UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Loss of Ezh2 function remodels the DNA replication initiation landscape

Prorok, Paulina; Forouzanfar, Faezeh; Murugarren, Nerea; Peiffer, Isabelle; Charton, Romain; Akerman, Ildem; Méchali, Marcel

DOI:
10.1016/j.celrep.2023.112280

License:
Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Prorok, P, Forouzanfar, F, Murugarren, N, Peiffer, I, Charton, R, Akerman, I & Méchali, M 2023, 'Loss of Ezh2 function remodels the DNA replication initiation landscape', *Cell Reports*, vol. 42, no. 4, 112280. https://doi.org/10.1016/j.celrep.2023.112280

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

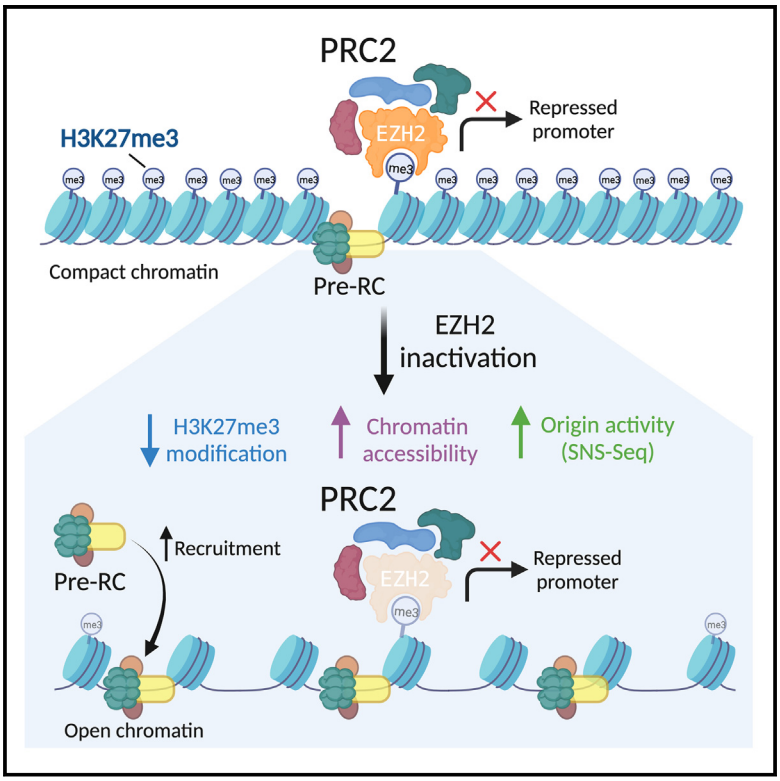
If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 29. Apr. 2024

Cell Reports

Loss of Ezh2 function remodels the DNA replication initiation landscape

Graphical abstract



Highlights

- Polycomb-repressed promoters are hotspots for DNA replication initiation
- *EZH2*$^{-/-}$ cells display a substantial increase in replication initiation at repressed promoters
- Increased replication initiation does not correlate with *de novo* H3K27ac or transcription
- Instead, replication initiates at sites of H3K27me3 loss and increased open chromatin

Authors

Paulina Prorok, Faezeh Forouzanfar, Nerea Murugarren, Isabelle Peiffer, Romain Charton, Ildem Akerman, Marcel Méchali

Correspondence

paulina.prorok@tu-darmstadt.de (P.P.), i.akerman@bham.ac.uk (I.A.), marcel.mechali@igh.cnrs.fr (M.M.)

In brief

Most promoters repressed by the Polycomb group of proteins also harbor DNA replication initiation in pluripotent stem cells. Prorok et al. find that EZH2 loss of function leads to marked increase in DNA replication initiation at these repressed promoters with loss of H3K27me3, irrespective of the activating histone H3K27ac mark or transcription.

Prorok et al., 2023, Cell Reports *42*, 112280
April 25, 2023 © 2023
https://doi.org/10.1016/j.celrep.2023.112280

Cell Reports

CellPress
OPEN ACCESS

Report

Loss of Ezh2 function remodels the DNA replication initiation landscape

Paulina Prorok,[1,4,5,*] Faezeh Forouzanfar,[1,5] Nerea Murugarren,[2,3,5] Isabelle Peiffer,[1] Romain Charton,[1] Ildem Akerman,[2,3,6,*] and Marcel Méchali[1,7,*]

[1]Institute of Human Genetics, CNRS-University of Montpellier, Montpellier 34090, France
[2]Institute of Metabolism and Systems Research (IMSR), University of Birmingham, Birmingham B152TT, UK
[3]Centre for Endocrinology, Diabetes and Metabolism, Birmingham Health Partners, Birmingham B152TT, UK
[4]Present address: Department of Biology, Technical University of Darmstadt, Darmstadt 64287, Germany
[5]These authors contributed equally
[6]Twitter: @ildemakerman
[7]Lead contact
*Correspondence: paulina.prorok@tu-darmstadt.de (P.P.), i.akerman@bham.ac.uk (I.A.), marcel.mechali@igh.cnrs.fr (M.M.)
https://doi.org/10.1016/j.celrep.2023.112280

SUMMARY

In metazoan cells, DNA replication initiates from thousands of genomic loci scattered throughout the genome called DNA replication origins. Origins are strongly associated with euchromatin, particularly open genomic regions such as promoters and enhancers. However, over a third of transcriptionally silent genes are associated with DNA replication initiation. Most of these genes are bound and repressed by the Polycomb repressive complex-2 (PRC2) through the repressive H3K27me3 mark. This is the strongest overlap observed for a chromatin regulator with replication origin activity. Here, we asked whether Polycomb-mediated gene repression is functionally involved in recruiting DNA replication origins to transcriptionally silent genes. We show that the absence of EZH2, the catalytic subunit of PRC2, results in increased DNA replication initiation, specifically in the vicinity of EZH2 binding sites. The increase in DNA replication initiation does not correlate with transcriptional de-repression or the acquisition of activating histone marks but does correlate with loss of H3K27me3 from bivalent promoters.

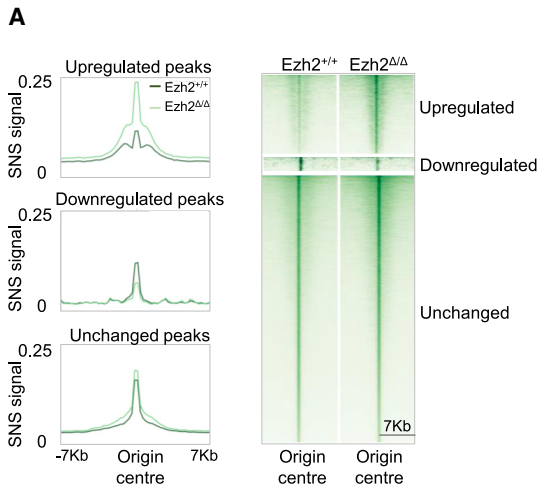
INTRODUCTION

In mammalian cells, genome duplication starts at replication initiation sites termed replication origins that are activated in a defined temporal order during each cell cycle. They are localized at specific sites, but their usage is flexible.[1–3] Thus, in a human or mouse cell, around 100,000 potential origins are present, but only a fraction of them (about 30%) are activated in each given cell (2–4), and their selection process remains poorly understood. Genome-wide mapping of origins did not reveal a strict DNA consensus sequence in metazoans, but some genetic and epigenetic (pertaining to chromatin and histone modifications) characteristics were identified. Among these features, the presence of G-rich elements, such as origin G-rich repeated elements (OGREs) localized upstream of initiation sites were found in more than 60% active origins in fly, mouse, and human cells, which can potentially form G-quadruplexes (G4s).[4–14]

Apart from sequence elements, independent assays have also associated DNA replication initiation with a set of chromatin marks that are mostly activating in nature.[7,15–17] However, we previously observed a strong association of polycomb group protein (PcG) binding sites with DNA replication origins in mouse embryonic stem cells (mESCs).[7] PcG proteins contribute to the regulation of pluripotency in mESCs and human ESCs by silencing differentiation-related genes.[18–20] Promoters repressed by the PcG proteins in mESCs retain bivalency (simultaneous presence of the repressive H3K27me3 and activating H3K4me3 histone marks), thereby retaining the ability to be activated as differentiation proceeds.[18] Two major PcG complex groups were identified: the Polycomb repressive complex 1 (PRC1) and PRC2. Depletion and/or mutation of PRC2 subunits leads to severe developmental defects at early stages of mouse embryogenesis.[18] A main component of PRC2 is the core subunit enhancer of zeste homolog 2 (EZH2), which has a Su(var) 3-9, Enhancer-of-zeste and Trithorax (SET) domain and contains the histone methyltransferase (HMTase) activity responsible for H3K27 trimethylation.[21]

Although EZH2 sites (around 4,000–10,000 in mESCs) represent a minor part of the total replication origins, 75%–90% of them are associated with replication origins[7] in mESCs. In contrast, only about half of H3K4me3 and a third of H3K27ac sites bear DNA replication initiation sites, highlighting that EZH2-bound regions are clearly hotspots of replication initiation (Figure S1A). Here, we used a transgenic inducible EZH2-deficient mESC line to functionally address how a PRC component may control the initiation of DNA replication. A stable genomic deletion of the EZH2 catalytic SET domain was generated by tamoxifen-induced Cre-Lox recombination[22] followed by genome-wide

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).



