A Functional Link between Nuclear RNA Decay and Transcriptional Control Mediated by the Polycomb Repressive Complex 2

Graphical Abstract

Highlights
- Depletion of ZFC3H1 in mouse ESCs results in differentiation defects
- PRC2 target genes are deregulated in Zfc3h1−/− cells
- Chromatin binding of PRC2 and H3K27me3 is reduced in Zfc3h1−/− cells
- Increased binding of RNA impairs PRC2 complex stability

Authors
William Garland, Itys Comet, Mengjun Wu, ..., Albin Sandelin, Kristian Helin, Torben Heick Jensen

Correspondence
helink@mskcc.org (K.H.), thj@mbg.au.dk (T.H.J.)

In Brief
ZFC3H1 targets pA+ RNA for decay by the nuclear RNA exosome. Garland et al. report a disruptive relationship between excess RNA and PRC2 upon depletion of ZFC3H1 in mouse ESCs. In such conditions, RNA is bound by PRC2 components, which show reduced binding to chromatin and fellow PRC2 proteins.
A Functional Link between Nuclear RNA Decay and Transcriptional Control Mediated by the Polycomb Repressive Complex 2


SUMMARY

Pluripotent embryonic stem cells (ESCs) constitute an essential cellular niche sustained by epigenetic and transcriptional regulation. Any role of post-transcriptional processes remains less explored. Here, we identify a link between nuclear RNA levels, regulated by the poly(A) RNA exosome targeting (PAXT) connection, and transcriptional control by the polycomb repressive complex 2 (PRC2). Knockout of the PAXT component ZFC3H1 impairs mouse ESC differentiation. In addition to the upregulation of bona fide PAXT substrates, Zfc3h1−/- cells abnormally express developmental genes usually repressed by PRC2. Such de-repression is paralleled by decreased PRC2 binding to chromatin and low PRC2-directed H3K27 methylation. PRC2 complex stability is compromised in Zfc3h1−/- cells with elevated levels of unspecific RNA bound to PRC2 components. We propose that excess RNA hampers PRC2 function through its sequestration from DNA. Our results highlight the importance of balancing nuclear RNA levels and demonstrate the capacity of bulk RNA to regulate chromatin-associated proteins.

INTRODUCTION

Embryonic stem cells (ESCs) are distinguished by their dual ability to self-renew and differentiate, both of which require tight regulatory control. ESC pluripotency is maintained by a complex molecular network centered around key transcription factors (TFs), including OCT4, SOX2, NANOG, and KLF2 (Morey et al., 2015; Takahashi and Yamanaka, 2006; Zhou et al., 2007). In addition, epigenetic mechanisms establish and maintain specialized chromatin through DNA methylation and histone modifications to allow the activation and repression of genes during development (Bibikova et al., 2008; Chen and Dent, 2014). In consequence, perturbation of DNA methyltransferases (DNMTs) or chromatin repressive complexes, for example, can disrupt mammalian development through the dysregulation of normal gene expression programs (Laugesen and Helin, 2014; Smith and Meissner, 2013).

The exit of ESCs from pluripotency requires concerted silencing of pluripotency factors and activation of lineage-specific genes (Loebel et al., 2003). A key player here is the polycomb repressive complex 2 (PRC2), which catalyzes the formation of facultative heterochromatin via trimethylation of lysine 27 on histone 3 (H3K27me3). PRC2 is dispensable for the maintenance of self-renewal in ESCs but functions to prevent inappropriate transcriptional activation of lineage-specific differentiation factors. These include highly conserved homeobox (HOX) factors, which are crucial for regulating axial patterning in development (Laugesen and Helin, 2014; Pearson et al., 2005), among other developmentally associated TFs that specify cell fate. In ESCs, these genes are repressed through a combination of transcriptional and epigenetic control requiring PRC2 and H3K27me3 (Mallo and Alonso, 2013). Furthermore, PRC2 is crucial for cell fate transitions during development from ESCs, where knockouts (KOs) of core complex components EZH2, SUZ12, and EED result in a block in differentiation (Chamberlain et al., 2008; Pasini et al., 2007; Shen et al., 2008).

Although research on ESC regulation has focused predominantly on transcriptional or epigenetic control, a role of post-transcriptional events, including their putative coupling to transcriptional control, has been less explored. A proper balance of RNA processing and decay ensures homeostasis, whereby different steps in gene regulation buffer one another to maintain a stable expression profile within cell types (for recent reviews see Schmid and Jensen, 2018; Timmers and Tora, 2018). Consequently, malfunction of nuclear RNA decay pathways are therefore linked to developmental disorders and human disease (Corbett, 2018). With improved sequencing technologies, the complexity of the non-protein-coding genome has been revealed (Carninci et al., 2005; Djebari et al., 2012). Functional roles
Figure 1. Zfcs1<sup>−/−</sup> Cells Display Defects in EB Differentiation

(A) Schematic representation of the nuclear exosome complex (EXO13) and its PAXT connection. Question marks denote yet-to-be-defined PAXT components.

(B) Western blotting analysis of WT and three independent Zfcs1<sup>−/−</sup> cell lines (#1–#3). Blots were probed with the indicated PAXT-related antibodies and actin (ACTB) as a loading control.

(C) qRT-PCR analysis of the indicated PAXT targets from total RNA isolated from WT and Zfcs1<sup>−/−</sup> cell lines. Primers were designed to span exon-exon junctions in order to amplify spliced host gene transcripts. Results are shown relative to Riplp0 mRNA (RP0) and normalized to average WT values. Columns represent average values of technical triplicates per sample, with error bars denoting SD. Individual data values from replicates are indicated as points.

(D) Phase contrast microscopy images of WT and Zfcs1<sup>−/−</sup> (#1–#3) colonies after 8 days of EB induction. Scale bars denote 200 μm.

(legend continued on next page)
have emerged for long non-coding RNAs (lncRNAs), which form ribonucleoprotein (RNP) complexes capable of regulating various stages of gene expression (Seiler and Coller, 2013; Rinn and Chang, 2012), including the maintenance of ESC pluripotency and differentiation (Guttman et al., 2011; Luo et al., 2016). Levels of lncRNAs in eukaryotic nuclei are tightly regulated by RNA decay systems, which consequently must affect biological processes regulated by such transcripts. Furthermore, it has been shown that many key regulators of gene expression also have RNA binding activity (Hendrickson et al., 2016; Khalil et al., 2009). PRC2, for example, binds promiscuously to RNA both in vitro and in vivo, with all core components (EZH2, SUZ12, and EED) contributing to varying degrees (Cifuentes-Rojas et al., 2014; Davidovich et al., 2013; Zhao et al., 2008, 2010). The function of PRC2-RNA binding is not fully understood, but results have suggested roles of both transcript-mediated recruitment and eviction of PRC2 to and from chromatin (Davidovich et al., 2013; Kaneko et al., 2013; Rinn et al., 2007; da Rocha et al., 2014).

The RNA exosome is an essential 3’-5’ ribonucleolytic complex involved in the regulation of the majority of nuclear transcripts (Kilchert et al., 2016; Mitchell et al., 1997; Schmid and Jensen, 2008). Assessing possible links between transcription regulation and the post-transcriptional balancing of RNA levels, the nuclear exosome stands out with its global activity in the processing of precursor RNAs and its ability to efficiently remove transcriptional by-products and otherwise nuclear retained RNA (Schmid and Jensen, 2018). To facilitate recognition and targeting of its plethora of transcript targets, the nuclear exosome associates with adaptor complexes; that is, two nucleoplasmic decay pathways are guided by the nuclear exosome targeting (NEXT) complex and the poly(A) exosome targeting (PAXT) connection, respectively (Lubas et al., 2011, 2015; Meola et al., 2016; Silla et al., 2018). NEXT and PAXT share a common subunit in the RNA helicase MTR4, which connects these adaptors to the exosome. PAXT is also composed of a large zinc finger protein, ZFC3H1, which bridges MTR4 to the nuclear poly(A) binding protein (PABPN1), aiding the targeting of polyadenylated (pA+) nuclear RNAs for exosome-mediated decay (Beaulieu et al., 2012; Bresson and Conrad, 2013; Meola et al., 2016; Ogami et al., 2017).

