8970–8986 *Nucleic Acids Research, 2023, Vol. 51, No. 17* *Published online 31 July 2023*
https://doi.org/10.1093/nar/gkad640

Dot1l cooperates with Npm1 to repress endogenous retrovirus MERVL in embryonic stem cells

Xin Zhao[1,†], Xiaomin Li[1,†], Haiyang Sun[1,†], Xuan Zhao[1], Tingting Gao[1], Panpan Shi[2], Fuquan Chen[2], Lin Liu[2] and Xinyi Lu[1,*]

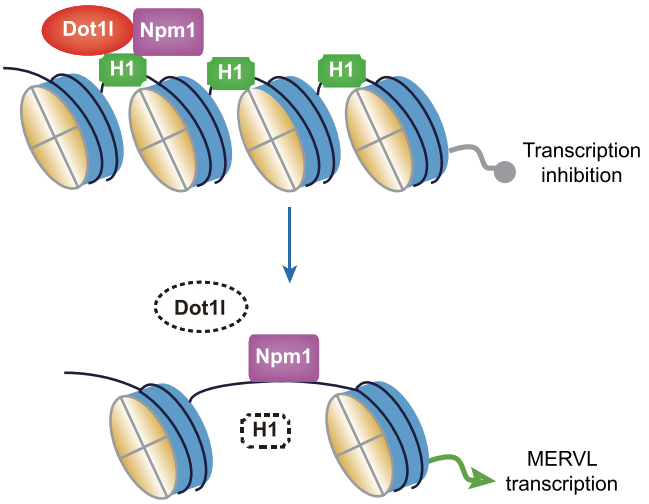
[1]State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin, Tianjin 300350, China and [2]State Key Laboratory of Medicinal Chemical Biology, College of Life Science, Nankai University, Tianjin, Tianjin 300071, China

Received May 10, 2022; Revised July 06, 2023; Editorial Decision July 16, 2023; Accepted July 20, 2023

ABSTRACT

Dot1l is a histone methyltransferase without a SET domain and is responsible for H3K79 methylation, which marks active transcription. In contradiction, Dot1l also participates in silencing gene expression. The target regions and mechanism of Dot1l in repressing transcription remain enigmatic. Here, we show that Dot1l represses endogenous retroviruses in embryonic stem cells (ESCs). Specifically, the absence of *Dot1l* led to the activation of MERVL, which is a marker of 2-cell-like cells. In addition, *Dot1l* deletion activated the 2-cell-like state and predisposed ESCs to differentiate into trophectoderm lineage. Transcriptome analysis revealed activation of 2-cell genes and meiotic genes by *Dot1l* deletion. Mechanistically, Dot1l interacted with and co-localized with Npm1 on MERVL, and depletion of *Npm1* similarly augmented MERVL expression. The catalytic activity and AT-hook domain of Dot1l are important to suppress MERVL. Notably, Dot1l-Npm1 restricts MERVL by regulating protein level and deposition of histone H1. Furthermore, Dot1l is critical for Npm1 to efficiently interact with histone H1 and inhibit ubiquitination of H1 whereas Npm1 is essential for Dot1l to interact with MERVL. Altogether, we discover that Dot1l represses MERVL through chaperoning H1 by collaborating with Npm1. Importantly, our findings shed light on the non-canonical transcriptional repressive role of Dot1l in ESCs.

GRAPHICAL ABSTRACT



INTRODUCTION

Embryonic stem cells (ESCs) are pluripotent cells derived from inner cell mass, which can differentiate into all adult cell types except extraembryonic tissues. A subset of ESCs was discovered to gain the ability to contribute to both inner cell mass and extraembryonic tissues. These cells exhibit an expression profile similar to that of 2-cell embryos and are marked by the presence of endogenous retrovirus MERVL, hence, they are named 2-cell-like cells (2CLCs) (1). Activation of MERVL alone in ESCs is sufficient to trigger the emergence of 2CLCs, suggesting a critical role of MERVL in the formation of 2CLCs (2). However, the detailed mechanism of MERVL activation still needs further study. One route to control MERVL is through epigenetic regulation. Previously, multiple epigenetic regulators have been characterized as repressors of MERVL in ESCs.

*To whom correspondence should be addressed. Tel: +86 22 85358316; Email: luxy@nankai.edu.cn
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.
Present address: Xin Zhao. Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Tongji Shanxi Hospital, Third Hospital of Shanxi Medical University, Taiyuan, Shanxi 030032, China.

One important class of epigenetic regulators of MERVL is H3K9 histone lysine methyltransferases with set domains. G9a, GLP, Setdb1 and Suv39h1 all contribute to the restriction of MERVL (3–5). However, whether other histone methyltransferases are involved in the modulation of retrotransposons is still unclear.

Among the histone lysine methyltransferases, one unique histone transferase is the disruptor of telomeric silencing 1-like (Dot1l) protein. Dot1l is the only histone lysine methyltransferase without a SET domain (6). It is also the only enzyme mediating the methylation of H3K79 (7,8), which is known as a marker of active euchromatin and expressed genes (9–11). Dot1l participates in the transcription elongation process by associating with RNA pol II and elongation factors (12,13). Paradoxically, Dot1l is essential for heterochromatin formation. Dot1l is specifically involved in the maintenance of telomere heterochromatin and gene silencing in yeast (14). In ESCs, Dot1l mutation results in the reduction of constitutive heterochromatin marked by H3K9me2 and H4K20me3 at centromeres and telomeres (7). Dot1l is absent in 2-cell embryo but the overexpression of Dot1l or its mutant without methyltransferase activity is still able to organize chromocenter-like structures in mouse 2-cell embryos (15). It is still unknown whether Dot1l is involved in the silencing of other genomic regions such as retrotransposons and how Dot1l performed its repressive role in ESCs.

Here, we show that Dot1l actively contributes to the silencing of retrotransposons. Loss of *Dot1l* activated MERVL, meiosis-related genes, and 2CLC state in ESCs. To better understand the mechanistic basis of these observations, we demonstrated that Dot1l interacted with histone chaperone Npm1 through its AT-hook domain. We further discovered that Npm1 and Dot1l co-localized on MERVL, and contributed to the restriction of MERVL in ESCs by chaperoning the histone H1 and its variants. Taken together, these findings expand our knowledge regarding the repressive action of Dot1l in transcription and the restriction of the MERVL-marked 2CLC state.

MATERIALS AND METHODS

Cell culture and inhibitor treatment

Mouse E3 ESCs were cultured on the feeder layer. E14 and J1 ESCs were cultured using 10 ng/ml recombinant LIF (Z03077, GenScript) on 0.2% gelatin-coated plates (353043, FALCON), in Dulbecco's modified Eagle's medium (DMEM, SH30243.01m, Hyclone) supplemented with 15% fetal bovine serum (FBS, SH30070.03, Hyclone), 2 mM L-glutamine (G0200, Solarbio),1% nonessential amino acids (N1250, Solarbio), 1% Penicillin–Streptomycin (P1400, Solarbio), and 0.1 mM β-mecaptoethanol (M3148-250, Sigma) in a humidified incubator at 37°C in 5% (v/v) CO_2. HEK 293T cells were cultured in DMEM (12100–046, Gibco) supplemented with 10% FBS (04-001-1A, Biological Industries), 1% L-glutamine (G0200, Solarbio), 1% Penicillin–Streptomycin (P1400, Solarbio) on 6-well plates (703001, NEST Biotechnology). For inhibitor treatment, ESCs were treated with 1 μM Dot1l inhibitor EPZ5676 (T3099, TargetMol) or 0.1 μM NAE inhibitor MLN4924 (S81085, MedMol) during cell passage and treated continuously for 48 h.

shRNA knockdown

For gene knockdown, the designed shRNAs were cloned into pSuper-puro vector and sequenced. During ESC passage, 1 μg shRNA plasmid was transfected with Polyjet (SL100688, SignaGen), according to the manufacturer's protocol. The transfected ESCs were selected with 1 μg/ml puromycin for 3 days and harvested. shRNAs sequences are available in Supplementary Table S1.

ESC differentiation

ESCs were differentiated to TSCs as previously reported (16). Briefly, after 24 h of ESCs passage, the medium was changed to TSCs differentiation medium, which contains: Roswell Park Memorial Institute (RPMI) 1640 medium (01–100-1ACS, Biological Industries) supplemented with 20% FBS (SH30070.03, Hyclone), 1 mM Sodium Pyruvate (SP0100, Solarbio), 2 mM L-glutamine (G0200, Solarbio), 1% Penicillin–Streptomycin (P1400, Solarbio), 0.1 mM β-mecaptoethanol (M3148-250, Sigma), 25 ng/ml human recombinant FGF4 (Z02984, GenScript) and 1 μg/ml heparin (S12004, Yuanye Biotechnology). The cells were harvested after inducing differentiation for 3 days.

Generation of *Dot1l*-knockout cell lines

Dot1l-knockout (*Dot1l*$^{-/-}$) ESCs were generated using CRISPR/Cas9 technology. Two specific sgRNAs targeting *Dot1l* genes were designed according to the published protocols. The sgRNAs were cloned into the pX459 vector. Transfection was carried out using Polyjet transfection reagent (SL100688, SignaGen). The transfected cells were selected with 2 μg/ml puromycin for 24h and then transferred to a 10 cm plate. One week later, the single-cell-derived clones were selected and the knockout was confirmed by Sanger sequencing of regions flanking deleted DNA. SgRNAs sequences are listed in Supplementary Table S1.

Establishment of overexpression ESC line

For the generation of *Dot1l* overexpression cell line, the full-length sequence of *Dot1l* was constructed into the pCAG-3HA vector. E14 ESCs and *Dot1l*$^{-/-}$ ESCs were transfected with Dot1l-HA vector or empty vector were selected in 800 μg/ml hygromycin B for ∼14 days. For the generation of truncated Dot1l mutant-overexpression cell line, the truncated sequence of *Dot1l* was cloned into the over vector flag vector respectively. Dot1l-ΔC-terminal mutant contained amino acid (a.a.) 1 to 660 of Dot1l. Dot1l-C-terminal contained a.a. 661 to 1540 of Dot1l. Dot1l-DOT1 & AT-hook contained a.a. 1 to 480 of Dot1l. Dot1l-DOT1 & Rootletin contained a.a. 1 to 320 and a.a. 481 to 660 of Dot1l. Dot1l-AT-hook & Rootletin contained a.a. 321 to 660 of Dot1l. Dot1l-DOT1 mutant contained a.a. 1 to 320 of Dot1l. Dot1l- AT-hook mutant contained a.a. 321 to 480 of Dot1l. Dot1l- Rootletin mutant contained a.a.

481 to 660 of Dot1l. *Dot1l*$^{-/-}$ ESCs were transfected with constructed vector or empty vector and were selected in 1 μg/ml puromycin for ∼10 days.