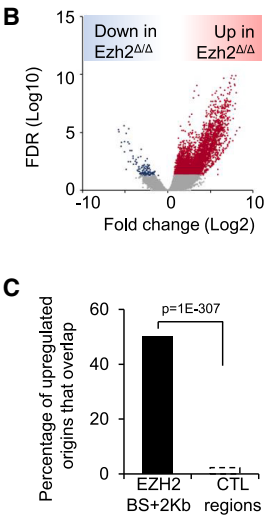


Figure 1. EZH2 KO induces upregulation of some DNA replication origins

(A) Left panel shows read profile plots of SNS-seq reads that fall on upregulated, downregulated, or unchanged origins in *Ezh2*$^{+/+}$ (dark green lines) and Ezh2$^{\Delta/\Delta}$ (light green lines) cells. Right panel shows corresponding heatmaps of read intensity of the same regions for *Ezh2*$^{+/+}$ (left) and Ezh2$^{\Delta/\Delta}$ (right) with the middle point of origins centered in the position zero. Both plots extend the origin center by 7 kb on each side. The y axis represents the signal intensity in counts per million (CPMs). n = 3, biological replicates (BRs).

(B) Volcano plot of differential activity analysis of SNS-seq samples from *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ cells. The y axis represents the false discovery rate (FDR; in log10), and the x axis represents fold change (in log2). In red are DNA replication origins that are statistically significantly upregulated in Ezh2$^{\Delta/\Delta}$ cells. n = 3, BRs.

(C) Bar plot shows the percentage of upregulated origins that overlap a consistent EZH2 binding site (black bar) within 2 kB. Dotted lines show the percentage of upregulated origins that overlap randomized control genomic regions of the same size and number as EZH2 binding sites (control [CTL] regions). p values obtained by chi-squared goodness-of-fit test.

profiling of replication initiation and transcriptional and epigenetic landscapes in these cells (Figure 1A). We show that loss of EZH2 activity reshapes DNA replication initiation in a specific manner at EZH2-bound repressed promoters.

RESULTS

Catalytically inactive EZH2 is degraded with no detectable effect on the cell cycle or checkpoint activation

In order to investigate the potential influence of PRC2 activity on DNA replication repertoire, we took advantage of a genetically engineered mESC line in which the catalytic domain of EZH2, SET, is permanently deleted by an inducible Cre-Lox recombination system[22] (Figure S1B). As previously reported,[22] tamoxifen-induced excision leads to efficient deletion of exons 14 and 15, generating a premature STOP codon prior to the SET domain coding exons of Ezh2 (Figure S1B). Exhaustive genomic deletion in the Ezh2 locus 48 h after addition of tamoxifen was confirmed (Figure S1C), resulting in truncated and degraded protein products at 96 h after the initial addition of tamoxifen (Figure S1D), which is the time point for all experimental assays.

Growth curves for *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ cells were similar (Figure S1E). Cell-cycle profile analysis by flow cytometry of cells pulsed with the thymidine analogue BrdU showed similar profiles and BrdU incorporation, reflecting similar percentages of cells in S phase (Figure S2A). These results confirm the lack of proliferation defects observed in a similar EZH2 loss-of-function cell line.[23] We did not observe changes in expression of Oct4 or other markers of cellular pluripotency (Figures S1D and S2B). Finally, we did not observe replication checkpoint activation in Ezh2$^{\Delta/\Delta}$ cells as observed in control cells treated with campthothecin, a topoisomerase 2 inhibitor (Figure S2C). We concluded that the deletion of the catalytic histone methylation activity of EZH2 did not affect cell growth and cell-cycle dynamics. However, as PcG sites represent a minor part of total potential replication origins, we asked whether the DNA replication initiation landscape is specifically impacted at direct EZH2 binding sites.

Loss of PRC activity leads to upregulation of DNA replication initiation

We identified DNA replication origins using RNA-primed nascent DNA analysis (short nascent strand sequencing [SNS-seq]) coupled with high-throughput sequencing in *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ mESCs from three independent replicates (Experimental Schematic, Figure S2D). Replication origins were called using two complementary software tools as previously described.[7] Our analysis identifies a total of 86,829 and 117,667 DNA replication initiation sites (ISs) in *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ mESC lines, respectively (Table S1).

We next analyzed the differential activity of DNA replication origins between SNS-seq samples from *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ cells using DiffBind, a computational tool that identifies statistically different accumulation of sequencing reads (Table S2). Comparison of DNA replication activity in *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ cells revealed statistically significant upregulation of activity in ~5,000 regions (false discovery rate [FDR] < 0.05, 7% of all origins), with the remaining origins either downregulated (0.2%) or unchanged (~93%) (Figures 1A and 1B). Figure 1A shows a heatmap of replication initiation activity, revealing a substantial upregulation of activity in the origins statistically upregulated in the Ezh2$^{\Delta/\Delta}$ cell category. Taken together, our data suggest that loss of Ezh2 function does not lead to detectable effects on DNA synthesis but to a predominant upregulation of activity for a select group of ~5,000 DNA replication origins.

Increased origin activity in Ezh2$^{\Delta/\Delta}$ cells correlates with EZH2 binding sites

We asked whether the observed changes in DNA replication activity was indeed related to EZH2 binding sites. Regions that reproducibly display an enrichment in two separate EZH2 chromatin immunoprecipitation (ChIP)-seq datasets[24,25] in mESCs were used as consistent binding sites for EZH2 (Table S1). Half

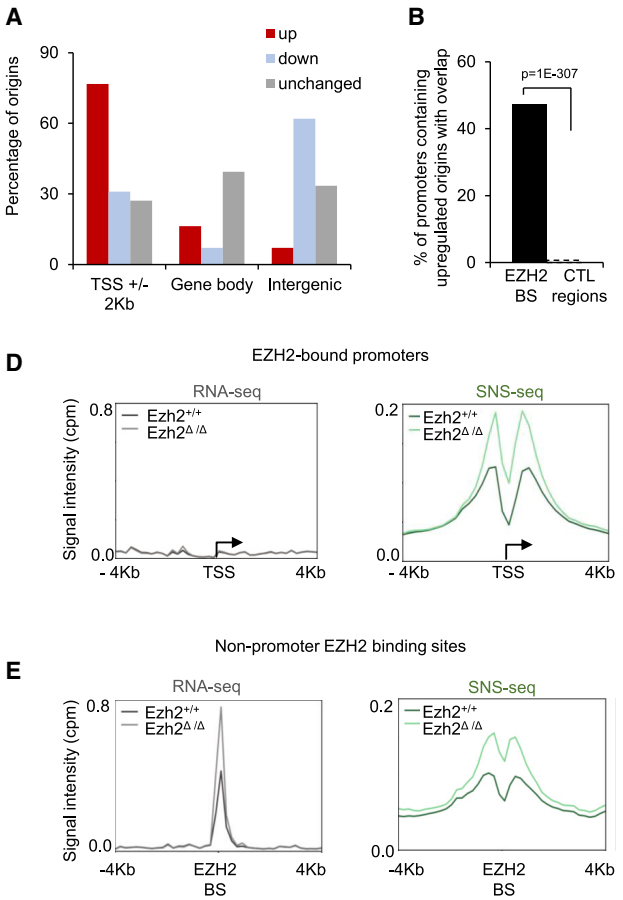


Figure 2. Increased DNA replication activity does not correlate with transcriptional activation

(A) Bar plots show overlaps of different classes of origins (upregulated in red, downregulated in blue, unchanged in gray) with ENSEMBL promoter regions (TSS ± 2 Kb), gene body (excluding TSS ± 2 Kb) and intergenic regions (excluding TSS ± 2 Kb).

(B) Bar plots showing the percentage of promoter regions (ENSEMBL gene TSS ± 2 Kb) containing upregulated origins that overlap EZH2 binding sites or randomized control genomic regions of the same size and number as EZH2 binding sites (CTL regions). p value obtained by chi-squared goodness-of-fit test.

(C) Pie chart shows the percentage of promoters containing upregulated origins that are transcriptionally repressed (TPM < 3).

(D) Read profile plots of RNA sequencing (RNA-seq; left panel, in gray) or SNS-seq (right panel, in green) for EZH2-bound promoters. The plots are centered at TSS and normalized by CPMs.

(E) Read profile plots of RNA-seq (left panel, in gray) or SNS-seq (right panel, in green) for EZH2-bound regions that do not overlap promoters. The plots are centered at the middle of the EZH2 peaks and normalized by CPMs.

In this figure, both SNS-seq and RNA-seq have n = 3 BRs, and lines represent merged read pileups.

of all upregulated SNS-seq peaks were found within 2 kb of consistent binding sites for EZH2, and this figure was 80% when we considered any EZH2 binding site (see STAR Methods; Figures 1C and S3A). We concluded that Ezh2$^{\Delta/\Delta}$ cells display an increase in replication initiation, specifically at EZH2 binding sites. This result led us to ask whether epigenetic marks deposited by EZH2 and/or the transcriptional program controlled by EZH2 is involved in this increase in replication origin activity.

Increased origin activity in Ezh2 knockout cells is not linked to a de-repression of transcription at EZH2 promoters

Approximately 77% of upregulated SNS-seq peaks reside within ENSEMBL gene promoters (transcription start site [TSS] ± 2 Kb) (Figure 2A, in red). This is a significant enrichment over origins that are downregulated or remain unchanged or a control set with a similar activity level (Figures 2A and S3B). This result confirmed that DNA replication initiation was specifically upregulated around promoter regions, and strikingly, a substantial portion of these promoters also contain consistent EZH2 binding sites (47%; Figure 2B) and are repressed (Figure 2C).

We next investigated whether upregulated origins are positioned at repressed gene promoters because they become transcriptionally active upon EZH2 knockout. To do this, we performed differential gene expression analysis on three replicates of RNA-sequencing samples obtained from *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ cells to identify genes that increase in transcriptional activity (Figure S2D). As previously reported,[22] Ezh2$^{\Delta/\Delta}$ cells display mild transcriptional dysregulation compared with *Ezh2*$^{+/+}$ cells (Table S2, 524 up and 526 down genes at $p < 0.05$). The overwhelming majority (>96%, ~7,000 genes; Table S1) of genes that are bound by EZH2 at their promoters remain transcriptionally silent (Figures 2D, left panel, and S3C) but still experience an upregulation of DNA replication activity (Figure 2D, right panel). We conclude that EZH2-bound gene promoters remain transcriptionally repressed following loss of function of EZH2 but are associated with upregulation of TSS-associated replication origin activity.

