were harvested 72 h post transfection. The NTERA-2 cells were transfected for 72 h in a 96-well format containing $6x10^3$ cells and were reverse transfected with 25 ng esiRNA together with 0.15 μL RNAiMax.

Functional esiRNA screen for H3K27me2 and H3K27me3

The esiRNA-screen was performed in 96-well format in four biological replicates, with each biological replicate containing two technical replicates. The NTERA-2 cells were reverse transfected with Lipofectamine RNAiMAX and fixed 72 h following transfection for immunofluorescence and Hoechst 33342 (Sigma-Aldrich) staining. Nuclei were stained with Hoechst 33342 and this signal was used to mask the nuclei of cells. In order to exclude apoptotic and mitotic cells from the analysis, gating parameters were carefully set, based on variations of intensity within each cell, and in the cell size and shape. Imaging was performed on a CellInsight Personal Image Cytometer (Thermo Scientific), measuring cell count, histone H3 and global H3K27me1/2/3 intensities per cell. All H3K27 methylation intensities were normalized to the total histone H3 to rule out perturbations which affected histone abundance rather than specifically altering methylation levels. All data was standardized using Z-scoring and subsequently p values were determined to measure the significance of the deviation of each sample's mean from the mean of the pool of controls. Bonferroni correction was used to adjust for multiple comparisons.

Western blotting

For western blotting of total protein, whole cell lysates were generated by lysing cells in IPH buffer containing protease inhibitors (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.5% NP40 and protease inhibitors 2 μg/ml aprotinin, 1 μg/ml leupeptin, 10 μg/ml PMSF, and 0.5 mM DTT). To obtain chromatin associated proteins, the lysates were incubated with Benzonase

(125 U/mg protein) after addition of 7 mM $MgCl_2$ (final 2 mM) at 4C. For Western Blotting of chromatin fractions, cells were harvested and processed as for CHIP analysis and the sonicated chromatin lysates were used for western blots. The protein lysates were separated on SDS–PAGE gels and transferred to nitrocellulose membranes. Membranes were subsequently probed using the relevant primary and secondary antibodies, and relative protein levels were determined by chemiluminescence. For Figures 1 and 2 (and Supplemental Information), western blots were exposed to X-ray films. For Figures 3 and 4 (and Supplemental Information), western blot chemoluminescence signals were captured with the ChemiDoc Imaging Systems (BioRad).

Peptide dot blots

Modified peptides representing Histone H3 N-terminal amino acids 20-40 (LATKAARKSAPATGGVKKPHR) were purchased from Peptide 2.0 1 μg of peptides were spotted onto nitrocellulose membrane and allowed to dry. Membranes were blocked in 5% skimmed milk in PBS-T (0.1%) for 1 hour at room temperature followed by probing with the relevant primary and secondary antibodies. Proteins were visualized by chemiluminescence.

Immunoprecipitations and proteomic analysis

Purification of endogenous multiprotein complexes was performed essentially as previously described (Oliviero et al., 2016). Briefly, nuclear pellets were lysed in high salt containing buffer C (20 mM HEPES at pH 7.6, 20% [v/v] glycerol, 0.42 M NaCl, 1.5 mM MgCl2, 0.2 mM EDTA, aprotinin 1 μg mL−1, leupeptin 10 μg mL−1, PMSF 1mM) and subsequently dialyzed against buffer C-100 (20 mM HEPES at pH 7.6, 20% [v/v] glycerol, 0.2 mM EDTA, 100 mM KCl, 1.5 mM MgCl2, 0.2 mM EDTA). 10 μg of antibodies were chemically cross-linked with dimethyl pimelimidate (Sigma-Aldrich) to 40 μg of Sepharose A beads (Sigma-Aldrich) and immunoprecipitations were performed in the presence of 250 Units Benzonase per mg protein (Sigma-Aldrich, E1014-5KU). 2 mg nuclear lysates were used for each immunoprecipitation. After 3 hours incubation at 4C, immuno-complex bound beads were washed five times with C-100 buffer supplemented with 0.02% NP-40 and with one time in Phosphate buffered saline (PBS). Proteins were either eluted with 2xSDS sample buffer for western blot analyses or digested with Trypsin for mass spectrometry analyses.