At steady state, IncRNA levels in mammalian cells are generally low, with estimates suggesting that less than 1,000 lncRNAs are present in more than one copy per cell (Djebali et al., 2012; Seiler et al., 2017). Thus, most lncRNAs are stoichiometrically inferior to their putative protein effectors, often weakening the associated mechanistic models claiming function of individual lncRNAs. The abundance of IncRNAs is regulated through nuclear RNA decay pathways and is considerably enriched upon removal of exosome components (Lubas et al., 2011; Meola et al., 2016; Silla et al., 2018). Manipulating PAXT activity, through the depletion of ZFC3H1, therefore allows an approach to study the general effects of excess pA+ RNA in the nucleus.

Here, we establish a functional link between PAXT activity and transcriptional control mediated by PRC2. Zfc3h1KO cells are unable to differentiate and exhibit phenotypes reminiscent of cells deficient for PRC2 activity. Consistently, normal PRC2 function is impaired in Zfc3h1KO cells and we provide evidence that this is due to PRC2 binding to stabilized RNAs. Our results highlight the importance of controlling nuclear RNA levels during key regulatory stages of ESC development and imply that modulation of bulk RNA levels is a potent way of lncRNA-mediated transcription regulation.

RESULTS

**Zfc3h1KO Cells Exhibit Defective Embryoid Body Differentiation**

To assess the role of exosome-mediated decay of nuclear pA+ RNAs in ESC pluripotency and differentiation, we used CRISPR/Cas9 to generate homozygous KOs of the PAXT component ZFC3H1 (Figure 1A). Zfc3h1 was specifically targeted because MTR4 and PABPN1 also reside in alternative nuclear complexes. Three biologically independent Zfc3h1KO ESC lines with disrupted Zfc3h1 ORFs were derived from single-cell KO clones (Figure S1A). In agreement with our previous observations in human cells, the expression of other known PAXT-related (Figure 1B) and exosome-related (Figure S1B) proteins was unaffected by ZFC3H1 depletion (Meola et al., 2016). Still, PAXT-mediated RNA decay was disrupted, which resulted in an approximately 2-fold accumulation of total nuclear pA+ RNA (Figure S1C), including spliced small nucleolar RNA (snoRNA) host gene (Snhg) IncRNAs (Meola et al., 2016; Figure 1C).

Zfc3h1KO cells were viable under 2i+LIF growth conditions, which selects against cellular differentiation, and appeared morphologically similar to wild-type (WT) cells (Figure S1D). Furthermore, expression of the key pluripotency TFs NANOG, ESRRB, KLF2, SOX2, and OCT4 was not perturbed (Figure S1E), which suggests that the absence of ZFC3H1 does not affect the self-renewal properties of ESCs. Strikingly, however, when cells were transferred from 2i+LIF conditions to serum-LIF media on low-attachment plates to allow the spontaneous development of embryoid bodies (EBs) (Figure S1F), Zfc3h1KO cells retained an ESC-like morphology, with smaller rounded colonies and a lack of differentiated cell types (Figure 1D). The cells also showed a lack of cystic EBs, cavities that normally form during differentiation because of programmed cell death, but instead remained as solid aggregates, apparent as dark spots on microscopy images (Figure 1D). Finally, between 7 and 8 days of differentiation, WT cells normally develop...
Figure 2. PRC2 Target Genes Are Upregulated in Zfc3h1−/− ESCs

(A) Differential gene expression analysis of RNA-seq data from Zfc3h1−/− versus WT cells after 7 days of EB induction shown as an MA plot. The y axis shows Zfc3h1−/− versus WT RNA-seq log2 FC. The x axis shows average normalized expression as log2 CPM values across all biological replicates. Each dot indicates a gene, and blue/yellow color denotes significant differential expression (edgeR FDR < 0.05). Red lines denote log2 FC > 1, < −1. Upregulated genes involved in pluripotency and downregulated genes involved in differentiation are highlighted.

(B) MA plot as in (A) but for undifferentiated (D0) Zfc3h1−/− versus WT cells. Values are averages of two WT and three Zfc3h1−/− biological replicates. Upregulated genes highlighted indicate PRC2 target genes involved in developmental processes.

(C) Bar plot of upregulated (log2 FC > 0.5) gene types in D0 Zfc3h1−/− cells. GENCODE RNA biotypes are color coded as indicated, and numbers of affected genes are shown in parenthesis.
spontaneous beating colonies, indicative of cardiomyocyte formation (Doetschman et al., 1985), which was not observed up to 10 days of EB induction of Zfc3h1−/− cells.

We then collected samples across the EB differentiation time course to analyze the expression of cell type-specific markers (e.g., pluripotency TFs normally decrease rapidly when cells differentiate). In contrast, Zfc3h1−/− cells retained high expression of the pluripotency TFs OCT4, SOX2, ESRRB, and KLF2 at later stages of the time course as assessed by RNA and protein analyses (Figures 1E and 1F). qRT-PCR primers spanning exon-intron (ExIn) borders of Oct4 and Nr0b1 pre-mRNAs showed that intronic sequences were elevated in Zfc3h1−/− cells (Figure S1G), suggesting that these genes might be transcriptionally upregulated. Finally, the expression of lineage-specific markers was impaired in Zfc3h1−/− cells, with little or no expression of endodermal, mesodermal, and ectodermal TFs (Figure 1G). Taken together, we conclude that Zfc3h1−/− cells show EB differentiation phenotypes and retain an expression profile reminiscent of undifferentiated cells.

**Transcripts from PRC2-Targeted Genes Are Enriched in Zfc3h1−/− Cells**

To obtain a global impression of the disparities between WT and Zfc3h1−/− cells, sequencing of rRNA-depleted total RNA (RNA-seq) was performed on samples harvested after 7 days (D7) of EB induction. This time point was chosen because it showed large gene expression differences between WT and Zfc3h1−/− cells as measured by qRT-PCR (Figures 1E and 1G). In addition, total RNA from ESC (D0) samples was sequenced to assess the undifferentiated starting point of the experiment. Biological replicates of RNA-seq libraries were overall highly correlated (Figures S2A and S2B). Differential expression (DE) analysis of D7 samples mirrored our qRT-PCR analyses: Zfc3h1−/− cells displayed elevated expression of pluripotency-associated TFs and decreased expression of germ layer-specific transcripts (Figure 2A). This, in conjunction with the morphological phenotypes, suggested that Zfc3h1−/− cells are unable to exit from the embryonic stem (ES) state and activate the normal genes required to initiate differentiation. We surmised that this inability to exit from pluripotency might be due to gene expression aberrancies in the D0 ES state disrupting normal progression into differentiation pathways. Consistent with this notion, DE analysis revealed higher levels of mRNAs encoding lineage-specific factors normally associated with differentiated cells in Zfc3h1−/− versus WT D0 samples (Figure 2B). A large number of these upregulated lineage markers comprised TFs, including HOX genes, which are involved in early developmental processes (Pearson et al., 2005). At first glance, such an expression profile seemingly contrast our observation that Zfc3h1−/− ESCs appear morphologically similar to WT ESCs. However, although exit from pluripotency requires the expression of developmental TFs, this must occur concomitant with suppression of pluripotency TFs, which predominantly define the cellular state of ESCs (Young, 2011). Moreover, the selective pressure of the 2i/LIF culture condition maintains the pluripotent state by blocking MEK and GSK activity and activating the STAT3 pathway (Wray et al., 2010; Ying et al., 2008).

Nevertheless, abnormal expression of developmental genes in the Zfc3h1−/− ESCs indicated a general deregulation of differentiation-associated genes. In WT ESCs, these genes are normally repressed in an inactive chromatin environment highly decorated with H3K27me3 (Mallo and Alonso, 2013). Furthermore, the repression of pluripotency genes during differentiation is regulated by H3K27me3 (Ober et al., 2015; Pasini et al., 2007). As this histone mark is solely catalyzed by the PRC2 complex, we compared the gene expression profile of Zfc3h1−/− ESCs with published PRC2 KO (Ezh1−/−/Ezh2−/−) RNA-seq data derived from the same parental mouse ESC line (Højfeldt et al., 2018). Of the 1,804 upregulated transcripts in the Zfc3h1−/− D0 samples (log2 fold change [FC] > 0.5, false discovery rate [FDR] < 0.05, edgeR), approximately 25% could be designated as non-coding RNAs (ncRNAs) (Figure 2C), including known PAXT targets upregulated because of their diminished decay, while ~75% of cases were protein coding. Almost one-third of the latter transcripts showed a significant (p < 1.1e-18, hypergeometric test) overlap with transcripts upregulated in PRC2 KO cells (Figure 2D). For these shared transcripts, exon and intronic reads were evenly upregulated in the Zfc3h1−/− D0 data (Figure 2E), suggesting that increased mRNA levels were based on increased transcription. This was validated by qRT-PCR analysis of upregulated HOX transcripts using ExIn-specific primers on chromatin-associated RNA to enrich for pre-mRNA (Figure 2F). We conclude that Zfc3h1−/− ESCs have higher levels of transcripts derived from a subset of PRC2 target loci, and these genes appear to be more transcriptionally active in the absence of PAXT.