We also noted that a small minority of EZH2-bound sites do not reside at known gene promoters (~1,900 regions; Table S1). Upon loss of function of EZH2, these specific sites experience an increase in non-genic transcriptional output, which corresponds to an increase in DNA replication initiation activity (Figure 2E). PRC is known to bind and repress promoters, but PRC activity is also associated with inactive enhancer regions,[26] also known as poised enhancers. Approximately 30% of these EZH2-bound non-promoter regions represent poised enhancers that were previously defined[27] (Figure S3D).

We conclude that the prominent increase in DNA replication initiation activity at EZH2-bound promoters in Ezh2$^{\Delta/\Delta}$ cells is not linked to transcriptional de-repression. However, *de novo* transcription from EZH2-bound non-promoter sites, possibly enhancers, is associated with increased DNA replication initiation activity.

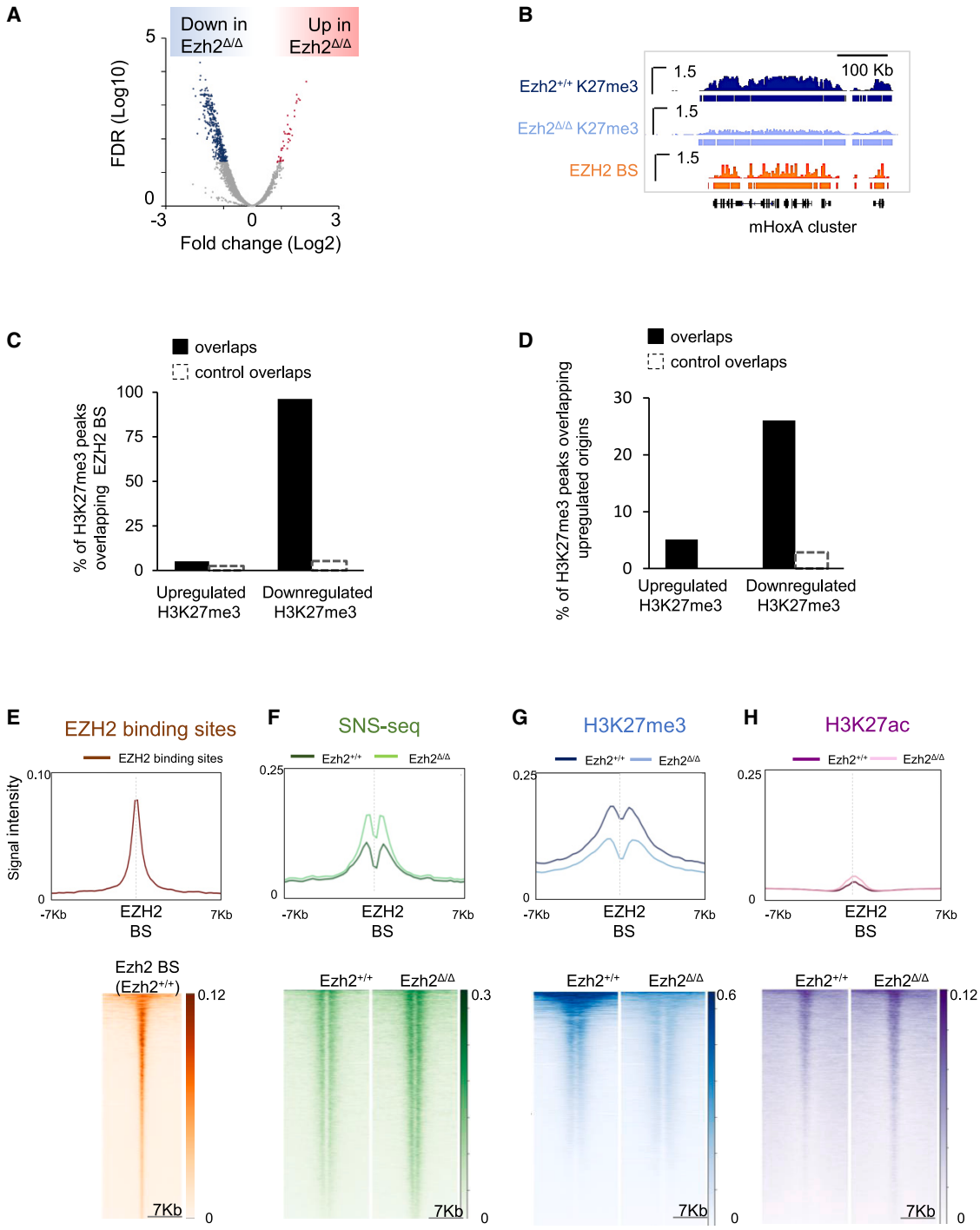


Figure 3. Loss of H3K27me3 is correlated with an increase in DNA replication initiation activity

(A) Volcano plot of differential activity analysis of H4K27me3 ChIP-seq samples from *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ cells. The y axis represents the FDR (in log10), and the x axis represents fold change (in log2). In blue are H4K27me3 regions that are statistically significantly decreased/downregulated in Ezh2$^{\Delta/\Delta}$ cells. n = 3, BRs.

(B) Chromatin landscape at a known PRC target, the HoxA cluster. IGV genome browser snapshots of merged read pileups from H3K27me3 ChIP-seq experiments (n = 3) from *Ezh2*$^{+/+}$ (navy) and Ezh2$^{\Delta/\Delta}$ (light blue) cells. Also shown are EZH2 binding sites in mESCs (orange). The y axis unit is CPMs.

(C) Bar plots show percentage of overlap of upregulated or downregulated H3K27me3 peaks with EZH2 binding sites (black bars). Dotted lines represent overlaps with randomized genomic regions (control overlaps) of the same size and number as EZH2 binding sites.

(D) Bar plots show percentage of overlap of upregulated or downregulated H3K27me3 peaks with upregulated SNS-seq peaks (origins, black bars). Dotted lines represent overlaps with randomized genomic regions (control overlaps) of the same size and number.

(legend continued on next page)

Upregulated origins at EZH2 binding sites correlate with a drop in H3K27me3

Ezh2 is a methyltransferase that mediates histone H3K27me3 deposition on target gene promoters inducing a transcriptional repressive state. Loss of EZH2 or its SET domain results in a loss of H3K27me3 deposition (reviewed in Yu et al.[21]). This loss might be hypothesized to lead to a gain in H3K27 acetylation,[28] in agreement with the proposal that H3K27 trimethylation prevents H3K27 acetylation at PRC2 target genes. In addition, the association of the PcG proteins and H3K27 methylation also inhibit the spreading of acetylation to other H3 residues.[28–30] We therefore examined the landscape of H3K27me3 and H3K27ac by ChIP-seq in relation to the increased DNA replication origin activity in *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ cells.

Differential peak intensity analysis using Diffbind suggests that deposition of the H3K27me3 is statistically significantly reduced at numerous genomic regions as well as known PRC2 target loci (419 sites at FDR < 0.05; Figures 3A and 3B; Table S2). Moreover, of the regions statistically significantly downregulated for the H3K27me3 mark in Ezh2$^{\Delta/\Delta}$ cells, almost all contain direct binding sites for EZH2 (96%; Figure 3C), and many contain significantly upregulated origins (26%; Figure 3D), suggesting that loss of H3K27me3 deposition is associated with an increase in origin activity.

Similarly, EZH2-bound sites experience a notable drop in H3K27me levels, which is associated with a similar increase in origin activity (Figures 3E–3G). Positive control regions, where all SNS and ChIP-seq peaks are plotted, display a similar scale of activity to EZH2-bound regions (Figures S4A–S4C, top panels), whereas negative control regions (randomized genomic regions of the same size and number) drawn to the same scale do not experience any changes in activity (Figures S4A–S4C, bottom panels).

Reductions in the repressive mark H3K27me3 are accompanied by a milder, but widespread, increase in the activating histone mark H3K27ac genome wide (7,157 sites; Figure S4E). However, we find that the regions displaying upregulation in H3K27ac levels are loosely associated with EZH2-bound regions (<10% contain an EZH2 binding site) or with upregulated origin activity (6% contain upregulated origins; Figures S4F and S4G). These results would be consistent with spreading of the H3K27ac signal observed in the absence of EZH2 activity.[28] As expected, there is very little H3K27ac signal at EZH2-bound sites in *Ezh2*$^{+/+}$ cells that typically represent repressed promoters (Figure 3H). The residual H3K27ac at EZH2-bound sites is not significantly augmented upon loss of EZH2 function compared with regions that are marked by H3K27ac (Figure 3H versus Figure S4D). We concluded that depletion of EZH2 results in an increased replication origin activity at EZH2 binding sites, which is associated with a drop in H3K27 methylation but is uncorrelated to changes in H3K27ac levels.

Replication initiation at bivalent domains

Our data showed that an increase in replication origin activity at EZH2 binding sites strongly correlated with a corresponding decrease in the repressive H3K27me3 mark. Nevertheless, we do not find transcriptional de-repression at EZH2-bound promoters or *de novo* deposition of H3K27 acetylation. We therefore considered that this transcription-independent upregulation of replication origin activity may be linked to the bivalent nature of polycomb-repressed genes.