Mass spectrometry

All in-solution tryptic digestions were performed as described previously (Wiśniewski et al., 2009). The peptides were analyzed with a Q Exactive mass spectrometer coupled with an EASY-nLC HPLC system (Thermo Fisher) and an in-house packed C18 column (New Objective). Parent ion spectra (MS1) were measured at resolution 70,000, AGC target 3e6. Tandem mass spectra (MS2, up to 10 scans per duty cycle) were obtained at resolution 17,500, AGC target 5e4, collision energy of 25. All mass spectrometry data were processed using the MaxQuant software, version 1.3.0.5 (Cox and Mann, 2008) using the mouse UniProt database. The following search parameters were used; Fixed Mod: carbamidomethylation, Variable Mods: methionine oxidation, Trypsin/P digest enzyme, Precursor mass tolerances 6 ppm, Fragment ion mass tolerances 20 ppm, Peptide FDR 1%, Protein FDR 1%. Volcano plots were generated using Perseus software, version 1.4 (Tyanova et al., 2016). Volcano plots were generated using the parameters FDR 0.05, and S 0.5. To generate the overlap between the Ezh2 and Suz12 IPs (Figure 1B), all proteins detected with a cut-off of at least 0.4 t test LFQ intensities in the Suz12 and Ezh2 IP over IgG were used. For wild-type compared to knockout ES cells (Figure 1F), those proteins detected with > 0.05 t test difference in LFQ intensities in E36 and in E14 wild-type cells over the respective knockout cell line were used to generate the Venn diagrams. All 40S and 60S ribosomal proteins were not considered and removed from the list.

ChIPs, ChIP-Rx and library preparation

All ChIP experiments were performed as described previously (Bracken et al., 2006). We also performed quantitative chromatin immunoprecipitation 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 this, a total of 5% of *Drosophila* chromatin at a sheared size of 200-300 bp was added to each ChIP reaction. For ChIP-RX experiments, the precipitated DNA was quantified using the Qubit dsDNA HS Assay Kit (Thermo Scientific, Q32854). A total of 1-5 ng of DNA from each ChIP-RX experiment was used for library preparation using the Illumina ChIP-Seq Sample Prep Kit (IP-102-1001) and Multiplexing Sample Preparation Oligonucleotide Kit (PE-400-1001). Following adaptor ligation, the DNA was PCR amplified for 15 cycles. DNA purification steps were then performed using DNA SPRI bead (Beckman Coulter, B23318) cleanup. The quality of the DNA libraries was assessed using a high sensitivity Bioanalyzer Chip (Agilent). The resultant libraries were then used for cluster generation and sequencing using HiSeq 2000 (Illumina) and 50 bp read length.

Next generation sequencing analysis

Sequencing reads were aligned to the mouse reference genome (mm9) using Bowtie (Langmead et al., 2009). Only unique alignments were retained for downstream analyses. For ChIP-Rx experiments, 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 tracks using the UCSC genome browser (Kent et al., 2002). Genome-wide differences in the wild-type and Nsd1 depleted cells for H3K27me3 and