**SUZ12 Chromatin Occupancy and H3K27me3 Levels Are Decreased in Zfc3h1−/− Cells**

As Zfc3h1−/− cells displayed increased expression of transcripts also upregulated in the absence of PRC2, we investigated the status of the PRC2 complex in the Zfc3h1−/− background. Expression of its core components, EZH2, SUZ12, and EED, was unaffected by the absence of ZFC3H1 (Figure S3A). Therefore, we decided to conduct chromatin immunoprecipitation sequencing (ChIP-seq) on the three Zfc3h1−/− cell lines, along

---

**Notes:**

(D) Venn diagram showing overlap of upregulated (log2 FC > 0.5, FDR < 0.05) protein-coding genes in Zfc3h1−/− and Ezh1−/−/Ezh2−/− undifferentiated (D0) cells versus WT cells. The numbers of genes in each set are shown.

(E) Relation between exonic and intronic read counts in Zfc3h1−/− and Ezh1−/−/Ezh2−/− upregulated genes. Axes show log2 FC of normalized RNA-seq reads calculated using edgeR in introns (x) and exons (y) of 331 intron-containing genes upregulated in both Zfc3h1−/− and Ezh1−/−/Ezh2−/− RNA-seq datasets (from the intersect in D), with a diagonal reference line (y = x). Dark green dots indicate FDR < 0.05, and light green dots indicate FDR < 0.05 for intronic reads.

(F) Upper panel: Genome Browser views of RNA-seq data from WT and Zfc3h1−/− D0 (1) libraries at the indicated HOX gene loci. Tracks show WT and Zfc3h1−/− read densities at + and – strands as indicated. RefSeq gene models and mm10 genome coordinates are shown, together with ExIn qRT-PCR primer amplicons used in the lower panel. Lower panel: qRT-PCR analysis of HOX gene pre-mRNAs from chromatin-associated RNA isolated from WT and Zfc3h1−/− (1) D0 cells. Results are shown relative to GAPDH mRNA values and normalized to average WT values. Results are displayed as in Figure 1C.
Figure 3. SUZ12 DNA Occupancy and H3K27me3 Levels Are Decreased in Zfc3h1−/− Cells
(A) MA plots showing global changes in H3K27me3 and SUZ12 ChIP-seq densities in Zfc3h1−/− versus WT cell lines. Dots indicate sliding genome windows, where color intensity indicates the density of overlapping points. Y axes show log2 FC Zfc3h1−/− versus WT normalized ChIP signal, and x axes show average log2 ChIP signal, all from three independent biological replicates.
(B) Boxplot distributions of log2 FC Zfc3h1−/− versus WT normalized H3K27me3 and SUZ12 ChIP signals on regions centered on 3678 SUZ12 peaks identified in WT cells.
(C) Average SUZ12 ChIP-seq signals for Zfc3h1−/− and WT cell lines in regions centered on the SUZ12 peaks as in (B). Zfc3h1−/− ChIP data are the average of three biological replicates.
(D) As in (C) but for H3K27me3 signal from regions centered on the SUZ12 peaks from (B).
(E) Heatmap representation of the relation between SUZ12 ChIP-seq, H3K27me3 ChIP-seq, and RNA-seq log2 FC in Zfc3h1−/− versus WT cells. Columns represent SUZ12 peaks identified in WT cells (N = 1,485), sorted after SUZ12 Zfc3h1−/− versus WT log2 FC. Rows correspond to SUZ12 ChIP-seq, H3K27me3 ChIP-seq, and RNA-seq Zfc3h1−/− versus WT log2 FC values.

(legend continued on next page)
with the parental WT cell line, using antibodies specific for H3K27me3 and SUZ12. Zfc3h1−/− cells displayed reduced H3K27me3 levels and SUZ12 DNA occupancy in comparison with WT cells (Figures 3A, S3B, and S3C). These reductions were generally found in regions with high average signal, and the effect on H3K27me3 levels was substantially stronger than the effect on SUZ12 binding (Figure 3B). Still, while H3K27me3 levels were clearly decreased in Zfc3h1−/− cells, they were not abolished (average reduction to ~75% versus WT, Figure S3D; as opposed to previously reported KOs of different PRC2 components in which H3K27 methylation is absent, Höjfeldt et al., 2018; Pasini et al., 2007; Shen et al., 2008; see Discussion). PRC2-bound regions, defined by SUZ12 ChIP-seq peaks in WT cells, had significantly (p < 2.2e-16, one-sided Mann-Whitney test) lower average SUZ12 and H3K27me3 occupancy in Zfc3h1−/− conditions (Figures 3B-3D). Moreover, regions depleted for PRC2 and H3K27me3 in Zfc3h1−/− cells showed a concomitant increase in RNA expression (Figure 3E), which was concomitantly elevated whether exonic or intronic reads were interrogated (Figure 3F).

ChIP-seq analyses of active and primed chromatin modifications (H3K4me1, H3K4me3, H3K27ac) were also carried out in WT and Zfc3h1−/− cells. Genes depleted for SUZ12 and H3K27me3 ChIP signals had a marked increase in H3K27ac levels, concordant with increased RNA transcription (Figures 3G and S3F). We also obtained low-sequence depth RNA polymerase II (RNAPII) ChIP-seq data, which showed increased occupancy at select PRC2 target genes in Zfc3h1−/− cells (Figure S3H). Altogether, this fits the notion that PRC2 loci are more transcriptionally active in the absence of normal polycomb-mediated repression via H3K27me3. We conclude that recruitment of SUZ12 to PRC2 target genes is reduced in Zfc3h1−/− cells, resulting in loss of H3K27me3 at these regions and abnormal RNA expression due to increased transcription.

Decreased PRC2 Complex Integrity in Zfc3h1−/− Cells

The discovered correlation between PAXT-mediated RNA decay and PRC2-mediated transcriptional repression is unprecedented. We therefore sought to address how impaired PRC2 function in Zfc3h1−/− cells relates to the primary phenotype of stabilized nuclear pA+ RNAs. As mentioned above, steady-state levels of PRC2 components remained unchanged in Zfc3h1−/− cells (Figures S3A, 4A, and 4B, “Inputs”). However, co-immunoprecipitation (coIP) analyses of SUZ12 (Figures 4A and S4A) or EZH2 (Figures 4B and S4B) from whole-cell lysates revealed their reduced binding to the remaining PRC2 core. Similar effects were observed when conducting IPs from nuclear extracts (Figures 4C and S4D). Such weakened PRC2 complex formation in Zfc3h1−/− cells was further supported by analyzing the sedimentation of PRC2 components through glycerol gradients: Although the overall protein distribution and sedimentation of loading controls remained unchanged (Figures S4E and S4F), SUZ12 and EZH2 showed a marked shift from high-molecular weight fractions toward lower molecular weight fractions of the gradient derived from Zfc3h1−/− lysates (Figures 4C and 4D). The distribution of EED was more dispersed throughout the gradient from WT lysates but still showed a shift toward lower fractions in Zfc3h1−/− samples (Figure S1F). Despite these shifts, a fraction of PRC2 sedimented normally in Zfc3h1−/− samples and consistently remained bound in coIPs (Figures 4A-4D and S4F). Together, these analyses suggested that the PRC2 complex is partially compromised in Zfc3h1−/− cells, which could be further elaborated by subcellular fractionation of WT and Zfc3h1−/− cells, giving rise to a slight, but significant, shift in the distribution of SUZ12, EZH2 and EED from chromatin-bound to nucleoplasmic fractions (Figures S4G and S4H). Previous reports demonstrated that the complete loss of SUZ12 or EED results in reduced levels of the remaining PRC2 core components (Höjfeldt et al., 2018; Montgomery et al., 2005; Pasini et al., 2007), suggesting that PRC2 complex stability depends on the interactions of all core proteins. In contrast, here we found lower levels of the PRC2 complex in Zfc3h1−/− cells despite normal expression of SUZ12, EZH2, and EED.