RNA purification and RT-qPCR

Total RNA from 1 to 2 million ESCs was harvested and extracted using RNAiso Plus (TaKaRa Bio) immediately after experiments and treated with 1 U DNase I (Thermo Fisher Scientific) at 37°C for 30 min in the presence of 40 U RNase inhibitor according to the manufacturer's protocols. The concentration and purity of total RNA was determined by NanoDrop 2000 (Thermo Scientific). 1 μg of DNase-treated RNA was used for reverse transcription with the Transcriptor First Strand cDNA Synthesis Kit (4897030001, Roche) and 2.5 μM Oligo-dT primers in a 10 μl reaction containing 5 U reverse transcriptase. The reverse transcription condition was: 25°C for 10 min, 50°C for 60 min, 85°C for 5 min and hold at 4°C. 10-fold diluted cDNA was used for qPCR analysis with 1 μM forward or reverse primer using Hieff qPCR SYBR Green Master Mix (11202ES08, Yeasen). Real-time qPCR was performed on a CFX384 Touch Real-Time PCR Detection Systems (Bio-Rad). Experiments were conducted for three independent samples harvested at different times with each data point run in duplicate. PCR cycling conditions were: 95°C for 10 min, 40 cycles of 95°C for 15 s, 60°C for 15 s and 72°C for 30 s. A melting curve of amplified DNA was subsequently acquired to assess the specificity of the primers. The comparative cycle threshold (Ct) method ($2^{-\Delta\Delta Ct}$) was used to quantify gene expression changes relative to Ct value in the control samples after normalization to the Ct of *Gapdh*. Primers with minimal off-target for genes were designed with Primer-Blast (17) to target exons of genes and amplify amplicon between 100 and 200 bp. For repeat sequences, primers were designed against the consensus sequence of repeats. Primers were synthesized by Azenta Life Sciences and sequences for qPCR analysis are listed in Supplementary Table S1.

Western blot analysis

Mouse ESCs were lysed in RIPA lysis buffer (R0010, Solarbio) supplemented with protease inhibitors PMSF (P0100, Solarbio) for 30 min on ice. Protein samples were separated by 8–15% SDS-PAGE gel and then transferred to PVDF membranes (Millipore). After blocking with 5% nonfat milk in PBS buffer containing 0.1% Tween-20 (PBST) for 2 h at room temperature, transferred membranes were incubated overnight at 4°C with different primary antibodies. After incubation with HRP-conjugated secondary antibodies, bands were visualized using a Chemiluminescence Imaging System (Tanon). Antibodies were used according to manufacturer's specifications. Primary antibodies are available in Supplementary Table S2. Secondary antibodies used are goat anti-rabbit IgG-HRP (sc-2004, Santa Cruz) and anti-mouse IgG-HRP (sc-516102, Santa Cruz).

Immunostaining and flow cytometry analysis

Immunostaining was performed as previously described (18). Briefly, cells were fixed with pre-cooled 80% ethanol and permeabilized with 0.5% Triton X-100 for 15 min on ice. Blocked cells were incubated with primary antibody against MERVL-gag (A-2801, EpiGentek) for 2 h on ice. The cells were then washed three times with PBST and incubated with secondary antibody Alexa Fluor 594-conjugated goat anti-rabbit IgG (ZF-0516, ZSGB-BIO) for 1.5 h on ice. The immunostained cells were analyzed by flow cytometry (BD LSRFortessa).

Protein co-immunoprecipitation (co-IP)

ESCs overexpressed Dot1l-HA or pCAG-3HA from 10-cm plate were washed with PBS twice and lysed with 1 ml of co-IP lysis buffer (20 mM Tris–HCl [pH 7.5], 20 mM KCl, 150 mM NaCl, 1.5 mM $MgCl_2$, 1% Glycine, 0.5% TritonX-100), containing protease inhibitors cocktail (B14001, Bimake). Cell extracts were centrifuged at 13 000 rpm for 15 min to remove insoluble debris. Precleared cell lysates were incubated overnight with anti-HA magnetic beads (B26202, Bimake) at 4°C. Beads were then washed three times with lysis buffer and twice with wash buffer. For co-IP in HEK293T cells, the truncated Dot1l domain-flag and Npm1-HA plasmids were co-transfected into HEK293T cells via Polyjet on 6-well plates. Two days after transfection, anti-Flag magnetic beads (B26102, Bimake) were used for immunoprecipitation. The rest of the experimental operation is the same as in ESCs. Bound proteins were eluted by SDS-PAGE loading buffer. Proteins were resolved on SDS-PAGE gel and analyzed by western blot.

Chromatin immunoprecipitation (ChIP)

ChIP was performed as previously described (18). Briefly, cells were cross-linked with 1% formaldehyde and terminated by adding glycine to a final concentration of 0.2 M. Cells were then lysed with ChIP lysis buffer and chromatin was fragmented by sonication using the Ultrasonic Cell Crusher. Next, the pretreated chromatin was incubated with 3 μg anti-Dot1l (D4O2T, Cell Signaling Technology), anti-Npm1 (sc-271737, Santa Cruz), anti-Histone H1.0 (A3298, ABclonal), anti-Histone H1.2 (A0646, ABclonal) antibodies loaded protein G MagBeads (L00274, GenScript) overnight at 4°C. After washing, elution and DNA purification, immunoprecipitated DNA were analyzed by qPCR.

Sequential chromatin immunoprecipitation (ChIP-reChIP)

For ChIP-reChIP experiments, chromatins isolated from E14 ESCs overexpressing Dot1l-HA were subjected to two rounds of ChIP. In the first ChIP, chromatin was equally divided for immunoprecipitation with control anti-Flag magnetic beads (B26102, Bimake) or anti-HA magnetic beads (B26202, Bimake). Immunoprecipitated chromatin was eluted with 80 μl elution buffer and diluted 10 times with ChIP dilution buffer, and a second ChIP was carried out with 3 μg of control IgG or anti-Npm1(sc-271737, Santa Cruz) antibodies. Purified DNA was analyzed by qPCR using primers for the indicated regions. All amplifications were performed in triplicate.

RNA-seq and data analysis

Total RNA was extracted from wild-type (WT) ESCs and *Dot1l*$^{-/-}$ ESCs with RNAiso Plus (TaKaRa Bio) and treated with DNase I (Thermo Fisher Scientific). At least 4 μg of total RNA was prepared for unstranded polyA-enriched RNA-seq libraries construction and sequencing. Each library was paired-end sequenced (150bp per end) to obtain 6 Gb data on Illumina Nova-seq by Azenta Life Sciences (Genewiz Incorporation). Cutadapt was used to remove adapter sequences of RNA-seq data and low-quality reads. Hisat2 was used to map RNA-seq reads to the mm10 mouse genome (19). Genes were annotated according to GENCODE database (20). Transposable elements (TEs) were annotated from UCSC Genome Browser (RepeatMasker). The number of reads mapped to genes was determined with featureCounts (21). The number of reads mapped to TEs and loci was determined with TEtranscripts and TElocal, respectively (22). Differential expression analysis and principal component analysis (PCA) was conducted by DESeq2 (23). Differentially expressed genes or repeats were selected with fold change > 1.5 and adjusted *P*-value < 0.05. The gene ontology (GO) analysis and KEGG pathway analysis were carried out with R package clusterProfiler (24). The Reactome pathway analysis was carried out with R package ReactomePA (25). Gene set enrichment analysis (GSEA) was done by the Python module gseapy. Hierarchical clustering analysis was performed for the different expressed genes using the Pheatmap package in R language. For analysis of MERVL-fusion transcripts, StringTie (26) was used to assemble RNA-seq reads into potential transcripts with default parameters. The visualization of RNA-seq tracks was performed using integrative genome visualizer (IGV) (27).

ChIP-seq and data analysis

For ChIP-seq, 2 ng of ChIP DNA was required for library preparation. ChIP-seq libraries were constructed with the ATseq kits according to the manufacturer's instructions. The libraries were sequenced on an Illumina Hiseq X Ten platform by Novogene Corporation (Tianjin, China) and 150 bp paired-end reads were generated. For ChIP-seq data analysis, and low-quality reads were filtered and the adapter sequences were removed with Cutadapt. Bowtie2 (28) was used to align reads to the mouse mm10 genome with parameters '–very-sensitive -k 5 –no-mixed –no-discordant', and then optimized the multi-read allocation by using CSEM (29). ChIP-seq signal heatmap and binding profile plot were generated by Deeptools (30). Macs2 (31) was used to call Dot1l binding peaks with parameters '–keep-dup auto -p 0.001 –broad –broad-cutoff 0.001'. Homer (32) and ChIPseeker (33) were used to annotate the peaks.

Statistical analysis

RT-qPCR and ChIP-qPCR results were analyzed by the Student's *t*-test. GraphPad Prism was used for the statistical analysis of data. $P < 0.05$ was considered statistically significant; these significant differences were defined as $^{*}P < 0.05$, $^{**}P < 0.01$ or $^{***}P < 0.001$. Differentially expressed genes and TEs were determined by Wald test; Benjamini-Hochberg procedure/false discovery rate (FDR) was used to adjust the *P*-value.

RESULTS

Dot1l restricts endogenous retrovirus MERVL in ESCs

To study the function of Dot1l in regulating retrotransposon MERVL, we first knocked down *Dot1l* with two independent shRNAs. Both shRNAs efficiently depleted *Dot1l* expression to 20–30% (Figure 1A). *Dot1l* decrement triggered the expression of MERVL (Figure 1B) without affecting the expression of pluripotency genes *Oct4*, *Sox2* and *Nanog* (Supplementary Figure S1A). To confirm the role of Dot1l, we established a *Dot1l*-knockout ESC line using Cas9 and gRNAs targeting exon 1 and exon 5 of *Dot1l* (Figure 1C; Supplementary Figure S1B). We verified the excision of first 5 exons of *Dot1l* with PCR (Supplementary Figure S1C and D). Our results from qPCR and western blot verified the loss of Dot1l protein (Figure 1D and E; Supplementary Figure S1E). *Dot1l* deletion did not disrupt ESC morphology nor the expression of pluripotency genes at both mRNA and protein levels (Supplementary Figure S1F-H). Since Dot1l is the only enzyme responsible for H3K79 methylation, we observed the loss of H3K79me1/2/3 after the deletion of *Dot1l* in ESCs (Figure 1F). In agreement with *Dot1l* depletion, *Dot1l*$^{-/-}$ ESCs activated MERVL, along with other retrotransposons (Figure 1G). Together, these results suggest that *Dot1l* is essential to repress retrotransposon MERVL.

Dot1l represses 2CLC state in ESCs

Given that Dot1l represses 2CLC marker MERVL, we asked whether Dot1l was diminished during the transition between ESCs and 2CLCs. Analysis of published RNA-seq data (34,35) revealed decreased *Dot1l* during 2CLC induction and upregulated *Dot1l* upon exit from 2CLC state (Figure 2A and B). We next examined the expression of MERVL encoded gag protein. We found that MERVL-gag was activated in *Dot1l*$^{-/-}$ ESCs (Figure 2C; Supplementary Figure S2A). The MERVL-gag^{+} 2CLC population also increased in *Dot1l*$^{-/-}$ ESCs (Figure 2D). In addition, *Dot1l*$^{-/-}$ ESCs exhibited stronger activation of trophectodermal genes (*Cdx2*, *Eomes*, *Plet1*, and *Ascl2*) than wild-type ESCs upon directed differentiation of ESCs to trophoblast stem cells (TSCs) (Figure 2E; Supplementary Figure S2B). Furthermore, *Dot1l*$^{-/-}$ ESCs augmented the expression of 2CLC genes (Figure 2F), confirming the acquisition of the 2CLC state in *Dot1l*$^{-/-}$ ESCs. We further re-introduced Dot1l in *Dot1l*$^{-/-}$ ESCs and observed the restoration of *Dot1l* expression at both protein and RNA levels (Figure 2G and H), which was accompanied by the decrement of MERVL expression but not pluripotency genes (Figure 2I; Supplementary Figure S2C). These evidences support that Dot1l restricts 2CLC state in ESCs.