PRC2 complexes in mESCs are associated with bivalent chromatin domains. The integrity of these domains is important for silencing of developmentally related gene expression, preventing stem cell differentiation. Bivalent promoters are enriched with H3K4me3 and H3K27me3, where H4K3me3 is an activating chromatin mark deposited by mixed lineage leukemia (MLL) complexes and H3K27me3 a repressive histone modification. While the H3K27me3 domains usually mark large genomic regions spanning multiple genes, bivalent domains are restricted to gene promoters where the H3K4me3 mark resides.[18] More than 95% of H3K27me3 marks present at a promoter in stem cells also have a H3K4me3 mark and thus are bivalent.[31] Since the majority of EZH2 binding sites are at promoter regions (Tables S1B and S1C), we asked whether EZH2 depletion impacts bivalent domain integrity and whether this is associated with DNA replication initiation activity. The high-confidence bivalent promoters were previously defined by Mantsoki et al.[31] based on the overlap of ChIP-seq datasets for H3K27me3 and H4K3me3 histone modifications from several mESCs. Our analysis suggests that upon loss of function of EZH2, DNA replication initiation activity is increased from bivalent promoters (Figure 4A, left panel). This is accompanied by a prominent loss of H3K27me3 from these regions, potentially resulting in an imbalance of the H3K27me3 and H3K4me3 ratio. Similar to EZH2-bound regions, residual amounts of H3K27ac are found at bivalent promoters and do not increase in knockout (KO) conditions (Figure 4A).

Interestingly, re-analysis of data from a published study[32] suggests that loss of EZH2 function in pluripotent hematopoietic stem cell lines does not alter the levels of H3K4me3 at bivalent promoters (Figure 4B), further confirming that increased DNA replication is not associated with increased levels of activating histone marks (H3K4me3 or H3K27ac) but rather with the loss of the repressive H3K27me3 mark. In these cells, loss of EZH2 resulted in an increased chromatin accessibility at repressed promoters, as assessed by assay for transposase-accessible chromatin and sequencing (ATAC-seq) (Figure 4B). We also observed that EZH2-bound TSSs retained a level of chromatin accessibility that was comparable to the levels observed in lowly expressed genes (3–10 transcripts per million (TPM); Figures 4C and 4D). Moreover, we observe that in mouse embryonic fibroblasts, loss of EZH2 results in altered nucleosome positioning on well-established polycomb targets (Figure S4H).

(E) Read profile (top) and heatmap (bottom) plots of EZH2 ChIP-seq experiments over EZH2 binding sites. BRs, n = 3 from 2 studies.
(F) Read profile and heatmap plots of SNS-seq experiments performed in *Ezh2*$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ cells plotted centered on EZH2-bound regions.
(G) Same as (F) but for H3K27me3 ChIP-seq data.
(H) Same as (F) but for H3K27ac ChIP-seq data. The plots are centered at the middle of the EZH2 peaks and normalized by CPMs.
In (E)–(H), merged read pileups plotted for BRs, n = 3.

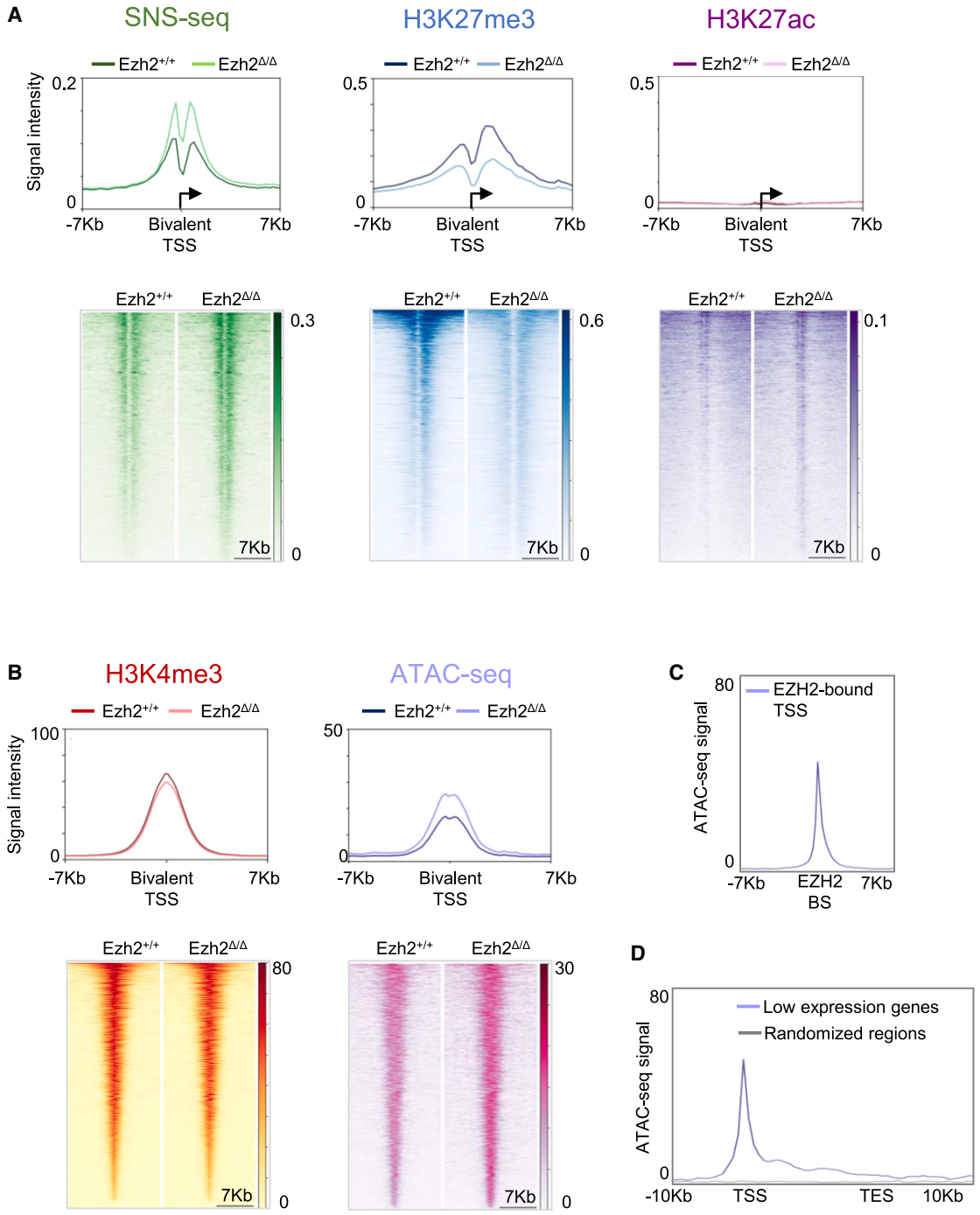


Figure 4. Replication initiation at bivalent domains

(A) Read profile (top) and heatmap (bottom) plots of SNS-seq and H3K27me3 and H3K27ac ChIP-seq experiments over mESC bivalent promoters. The plots are centered at the TSS of the bivalent genes and normalized by CPMs. n = 3, BRs.

(B) Read profile (top) and heatmap (bottom) plots of H3K4me3 (n = 2, BRs) and ATAC-seq (n = 4 for wild type [WT], n = 3 for KO, BRs) experiments over bivalent promoters in pluripotent hematopoietic stem cells. The plots are centered at the TSS of the bivalent genes and normalized by CPMs.

(C) Read profile plots of ATAC-seq experiments over promoters bound by EZH2 in mouse stem cells. The plots are centered at the TSS of the genes and normalized by CPMs.

(D) Read profile plots of ATAC-seq experiments over gene promoters expressed at low levels (3–10 TPMs, in blue) or randomized genomic regions (in gray). The plots are centered at the gene body (TSS to TES) and normalized by CPMs. n = 4 for WT, n = 3 for KO, BRs, merged read pileups plotted.

DISCUSSION

Here, we comprehensively characterized the consequences of EZH2 depletion on the activity of DNA replication origins in mESCs. We used tamoxifen-inducible EZH2-SET domain-deficient cells to address the involvement of the Polycomb complex (and its associated H3K27me3 mark) in the activity of DNA replication origins. We identified DNA replication origins using the RNA-primed nascent DNA analysis coupled with high-throughput sequencing and revealed those affected by EZH2 KO induction. In parallel, epigenetic marks (H3K27me3 and H3K27ac) as well as transcription profiles were analyzed. Our data are consistent with previous observations that loss of EZH2 function leads to loss of the repressive mark H3K27me3 and a further genome-wide spreading of the H3K27ac signal.

Our data showed that abrogation of EZH2 or its catalytic activity led to increased activity of DNA replication origins. Upregulated DNA replication origins are preferentially associated with repressed, EZH2-bound promoter regions. Despite a notable loss of the H3K27me3 mark at these promoters, almost all EZH2-bound promoters remain transcriptionally silent; thus, upregulation of origin activity was not linked to active transcription at these promoters. Previous studies linked increased DNA replication activity to transcriptional activity.[4,33] Interestingly, our results suggest that an increase in DNA replication activity can be uncoupled from active mRNA transcription at repressed promoters. This result is also in line with our observation that the presence of a G-rich element at promoters is sufficient to confer replication origin activity, independently of mRNA transcription[4] and with the affinity of PRC2 complexes with G4 structures.[34,35]

We found a loose association of upregulated origins with the regions that contain upregulated H3K27ac peaks, suggesting that this activating mark on its own might not be the signal recruiting DNA replication initiation activity at these sites. Taken together, we link upregulated DNA replication activity in the absence of EZH2 to lack of H3K27me3 deposition and not to transcriptional de-repression or the unspecific spreading of the activating H3K27ac mark.