Given the RNA accumulation observed in the absence of PAXT (Meola et al., 2016; Figures 1C, 2C, and S1C) and the RNA-binding properties of PRC2, we hypothesized that accumulation of PAXT targets could contribute to PRC2 disruption, linking the primary Zfc3h1−/− phenotype with a secondary PRC2 phenotype. Hence, we carried out native IPs of SUZ12 and EZH2, using lysates prepared from WT and Zfc3h1−/− ESCs, and measured the amount of co-isolated RNA. From WT cells, EZH2 IPs showed a greater enrichment of RNA over SUZ12 IPs and the IgG control (Figure 4E), which is in agreement with EZH2 being the PRC2 component with the strongest capacity for RNA binding (Cifuentes-Rojas et al., 2014). However, elevated RNA levels were isolated from both SUZ12 and EZH2 IPs of Zfc3h1−/− versus WT cell lysates. Thus, concomitant with complex disruption, more RNA was bound to PRC2 components upon PAXT depletion. RNA isolated from SUZ12 and EZH2 IPs was sequenced (RIP-seq) to interrogate any differences in PRC2-bound transcripts from WT versus Zfc3h1−/− ESCs. Consistent with previous native RIP-seq analysis of PRC2 components (Davidovich et al., 2013; Khalil et al., 2009; Zhao et al., 2010), we found that EZH2 and SUZ12 showed promiscuous binding to both protein-coding RNA and ncRNA. In Zfc3h1−/− samples, both SUZ12 and EZH2 RPs revealed PAXT targets, including pA+ nuclear IncRNAs (Figure S5A), along with transcripts from de-repressed PRC2 target genes (Figure S5B). However, for transcripts upregulated in Zfc3h1−/− ESCs, there was no significant enrichment of coding RNAs or ncRNAs over the input samples (Figure S5C). Furthermore, there was no specific

(F) Correlation of Zfc3h1−/− versus WT log2 FC RNA-seq signal in intronic versus exonic regions. Points displayed are intron-containing transcripts with log2 FC > 0 in Zfc3h1−/− versus WT overlapping with SUZ12 peaks that show log2 FC > 0 in Zfc3h1−/− ChIP-seq data (N = 365). Axes show Zfc3h1−/− versus WT log2 FC RNA-seq in exonic (y) and intronic (x) regions of these transcripts.

(G) Genome Browser views of four PRC2 target genes (Hoxd8, Sox21, Gsc, and Nefm). Displayed tracks include H3K4me1, H3K4me3, H3K27ac, H3K27me3, and SUZ12 ChIP-seq data as well as RNA-seq data from WT and Zfc3h1−/− (#1) cell lines. RNA-seq tracks on both strands are shown (+ and −, respectively). Gene models are based on RefSeq. Genome coordinates (mm10) are indicated for each panel.
enrichment of RNAs upregulated in Zfc3h1−/− ESCs only, in both Zfc3h1−/− and Ezh1−/−/Ezh2−/− ESCs (Højfeldt et al., 2018) or for transcripts upregulated at SUZ12-depleted loci (Figure S5D). This suggests that increased nuclear RNA levels in Zfc3h1−/− ESCs leads to more RNA binding to PRC2 components with no general specificity.

To address whether the disruption of PRC2 in Zfc3h1−/− cells was indeed dependent on increased RNA levels, we repeated the SUZ12 coIP analyses using cell extracts that were either mock or RNaseA treated. In the presence of RNase, SUZ12 IPs from Zfc3h1−/− extracts recapitulated the coIP efficiency of equivalent WT extracts (Figures 4F and S5E). More bait protein was generally pulled down in the RNase-treated samples, which suggested that RNA already interfered with the IP in WT extracts, but to a greater extent in Zfc3h1−/− extracts. Taken together, we propose that increased RNA binding weakens PRC2 complex formation, thereby decreasing its normal function, stability, and recruitment to chromatin (Figure 4G; see Discussion).

**DISCUSSION**

The molecular decisions that govern correct progression through cellular differentiation require a complex agreement of checks and balances. Here, the transcriptional and epigenetic profiles at the ESC stage are important starting points, whose dysregulation may be deleterious for development. In the present study, we discovered a link between excess nuclear RNA and PRC2-mediated transcriptional control in ESCs, suggesting an essential role of nuclear RNA turnover in cellular commitment to differentiation.

Deletion of the nuclear exosome adaptor ZFC3H1 to some extent phenocopies ESCs depleted for PRC2 components; cells...
self-renew and appear morphologically similar to WT cells but show a loss of H3K27me3, deregulation of PRC2 target genes, and difficulty initiating differentiation (Boyer et al., 2006; Chamberlain et al., 2008; Pasini et al., 2007; Shen et al., 2008). In contrast to PRC2-depleted cells, Zfc3h1−/− ESCs retain a degree of SUZ12 binding and H3K27me3 modification, which presumably explains the less severe deregulation of PRC2 target genes. Still, the defect appears to be sufficient to restrict the progression into differentiation. In agreement with previous studies, we find that deregulation of PRC2 target genes results in abnormal transcription of developmental genes (Boyer et al., 2006; Lee et al., 2008). It would seem counterintuitive for ESCs expressing developmental genes to retain self-renewal ability and to lose the ability to differentiate upon induction of EB formation. However, it has been suggested that the expression of pluripotency TFs is sufficient for self-renewal and overrides any abnormal expression of lineage markers (Chamberlain et al., 2008). As cells exit pluripotency, PRC2 functions to silence key maintenance factors such as Oct4, Sox2, and Nanog by depositing H3K27me3 at their loci (Obier et al., 2015). With PRC2 function decreased in Zfc3h1−/− cells, this presumably explains their retained transcriptional activity of pluripotency TFs and block in differentiation (Figures 1E, 1F, and S1G). Together this reiterates that PRC2 and, in turn, ZFC3H1 are dispensable for self-renewal but are essential for the progression into differentiation (Chamberlain et al., 2008; Pasini et al., 2007; Shen et al., 2008).

An RNA-binding ability of the PRC2 complex has been widely documented, with suggested models for transcript-mediated recruitment or eviction of PRC2 to or from DNA (Davidovich et al., 2013; Kaneko et al., 2013; Rinn et al., 2007; da Rocha et al., 2014). A primary phenotype of ZFC3H1 depletion is the stabilization of nuclear pA+ RNAs (Meola et al., 2016; Ogami et al., 2017; this study), which are, by nature, unstable and typically present only in trace amounts under normal conditions. Through removing the targeting mechanisms for decay, this increases the concentration of pA+ RNAs (Silla et al., 2018). Taken together with increased binding of RNA to EZH2 and SUZ12 in Zfc3h1−/− cells, we therefore propose that increased transcript levels negatively affect PRC2 function through its increased RNA binding (Figure 4G). Previous studies initially suggested that excess RNA can inhibit the methyltransferase activity of EZH2 in vitro (Cifuentes-Rojas et al., 2014; Kaneko et al., 2014), which was further elaborated to suggest that decreased catalytic activity was due to RNA titrating PRC2 off nucleosomes (Wang et al., 2017). This was supported by observations that DNA- and RNA-binding capabilities of PRC2 are mutually exclusive in vitro (Beltran et al., 2016; Wang et al., 2017). More recently, an RNA-binding region was identified at an allosteric regulatory region of PRC2 in close proximity to the methyltransferase region of EZH2, which is subsequently inhibited by RNA binding (Zhang et al., 2019). It is therefore plausible that increased nuclear RNA levels dually affect PRC2 function by decreasing its catalytic activity as well as its DNA-binding capacity. We also find that the interaction between PRC2 subunits is compromised in Zfc3h1−/− cells with reduced binding between core subunits in colP and glycerol gradient assays (Figures 4A–4D). Such decreased interaction between SUZ12 and EZH2 may contribute to an explanation of the stronger loss of H3K27me3 compared with SUZ12 ChIP-seq signal in Zfc3h1−/− cells (Figure 4B); that is, residual SUZ12, uncoupled from EZH2, may still bind DNA. In line with this, our ChIP and subcellular fractionation data showed that a fraction of PRC2 components are still associated with chromatin. This partial phenotype allows us only to speculate at this point and will require further investigation to understand the status of PRC2 complex proteins that are still able to bind DNA. However, in support of this possibility, the N-terminal region of SUZ12 recapitulates SUZ12-binding patterns but lacks EZH2 interaction and thereby does not rescue H3K27me3 activity (Højfeldt et al., 2018). The use of such mutants, in combination with Zfc3h1−/− cells and extended ChIP datasets, might allow unpacking of the molecular basis behind these observations.