Dot1l regulates transcriptome change in ESCs

Next, we examined the transcriptome of ESCs after the loss of *Dot1l*. The expression of 1044 genes increased while

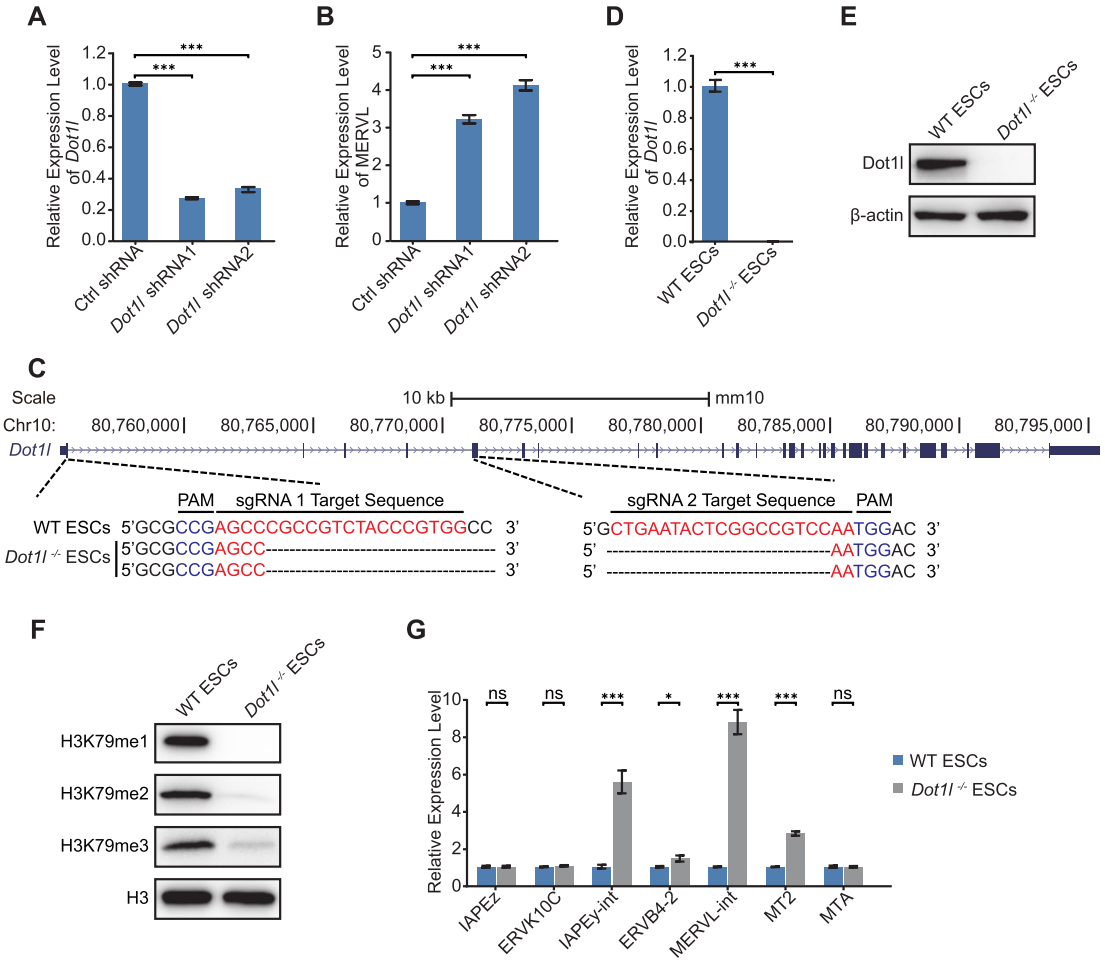


Figure 1. Dot1l deficiency activates retrotransposons. (**A**) qPCR analysis of the expression of *Dot1l* in ESCs treated with control (Ctrl) shRNA or *Dot1l* shRNAs. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**B**) qPCR analysis of the expression of MERVL after the depletion of *Dot1l* in ESCs. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**C**) Schematic of mutation sites in *Dot1l*$^{-/-}$ ESCs. Black dashes: deleted bases; red bases: sgRNA target sequences; blue bases: protospacer adjacent motif (PAM) sequences. (**D**) qPCR analysis of the expression of *Dot1l* in WT ESCs and *Dot1l*$^{-/-}$ ESCs. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**E**) Western blot analysis of Dot1l protein in WT ESCs and *Dot1l*$^{-/-}$ ESCs. β-Actin was used as a loading control. (**F**) Western blot analysis of H3K79me1/2/3 level in WT ESCs and *Dot1l*$^{-/-}$ ESCs. H3 was used as a loading control. (**G**) qPCR analysis of the expression of retrotransposons in WT ESCs and *Dot1l*$^{-/-}$ ESCs. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). ns: non-significant, $^{*}P < 0.05$, $^{***}P < 0.001$ in Student's *t*-test.

only 244 genes were downregulated in *Dot1l*$^{-/-}$ ESCs (Figure 3A). For TEs, MERVL/MT2, IAPEy-int, MamGyp-int, RLTR45-int and LTR31 were upregulated in *Dot1l*$^{-/-}$ ESCs (Figure 3B; Supplementary Figure S3A). In contrast, only RLTRETN were downregulated after the loss of *Dot1l* (Figure 3B). The number of activated loci was the highest for MERVL/MT2 among all TEs with increased expression (Figure 3C). Consistent with qPCR data, 2-cell genes were enriched among upregulated genes in *Dot1l*$^{-/-}$ ESCs (Figure 3D). In addition, we observed activation of MERVL/MT2-fusion genes in *Dot1l*$^{-/-}$ ESCs (Figure 3E; Supplementary Figure S3B). Principal component analysis (PCA) revealed that gene expression of *Dot1l*$^{-/-}$ ESCs is more closely correlated with that of 2CLCs than previously identified chemical-induced totipotent cells (36,37) (Supplementary Figure S3C). We subsequently examined the classes of genes affected by gene ontology (GO) analysis. Interestingly, meiosis and reproduction-related genes were boosted (Figure 3F; Supplementary Figure S3D) whereas genes related to developmental processes were downregulated (Figure 3G). KEGG and Reactome enrichment analysis revealed upregulation of metabolism and cell junction-related genes as well as downregulation of signalling pathways regulating pluripotency and FGF signalling in *Dot1l*$^{-/-}$ ESCs (Supplementary Figure S3E–H). Together, these results show that Dot1l suppresses 2-cell gene expression and MERVL at the transcriptome level.

Dot1l collaborates with Npm1 to suppress MERVL

Dot1l was often associated with the activation of gene expression, however, we found that Dot1l repressed retrotransposons such as MERVL in ESCs. To investigate how Dot1l suppresses MERVL, we examined the potential interactors of Dot1l. It was previously reported that Mllt10 (AF10), Mllt6 (AF17) and Mllt3 (AF9) formed a complex with Dot1l (38). We expect that to allow the expression of

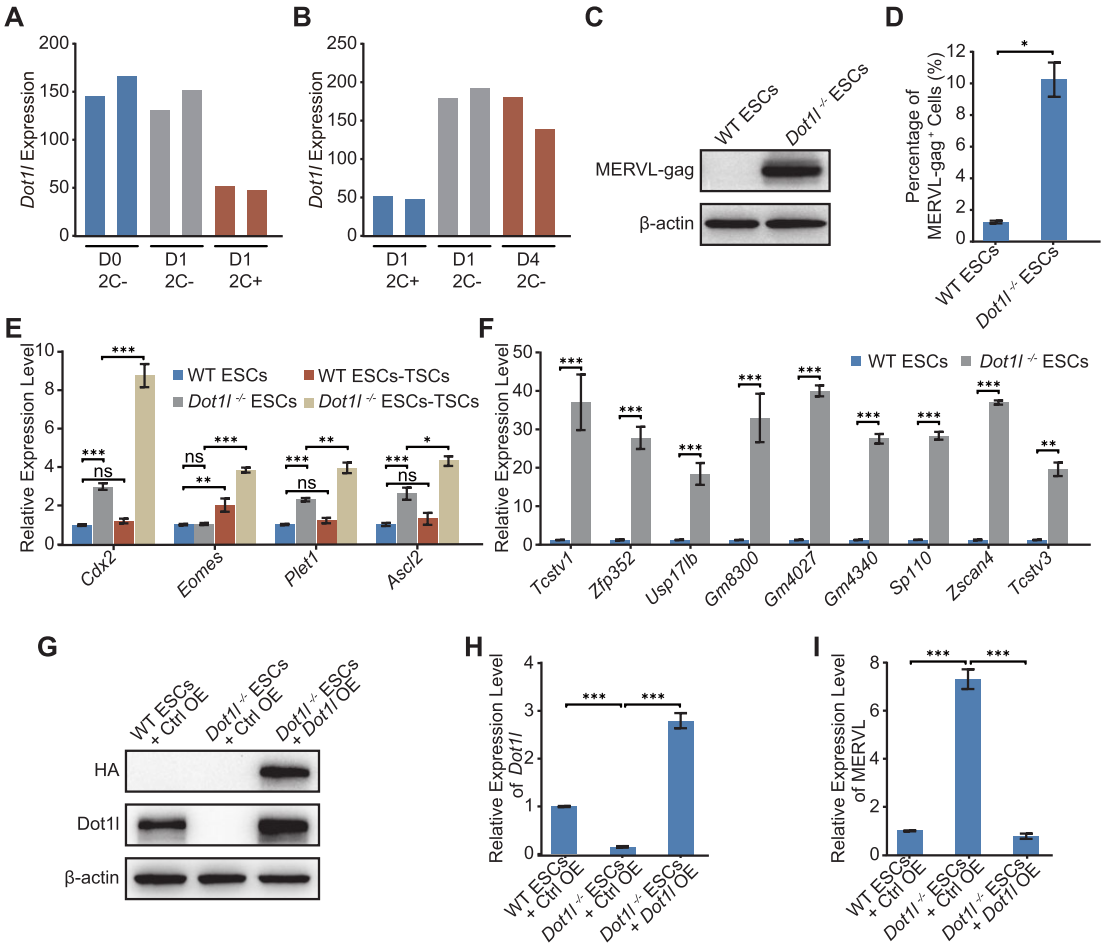


Figure 2. Loss of *Dot1l* transformed ESCs to 2CLCs. (**A**) The expression of *Dot1l* in the entry to 2CLC state. D0 2C$^-$ refers to MERVL-negative population before Dux induction, D1 2C$^-$ and D1 2C$^+$ represent MERVL-negative and MERVL-positive populations after 1 day Dux induction respectively. (**B**) The expression of Dot1l during the exit from 2CLC state. D1 2C$^+$ represents 2C-like cells after 1 day of Dux induction. D4 2C$^-$ represents D1 2C$^+$ cells cultured for 3 days to exit the 2C-like state. 2C$^-$, MERVL-negative cells; 2C$^+$, MERVL-positive 2CLCs. (**C**) Western blot analysis of MERVL-gag level in WT ESCs and *Dot1l*$^{-/-}$ ESCs. β-Actin was used as a loading control. (**D**) Flow cytometry analysis of the MERVL-gag$^+$ population in WT ESCs and *Dot1l*$^{-/-}$ ESCs. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**E**) qPCR analysis of the expression of TSC genes in WT ESCs and *Dot1l*$^{-/-}$ ESCs differentiated into TSCs respectively. qPCR data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**F**) qPCR analysis of the expression of 2-cell embryo genes in WT ESCs and *Dot1l*$^{-/-}$ ESCs. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**G**) Immunoblot analysis of the expression of Dot1l after overexpression of HA-Dot1l in WT ESCs and *Dot1l*$^{-/-}$ ESCs. β-Actin was included as a loading control. Ctrl OE: control vector overexpression. (**H, I**) qPCR analysis of the expression of *Dot1l* (H) and MERVL (I) after overexpression of *Dot1l* in *Dot1l*$^{-/-}$ ESCs. Ctrl OE: control vector overexpression. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). ns: non-significant, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ in Student's *t*-test.