In mESCs, almost all promoters that carry a H3K27me3 mark are bivalent. Bivalent domains have a complex chromatin structure, and the H3K27me3 mark is not the only chromatin mark present at these sites. In fact, re-analysis of similar EZH2 loss-of-function data in multi-potent hematopoietic stem cell lines[32] reveals that bivalent promoters have similar levels of H3K4me3 signal after loss of EZH2 but display increased chromatin accessibility as assessed by ATAC-seq (Figure 4B). Interestingly, despite transcriptional repression, PcG-repressed promoters maintain a small amount of open chromatin, which is increased upon loss of function of EZH2 (Figures 4B and 4C). In mouse embryonic fibroblasts, loss of EZH2 also coincides with loss of a nucleosome in the vicinity of the TSS at well-established polycomb target genes (Figure S4H). We previously reported that in mESCs, a nucleosome-free region is present at the origin G-rich element,[7] while nucleosomes are positioned on both sides. Positioned nucleosomes are also present at yeast and *Drosophila* origins.[36,37] Open chromatin and loss or re-positioning of nucleosomes have previously been associated with DNA replication initiation activity and may underlie high DNA replication initiation activity at PcG-repressed promoters. Upon loss of EZH2 function, EZH2-repressed promoters remain transcriptionally silent but have increased chromatin accessibility (likely through nucleosome re-positioning). Our data are consistent with a model in which the exposed DNA forms a "platform" for origin recognition complex (ORC) binding at the EZH2 binding site, resulting in increased DNA replication initiation on either side of this platform.

Limitations of the study

EZH2 is the catalytic subunit of PRC2, and it is known that the absence of EZH2 does not inhibit the formation of the PRC2 complex. Future studies are needed to discern if the PRC2 scaffold on other subunits of PRC2 has an intrinsic ability to recruit DNA replication initiation (which goes unhindered in the absence of EZH2). Due to the nature of next-generation sequencing (NGS), both SNS-seq and ChIP-seq experiments are not fully quantifiable; therefore, we likely underestimated the amount of increased DNA replication initiation and loss of the H3K27me3 mark in *EZH2*$^{-/-}$ cells. Finally, our study examined chromatin marks, accessibility, and nucleosome positioning in *EZH2*$^{-/-}$ cells; however, the impact of loss of function for EZH2 may include other aspects of chromatin landscape, which impacts DNA replication initiation.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines and cell culture
- METHOD DETAILS
  - Excision of the EZH2 SET domain
  - Growth curves for Ezh2$^{+/+}$ and Ezh2$^{\Delta/\Delta}$ cells
  - Flow cytometry analysis
  - RNA-primed short nascent strand (SNS) isolation and sequencing
  - RNA extraction and sequencing
  - Chromatin extraction and chromatin immunoprecipitation (ChIP) and sequencing
  - Alignment and quality control of raw sequencing data
  - Peak calling
  - Quantification and differential peak activity (SNS-seq and ChIP-seq)
  - Differential gene expression analysis
  - Overlap of genomic regions and control regions
  - Data visualization
  - Nucleosome positioning
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112280.

ACKNOWLEDGMENTS

We warmly thank Amanda Fisher for generously providing us with the EZH2 KO mESC line and Jian Xu for providing details on his data. We are also thankful to Fiona Docherty and Giacomo Cavalli for useful discussions on our manuscript. P.P. was supported by the ANR14-CE10-0019 OriChoice and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, Project-ID 393547839 – SFB1361, CA198/9-2, CA 198/12-1). This research was supported by the Birmingham Fellowship Programme, RD Lawrence Fellowship (Diabetes UK, 20/0006136), and Academy of Medical Sciences Springboard (SBF006\1140) to I.A. This research also received funding to M.M. from the Association Française contre les Myopathies (AFM); the ARC Foundation; the ANR14-CE10-0019; the MSD AVENIR Fund GENE-IGH; the Labex EpiGenMed (reference ANR-10-LABX-12-0); and Pre-maturation support from the Region Occitanie. The graphical abstract was made using BioRender. The bioinformatics on this project was supported by the high-power computing cluster of University of Birmingham (CastLes and BlueBear) and the Lakitu server of IGH.

AUTHOR CONTRIBUTIONS

P.P. and M.M. conceived the study, and I.A. led the bioinformatic analysis of the data. P.P. performed all experiments with the help of I.P. and R.C., as well as preliminary bioinformatic analysis. F.F. performed experiments that supported this study. N.M. performed bioinformatic analysis on public and generated datasets under the supervision of I.A. P.P., I.A., and M.M. wrote the paper with input from N.M., F.F., and R.C.

DECLARATION OF INTERESTS

I.A. is a partner at mireX Genomics.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in their field of research or within their geographical location.

Received: July 18, 2022
Revised: December 13, 2022
Accepted: March 3, 2023
Published: March 29, 2023

REFERENCES

1. Méchali, M. (2010). Eukaryotic DNA replication origins: many choices for appropriate answers. Nat. Rev. Mol. Cell Biol. *11*, 728–738. https://doi.org/10.1038/nrm2976.
2. Fragkos, M., Ganier, O., Coulombe, P., and Méchali, M. (2015). DNA replication origin activation in space and time. Nat. Rev. Mol. Cell Biol. *16*, 360–374. https://doi.org/10.1038/nrm4002.
3. Ekundayo, B., and Bleichert, F. (2019). Origins of DNA replication. PLoS Genet. *15*, e1008320. https://doi.org/10.1371/journal.pgen.1008320.
4. Akerman, I., Kasaai, B., Bazarova, A., Sang, P.B., Peiffer, I., Artufel, M., Derelle, R., Smith, G., Rodriguez-Martinez, M., Romano, M., et al. (2020). A predictable conserved DNA base composition signature defines human core DNA replication origins. Nat. Commun. *11*, 4826. https://doi.org/10.1038/s41467-020-18527-0.
5. Cayrou, C., Coulombe, P., Vigneron, A., Stanojcic, S., Ganier, O., Peiffer, I., Rivals, E., Puy, A., Laurent-Chabalier, S., Desprat, R., and Méchali, M. (2011). Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. Genome Res *21*, 1438–1449. https://doi.org/10.1101/gr.121830.111.
6. Cayrou, C., Coulombe, P., Puy, A., Rialle, S., Kaplan, N., Segal, E., and Méchali, M. (2012). New insights into replication origin characteristics in metazoans. Cell Cycle *11*, 658–667. https://doi.org/10.4161/cc.11.4.19097.
7. Cadoret, J.C., Meisch, F., Hassan-Zadeh, V., Luyten, I., Guillet, C., Duret, L., Quesneville, H., and Prioleau, M.N. (2008). Genome-wide studies highlight indirect links between human replication origins and gene regulation. Proc. Natl. Acad. Sci. USA *105*, 15837–15842. https://doi.org/10.1073/pnas.0805208105.
8. Besnard, E., Babled, A., Lapasset, L., Milhavet, O., Parrinello, H., Dantec, C., Marin, J.M., and Lemaitre, J.M. (2012). Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. Nat. Struct. Mol. Biol. *19*, 837–844. https://doi.org/10.1038/nsmb.2339.
9. Cayrou, C., Ballester, B., Peiffer, I., Fenouil, R., Coulombe, P., Andrau, J.C., van Helden, J., and Méchali, M. (2015). The chromatin environment shapes DNA replication origin organization and defines origin classes. Genome Res *25*, 1873–1885. https://doi.org/10.1101/gr.192799.115.
10. Costas, C., de la Paz Sanchez, M., Stroud, H., Yu, Y., Oliveros, J.C., Feng, S., Benguria, A., López-Vidriero, I., Zhang, X., Solano, R., et al. (2011). Genome-wide mapping of Arabidopsis thaliana origins of DNA replication and their associated epigenetic marks. Nat. Struct. Mol. Biol. *18*, 395–400. https://doi.org/10.1038/nsmb.1988.
11. Comoglio, F., Schlumpf, T., Schmid, V., Rohs, R., Beisel, C., and Paro, R. (2015). High-resolution profiling of Drosophila replication start sites reveals a DNA shape and chromatin signature of metazoan origins. Cell Rep *11*, 821–834. https://doi.org/10.1016/j.celrep.2015.03.070.
12. Langley, A.R., Gräf, S., Smith, J.C., and Krude, T. (2016). Genome-wide identification and characterisation of human DNA replication origins by initiation site sequencing (ini-seq). Nucleic Acids Res. *44*, 10230–10247. https://doi.org/10.1093/nar/gkw760.
13. Prorok, P., Artufel, M., Aze, A., Coulombe, P., Peiffer, I., Lacroix, L., Guédin, A., Mergny, J.L., Damaschke, J., Schepers, A., et al. (2019). Involvement of G-quadruplex regions in mammalian replication origin activity. Nat. Commun. *10*, 3274. https://doi.org/10.1038/s41467-019-11104-0.
14. Valton, A.-L., Hassan-Zadeh, V., Lema, I., Boggetto, N., Alberti, P., Saintomé, C.Riou, J.-F., and Prioleau, M.-N., (2014). G4 motifs affect origin positioning and efficiency in two vertebrate replicators. EMBO J. 33, 732–746. https://doi:10.1002/embj.201387506
15. Smith, O.K., Kim, R., Fu, H., Martin, M.M., Lin, C.M., Utani, K., Zhang, Y., Marks, A.B., Lalande, M., Chamberlain, S., et al. (2016). Distinct epigenetic features of differentiation-regulated replication origins. Epigenet. Chromatin *9*, 18. https://doi.org/10.1186/s13072-016-0067-3.
16. Kurat, C.F., Yeeles, J.T.P., Patel, H., Early, A., and Diffley, J.F.X. (2017). Chromatin controls DNA replication origin selection, lagging-strand synthesis, and replication fork rates. Mol. Cell *65*, 117–130. https://doi.org/10.1016/j.molcel.2016.11.016.
17. Smith, O.K., and Aladjem, M.I. (2014). Chromatin structure and replication origins: determinants of chromosome replication and nuclear organization. J. Mol. Biol. *426*, 3330–3341. https://doi.org/10.1016/j.jmb.2014.05.027.
18. Schuettengruber, B., Bourbon, H.M., Di Croce, L., and Cavalli, G. (2017). Genome regulation by polycomb and trithorax: 70 Years and counting. Cell *171*, 34–57. https://doi.org/10.1016/j.cell.2017.08.002.
19. Zijlmans, D.W., Talon, I., Verhelst, S., Bendall, A., Van Nerum, K., Javali, A., Malcolm, A.A., van Knippenberg, S.S.F.A., Biggins, L., To, S.K., et al. (2022). Integrated multi-omics reveal polycomb repressive complex 2 restricts human trophoblast induction. Nat. Cell Biol. *24*, 858–871. https://doi.org/10.1038/s41556-022-00932-w.
20. Loh, C.H., van Genesen, S., Perino, M., Bark, M.R., and Veenstra, G.J.C. (2021). Loss of PRC2 subunits primes lineage choice during exit of pluripotency. Nat. Commun. *12*, 6985. https://doi.org/10.1038/s41467-021-27314-4.
21. Yu, J.R., Lee, C.H., Oksuz, O., Stafford, J.M., and Reinberg, D. (2019). PRC2 is high maintenance. Genes Dev. *33*, 903–935. https://doi.org/10.1101/gad.325050.119.