H3K36me2 were visualized by partitioning the genome into bins at a 50kb resolution and subtracting wild-type ChIP-Rx normalized read counts in each bin from the corresponding counts in the knock-downs. Data were scaled for visualization purposes. ChIP enriched regions (peaks) were computed using macs2 (Zhang et al., 2008) for the H3K27me3 wild-type and Nsd1.1 and Nsd1.2 depleted cells at a FDR of 0.1. Average profiles were constructed for these regions (midpoint of peak ± 40kb) using ngsplot (Shen et al., 2014) for H3K27me3 and their corresponding sites in H3K36me2. Genome-wide correlations between histone modifications were computed by partitioning the mouse genome into 10kb bins and comparing the normalized reads counts within each bin. Normalized read counts were scaled to a 0-1 range. Furthermore, we isolated those 10kb bins that had a gain in H3K27me3 greater than 0.2 reads per reference per million (RPM) and a loss of less than 0.2 RPM in H3K36me2. Promoters (TSS ± 2.5kb), gene bodies and intergenic regions overlapping these bins were obtained and the ChIP-Rx normalized read counts of these regions represented as boxplots. Additionally, principal component analysis (PCA) using FactoMineR (Lê et al., 2008) was performed to understand the degree of correlation between the different histone modifications.

Quantitative real-time PCR

Total RNA was extracted from cells using the RNeasy kit (QIAGEN) according to manufacturer's protocol. RNA was used to generate cDNA by reverse transcriptase PCR using the TaqMan Reverse Transcription kit (Applied Biosytems). Relative mRNA expression levels were determined using the SYBR Green I detection chemistry (Applied Biosystems) on the ABI Prism 7500 Fast Real-Time PCR System. The ribosomal constituent Rplpo or Gapdh were used as a control gene for normalization. For Box-Whisker-Plots, the online tool BoxPlotR (http://shiny.chemgrid.org/boxplotr) was used. The Primer sequences are available in Table S4.

QUANTIFICATION AND STATISTICAL ANALYSIS

All qPCR data are presented as the mean and standard deviation from either technical replicates or independent biological replicates (n represents biological replicates). The information is indicated in the Figure legends, accordingly. In Figure 1, for peptide mapping, the MaxQuant software, version 1.3.0.5 (Cox and Mann, 2008) was used with an FDR of 1% for peptides and proteins. Volcano plots relating to Figures 1 and S1, Tables S1 and S2, were generated using Perseus software, version 1.4, as previously described, with FDR 0.05, and S 0.5. In the esiRNA screen in Figure 2 and Table S3, immunofluorescence intensities were measured and determined by the built-in software of the CellInsight Personal Image Cytometer (Thermo Scientific). All immunofluorescence intensities were subsequently standardized using Z-scoring, by the following formula: $z = x - \mu / \sigma$, where x is the well intensity, μ is the mean control intensity for the plate (n = 8 control wells per plate), and σ is the standard deviation of the control wells in the plate. Prism software (version 6.07, GraphPad) was used to perform one-way ANOVA analysis of the standardized intensities of all plates (4 independent replicates, with 2 technical replicates each), comparing the mean of each sample to the mean of the negative controls. Bonferroni correction was used to control for multiple comparisons and confidence intervals were computed for 95% confidence with statistical significance defined using an alpha of 0.05 after correction. In Table S3, significance is denoted as follows: $p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = ***, and $p < 0.0001$ = ****. Finally, for the ChIP-Rx statistical analysis in Figure 4B, the Pearson correlation coefficients and p values were computed using the cor.test function in the R programming language. Further details can be found in the figure legends.

DATA AND SOFTWARE AVAILABILITY

The accession number for the H3K27me3 ChIP-Rx, H3K36me2 ChIP-Rx, and ChIP-seq data reported in this paper is GEO: GSE107773. The accession numbers for the publicly available H3K27me3 and H3K36m3 data reported in this paper are GEO: GSM1234538 and GSM1234540, respectively. The accession numbers for the publicly available RNA-seq data from ESC differentiation into embryoid bodies reported in this paper are GEO: GSE51006, GSM1249334, and GSM1249336. The accession number for the proteomics data reported in this paper is PRIDE: PXD005381. Original images have been deposited to Mendeley Data and are available at: https://data.mendeley.com/datasets/4pzh3gjkm9/draft?a=0297530d-c12d-45ee-bba1-867c3103c13e.