MERVL, the co-factors of Dot1l should be lowly expressed at 2-cell stage compared to other stages of pre-implantation development, since Dot1l is absent at 2-cell stage (15,36). However, we did not observe a specific downregulation of *Mllt10*, *Mllt6* and *Mllt3* at 2-cell stage (Supplementary Figure S4A). Depletion of these interactors of Dot1l did not activate MERVL (Figure 4A and B). Surprisingly, knockdown of *Npm1*, another major interacting protein of Dot1l (39), activated the expression of MERVL, while its expression was also maintained at low level in 2-cell embryos (Figure 4C–E). This can be validated in another ESC line (J1) (Supplementary Figure S4B and C). In addition, *Npm1* deficiency elevated the expression of 2CLC genes (Figure 4F). Thus, we further examined the chromatin localization of Dot1l and Npm1 by chromatin immunoprecipitation (ChIP). Our ChIP-qPCR results found that both Dot1l and Npm1 were bound to MERVL (Figure 4G and H). ChIP-seq revealed that most Dot1l and Npm1 were bound to intergenic and intronic regions (Supplementary Figure S4D and E). Dot1l and Npm1 displayed similar enrichment patterns on MERVL (Figure 4I and J). In addition, Dot1l and Npm1 have a similar binding pattern on genes (Supplementary Figure S4F). Both Dot1l and Npm1 were also enriched on other upregulated TEs (IAPEy-int and RLTR45-int) and downregulated RLTRETN (Supplementary Figure S4G–I). Moreover, the Npm1 binding is enriched on Dot1l ChIP-seq peaks, indicating their co-localization on chromatin (Figure 4K). Furthermore, sequential ChIP–qPCR (ChIP-reCHIP) experiments elucidated the co-localization of Npm1 and Dot1l on MERVL (Figure 4L). These results demonstrate that Dot1l interacts with Npm1 to co-repress MERVL.

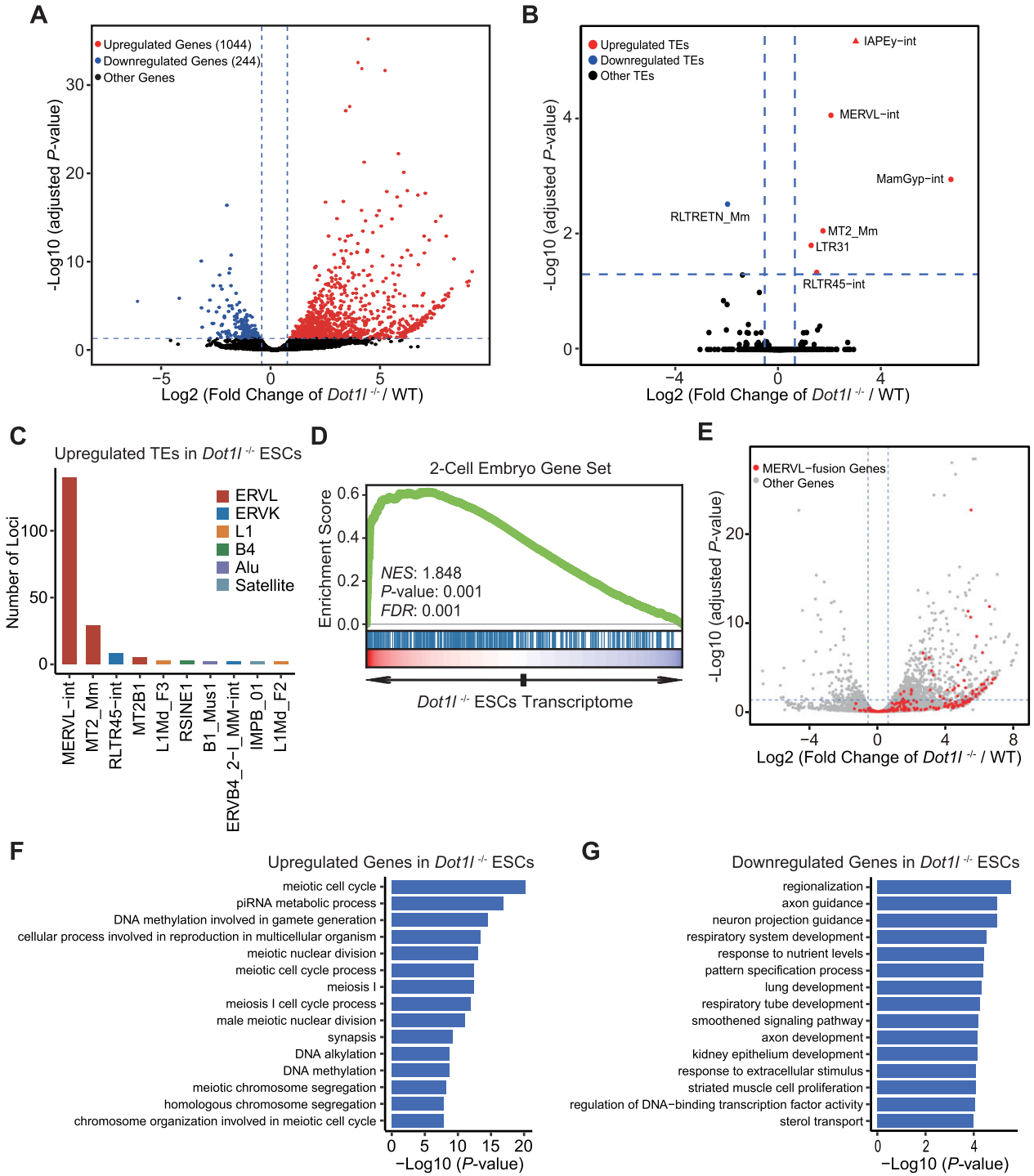


Figure 3. Transcriptome regulated by Dot1l in ESCs. (**A**) The volcano plot of gene expression in *Dot1l*$^{-/-}$ ESCs versus WT ESCs. Red, up-regulated genes; blue, down-regulated genes; black, other genes. Genes with expression change ≥1.5-fold and adjusted $P < 0.05$ are shown. (**B**) The volcano plot shows transcriptome analysis of TEs expression after *Dot1l* knockout. Red, upregulated TEs; blue, downregulated TEs; black, other TEs; adjusted $P < 0.05$, Wald test. (**C**) The TEs with the highest number of loci upregulated in *Dot1l*$^{-/-}$ ESCs. (**D**) Gene set enrichment analysis (GSEA) of 2-cell genes in the transcriptome of *Dot1l* $^{-/-}$ ESCs. Red, up-regulated genes; blue, down-regulated genes; *NES*, normalized enrichment scores; *FDR*, false discovery rate. The Kolmogorov–Smirnov statistic was used for the calculation of the *P*-value. (**E**) The volcano plot of all expressed genes in WT ESCs and *Dot1l*$^{-/-}$ ESCs. Genes fused with MERVL are labelled in red. (**F, G**) Gene ontology (GO) analysis of upregulated genes (F) and downregulated genes (G) after *Dot1l* knockout. The analysis was done with clusterProfiler.