22. Pereira, C.F., Piccolo, F.M., Tsubouchi, T., Sauer, S., Ryan, N.K., Bruno, L., Landeira, D., Santos, J., Banito, A., Gil, J., et al. (2010). ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency. Cell Stem Cell *6*, 547–556. https://doi.org/10.1016/j.stem.2010.04.013.

23. Su, I.H., Basavaraj, A., Krutchinsky, A.N., Hobert, O., Ullrich, A., Chait, B.T., and Tarakhovsky, A. (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat. Immunol. *4*, 124–131. https://doi.org/10.1038/ni876.

24. Kaneko, S., Son, J., Shen, S.S., Reinberg, D., and Bonasio, R. (2013). PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1258–1264. https://doi.org/10.1038/nsmb.2700.

25. Fang, D., Gan, H., Cheng, L., Lee, J.H., Zhou, H., Sarkaria, J.N., Daniels, D.J., and Zhang, Z. (2018). H3.3K27M mutant proteins reprogram epigenome by sequestering the PRC2 complex to poised enhancers. Elife *7*, e36696. https://doi.org/10.7554/eLife.36696.

26. 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. https://doi.org/10.1016/j.molcel.2013.10.030.

27. Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res. *21*, 1273–1283. https://doi.org/10.1101/gr.122382.111.

28. Lavarone, E., Barbieri, C.M., and Pasini, D. (2019). Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. Nat. Commun. *10*, 1679. https://doi.org/10.1038/s41467-019-09624-w.

29. Pasini, D., Malatesta, M., Jung, H.R., Walfridsson, J., Willer, A., Olsson, L., Skotte, J., Wutz, A., Porse, B., Jensen, O.N., and Helin, K. (2010). Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. Nucleic Acids Res. *38*, 4958–4969. https://doi.org/10.1093/nar/gkq244.

30. Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., and Harte, P.J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development *136*, 3131–3141. https://doi.org/10.1242/dev.037127.

31. Mantsoki, A., Devailly, G., and Joshi, A. (2015). CpG island erosion, polycomb occupancy and sequence motif enrichment at bivalent promoters in mammalian embryonic stem cells. Sci. Rep. *5*, 16791. https://doi.org/10.1038/srep16791.

32. Gu, Z., Liu, Y., Cai, F., Patrick, M., Zmajkovic, J., Cao, H., Zhang, Y., Tasdogan, A., Chen, M., Qi, L., et al. (2019). Loss of EZH2 reprograms BCAA metabolism to drive leukemic transformation. Cancer Discov. *9*, 1228–1247. https://doi.org/10.1158/2159-8290.CD-19-0152.

33. Sequeira-Mendes, J., Díaz-Uriarte, R., Apedaile, A., Huntley, D., Brockdorff, N., and Gómez, M. (2009). Transcription initiation activity sets replication origin efficiency in mammalian cells. PLoS Genet. *5*, e1000446. https://doi.org/10.1371/journal.pgen.1000446.

34. Beltran, M., Tavares, M., Justin, N., Khandelwal, G., Ambrose, J., Foster, B.M., Worlock, K.B., Tvardovskiy, A., Kunzelmann, S., Herrero, J., et al. (2019). G-tract RNA removes Polycomb repressive complex 2 from genes. Nat. Struct. Mol. Biol. *26*, 899–909. https://doi.org/10.1038/s41594-019-0293-z.

35. Wang, X., Goodrich, K.J., Gooding, A.R., Naeem, H., Archer, S., Paucek, R.D., Youmans, D.T., Cech, T.R., and Davidovich, C. (2017). Targeting of polycomb repressive complex 2 to RNA by short repeats of consecutive guanines. Mol. Cell *65*, 1056–1067.e5. https://doi.org/10.1016/j.molcel.2017.02.003.

36. Eaton, M.L., Galani, K., Kang, S., Bell, S.P., and MacAlpine, D.M. (2010). Conserved nucleosome positioning defines replication origins. Genes Dev. *24*, 748–753. https://doi.org/10.1101/gad.1913210.

37. MacAlpine, H.K., Gordân, R., Powell, S.K., Hartemink, A.J., and MacAlpine, D.M. (2010). Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading. Genome Res. *20*, 201–211. https://doi.org/10.1101/gr.097873.109.

38. Liu, Y., Gu, Z., Cao, H., Kaphle, P., Lyu, J., Zhang, Y., Hu, W., Chung, S.S., and Dickerson, K.E. (2021). *Convergence of oncogenic cooperation at single-cell and single-gene levels drives leukemic transformation*. Nat Commun *12* (*1*), 6323. https://doi.org/10.1038/s41467-021-26582-4.

39. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

40. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

41. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

42. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

43. Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W187–W191. https://doi.org/10.1093/nar/gku365.

44. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics. *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

45. 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. https://doi.org/10.1186/gb-2008-9-9-r137.

46. Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics. *25*, 1952–8. https://doi.org/10.1093/bioinformatics/btp340.

47. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

48. Stark, R., and Brown, G.. *DiffBind: differential binding analysis of ChIP-Seq peakdata*. http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf.

49. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527. https://doi.org/10.1038/nbt.3519.

50. Wickham, H. (2016). ggplot2; Elegant Graphics for Data Analysis, Second Edition (Springer Cham).

51. Kolde, R. (2012). Pheatmap: pretty heatmaps. R package version *1* (*2*), 726.

52. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26. https://doi.org/10.1038/nbt.1754.

53. R Development Core Team (2013). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

54. Rossum, G.v. (1995). *Python reference manual*, in *Department of Computer Science [CS]*. CWI.

55. Schep, A.N., Buenrostro, J.D., Denny, S.K., Schwartz, K., Sherlock, G., and Greenleaf, W.J. (2015). Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. Genome Res. *25*, 1757–1770. https://doi.org/10.1101/gr.192294.115.

56. Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. Science *327*, 996–1000. https://doi.org/10.1126/science.1184208.

57. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

58. Xu, S., Grullon, S., Ge, K., and Peng, W. (2014). Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. Methods Mol. Biol. *1150*, 97–111. https://doi.org/10.1007/978-1-4939-0512-6_5.

59. Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389–393. https://doi.org/10.1038/nature10730.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
mouse anti-BrdU	Becton Dickinson	Cat# 347580; AB_10015219
mouse anti-PCNA	Sigma-Aldrich	Cat# P8825; AB_477413
mouse anti-Actin	Sigma-Aldrich	Cat# A4700; AB_476730
rabbit anti-Oct4	Abcam	Cat# ab19857; AB_445175
mouse anti-Chk1	Santa-cruz	Cat# sc-8408; AB_627257
rabbit anti-pChk1	Cell signaling	Cat# 2341
rabbit anti-Ezh2	Sigma-Aldrich	Cat# E6906; AB_1078770
rabbit anti-H3K27ac	Abcam	Cat# ab4729; AB_2118291
rabbit anti-H3K27me3	Cell signaling	Cat# C36B11
secondary FITC-coupled donkey anti-mouse igG	Jackson Laboratories	Cat# 715-096-151; AB_2340796
secondary anti-mouse igG HRP conjugated	GE Healthcare	Cat# NA931V
secondary anti-rabbit igG HRP conjugated	GE Healthcare	Cat# NA934V
Chemicals, peptides, and recombinant proteins		
AMPure XP Reagent	Beckman coulter	A63880
BioPrime™ Array CGH Genomic Labeling System	Invitrogen	18095011
5-bromo-2′-deoxyuridine	Sigma-Aldrich	B5002
BSA	Sigma-Aldrich	A2153
Cytiva Amersham™ Hyperfilm™ MP	GE Healthcare	10218674
DNA polymerase I (Klenow fragment)	NEB	M0210S
dNTP	NEB	N0446S
Pierce™ ECL Western Blotting-Substrat	ThermoFisher Scientific	32106
EDTA	Sigma-Aldrich	E9884
EGTA	Millipore	324626
Ethanol	Sigma-Aldrich	493511
Fetal bovine serum	Biowest	S1810
DynabeadsTM Protein G for Immunoprecipitation	Thermo Fisher	10004D
Glycerol	Sigma-Aldrich	G9012
HCl	Sigma-Aldrich	H1758
HEPES	Sigma-Aldrich	H3375
KO-DMEM	Gibco-Thermo Fisher	10829018
IllustraTM CyScribe GFXTM purification kit	GE Healthcare	28-9034-70
Illumina TruSeq Stranded mRNA Sample Preparation Kit	Illumina	20020594
Leukemia Inhibitory Factor (recombinant mouse LIF)	Isokine	01-A1140-0010
Milk	Régilait	FR 71.448.001 CE
N-lauroylsarcosine	Sigma-Aldrich	L9150
Na-Borate	Sigma-Aldrich	202215
NaCl	Sigma-Aldrich	S7653
Non-essential amino acids (MEM NEAA)	Invitrogen	11140035
Nonidet™ P 40 Substitute	Sigma-Aldrich	74385