Finally, we show that the disruption between core PRC2 subunits in Zfc3h1−/− cell extracts can be rescued upon RNase treatment (Figure 4F). These results echo previous data showing that RNase treatment increases chromatin association of PRC2 in cells and, reciprocally, recombinant PRC2 can be titrated off nucleosomes by increasing nuclear RNA levels (Beltran et al., 2016). We propose that this antagonism is a general effect of increased nuclear RNA, as we do not see specific enrichment of particular transcripts in RIP-seq experiments: RNAs that are upregulated in the Zfc3h1−/− inputs are also upregulated in the RIP data. This appears in line with previous findings, that the RNA-binding capability of PRC2 is non-specific and promiscuous in nature.

Although PRC2 has garnered considerable attention, RNA-binding capacities have also been reported for other chromatin regulators, including DNMTs, histone deacetylases (HDAC1), chromatin remodeling proteins (ATRX), DNA demethylases (TET1/2), and other histone methyltransferases (G9a) (Castellanos-Rubio et al., 2016; He et al., 2016; Hendrickson et al., 2016; Holz-Schietinger and Reich, 2012; Li et al., 2018; Di Ruscio et al., 2013; Sarma et al., 2014). Interestingly, RNA has been suggested to have a regulatory function in a number of models, either locally at specific loci or by more global mechanisms; that is, similar to PRC2, RNA binding has been proposed to sequester DNMT1 from transcriptionally active regions as a regulatory mechanism to prevent DNA methylation at these loci (Di Ruscio et al., 2013). Indeed, RIP experiments demonstrate that SUZ12 and DNMT1 both tend to associate with the 5′ ends of RNA (Hendrickson et al., 2016), and both proteins also show a greater affinity for RNA over DNA (Di Ruscio et al., 2013; Wang et al., 2017). RNA sequestration is also suggested in some disease models in which IncRNAs are overexpressed and affect chromatin modifiers through their abnormal titration (Gupta et al., 2010; Li et al., 2018; Merry et al., 2015; Prensner et al., 2013). Taken together with our results, this highlights the importance of maintaining a stable nuclear transcriptome through active RNA decay to prevent off-target effects as a result of RNA accumulation. Moreover, our results demonstrate the capacity of changed bulk RNA levels to affect cellular transcription programs. Although dysregulation of PRC2 on a global level is highlighted here, it is equally feasible that an unbalanced transcriptome might affect the function of other bivalent chromatin/RNA-binding proteins either locally or globally.
STAR METHODS

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

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - mES cell culture and differentiation
- METHOD DETAILS
  - CRISPR/Cas9 KOs
  - RNA isolation
  - pA+ RNA purification
  - qRT-PCR analysis
  - RNA-seq library preparation
  - Western blotting analysis
  - IP experiments
  - Glycerol gradient sedimentation analysis
  - RNaseA treatment
  - Subcellular fractionation
  - RIP experiments
  - RIP-seq
  - ChIP experiments
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Processing and analysis of RNA-seq data
  - Processing and analysis of RIP-seq data
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.10.011.

ACKNOWLEDGMENTS

We thank Nadia Laurs Schmidt, Dorthe Riishøj, and Anna Fossum for excellent technical assistance; Manfred Schmid and Peter Refsing Andersen for critical comments; and members of the T.H.J., K.H., and A.S. groups for advice and technical assistance; Manfred Schmid and Peter Refsing Andersen for critical comments.

The authors declare no competing interests.

REFERENCES


is implicated in the initial Xist-induced targeting of PRC2 to the inactive X chromosome. Mol. Cell 53, 301–316.


### STAR★METHODS

#### KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSRC2 (ZFC3H1) antibody</td>
<td>Novus Biologicals</td>
<td>Cat# NB100-68267; RRID:AB_2218116</td>
</tr>
<tr>
<td>ZFC3H1 antibody</td>
<td>Sigma-Aldrich</td>
<td>Cat# HPA-007151; RRID:AB_1846133</td>
</tr>
<tr>
<td>MTR4 (SKIV2L2) antibody</td>
<td>Abcam</td>
<td>Cat# ab70551; RRID:AB_1270701</td>
</tr>
<tr>
<td>PABPN1 antibody</td>
<td>Abcam</td>
<td>Cat# ab75855; RRID:AB_1310538</td>
</tr>
<tr>
<td>EXOSC4 (RRP41) antibody</td>
<td>In house (Pruijn lab)</td>
<td>Brouwer et al., 2001</td>
</tr>
<tr>
<td>DIS3 antibody</td>
<td>Sigma-Aldrich</td>
<td>Cat# HPA039281; RRID:AB_10795583</td>
</tr>
<tr>
<td>ZCCHC8 antibody</td>
<td>Abcam</td>
<td>Cat# ab68739; RRID:AB_1271512</td>
</tr>
<tr>
<td>ACTB (β-ACTIN) antibody</td>
<td>Sigma-Aldrich</td>
<td>Cat# A2228; RRID:AB_476697</td>
</tr>
<tr>
<td>NANO4 antibody</td>
<td>Cell Signaling</td>
<td>Cat# 8822; RRID:AB_11217637</td>
</tr>
<tr>
<td>ESRRB antibody</td>
<td>R+D Systems</td>
<td>Cat# PP-H6705-00; RRID:AB_2100412</td>
</tr>
<tr>
<td>KLF2 antibody</td>
<td>Millipore</td>
<td>Cat# 09-820; RRID:AB_10807287</td>
</tr>
<tr>
<td>SOX2 antibody</td>
<td>Santa Cruz</td>
<td>Cat# sc-17320; RRID:AB_2286684</td>
</tr>
<tr>
<td>OCT4 antibody</td>
<td>Abcam</td>
<td>Cat# ab19857; RRID:AB_445175</td>
</tr>
<tr>
<td>VCL (VINCULIN) antibody</td>
<td>Sigma-Aldrich</td>
<td>Cat# V9131; RRID:AB_477629</td>
</tr>
<tr>
<td>SUZ12 antibody</td>
<td>Cell Signaling</td>
<td>Cat# 3737; RRID:AB_2196850</td>
</tr>
<tr>
<td>EZH2 antibody</td>
<td>Cell Signaling</td>
<td>Cat# 5246; RRID:AB_10694683</td>
</tr>
<tr>
<td>EZH2 antibody</td>
<td>In house (Helin lab)</td>
<td>AC22</td>
</tr>
<tr>
<td>EZH2 antibody</td>
<td>Abcam</td>
<td>Cat# ab195409</td>
</tr>
<tr>
<td>EED antibody</td>
<td>In house (Helin lab)</td>
<td>Bracken et al., 2003</td>
</tr>
<tr>
<td>SFPQ antibody</td>
<td>Abcam</td>
<td>Cat# ab50935; RRID:AB_882523</td>
</tr>
<tr>
<td>TUBA1B (α-TUBULIN) antibody</td>
<td>Rockland</td>
<td>Cat# 600-401-880; RRID:AB_2137000</td>
</tr>
<tr>
<td>RPBI NTD antibody</td>
<td>Cell Signaling</td>
<td>Cat# 14958; RRID:AB_2687876</td>
</tr>
<tr>
<td>H3K27me3 antibody</td>
<td>Cell Signaling</td>
<td>Cat# 9733; RRID:AB_2616029</td>
</tr>
<tr>
<td>H3K27ac antibody</td>
<td>Cell Signaling</td>
<td>Cat# 8173; RRID:AB_10949503</td>
</tr>
<tr>
<td>H3K4me1 antibody</td>
<td>Cell Signaling</td>
<td>Cat# 5326; RRID:AB_10695148</td>
</tr>
<tr>
<td>H3K4me3 antibody</td>
<td>Cell Signaling</td>
<td>Cat# 9751; RRID:AB_2616028</td>
</tr>
<tr>
<td>H3 antibody</td>
<td>Abcam</td>
<td>Cat# ab1791; RRID:AB_302613</td>
</tr>
<tr>
<td>H4 antibody</td>
<td>Millipore</td>
<td>Cat# 05858; RRID:AB_390138</td>
</tr>
<tr>
<td>IgG</td>
<td>Millipore</td>
<td>Cat# PP64; RRID:AB_97852</td>
</tr>
<tr>
<td><strong>Bacterial and Virus Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5a Chemically Competent Cells</td>
<td>Prepared in lab</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSK3 inhibitor (CHIR99021)</td>
<td>Sigma-Aldrich</td>
<td>Cat# SML1046</td>
</tr>
<tr>
<td>MEK1/2 inhibitor (PD0325901)</td>
<td>Sigma-Aldrich</td>
<td>Cat# PZ0162</td>
</tr>
<tr>
<td>N-2 Supplement</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 17502048</td>
</tr>
<tr>
<td>B-27 Supplement</td>
<td>Thermo-Fisher Scientific</td>
<td>Cat# 17504044</td>
</tr>
<tr>
<td>TRIZOL Reagent</td>
<td>Thermo-Fisher Scientific</td>
<td>Cat# 15596018</td>
</tr>
<tr>
<td>Lipofectamine 2000 Transfection Reagent</td>
<td>Thermo-Fisher Scientific</td>
<td>Cat# 11668019</td>
</tr>
<tr>
<td>Protein A Dynabeads</td>
<td>Thermo-Fisher Scientific</td>
<td>Cat# 10008D</td>
</tr>
<tr>
<td>Agencourt AMPure XP Beads</td>
<td>Thermo-Fisher Scientific</td>
<td>Cat# 10136224</td>
</tr>
<tr>
<td>Benzonase nuclease</td>
<td>Millipore</td>
<td>Cat# 70746</td>
</tr>
<tr>
<td>RNaseA</td>
<td>Thermo-Fisher Scientific</td>
<td>EN0531</td>
</tr>
</tbody>
</table>