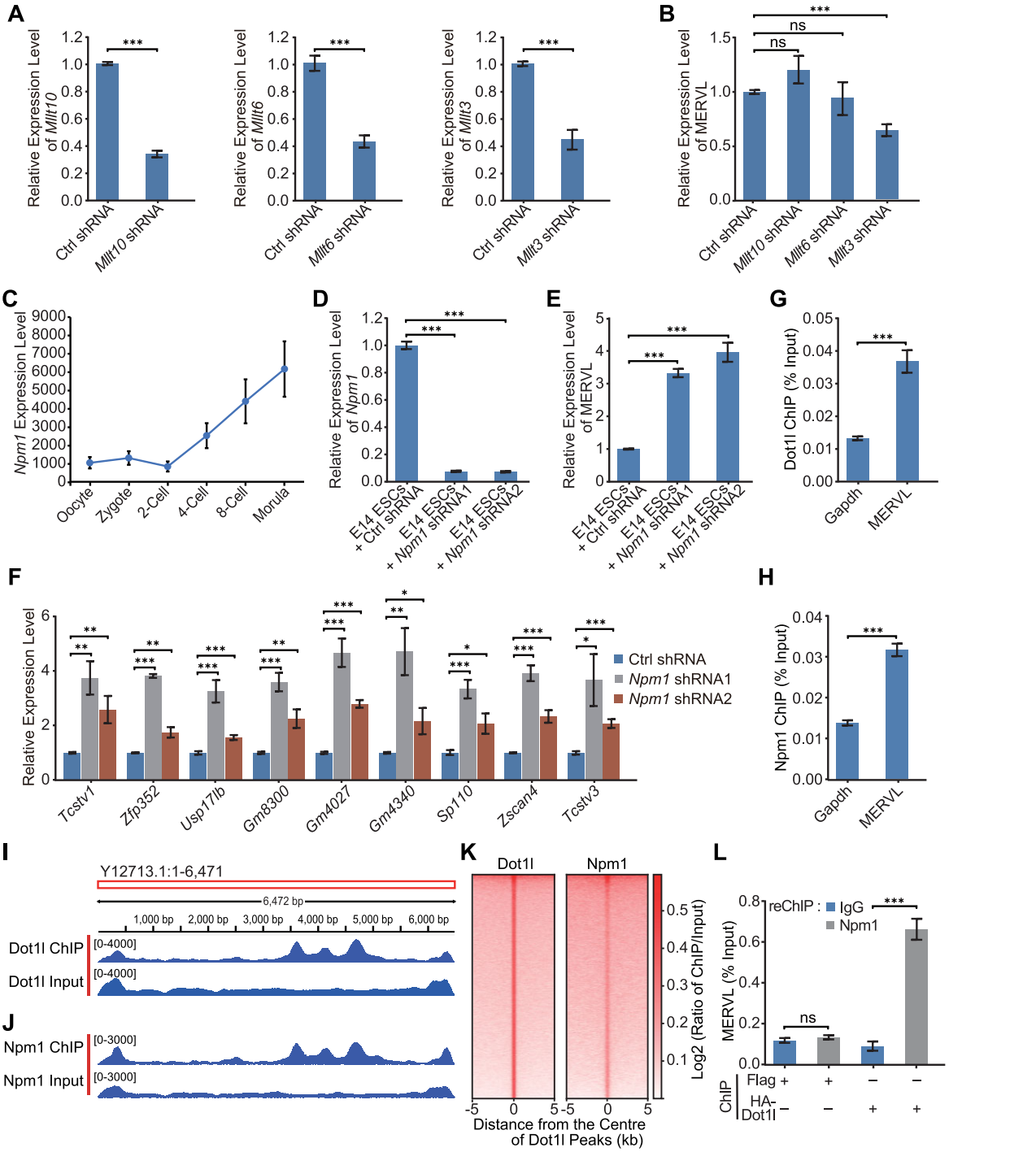


Figure 4. Dot1l interacts with Npm1 to repress MERVL. (**A**) qPCR analysis of the expression of *Mllt10*, *Mllt6* and *Mllt3* after transfected with control (Ctrl) shRNA and shRNAs against *Mllt10*, *Mllt6* and *Mllt3*. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**B**) qPCR analysis of the expression of MERVL after transfected with control (Ctrl) shRNA and shRNAs against *Mllt10*, *Mllt6* and *Mllt3*. qPCR data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**C**) The expression level of *Npm1* during early embryogenesis according to published RNA-seq data (71). (**D**) qPCR analysis of the expression of *Npm1* in E14 ESCs treated with control (Ctrl) shRNA or *Npm1* shRNAs. qPCR data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**E**) qPCR analysis of the expression of MERVL after the depletion of *Npm1* in E14 ESCs. qPCR data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**F**) qPCR analysis of the expression of 2-cell embryo genes in E14 ESCs treated with control (Ctrl) shRNA or *Npm1* shRNAs. qPCR data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**G, H**) ChIP-qPCR analysis of Dot1l (G) and Npm1 (H) binding on MERVL. ChIP-qPCR data were normalized to input and that of the control region. Data are presented as mean ± s.e.m. (n = 3 independent experiments). (**I, J**) Visualization the Dot1l (I) and Npm1 (J) ChIP-seq peaks mapped to MERVL consensus sequence in IGV. (**K**) Enrichment heatmap of Npm1 around the binding regions of Dot1l. The ChIP-seq signal was calculated as the $\log_2$ ratio of normalized reads relative to the input. (**L**) A ChIP-reChIP assay was performed to detect the co-occupancy of HA-Dot1l and Npm1 on MERVL. ChIP DNA was analyzed by qPCR using primers of MERVL. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). ns: non-significant, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ in Student's *t*-test.

The methyltransferase activity and AT-hook domain are vital to the transcriptional repression by Dot1l

To further study which domain(s) of Dot1l is responsible for Dot1l to repress transcription, we overexpressed Dot1l mutants with each domain deleted and individual domains of Dot1l to rescue *Dot1l* knockout (Figure 5A; Supplementary Figure S5A–E). The removal of the C-terminal domain of Dot1l did not disrupt its ability to repress MERVL (Figure 5A and B). None of the individual domains of Dot1l exhibited repression activity against MERVL (Figure 5A and B). Overexpression of Dot1l with the deletion of rootletin domain and C-terminal domain at the same time, leaving only DOT1 and AT-hook domain, can efficiently rescue MERVL in *Dot1l*$^{-/-}$ ESCs, suggesting that rootletin domain and C-terminal domain are unnecessary whereas DOT1 plus AT-hook domain is sufficient for the transcriptional repression by Dot1l (Figure 5A and B). In contrast, the absence of either DOT1 or AT-hook domain sabotaged the ability of Dot1l to repress MERVL. These findings prove that the DOT1 domain and AT-hook domain together are required to repress MERVL (Figure 5A and B). Given that the DOT1 domain carries the methyltransferase activity of Dot1l, we asked whether the catalytic role of Dot1l was necessary for Dot1l to repress MERVL. Interestingly, mutations of the Dot1l catalytic site disabled Dot1l to repress MERVL (Figure 5C and D; Supplementary Figure S5F) without disturbing the expression of pluripotency genes (Supplementary Figure S5G). The importance of Dot1l methyltransferase activity was ascertained by the treatment of different ESC lines with Dot1l inhibitors (Figure 5E–G). Impeding Dot1l activity elevated MERVL expression in three different ESC lines (E14, J1 and E3), mimicking the effect of *Dot1l* deletion (Figure 5E–G). We further examined which domain was necessary for Dot1l to interact with Npm1. Npm1 interacted with the full-length Dot1l and mutants with the deletion of DOT1 domain, rootletin domain, or C-terminal domain (Figure 5H and I). However, the removal of the AT-hook domain completely abolished the interaction between Dot1l and Npm1 (Figure 5H and I). These results evidence that Dot1l activity and AT-hook domain, which interacts with Npm1, are imperative to the repression role of Dot1l.

Dot1l and Npm1 restrict MERVL through histone H1

Npm1 is a multifunctional protein with chaperone activity towards histone H1 (40,41). It can assist in the deposition of histone H1 to DNA (41). Besides regulating histone deposition, histone chaperones play critical parts in the regulation of histone stability and recycling (42). Hence, we examined whether *Npm1* knockdown affected histone H1 expression. Neither reduced *Npm1* nor the absence of *Dot1l* influenced mRNA expression of histone H1.0 and its variant H1.2 (Figure 6A and B). Interestingly, *Npm1* depletion reduced the protein level of H1.0 and H1.2 (Figure 6C). Additionally, previous discoveries prove that Dot1l interacts with histone H1 family proteins (39). Therefore, we asked whether Dot1l repressed MERVL by working with Npm1 to influence histone H1 protein expression and deposition. Both *Npm1* depletion and *Dot1l* loss reduced the protein level of histone H1 and its variant H1.2 (Figure 6C and D). Furthermore, the deletion of *Dot1l* resulted in the reduction of histone H1.0 and H1.2 enrichments on MERVL (Figure 6E and F). Histone H1 and H1.2 were recently shown to be crucial repressors of 2-cell genes and 2CLC state (43). Our analysis of published RNA-seq data (43) after the loss of histone H1 variants (H1.2, H1.3 and H1.4) confirmed that histone H1 variants were also critical to the repression of retrotransposons such as MERVL (Figure 6G and H; Supplementary Figure S6A). Chromatin on MERVL region became more open upon the loss of histone H1 variants (Supplementary Figure S6B). Therefore, it is possible for Dot1l to regulate MERVL through histone H1. We also attempted to identify whether it is possible for Dot1l to regulate MERVL expression indirectly through known repressors. We first examined the RNA-seq expression of 100 reported MERVL repressors after *Dot1l* deletion (44–52) (Supplementary Figure S6C). We further used qPCR to validate genes which were downregulated for > 1.5-fold, and Hdac4/5. We found that only *Hdac4* and *Hdac5* were slightly downregulated after *Dot1l* loss (Supplementary Figure S6C and D). *Hdac4* and *Hdac5* were recently found to cooperatively repress MERVL and 2CLC state (52). Hence, their simultaneous downregulation implies that Dot1l is also able to partially repress MERVL indirectly.

Dot1l acts through Npm1 to regulate histone H1

We next assessed how histone H1 protein was affected by *Dot1l* deletion, using H1.2 as an example. We initially focused on whether the catalytic activity of Dot1l is necessary for H1.2 loading and stability. We found that a catalytic domain-inactivated Dot1l was unable to rescue H1.2 protein level (Figure 7A) or its depositions on MERVL (Figure 7B; Supplementary Figure S7A), which indicated a crucial role of Dot1l catalytic activity. The catalytic inactivation of Dot1l neither affected the binding of itself nor Npm1 to MERVL (Figure 7C and D; Supplementary Figure S7B and C). However, we observed that the interaction between H1.2 and Npm1 was affected by *Dot1l* deletion, even after taking into consideration the reduction of H1.2 in *Dot1l*$^{-/-}$ ESCs (Figure 7E). Since Npm1 is a histone chaperone that can mediate the recruitment of histone H1 to chromatin (41), the reduced interaction between Npm1 and H1.2 would result in decreased loading of H1.2 to MERVL. In addition, simultaneous depletion of *Dot1l* and *Npm1* did not further augment MERVL than *Dot1l* or *Npm1* depletion alone, strongly indicating that Npm1 and Dot1l act in the same pathway to repress MERVL (Figure 7F; Supplementary Figure S7D). Furthermore, *Npm1* depletion disrupted the interaction between Dot1l and MERVL (Figure 7G), suggesting that Npm1 is required for Dot1l recruitment to MERVL. These results reveal that Dot1l and Npm1 function together to repress MERVL transcription.

Dot1l regulates histone H1 stability via ubiquitination

As histone H1 stability can be regulated by histone ubiquitination, we examined the ubiquitination of H1.2 after *Dot1l* loss. We discovered that the ubiquitinated H1.2 level increased in *Dot1l*$^{-/-}$ ESCs (Figure 8A). Polyubiquitinated