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Nylon filter (FACS analysis) (cell stainer)	Falcon	352350
Gibco™ DPBS (1X)	Gibco	14190144
PhusionTM High-Fidelity DNA Polymerase	NEB	M0530S
propidium iodide (PI)	Sigma-Aldrich	P4864
cOmplete™, EDTA-free Protease Inhibitor Cocktail	Roche	4693132001
Proteinase K	Sigma-Aldrich	P6556
QIAquick PCR purification columns	Qiagen	28104
Random Primer 6	NEB	S1230S
RNase A	Sigma-Aldrich	R6513
RNeasy Kit	Qiagen	74104
SDS 20%	Biosolve (Dutscher)	00198112323BS
Dodium deoxycholate	Sigma-Aldrich	D6750
Sodium pyruvate	Gibco	11360070
Sucrose	Sigma-Aldrich	S0389
LightCycler® 480 SYBR Green I Master	Roche	4887352001
T4 Polynucleotide Kinase	NEB	M0201S
Tamoxifen ((Z)-4-Hydroxytamoxifen)	Sigma-Aldrich	H7904
Triton X-100	Sigma Aldrich	X100
TruSeq ChIP Sample Prep Kit	Illumina	IP-202-1012
Trypsin-EDTA (0.05%), phenol red	Gibco	25300096
λ-Exonuclease (λ-Exo)	Thermo Fisher	synthesis on request (50U/μl)
β-mercaptoethanol	Thermo Fisher	31350010
DNAzol	Thermo Fisher	10503027
Deposited data		
H3K27me3 and H3K27ac ChIP-Seq, RNA-Seq, SNS-Seq of $Ezh2^{\Delta/\Delta}$ and $Ezh2^{+/+}$ genotypes, mESC	This paper	GSE205756
EZH2 ChIP-Seq, E14 mESC	Kaneko et al.[24]	PMID: 24141703, GSE49431
EZH2 ChIP-Seq, 129Sv/J mESC	Fang et al.[25]	PMID: 29932419, GSE94834
ATAC-Seq, LSK C57BL/G	Gu et al.[32]	PMID: 31189531, GSE132077
Bivalent domains (H3K27me3 and H3K4me3 ChIP-Seq), LSK C57BL/6	Liu et al.[38]	PMID: 34732703, GSE132078
Bivalent promoters (list), mESC various	Mantsoki et al.[31]	PMID: 26582124
Experimental models: Cell lines		
ES tamoxifen-inducible Ezh2Δ/Δ	Professor Amanda Fisher	PMID: 20569692
Oligonucleotides		
Ezh2 3' - CTGCTCTGAATGGCAACTCC	Eurofins (custom synthesis)	N/A
Ezh2 5' - TTATTCATAGAGCCACCTGG	Eurofins (custom synthesis)	N/A
5loxP - ACGAAACAGCTCCAGATTCAGGG	Eurofins (custom synthesis)	N/A
Software and algorithms		
SNS-seq Analysis Pipeline	This study and Akerman et al. (PMID: 32958757)	https://doi.org/10.5281/zenodo.7655100
Bcl2fastq (v2.17)	Illumina	https://emea.support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html
FastQC (v0.11.9)	Andrews[39]	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Trimmomatic (v0.39)	Bolger et al.[40]	http://www.usadellab.org/cms/?page=trimmomatic

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bowtie2 (v2.3.5.1)	Langmead and Salzberg[41]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
SAMtools (v1.10)	Li et al.[42]	http://www.htslib.org/
deepTools (v3.5.0)	Ramírez et al.[43]	https://deeptools.readthedocs.io/en/develop/
STAR (v2.7.2b-GCC-8.3.0)	Doblin et al.[44]	https://github.com/alexdobin/STAR
MACS2 (v2.2.1, v2.2.7.1)	Zhang et al.[45]	https://github.com/macs3-project/MACS/wiki/Install-macs2
SICER (v1.1)	Zang et al.[46]	https://bioweb.pasteur.fr/packages/pack@SICER@1.1
BEDtools (v2.3029.02)	Quinlan et al.[47]	https://bedtools.readthedocs.io/en/latest/
Diffbind (v3.9)	Stark and Brown[48]	https://bioconductor.org/packages/release/bioc/html/DiffBind.html
Kallisto (v0.48.0)	Bray et al.[49]	https://github.com/pachterlab/kallisto
BiocManager	R	https://cran.r-project.org/web/packages/BiocManager/vignettes/BiocManager.html
ggplot2 (v3.1.0)	Wickham[50]	https://ggplot2.tidyverse.org/
Pheatmap (v1.0.12)	Kolde[51]	https://cran.rproject.org/web/packages/pheatmap/index.html
IGV (v2.10)	Robinson et al.[52]	https://software.broadinstitute.org/software/igv/
R (v4.1.0)	R Core Team[53]	https://www.r-project.org/
Python (v3.8.2)	Rossum,[54] Python Software Foundation	https://www.python.org/downloads/
Rstudio	Rstudio Team	http://www.rstudio.com/
BioRender.com	BioRender	https://biorender.com/
NucleoATAC	Schep et al.[55]	https://nucleoatac.readthedocs.io/en/latest/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Marcel Mechali (marcel.mechali@igh.cnrs.fr).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All raw data generated in this study is available on GEO data with accession GSE205756. All called peaks in this publication are given in Table S1. Other publicly available data sources used in this study are given in Table S4.
- All original code has been deposited at Zenodo/Github and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to re-analyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines and cell culture

The tamoxifen-inducible Ezh2$^{\Delta/\Delta}$ mouse Embryonic Stem Cell (ESC) line was a gift from Amanda Fisher.[22] Cells were cultured on gelatin-coated dishes in KnockOut^{TM} Dulbecco's Modified Eagle's Medium (KO-DMEM, #10829018, Invitrogen) containing 10% fetal bovine serum (Biowest -S1810-Eurobio), 2 mM L-glutamine (25030-024- Invitrogen), 1% non-essential amino acids (MEM NEAA-11140035-Invitrogen), 0.1 mM β-mercaptoethanol (31350010- Invitrogen), 1mM Sodium Pyruvate and 1,000 U ml-1 recombinant Leukemia Inhibitory Factor (LIF, Eurobio).

METHOD DETAILS

Excision of the EZH2 SET domain

For excision of the SET domain, cells were incubated with 100 nM Tamoxifen (Sigma Aldrich) for 48h and grown for next 48h in cell culture medium without tamoxifen supplementation. The Ezh2-SET deletion of genomic locus was confirmed by a PCR on gDNA using following primers Ezh2 3' - CTGCTCTGAATGGCAACTCC, Ezh2 5' - TTATTCATAGAGCCACCTGG, Ezh2 left 5loxP - ACGAAACAGCTCCAGATTCAGGG. PCR reactions contained genomic DNA template, 1 μM of each primer, 0.2 mM dNTPs, 1 x Reaction buffer, Phusion™ High-Fidelity DNA Polymerase and cycling conditions were set to (10'' at 98°C, 15' at 60°C, 15' at 72°C) x35. In the absence of tamoxifen, the size of the PCR product was 450 bp in length and upon 48h tamoxifen treatment the excised product size was 370 bp. The expression of a truncated version of EZH2 was additionally confirmed in western blot analysis using antibodies against EZH2 (List of antibodies used in this study are provided in Table S3).

Growth curves for $Ezh2^{+/+}$ and $Ezh2^{\Delta/\Delta}$ cells

For the growth curve experiment, 240,000 $Ezh2^{+/+}$ mES were seeded per well in 6-well microtiter plates, in growth medium as previously described. Cells were treated with Tamoxifen ($Ezh2^{\Delta/\Delta}$) or left untreated ($Ezh2^{+/+}$, control) over the course of 2 days, and tamoxifen was removed at day2. Time points were taken at day 2, 3 and 4 from the point of addition of tamoxifen (48, 72 and 96 hours). Experiments were performed in triplicate. At each timepoint, cells were trypsinized and counted using a Malassez chamber.

Flow cytometry analysis

For labelling, 5-bromo-2′-deoxyuridine (BrdU, Sigma-Aldrich B5002) was added directly to the cell culture medium at 10 μM, and the cells incubated at 37°C in 5% CO_2 for 15-20 min. Cells were dissociated by treatment with 0.05% trypsin-EDTA, resuspended in cell medium, and centrifuged. Pelleted cells were washed once with PBS, then fixed with 70% ethanol, 30% PBS at -20C for at least 30 min. After being washed once with PBS, cells were treated with RNase A (Sigma-Aldrich R6513, 50 μg/ml in PBS) for 30-60 min. To denature genomic DNA, cells were incubated in 2N HCl at room temperature for 30 min. Cells were then centrifuged and the pellet resuspended in 0.1 M Na Borate (pH 8.5) for 2-3 min to neutralise the acid. The cells were washed once with PBS/0.5% NP-40, followed by incubation with mouse anti-BrdU antibody (Becton Dickinson #347580, 25 μg/ml, in PBS/0.5% NP-40/1% BSA) at room temperature for 45 min-2 h. Cells were washed with PBS/0.5% NP-40, then incubated with FITC-coupled donkey anti-mouse antibody (Jackson Laboratories, code 715-096-151, at 10 μg/ml in PBS/0.5% NP-40/1% BSA) in the dark at room temperature for 45 min-2 h. Cells were washed with PBS/0.5% NP-40 and incubated with PI (Sigma-Aldrich P4864, 25 μg/ml in PBS) at room temperature for 10 min. Cells were further purified of aggregates by passing them though a nylon filter (40 μm pore size). The samples were analysed by flow cytometry using a MACSQuant flow cytometer (Miltenyi Biotec Inc.) at the Montpellier Ressources Imagerie platform (https://www.mri.cnrs.fr/en/). Flow cytometry data were analysed, visualised and quantified using FlowJo software (Becton Dickinson).