(Continued on next page)
**LEAD CONTACT AND MATERIALS AVAILABILITY**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Torben Heick Jensen (thj@mbg.au.dk).

All unique/stable reagents generated in this study are available from the lead contact upon request.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

**mES cell culture and differentiation**

E14TG2a mouse ESCs (male genotype, XY) were cultured on 0.2% gelatin coated plates in 2i/LIF containing medium (1:1 mix of Neurobasal (GIBCO) and DMEM/F-12 (GIBCO) supplemented with 1x Pen-Strep (GIBCO), 2 mM Glutamax (GIBCO), 50 μM

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNeasy Mini Kit</td>
<td>QIAGEN</td>
<td>Cat# 74106</td>
</tr>
<tr>
<td>Dynabeads mRNA Purification Kit</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 61006</td>
</tr>
<tr>
<td>Taqman Reverse Transcription Reagents</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# N8080234</td>
</tr>
<tr>
<td>Truseq Stranded Total RNA Library Prep Gold Kit</td>
<td>Illumina</td>
<td>Cat# 20020598</td>
</tr>
<tr>
<td>Subcellular Protein Fractionation Kit for Cultured Cells</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 78840</td>
</tr>
<tr>
<td>NEBNext Ultra II DNA Library Prep Kit for Illumina</td>
<td>NEB</td>
<td>Cat# E7645S</td>
</tr>
<tr>
<td>EzH1&lt;sup&gt;−/−&lt;/sup&gt;/EzH2&lt;sup&gt;−/−&lt;/sup&gt; RNA-seq data</td>
<td>Gene Expression Omnibus (GEO)</td>
<td>GSE103685</td>
</tr>
<tr>
<td>Deposited RNA-seq, ChIP-seq, RIP-seq datasets</td>
<td>This study</td>
<td>GSE137491</td>
</tr>
<tr>
<td>ES-E14TG2a cell line</td>
<td>ATCC</td>
<td>N/A</td>
</tr>
<tr>
<td>ES-E14TG2a Zfc3h1&lt;sup&gt;−/−&lt;/sup&gt; #1</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>ES-E14TG2a Zfc3h1&lt;sup&gt;−/−&lt;/sup&gt; #2</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>ES-E14TG2a Zfc3h1&lt;sup&gt;−/−&lt;/sup&gt; #3</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>sgRNA primers</td>
<td>See Table S1</td>
<td>N/A</td>
</tr>
<tr>
<td>qRT-PCR primers</td>
<td>See Table S2</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Software and Algorithms**

<table>
<thead>
<tr>
<th>Name (Version)</th>
<th>Source</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ImageJ (1.51h)</td>
<td>NIH, Univ. of Wisc. Madison</td>
<td><a href="https://imagej.nih.gov/ij/">https://imagej.nih.gov/ij/</a></td>
</tr>
<tr>
<td>IGV (2.4.14)</td>
<td>Robinson et al., 2011</td>
<td><a href="https://software.broadinstitute.org/software/igv/">https://software.broadinstitute.org/software/igv/</a></td>
</tr>
<tr>
<td>cellSens Entry (1.16)</td>
<td>Olympus</td>
<td><a href="https://www.olympus-lifescience.com/en/software/cellsens/">https://www.olympus-lifescience.com/en/software/cellsens/</a></td>
</tr>
<tr>
<td>Graphpad Prism (7.0a)</td>
<td>Graphpad</td>
<td><a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a></td>
</tr>
<tr>
<td>HISAT (0.1.6.beta)</td>
<td>Kim et al., 2015</td>
<td><a href="http://www.ccb.jhu.edu/software/hasit/index.shtml">http://www.ccb.jhu.edu/software/hasit/index.shtml</a></td>
</tr>
<tr>
<td>bedtools (2.23.0)</td>
<td>Quinlan and Hall., 2010</td>
<td><a href="https://bedtools.readthedocs.io/en/latest/">https://bedtools.readthedocs.io/en/latest/</a></td>
</tr>
<tr>
<td>Rsubread (1.32.1)</td>
<td>Liao et al., 2013</td>
<td><a href="https://bioconductor.org/packages/release/bioc/html/Rsubread.html">https://bioconductor.org/packages/release/bioc/html/Rsubread.html</a></td>
</tr>
<tr>
<td>Trim Galore (0.4.4)</td>
<td>Babraham Institute</td>
<td><a href="https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/">https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/</a></td>
</tr>
<tr>
<td>EaSeq (1.111)</td>
<td>Lerdrup et al., 2016</td>
<td><a href="https://easeq.net/">https://easeq.net/</a></td>
</tr>
<tr>
<td>SAMtools (1.6.1)</td>
<td>Li et al., 2009</td>
<td><a href="http://samtools.sourceforge.net/">http://samtools.sourceforge.net/</a></td>
</tr>
<tr>
<td>deepTools (2.5.3)</td>
<td>Ramirez et al., 2014</td>
<td><a href="https://deeptools.readthedocs.io/en/develop/">https://deeptools.readthedocs.io/en/develop/</a></td>
</tr>
<tr>
<td>MACS2 (2.1.1.20160309)</td>
<td>Zhang et al., 2008</td>
<td><a href="https://github.com/taoliu/MACS/">https://github.com/taoliu/MACS/</a></td>
</tr>
</tbody>
</table>
β-mercaptoethanol (GIBCO), 0.1 mM nonessential amino acids (GIBCO), 1 mM sodium pyruvate (GIBCO), 0.5x N2 supplement (GIBCO), 0.5x B27 supplement (GIBCO), 3 μg GSK3i (CHIR99021), 1 μM MEKi (PD0325901) and Leukemia Inhibitory Factor (LIF; produced in house)). Cells were passaged every 2-3 days by aspirating medium, dissociating cells with 0.25% trypsin-EDTA (GIBCO) briefly at 37°C before the addition of an equal volume of 1x trypsin inhibitor (Sigma) and gentle disruption by pipetting. Cells were pelleted by centrifugation, washed in 2i-LIF to remove excess trypsin and pelleted again before resuspending and plating ~1x10⁶ cells/10 cm plate.