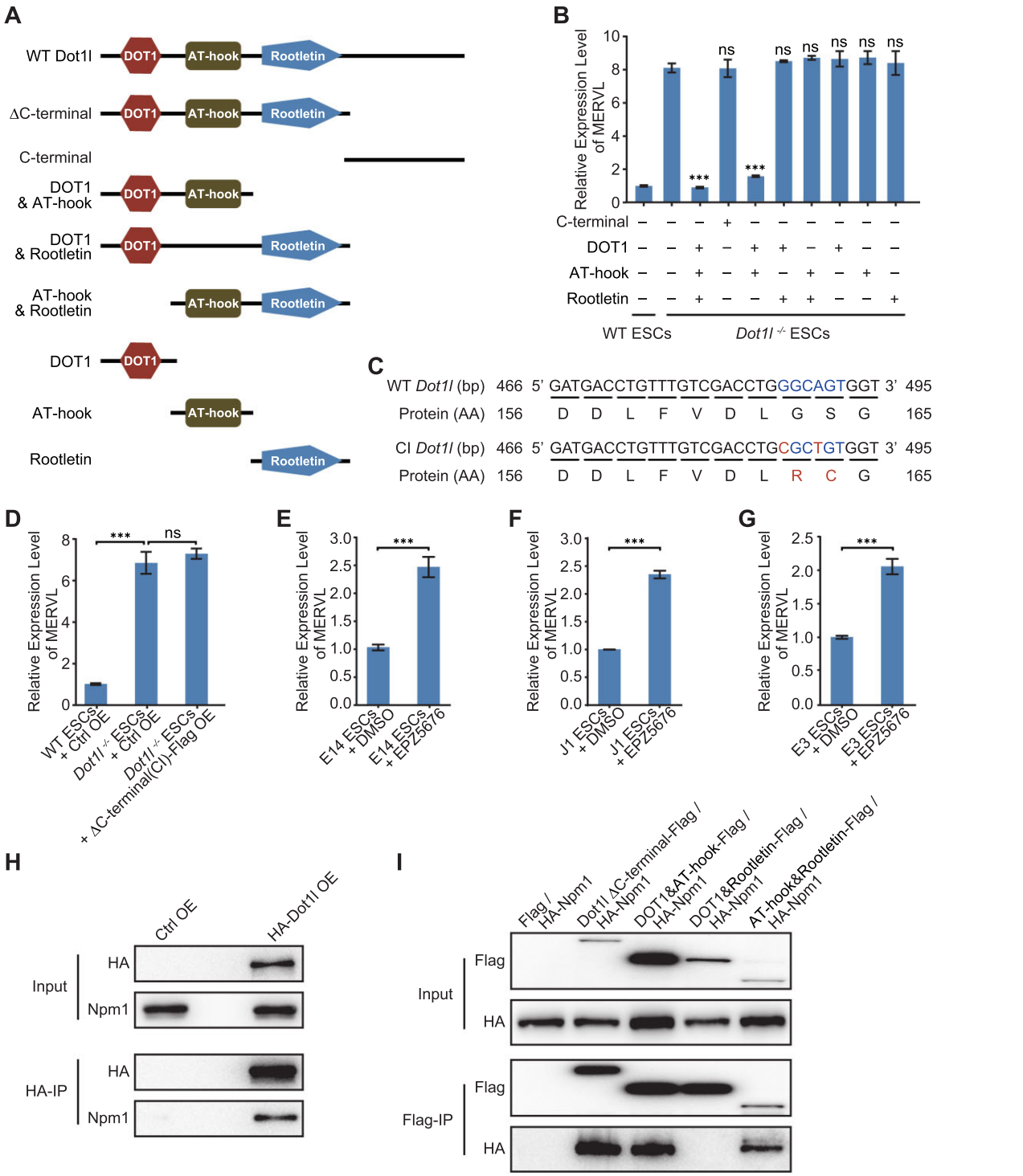


Figure 5. Distinct roles of individual Dot1l domains. (**A**) A schematic summary of Dot1l mutants used for functional rescue. Δ, deletion. (**B**) qPCR analysis of MERVL expression after the rescue with the overexpression of Dot1l mutants in *Dot1l*$^{-/-}$ ESCs. qPCR results were normalized to *Gapdh* and compared with the overexpressed control vector in *Dot1l*$^{-/-}$ ESCs to analyze whether the differences are statistically significant. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**C**) Sequences of WT (top) and CI (bottom) *Dot1l* alleles at exon 5. Mutated bases and amino acid residues are in red. WT: wild-type; CI: catalytic inactive; AA: amino acid. (**D**) qPCR analysis of MERVL expression after rescue of Dot1l ΔC-terminal (CI) mutant in *Dot1l*$^{-/-}$ ESCs. Ctrl OE: control vector overexpression; CI: catalytic inactive. qPCR data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**E–G**) qPCR analysis of the expression of MERVL in E14 ESCs (E), J1 ESCs (F) and E3 ESCs (G) treated with Dot1l inhibitor (EPZ5676). DMSO treated samples were included as a control. qPCR data are presented as mean ± s.e.m. ($n = 3$ independent experiments). ns: non-significant, ***$P < 0.001$ in Student's *t*-test. (**H**) Western blot analysis of HA-Dot1l/Npm1 co-immunoprecipitation in ESCs overexpressing control empty vector (Ctrl OE) or HA-tagged Dot1l (HA-Dot1l OE). IP was done with anti-HA magnetic beads. IP, immunoprecipitation. OE, overexpression. (**I**) Co-IP experiments to detect the interactions of Flag-tagged Dot1l mutants with HA-tagged Npm1 in HEK293T cells respectively, IP was done with anti-Flag magnetic beads. The resulting co-IP samples were subjected to western blot.

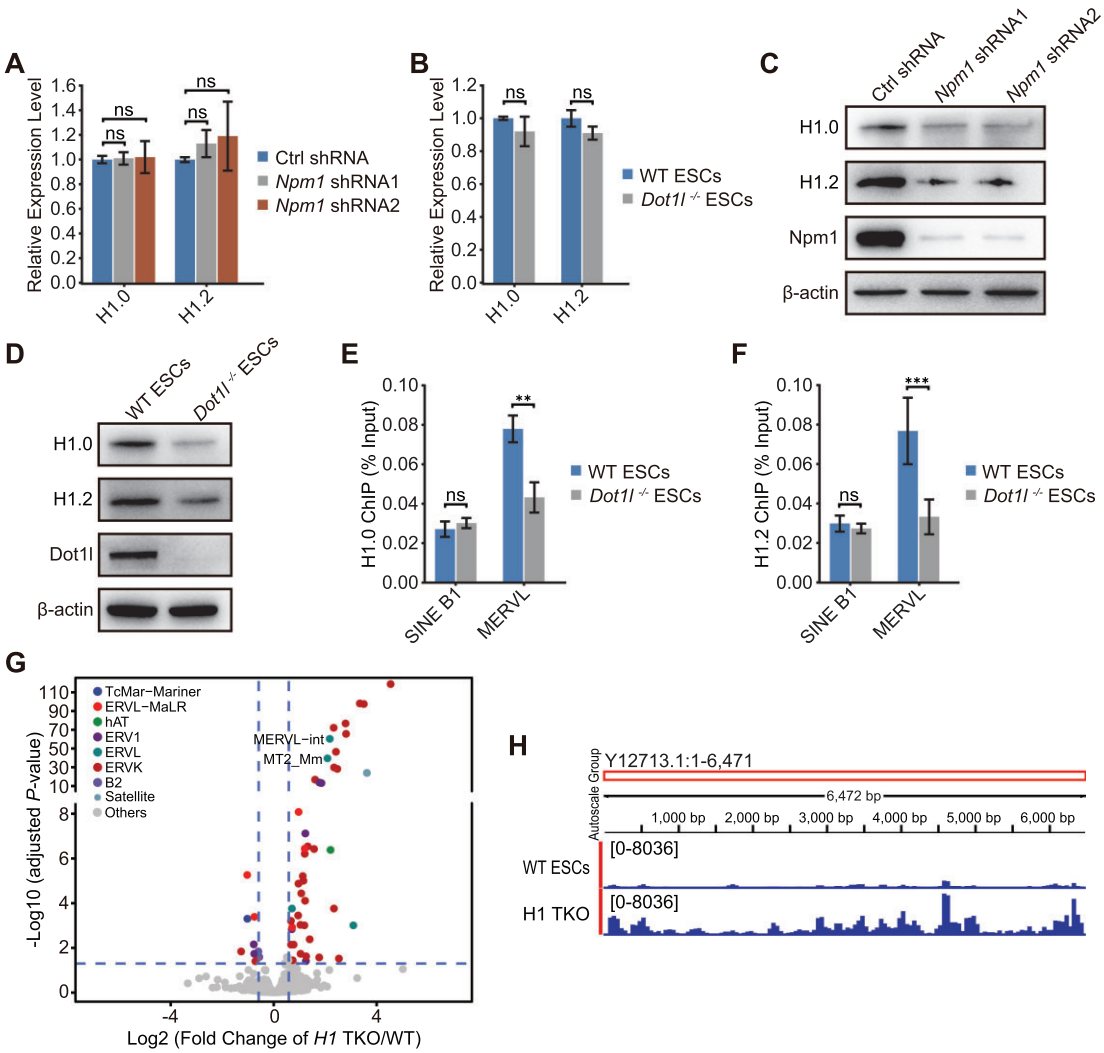


Figure 6. Dot1l and Npm1 regulate histone H1 protein level and its loading onto chromatin. (**A**) qPCR analysis of the expression of *H1.0* and *H1.2* in E14 ESCs treated with control (Ctrl) shRNA or *Npm1* shRNAs. qPCR data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**B**) qPCR analysis of the expression of *H1.0* and *H1.2* in WT ESCs and *Dot1l*$^{-/-}$ ESCs. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**C**) Western blot analysis of H1.0, H1.2 and Npm1 protein in WT ESCs treated with control (Ctrl) shRNA or *Npm1* shRNAs. β-Actin was used as a loading control. (**D**) Western blot analysis of H1.0, H1.2 and Dot1l protein in WT ESCs and *Dot1l*$^{-/-}$ ESCs. β-Actin was used as a loading control. (**E**, **F**) ChIP-qPCR analysis of H1.0 (E) and H1.2 (F) binding on different retrotransposons. ChIP-qPCR data were normalized to input and that of the control region. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). ns: non-significant, $^{**}P < 0.01$, $^{***}P < 0.001$ in Student's *t*-test. (**G**) The volcano plot shows transcriptome analysis of TEs expression after *H1* TKO (triple-knockout of *H1.2*, *H1.3* and *H1.4*). The result from DEseq2 was used to plot the diagram. Different colored dots represent different retroelement families. Colored dots indicate TEs with significant expression change (adjusted $P < 0.05$, Wald test). (**H**) The RNA-seq reads mapped to MERVL in WT ESCs and ESC with *H1* TKO (triple knockout of *H1.2*, *H1,3* and *H1.4*) were visualized using IGV.

proteins can be degraded by the proteasome. NEDD8-activating enzyme (NAE) controls the activity of the cullin RING subtype of ubiquitin ligases upstream of the proteasome (53). Thus, we explored whether NAE participated in this process. Notably, inhibiting NAE with MLN4924 impeded H1.2 ubiquitination and partially stabilized the protein level of H1.2 (Figure 8B and C), thereby restoring the expression of MERVL in *Dot1l*$^{-/-}$ ESCs (Figure 8D). This suggests a critical role of Dot1l in regulating MERVL expression via H1.2 ubiquitination. Subsequently, we screened H1.2 ubiquitination sites, which may affect chromatin conformation (54,55) and found that only overexpressing H1.2 carrying K64R mutation partially rescued MERVL (Figure 8E; Supplementary Figure S8A–E). We further confirmed that H1.2 ubiquitination decreased after overexpressing H1.2K64R (Figure 8F). These results suggest that Dot1l regulates H1.2 stability and the subsequent expression of MERVL through ubiquitination.

DISCUSSION

In summary, we propose a model that Dot1l acts through Npm1 to control the expression and deposition of histone H1 variants to repress MERVL and 2CLC state (Figure 8G). The relatively stable H3K79 methylation deposited by Dot1l is usually associated with transcription activation (10). Here, we illustrate that *Dot1l* participates in the repression of genes and retrotransposons (Figure 3A and B). A