RNA-primed short nascent strand (SNS) isolation and sequencing

SNSs were purified as previously described in detail.[7,10] Briefly, cells were lysed with DNAzol, then nascent strands (NS) were first separated from genomic DNA based on sucrose gradient size fractionation n=3, biological replicates. Fractions corresponding to 0.5-2 kb were isolated and phosphorylated by T4 polynucleotide kinase and digested with two consecutive overnight rounds of λ-Exonuclease (λ-Exo, Thermo Fisher, 100-150 units) to eliminate contaminating DNA. A third λ-Exo digestion (100 units) was performed for 2 h at 37°C. Three samples were purified from three independent cell cultures. The "RNaseA" controls were obtained by treating NS-containing fractions with 50–100 μg/mL RNase A before λ−Exo digestion to remove RNA primers at the 5′ ends of the NSs. The quality of origin enrichment in each sample was determined by qPCR amplification using the specific primers with the LightCycler 480 SYBR Green Master mix, in a LightCycler 480 II, Roche. The NS enrichment was calculated as the ratio of the signal scored at origin-specific and background regions. Single-stranded NS DNA and RNaseA treated control DNA was first purified using an IllustraTM CyScribe GFXTM purification kit (GE Healthcare), then converted into double-stranded DNA by random priming using DNA polymerase I (Klenow fragment) and the Array CGH kit (BioPrime). The Illumina TruSeq ChIP Sample Prep Kit was used for preparation of sequencing libraries following manufacturer's instructions. Samples were sequenced using the Illumina HiSeq 2500 (50 bp, single-end reads) at the Montpellier GenomiX (MGX) facility (https://www.mgx.cnrs.fr/).

RNA extraction and sequencing

Approximately 5 million mESCs from three independent cell cultures (n=3) each of $Ezh2^{WT}$ and $Ezh2^{\Delta/\Delta}$ were collected for total RNA extraction using the RNeasy Kit (Qiagen). Libraries were prepared using the Illumina TruSeq Stranded mRNA Sample Preparation Kit and sequenced using an Illumina HiSeq 2500 apparatus at the Montpellier GenomiX facility (50 bp, single-end reads).

Chromatin extraction and chromatin immunoprecipitation (ChIP) and sequencing

This was carried out according to the protocol referenced in.[56] Briefly, samples (n=3, biological replicates) were crosslinked for 2 min in 1% formaldehyde at room temperature, then quenched for 5 min at room temperature with 125 mM glycine. Chromatin extraction was performed with approximately 1 x 10^7 ESCs from three independent cell cultures, harvested by centrifugation and washed twice

in cold PBS. Pellets were resuspended by rotating in 1 ml of cold lysis Buffer A (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100) freshly supplemented with protease inhibitor cocktail. Lysate was centrifuged at 2500 rpm for 5 min, and the resulting supernatant discarded. 1 ml cold lysis Buffer B (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, with protease inhibitors) was added to the pellet, which was rotated at room temperature for 10 min, then centrifuged at 2500 rpm for 5 min, and the supernatant discarded. Extraction of chromatin was performed by adding 750 μl of Buffer C (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodium deoxycholate, 0.5% N-lauroylsarcosine, with protease inhibitors) to the remaining pellets. Then, purified chromatin was sonicated in an ultrasonic bath (Diagenode, Bioruptor) to generate DNA fragments of average length 200–500 bp. Insoluble material was pelleted and removed. Sonicated chromatin was immunoprecipitated overnight with 1 μg antibody (Table S3) in a final volume of 500 μl ChIP buffer (20mM Tris-HCl pH8, 2mM EDTA, 150mM NaCl, 0,1% SDS, 1% Triton, with protease inhibitors). Immunocomplexes were recovered with 30 μl of G-agarose beads (Dynabeads™ Protein G for Immunoprecipitation 10004D; Thermo Fisher Scientific) slurry incubated for 2 h at 4°C. Beads were washed as previously described and eluted in 100 ml of elution buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8) for 25 min at 65°C after vigorous vertexing. DNA complexes were de-crosslinked at 65°C overnight, treated with proteinase K and purified with QIAquick PCR purification columns (Qiagen), according to the manufacturer's instructions. Immunoprecipitated DNA was analyzed using SYBR Green (ABI) in a real-time qPCR machine (LightCycler 480 II, Roche). The Illumina TruSeq ChIP Sample Prep Kit was used for preparation of sequencing libraries. Samples were sequenced using the Illumina HiSeq 2500 (50 bp, single-end reads) at the Montpellier GenomiX facility.

Alignment and quality control of raw sequencing data

Fastq files were produced with a bcl2fastq version 2.17 from Illumina by the MGX facility. The quality of the produced reads (50 bp, single-end) from each ChIP-seq, SNS-seq and RNA-seq sample was assessed before and after trimming with FastQC (v0.11.9). Trimming was performed with Trimmomatic[40] (v0.39) (parameters 3:10:8:6 LEADING:5 TRAILING:5 SLIDINGWINDOW:4:15 HEADCROP:3 MINLEN:27). For ChIPip-seq and SNS-seq, the reads were then aligned to the mouse genome (GRCm38/mm10, UCSC, Dec. 2011) using Bowtie2[41] (v2.3.5.1), and the produced .sam files were converted to .bam and .bw files using SAMtools[42] (v1.10) and deepTools[43] (v. 3.5.0) (parameters –normalizeUsing CPM –binSize 10), respectively. For RNA-seq, the reads were aligned to the mouse genome (mm10) using STAR[57] (v2.7.2b-GCC-8.3.0) (parameters —runMode alignReads –outFilterMultimapNmax 20 –outReadsUnmapped Fastx_failed –outSAMtype BAM SortedByCoordinate). All our samples had high alignment rates, typically > 85%.

Peak calling

SNS-seq Peaks were called using two peak calling programs MACS2 (v2.2.1) and SICER (v1.1).[45,58] MACS2 (additional parameters –bw 500 -p 1e-5 -s 60 -m 10 30–gsize 2.7e9) was used to identify narrow peaks of replication initiation, the subsequent peak calling using SICER (parameters: redundancy threshold = 1, window size (bp) = 200, fragment size = 150, effective genome fraction = 0.85, gap size (bp) = 600 and FDR = 1e-3) permitted to identify broader zones of replication initiation. MACS2 narrow peaks intersecting with SICER peaks (using Bedtools[47] intersect v2.30.0) were considered as sites of DNA replication initiation. ChIP-seq peaks (H3K27me3, H3K27ac, EZH2) were called using MACS2 (v2.2.7.1) (default parameters plus –bw 500 -p 1e-5 -s 60 -m 10 30 –gsize 2.3e9) in obedience accordance with the ENCODE standards (broad peaks for H3K27me3 and narrow for H3K27ac). The generated MACS2 peaks were merged with BEDtools[47] (v2.3029.02) and intersected to obtain bed files of all peaks, wild-type peaks and knockout peaks (H3K27me3 and H3K27ac). EZH2 peaks were called using the common peaks between ref 24, 25 (consistent peaks) or merged between the references (all peaks), coordinates given in Table S1 (mm10).

Quantification and differential peak activity (SNS-seq and ChIP-seq)

Quantification of the SNS/ChIP-seq signals was done using the R-package Diffbind[59] (v3.9), using the peaks obtained previously and the TMM minus normalization. The same package was used to define differential signal in SNS-seq (Diffbind v2.14.0) and ChIP-seq samples (Diffbind (v3.9) in $Ezh2^{+/+}$ and $Ezh2^{\Delta/\Delta}$ cells using R Studio. Our scripts are available on github/iakerman/SNSseq.

Differential gene expression analysis

RNA-seq samples from $Ezh2^{+/+}$ and $Ezh2^{\Delta/\Delta}$ cells (n=3, biological replicates, BR)) were subjected to differential gene expression analysis using DeSeq2 (v.1.18.0 in R v. 4.1.0) using default parameters but removing genes with less than 10 reads in at least 3 samples. TPMs were calculated using kallisto[49] and its default gene annotations (ENSEMBL). Mouse gene names and coordinates (EnsDb.Mmusculus.v79) were obtained using the biomaRt library BiocManager in R. Curated Refseq genes were downloaded from UCSC Table browser and the kallisto/Patcher lab website.

Overlap of genomic regions and control regions

BEDTools (v2.30.0) was used to sort the data (sort -V -k1,1 -k2,2n), merge the peaks (merge), produce shuffled controls (shuffle -chrom), extend the peaks' size (flank -l -r) and overlap (bedtools intersect -wa) the data. Control regions for upregulated origins were produced using a custom R script that selected equally active (based on normalised Diffbind scores) origins from a list of unchanged origins. Significance of overlap was calculated in R, using Chi-square test.

Data visualization

Heatmap and box plots were generated using ggplot2[50] (v3.1.0), pheatmap (v1.0.12) or deepTools in R (v4.1.0). Bar and pie charts were generated in Microsoft Excel using data obtained with BEDTools, R, Diffbind. For read-pile up plots, matrixes were computed and density profile plots and heatmaps were obtained using with deepTools[43] (v3.5.0) and Python (v3.8.2). IGV[52] (MIT, v2.10) was used to visualize the sequencing data on mm10 genome regions.

Nucleosome positioning

Nucleosome positioning was calculated using NucleoATAC (v0.2.1)[55] based on data from Gu et al.[32] with default parameters (nucleoatac run) centered at transcription start sites of genes bound by EZH2.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details of experiments can be found in figure legends and method details section. All quantification and statistical analysis was performed in R or using dedicated software (i.e. Diffbind, DeSeq2, listed in Key Resources). Chi-square test was performed on the raw count numbers by testing if the the observed versus expected frequencies (of overlap) between two sample groups were the same. Significance was determined at $p < 0.05$. n represents number of biological replicates (BR). Where appropriate, error bars represent SEM.