For differentiation into EBs, dissociated cells were washed 2x in Serum-LIF (GMEM (Sigma) supplemented with 10% heat-inactivated fetal bovine serum, 1x Pen-Strep (GIBCO), 2 mM Glutamax (GIBCO), 0.1 mM nonessential amino acids (GIBCO), 1 mM sodium pyruvate (GIBCO), 50 μM β-mercaptoethanol (GIBCO)) before seeding 1.5x10⁶ cells into 10 cm Petri-dishes containing Serum-LIF media. Media was changed at days 2, 3, 5 and 7. At day 7, EBs were transferred to 0.2% gelatin coated plates and grown for a further 3 days.

Phase contrast microscopy images were captured using an Olympus IX73 inverted microscope using the cellSens Entry software (Olympus).

**METHOD DETAILS**

**CRISPR/Cas9 KOs**
KO cell lines were generated by CRISPR/Cas9 targeting of Zfc3h1 in WT ESC. Single guide (sg) RNAs (Table S1) were cloned into the pSPCas9(BB)-2A-GFP vector (pX458, Addgene plasmid ID: 48138) as previously described (Ran et al., 2013) and transfected into ES cells using Lipofectamine 2000 (Thermo). Single cell clones were isolated by GFP sorting using FACS into 0.2% gelatin coated 96 well plates containing 2i/LIF and expanded. KO clones were screened by western blotting analysis and validated by Sanger sequencing of amplified genomic DNA around the cut site. Three independent Zfc3h1−/− cell lines were derived from expanded single cell clones.

**RNA isolation**
Total RNA was isolated using the RNeasy Mini Kit (QIAGEN) according to the manufacturer’s instructions or by Trizol extraction (Thermo) using the standard protocol. For chromatin associated RNA, samples were prepared as previous described (Conrad and Ørom, 2017).

**pA+ RNA purification**
pA+ RNA was isolated from nuclear RNA samples using the Dynabeads mRNA Purification Kit (Thermo). For isolation of nuclei, 2x10^⁷ cells were resuspended in nuclear isolation buffer (NIB) (10 mM Tris pH 7.4, 150 mM NaCl, 0.15% Igepal CA-630) supplemented with protease inhibitors and lysed at 4°C on a rotating wheel for 5 minutes. Lysates were overlaid onto 1 mL Sucrose buffer (10 mM Tris pH 7.4, 150 mM NaCl, 24% sucrose) in a DNA LoBind tube (Eppendorf) and nuclei were pelleted for 10 minutes at 2000 x g. Nuclei were resuspended in 1 mL Trizol (Thermo) and RNA was extracted using the standard protocol. 50 μg of nuclear RNA extracts were heated to 65°C and cooled on ice before incubating with oligo dT(25) Dynabeads (Thermo). Bead complexes were washed twice before elution in 10 mM Tris pH 7.5 and recovered RNA were assessed using a NanoDrop Lite Spectrophotometer (Thermo).

**qRT-PCR analysis**
cDNA was prepared from 500 ng of total RNA with TaqMan Reverse Transcription reagents (Thermo) using random hexamers. qRT-PCR was performed using the LightCycler 480 SYBR Green I (Roche) in technical triplicates. Primers used in qRT-PCR are listed in Table S2.

**RNA-seq library preparation**
RNA-seq libraries were prepared from 1 μg of total RNA using the TruSeq Stranded Total RNA library prep kit with RiboZero Gold (Illumina) according to the manufacturer’s instructions. Three biological replicates from each sample were prepared. RNA integrity and library quality were assessed on a Bioanalyzer 2000 using RNA Nano and DNA 1000 chips (Agilent), respectively. Libraries were quantified and normalized for multiplexing using the KAPA library quantification Kit for Illumina (KAPA Biosystems) and sequenced on an Illumina NextSeq 550 (75-bp, paired-end).

**Western blotting analysis**
Protein lysates were prepared using TOPEX+ buffer (Riising et al., 2014) (300 mM NaCl, 50 mM Tris-HCl pH 7.5, 0.5% Triton X-100, 1% SDS) freshly supplemented with protease inhibitors, 1 mM DTT and 33.3 U/ml Benzonase (Novagen). SDS-PAGE and western blotting analysis were carried out according to standard protocols with the antibodies listed in the Key Resources Table and HRP conjugated secondary antibodies (Vector Laboratories and Agilent). Bands were visualized by Super Signal West Pico chemiluminescent ECL (Thermo) and exposed either on Amersham Hyperfilm ECL films (GE Healthcare) and developed (Ferrania Imagine Technologies) or digitally captured using an Amersham Image 600 (GE Healthcare). Images were processed and quantified using ImageJ (Schneider et al., 2012).
**IP experiments**

For whole cell IPs, 1x10^7 cells/IP were resuspended in HT150 extraction buffer (20 mM HEPES pH 7.4, 150 mM NaCl, 0.5% Triton X-100) supplemented with protease inhibitors and sheared mechanically using 22 G and 27 G needles sequentially with 6 strokes each. For nuclear IPs, 1x10^7 cells/IP were resuspended in nuclear isolation buffer (NIB) (10 mM Tris pH 7.4, 150 mM NaCl, 0.15% Igepal CA-630) supplemented with protease inhibitors and lysed at 4°C on a rotating wheel for 5 minutes. Lysates were overlaid onto 1 mL Sucrose buffer (10 mM Tris pH 7.4, 150 mM NaCl, 24% sucrose) in a DNA LoBind tube (Eppendorf) and nuclei were pelleted for 10 minutes at 2000 x g. Nuclei were resuspended in 250 µl/IP RNA isolation buffer (RIB) (25 mM Tris pH 7.4, 150 mM KCl, 0.5 mM DTT, 0.5% Igepal CA-630) supplemented with protease inhibitors and sheared mechanically using 22 G and 27 G needles sequentially with 6 strokes each.

Clariﬁed lysates were treated with DNasel (Thermo) for 20 minutes at 37°C before pre-clearing with rabbit IgG (Millipore) and Protein-A Dynabeads (Thermo) for 2 hours at 4°C. Supernatants were incubated with either IgG, SUZ12 (Cell Signaling) or EZH2 (made in house) antibodies overnight at 4°C with Protein-A Dynabeads. Beads were washed 3 times with the respective extraction buffer, transferring beads to a fresh tube on the final wash. Proteins were eluted by boiling in 1X NuPAGE loading buffer (Invitrogen). 10X reducing agent (Invitrogen) was added to the supernatants before denaturing for 10 minutes at 95°C and proceeding with western blotting analysis.

**Glycerol gradient sedimentation analysis**

The glycerol gradient sedimentation analysis was performed as previously described (Chu et al., 2014), with minor modifications. Briefly, whole cell extracts from ~2x10^10 cells were resuspended in BC100 buffer (5 mM HEPES pH 7.5, 100 mM NaCL, 1 mM MgCl₂, 0.5 mM EGTA, 0.1 mM EDTA, 10% v/v glycerol, 1 mM DTT) supplemented with protease inhibitors, lysed by sonication (3 x 5 s, amplitude 2) and centrifuged at 14,000 rpm for 20 minutes. Clarified lysates were loaded on 10%–50% (v/v) glycerol gradients prepared in BC100 buffer and centrifuged at 35,000 rpm for 28 hours using a SW41 rotor (Beckman). Gradients were separated into 18 fractions and protein content was assessed using the Bradford assay (Bio-Rad). Remaining fractions were TCA precipitated, acetone washed and resuspended in 1x NuPAGE loading buffer (Thermo). Samples were separated by SDS-PAGE and analyzed either by western blotting analysis or using the Blue Silver modiﬁed Neuhoﬀ’s colloidal Coomassie Blue G-250 stain (Candiano et al., 2004).

**RNaseA treatment**

Cells were treated with RNaseA as previously described (Beltrán et al., 2016). Cells were trypsinised and permeabilised with 0.05% Tween-20 (Sigma) in PBS for 10 minutes on ice. Cells were washed once, resuspended in PBS and either mock-treated or treated with 1 mg ml⁻¹ RNaseA (Thermo) for 30 minutes at RT with gentle agitation. Cells were washed twice with PBS before proceeding with lysis and IP.

**Subcellular fractionation**

Cells were separated into cytoplasmic, nuclear and chromatin fractions using the Subcellular Protein Fractionation Kit (Pierce) according to the manufacturer’s instructions. Cells were harvested in 1x10^7 aliquots, split into 2 samples for either whole cell extraction using TOPEX+ buffer or fractionation. Equivalent lysate volumes were separated by SDS-PAGE and analyzed by western blotting.