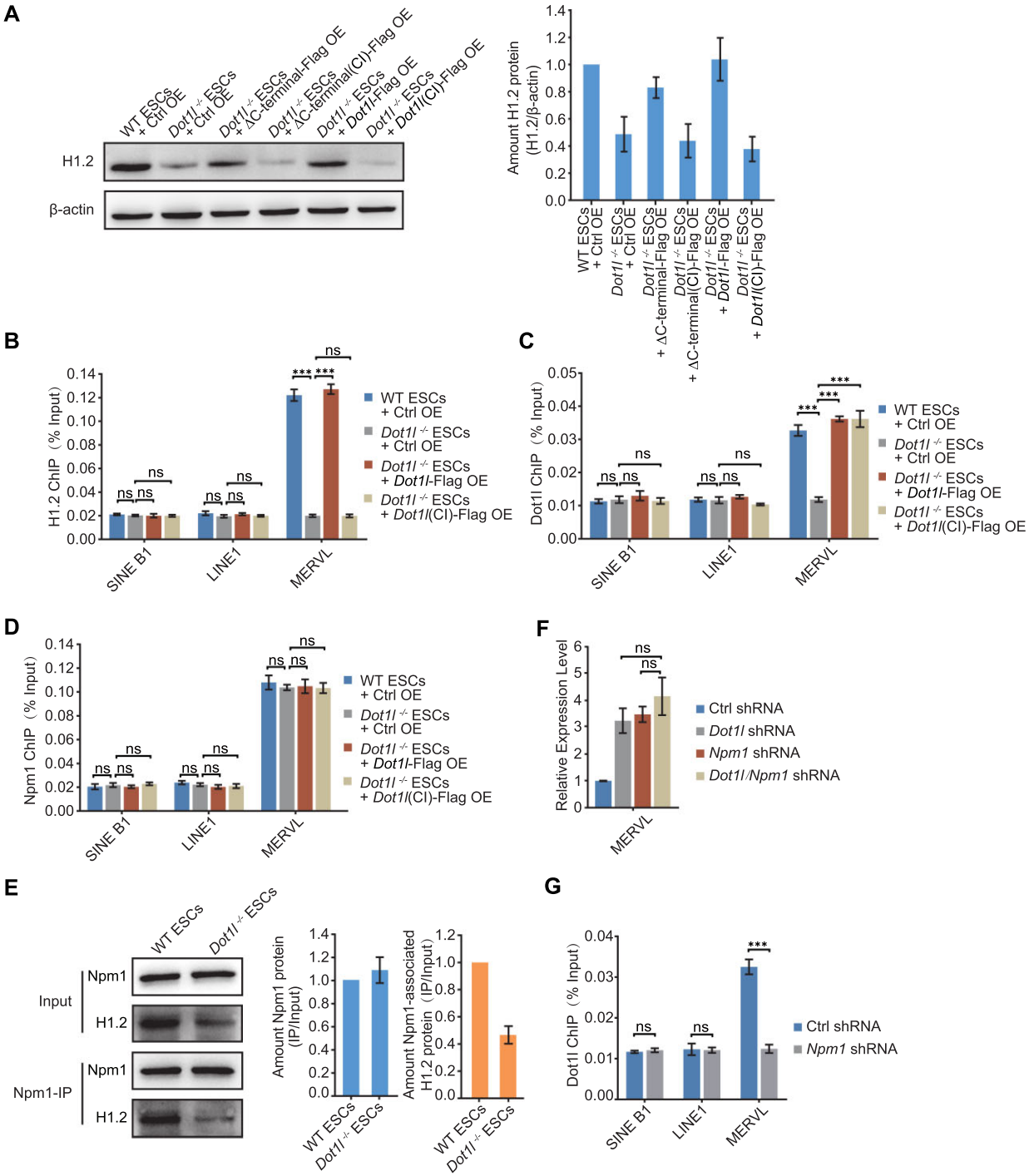


Figure 7. Dot1l regulates interaction between Npm1 and histone H1. (**A**) Western blot analysis of H1.2 protein after rescue with catalytic-active or inactive (CI) Dot1l mutants in *Dot1l*$^{-/-}$ ESCs (left) and quantification (right). β-Actin was used as a loading control. The ratio of H1.2/β-actin protein from *Dot1l*$^{-/-}$ samples was normalized to that from WT ESCs + Ctrl OE sample, which is set as 1. Data are presented as mean ± s.e.m. (n = 3 independent experiments). Ctrl OE: control vector overexpression; CI: catalytic inactive; Δ, deletion. (**B**) ChIP-qPCR analysis of H1.2 binding on MERVL. ChIP-qPCR data were normalized to input and that of the control region. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**C**) ChIP-qPCR analysis of Dot1l and Dot1l (CI)-Flag binding on MERVL. ChIP-qPCR data were normalized to input and that of the control region. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**D**) ChIP-qPCR analysis of Npm1 binding on MERVL. ChIP-qPCR data were normalized to input and that of the control region. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**E**) Western blot analysis of Npm1/H1.2 co-immunoprecipitation in WT ESCs or *Dot1l*$^{-/-}$ ESCs (left) and quantification (right). IP, immunoprecipitation. IP was done with anti-Npm1 antibody. 1.25% input was loaded as a control. The ratio of IP protein/input protein from *Dot1l*$^{-/-}$ sample was normalized to that from WT ESC sample, which is set as 1. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**F**) qPCR analysis of the expression of MERVL after the depletion of *Dot1l* and *Npm1* individually or simultaneously in ESCs. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). ns: non-significant. (**G**) ChIP-qPCR analysis of Dot1l binding on MERVL upon the *Npm1* depletion. ChIP-qPCR data were normalized to input and that of the control region. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). ns: non-significant, ***$P < 0.001$ in Student's *t*-test.

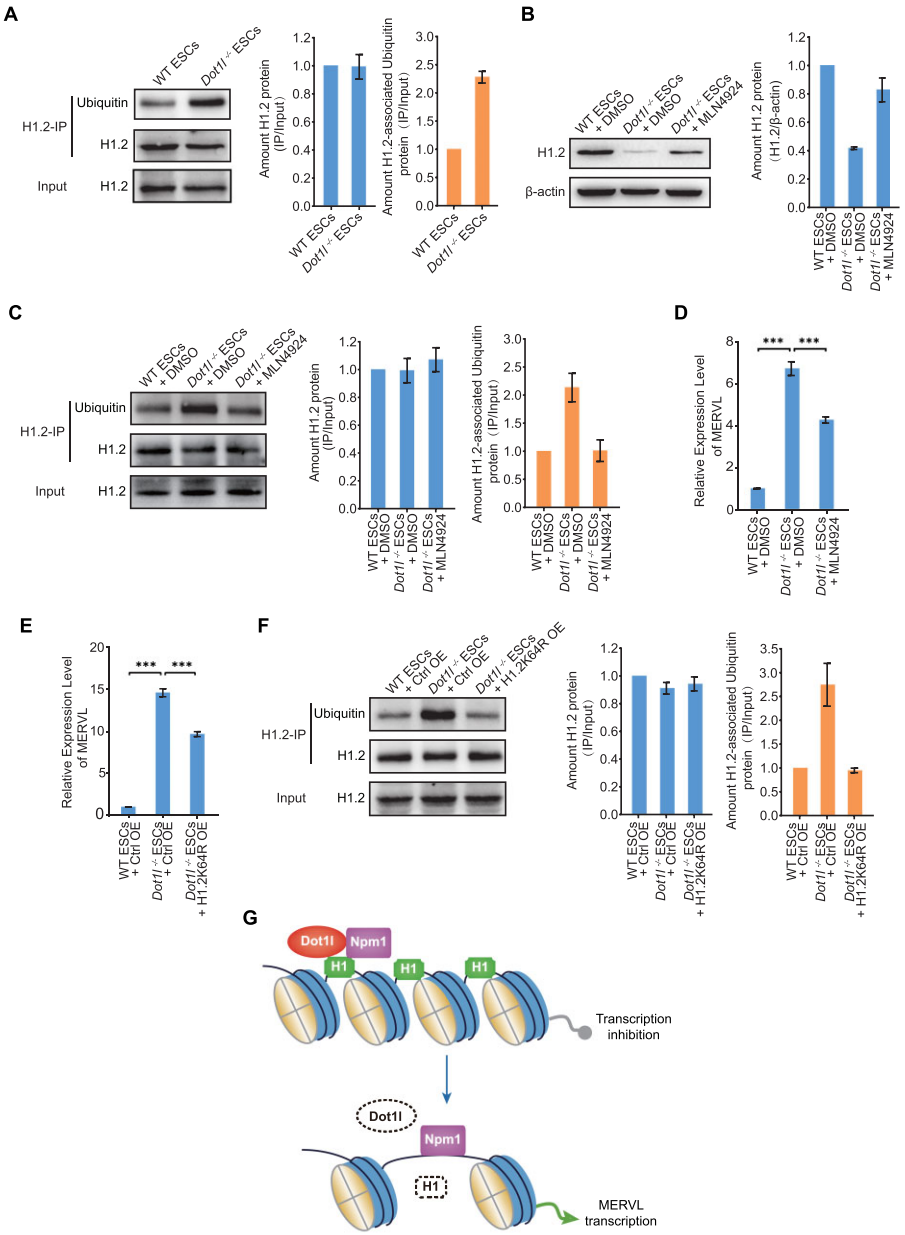


Figure 8. Histone H1.2 was degraded by ubiquitination in *Dot1l*$^{-/-}$ ESCs. (**A**) Western blot analysis of ubiquitination after immunoprecipitation with H1.2 antibody in WT ESCs or *Dot1l*$^{-/-}$ ESCs (left) and quantification (right). IP, immunoprecipitation. IP was done with anti-H1.2 antibody. 1.25% input was loaded as a control. H1.2 input protein amount of *Dot1l*$^{-/-}$ ESCs is adjusted to match that in WT ESCs and IP protein loading was adjusted proportionally according to the input. The ratio of IP protein/input protein from *Dot1l*$^{-/-}$ sample was normalized to that from WT ESC sample, which is set as 1. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**B**) Western blot analysis of H1.2 levels in *Dot1l*$^{-/-}$ ESCs treated with the inhibitor of NAE (MLN4924) (left) and quantification (right). DMSO treated sample was included as a control. β-Actin was used as a loading control. The ratio of H1.2/β-actin protein from *Dot1l*$^{-/-}$ samples was normalized to that from WT ESC + DMSO sample, which is set as 1. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**C**) Western blot analysis of ubiquitination after immunoprecipitation with H1.2 antibody in *Dot1l* $^{-/-}$ ESCs treated with the inhibitor of NAE (MLN4924) (left) and quantification (right). IP was done with anti-H1.2 antibody. 1.25% input was loaded as a control. H1.2 input protein amount of *Dot1l* $^{-/-}$ESCs and treated with NAE inhibitor (MLN4924) are adjusted to match that in WT ESCs. The ratio of IP/input protein from *Dot1l*$^{-/-}$ samples was normalized to that from WT ESC + DMSO sample, which is set as 1. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**D**) qPCR analysis of the expression of MERVL in *Dot1l*$^{-/-}$ ESCs treated with NAE inhibitor (MLN4924). DMSO treated sample was included as a control. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). ns: non-significant, ***$P < 0.001$ in Student's *t*-test. (**E**) qPCR analysis of the expression of MERVL after overexpression of H1.2K64R in *Dot1l*$^{-/-}$ ESCs. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). ***$P < 0.001$ in Student's *t*-test. Ctrl OE: control vector overexpression; K, Lysine; R, Arginine. (**F**) Western blot analysis of ubiquitination after immunoprecipitation with H1.2 antibody after overexpression of H1.2K64R in *Dot1l*$^{-/-}$ ESCs (left) and quantification (right). IP was done with anti-H1.2 antibody. 1.25% input was loaded as a control. H1.2 input protein amount of *Dot1l*$^{-/-}$ ESCs and H1.2K64R OE are adjusted to match that in WT ESCs and IP protein loading was adjusted proportionally according to the input. The ratio of IP protein/input protein from *Dot1l*$^{-/-}$ sample was normalized to that from WT ESC sample, which is set as 1. Data are presented as mean ± s.e.m. ($n = 3$ independent experiments). (**G**) Schematic of Dot1l repressing MERVL in ESCs. Histone H1 is loaded to chromatin by Dot1l and Npm1 to maintain the repression of MERVL. Conversely, in the absence of Dot1l, histone H1 dissociates from chromatin and its protein level decreases, resulting in de-repression of MERVL transcription.

most recent study also supports the transcription repressive role of Dot1l in the maintenance of heterochromatin at pericentromere (56). The distinct role of Dot1l in transcription regulation could be affected by the complex within which Dot1l performs its functions. Dot1l normally acts in a complex consisting of Mllt10 (AF10), Mllt6 (AF17), and Mllt3 (AF9) or Mllt1 (ENL) (57–59). Dot1l interacts with AF9 through three regions ranging from 628–900 a.a. (60) and interacts with RNA pol II through 618–627 a.a. (61). However, the introduction of truncated Dot1l containing only N-terminal 480 a.a. almost completely rescued MERVL activation caused by *Dot1l* knockout (Figure 5A and B). These data suggest that the interaction between Dot1l with AF9/ENL or RNA pol II is not required for the repression by Dot1l. This is supported by the unaffected expression of MERVL after the depletion of AF9 (Figure 5A and B). Published proteome analysis uncovers that Dot1l is also present in a complex containing Npm1 (39), which does not interact with AF9/ENL. In terms of the regulation of retrotransposons, we observed the activation of MERVL upon both depletion of *Dot1l* or *Npm1* (Figure 1B, G; Figure 4D and E), but not *Mllt10* (*AF10*), *Mllt6* (*AF17*), or *Mllt3* (*AF9*) (Figure 4A and B). These findings argue that Dot1l noncanonically works as a transcription repressor in the complex with Npm1, instead of AF10, AF17, and AF9. Another possible route to modulate the specificity of Dot1l is the recruitment of the Dot1l-Npm1 complex to specific genomic loci. While Dot1l serves as a transcription activator, it can be recruited to various genomic loci by transcription factors. For instance, Dot1l can be recruited by Zc3h10 to activate thermogenic gene expression (62). Hence, it is plausible that a specific transcription factor is involved in the targeting of Dot1l-Npm1 to MERVL. It will be intriguing to characterize the recruiters of Dot1l-Npm1 in the future.