**RIP experiments**

Whole cell IPs were performed as described above with alterations. All buffers were additionally supplemented with 100 U/ml Ribolock RNase Inhibitor (Thermo). Following overnight IP with IgG, SUZ12 or EZH2 antibodies, beads were washed 3 times with HT150 buffer, transferring beads to a fresh tube on the final wash. RNA was isolated from the IPs by the addition of 1 mL Trizol, homogenization and incubating for 1 hour on ice. Beads were removed magnetically and RNA was isolated from Trizol using the standard protocol.

**RIP-seq**

Strand speciﬁc libraries were prepared from ~200 ng RNA isolated in SUZ12 and EZH2 RIP experiments by BGI Tech Solutions (Europe) according to their IncRNA-seq library preparation protocol. Two biological replicates were prepared from each sample. RNA integrity was assessed using a BioAnalyzer 2000 (Agilent) using RNA Nano chips. Samples were ribodepleted using Ribo-Zero (Illumina) and libraries prepared using the TruSeq Stranded mRNA library prep kit (Illumina). Libraries were sequenced on a BGISEQ-500 (100 bp, paired end).

**ChIP experiments**

ChIP experiments were carried out according to standard protocols. Briefly, ES cells were cross-linked by the addition of 1% formaldehyde (Sigma) in the dish for 10 minutes at RT before quenching with glycine. DNA was sheared to ~200 bp fragments by sonication using a Biorupter (Diagenode) and validated by agarose gel electrophoresis. ChIPs were carried out using 200 µg of chromatin and 2-3 µg of the indicated antibodies (Key Resources Table). Libraries for ChIP-seq were prepared using NEBNext Ultra II DNA Library prep kit (NEB) using AmpureXP beads (Beckman) for size selection. Libraries were assessed on a Bioanalyzer 2000 (Agilent).
using High Sensitivity DNA chips and quantified using the Qubit dsDNA HS assay kit (Thermo). Libraries were sequencing on an Illumina NextSeq 550 (75 bp, single end).

**QUANTIFICATION AND STATISTICAL ANALYSIS**

**Processing and analysis of RNA-seq data**

Quality control of sequence reads was done using FastQC v0.11.2 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Illumina adaptors (as provided with the FastQC tool), low quality bases, the first 12 bases and reads shorter than 25 nt were removed with Trimmomatic v0.32, using settings “ILLUMINACLIP:<TrueSeq3_PE_2>:2:30:10 HEADCROP:12 LEADING:22 SLIDING WINDOW:4:22 MINLEN:25” (Bolger et al., 2014). Both paired and unpaired (due to the trimming) reads were mapped using HISAT v0.1.6.beta (Kim et al., 2015), against the mouse genome (mm10), where a list of GENCODE M12 (Frankish et al., 2019) annotated splice sites was also provided. HISAT was furthermore run with maximum fragment length set to 1000 and the--rf parameter (for the upstream/downstream mate orientation), otherwise default settings was used.

To generate visualization, properly paired reads mapping onto unique genomic locations were selected. Genomc cov from bedtools v2.23.0 (Quinlan and Hall, 2010) was used to calculate strand-specific per-base genome coverage in bedgraph format. Bedgraph files were converted into bigwig format for using the UCSC Genome Browser Utility ‘bedGraphToBigWig’ (Kent et al., 2010). Finally, the per base coverage was CPM normalized. Genome browser images are generated from IGV (Robinson et al., 2011; Thorvaldsdóttir et al., 2013).

Paired and uniquely mapped exonic reads for GENCODE M12 genes were counted using featureCounts from the R package Rsubread (1.32.1) (Liao et al., 2013) for individual RNA-seq libraries. RNA-seq read counts for PRC2 KO (Ezh1−/−/Ezh2−/−) and corresponding WT samples were obtained from the Gene Expression Omnibus (GEO), accession number: GSE103685 (Højfeldt et al., 2018). Differential expression analysis was performed using the R package edgeR (version 3.24.1) with default parameters (McCarthy et al., 2012; Robinson et al., 2010). Intronic regions of a gene were defined as regions in the gene body that do not overlap with an exon from any GENCODE M12 annotated transcript isoform. Intronic reads were counted and differential expression analysis was performed using the same method as for exonic reads.

**Processing and analysis of RIP-seq data**

Quality control of sequence reads was done using FastQC v0.11.2. Illumina adaptors, low-quality bases with Phred score lower than 20, and reads shorter than 25 bp were removed using Trim Galore (version: 0.4.4, https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/); reads were further trimmed using parameters--clip_R1 13–clip_R2 13–three_prime_clip_R1 1–three_prime_clip_R2 1. Trimmed reads were mapped as described for RNA-seq data processing. BAM alignment files were further processed to select uniquely mapped and properly paired reads using SAMtools (version 1.6.1) (Li et al., 2009). Duplicated reads were removed using MarkDuplicates (version: 2.8.1) from GATK (McKenna et al., 2010) with default settings and processed BAM files were used for downstream analysis. The strand specific genomic coverage was calculated using bamCoverage from deepTools (Ramírez et al., 2014). Raw reads for GENCODE M12 genes were counted using the same method as for RNA-seq, and differential analysis was performed using edgeR. To calculate the enrichment of transcripts pulled down by SUZ12 and EZH2 IP among Zfc3h1−/− compared to WT, contrast in the differential analysis was made as (Zfc3h1−/− IP – Zfc3h1−/− input) – (WT IP – WT input) and the enrichment score was defined as log2 fold change calculated from the contrast.

**Processing and analysis of ChiP-seq data**

Reads were mapped onto the mouse (mm10) genome with Bowtie (Langmead and Salzberg, 2012), selecting only hits with the best stratum with up to two mismatches in the seed and reporting up to four good alignments per read. Only two copies of identical reads were kept. The fragment sizes could be deduced using the ChiP-Cor tool (Ambrosini et al., 2016). Reads were shifted to the center of their fragment by half of the deduced fragment size. Genomc cov from bedtools v2.23.0 (Quinlan and Hall, 2010) was used to calculate strand specific per base genome coverage in bedgraph format. Bedgraph files were converted into bigwig format for using the UCSC Genome Browser Utility ‘bedGraphToBigWig’ (Kent et al., 2010). Finally, the per base coverage was CPM normalized.

For global MA plots, ChIP signals from all libraries were quantified on genomic windows of 2 kb sliding by 500 bp using the UCSC Genome Browser Utility bigWigAverageOverBed (Kent et al., 2010). SUZ12 peaks were called using MACS2 (Zhang et al., 2008) (version 2.1.1.20160309) with parameters--qvalue 0.05–broad–broad-cutoff 0.3. ENCODE blacklisted peaks (ENCODE Project Consortium, 2012) and low quality peaks (-log10(qvalue) ≤ 1) were filtered out. Consistent SUZ12 peaks between Zfc3h1−/− replicates were defined as peaks overlapping in at least two replicates. A single SUZ12 reference peak set for WT and Zfc3h1−/− was obtained by pooling SUZ12 peaks of WT and consistent SUZ12 peaks of Zfc3h1−/− and merging overlapping peak regions into a single region using mergeBed from bedtools (v2.23.0). Raw reads of H3K27me3 and SUZ12, for SUZ12 peaks in the reference peak set, were obtained using featureCounts from the R package Rsubread (1.32.1), a pseudocount of 1 was added when normalizing raw read counts to the library size. log2FC values were calculated between mean values of the normalized read counts from Zfc3h1−/− replicates and the normalized read counts from WT. For mean signal plots, genomic coverage from bigWig files was calculated using computeMatrix from deepTools (version 2.5.3) (Ramírez et al., 2014). Regions without read coverage in bigWig files were treated as 0, mean values from the replicates were calculated and mean values at each position were plotted. For the heatmap, a list of genes
overlapping with at least one SUZ12 peak from the reference peak set was obtained. H3K27me3 and SUZ12 were quantified in the gene bodies, log₂FC values of H3K27me3 and SUZ12 between Zfc3h1−/− and WT were calculated using the method described above, and log₂FC values of gene exonic expression were computed from the differential expression analysis of RNA-seq described above.

**DATA AND CODE AVAILABILITY**

All high-throughput RNA-seq, ChIP-seq and RIP-seq datasets generated during this study are available at the Gene Expression Omnibus (GEO) under accession code GSE137491.