Npm1 was previously discovered as an H1 histone chaperone (40,41). It can interact with histone H1 variants and facilitate the deposition of H1 onto DNA (40,41). Dot1l is able to interact with histone H1 as well (39). The AT-hook domain, which is responsible for Dot1l to interact with Npm1, is critical to silencing MERVL expression (Figure 5A and B). Dot1l itself can also act as histone chaperone to mediate the exchange of histone in yeast (63). These evidences further substantiate that Npm1 and Dot1l act together to chaperone histone H1. It is noted that methyltransferase activity of Dot1l is indispensable to restricting MERVL (Figure 5D–G), meaning that in ESCs, Dot1l represses MERVL differently from the 2-cell embryo in where Dot1l with the inactive catalytic domain is still able to organize heterochromatin (15). These findings imply that Dot1l potentially mediates the methylation of proteins within the Dot1l-Npm1 complex to exert its repressive action. Currently, Dot1l is best known for its ability to catalyze the methylation of histone H3 at the lysine 79 position (6). However, H3K79 methylation is often associated with activated gene expression (9–11). Thus, we cannot exclude the possibility that Dot1l within the Dot1l-Npm1 complex may methylate other proteins to repress MERVL, providing that Dot1l has the potential to methylate non-histone substrates such as androgen receptor (64). It warrants further investigation on the novel methylation substrates of Dot1l.

Dot1l is usually linked with the transcriptional elongation event marked by H3K79 methylation (12). However, our results and previous findings show that Dot1l deletion only causes change in a subset of genes (12) (Figure 3A), although its loss causes genome-wide disappearance of H3K79 methylation (7,12) (Figure 1F). In terms of the regulation of repeat sequence, we also found the binding of Dot1l and Npm1 to downregulated repeats and repeats without expression change (Supplementary Figure S4I). These observations suggest the presence of other unidentified factors that contribute to the regulation of gene expression by Dot1l besides the binding of Npm1 and Dot1l. One possible factor is the local epigenetic environment of repeats. For example, the presence of ubiquitinated H2B is able to stabilize Dot1l binding to the nucleosomes (65). The combination of epigenetic contexts may regulate the expression of specific retrotransposons (45). Among dysregulated repeats bound by Dot1l, only MERVL/MT2, IAPEy and RLTR45-int were among the most strongly activated repeats by triple-knockout of H1 variants (Figure 6G), suggesting that histone H1 contributes to the repression role of Dot1l-Npm1. Hence, Dot1l and Npm1 may only be able to repress repeats in specific epigenetic contexts. The presence of histone H1, such as the case for MERVL (Figure 6E and F), is probably one of the prerequisites. It will be interesting to characterize other co-factors that facilitate the regulation of gene expression by Dot1l and Npm1.

Previous studies have reported that both Dot1l and Npm1 play a part in leukemogenesis. Recruitment of Dot1l by MLL-AF9/AF4 induced transcriptional activation and transform hematopoietic progenitors (66,67). Cytoplasmic dislocation of the mutant NPM1 protein aberrantly activates HOX and FLT3 and resulted in acute myeloid leukaemia whereas Dot1l activity contributes to the activation of these genes (68). However, the relationship between Npm1 and Dot1l in leukaemia is unclear. Based on our discovery that Dot1l and Npm1 interact and repress retrotransposons in ESCs, it is likely that the loss of repression by the Npm1-Dot1l complex after the cytoplasmic mislocalization of Npm1 may contribute to the activation of Npm1-target genes in leukaemia cells. In addition, Npm1 depletion leads to premature ageing of hematopoietic stem cells whereas Dot1l can regulate senescence-associated secretory phenotype in aged cells (69,70). Therefore, it would be intriguing to explore whether transcription-repressive role of Dot1l and Npm1 is involved in ageing.

In conclusion, our study demonstrates that Dot1l and Npm1 cooperate to repress MERVL and totipotency via chaperoning histone H1 and variants. Our insights into the repressive function of Dot1l and Npm1 will inform our understanding of transcriptional regulation during mammalian development and in diseases.

DATA AVAILABILITY

High-throughput sequencing data for RNA-seq and ChIP-seq generated from this study have been deposited in Gene Expression Omnibus (GEO) under GSE201849. Published RNA-seq data and ATAC-seq in ESCs with a triple knockout of histone H1 variants are available in GEO under accession GSE153041 (43). RNA-seq for entry into and exit

from 2CLC state are from GSE121451 (35) and GSE133232 (34). RNA-seq of 2CLCs is from GSE33920 (1). Single cell RNA-seq data from early embryos are from GSE53386 (71) and GSE138760 (72). RNA-seq data of TLSCs are from GSE185030 (36) and that of TBLCs are from GSE168728 (37). Primordial germ cell expression is from GSE141181 (73).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Author contributions: Xin.Z., X.Li., T.G., F.C. and P.S. conducted the experiments; H.S. and Xuan.Z. performed bioinformatics analysis; Xin.Z, X.Li. and X.Lu designed the experiments and wrote the paper; L.L. discussed the project and supervised P.S.; X.Lu. conceived and supervised the project.

FUNDING

National Key Research and Development Program of China [2022YFA1103800 and 2018YFA0107000]; National Natural Science Foundation of China [32070858]. Funding for open access charge: National Natural Science Foundation of China [2022YFA1103800].
Conflict of interest statement. None declared.

REFERENCES

1. Macfarlan,T.S., Gifford,W.D., Driscoll,S., Lettieri,K., Rowe,H.M., Bonanomi,D., Firth,A., Singer,O., Trono,D. and Pfaff,S.L. (2012) Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature*, **487**, 57–63.
2. Yang,F., Huang,X., Zang,R., Chen,J., Fidalgo,M., Sanchez-Priego,C., Yang,J., Caichen,A., Ma,F., Macfarlan,T. *et al.* (2020) DUX-miR-344-ZMYM2-mediated activation of MERVL LTRs induces a totipotent 2C-like State. *Cell Stem Cell*, **26**, 234–250.
3. Maksakova,I.A., Thompson,P.J., Goyal,P., Jones,S.J., Singh,P.B., Karimi,M.M. and Lorincz,M.C. (2013) Distinct roles of KAP1, HP1 and G9a/GLP in silencing of the two-cell-specific retrotransposon MERVL in mouse ES cells. *Epigenetics Chromatin*, **6**, 15.
4. Wu,K., Liu,H., Wang,Y., He,J., Xu,S., Chen,Y., Kuang,J., Liu,J., Guo,L., Li,D. *et al.* (2020) SETDB1-Mediated cell fate transition between 2C-like and pluripotent states. *Cell Rep.*, **30**, 25–36.
5. Puri,D., Koschorz,B., Engist,B., Onishi-Seebacher,M., Ryan,D., Soujanya,M. and Montavon,T. (2021) Foxd3 controls heterochromatin-mediated repression of repeat elements and 2-cell state transcription. *EMBO Rep.*, **22**, e53180.
6. Feng,Q., Wang,H., Ng,H.H., Erdjument-Bromage,H., Tempst,P., Struhl,K. and Zhang,Y. (2002) Methylation of H3-lysine 79 is mediated by a new family of hmtases without a SET domain. *Curr. Biol.*, **12**, 1052–1058.
7. Jones,B., Su,H., Bhat,A., Lei,H., Bajko,J., Hevi,S., Baltus,G.A., Kadam,S., Zhai,H., Valdez,R. *et al.* (2008) The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet.*, **4**, e1000190.
8. Frederiks,F., Tzouros,M., Oudgenoeg,G., van Welsem,T., Fornerod,M., Krijgsveld,J. and van Leeuwen,F. (2008) Nonprocessive methylation by Dot1 leads to functional redundancy of histone H3K79 methylation states. *Nat. Struct. Mol. Biol.*, **15**, 550–557.
9. Ng,H.H., Ciccone,D.N., Morshead,K.B., Oettinger,M.A. and Struhl,K. (2003) Lysine-79 of histone H3 is hypomethylated at silenced loci in yeast and mammalian cells: a potential mechanism for position-effect variegation. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 1820–1825.
10. Barski,A., Cuddapah,S., Cui,K., Roh,T.Y., Schones,D.E., Wang,Z., Wei,G., Chepelev,I. and Zhao,K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823–837.
11. Schubeler,D., MacAlpine,D.M., Scalzo,D., Wirbelauer,C., Kooperberg,C., van Leeuwen,F., Gottschling,D.E., O'Neill,L.P., Turner,B.M., Delrow,J. *et al.* (2004) The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.*, **18**, 1263–1271.
12. Cao,K., Ugarenko,M., Ozark,P.A., Wang,J., Marshall,S.A., Rendleman,E.J., Liang,K., Wang,L., Zou,L., Smith,E.R. *et al.* (2020) DOT1L-controlled cell-fate determination and transcription elongation are independent of H3K79 methylation. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 27365–27373.
13. Wu,A., Zhi,J., Tian,T., Cihan,A., Cevher,M.A., Liu,Z., David,Y., Muir,T.W., Roeder,R.G. and Yu,M. (2021) DOT1L complex regulates transcriptional initiation in human erythroleukemic cells. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2106148118.
14. Singer,M.S., Kahana,A., Wolf,A.J., Meisinger,L.L., Peterson,S.E., Goggin,C., Mahowald,M. and Gottschling,D.E. (1998) Identification of high-copy disruptors of telomeric silencing in Saccharomyces cerevisiae. *Genetics*, **150**, 613–632.
15. Ooga,M., Suzuki,M.G. and Aoki,F. (2013) Involvement of DOT1L in the remodeling of heterochromatin configuration during early preimplantation development in mice. *Biol. Reprod.*, **89**, 145.
16. Ng,R.K., Dean,W., Dawson,C., Lucifero,D., Madeja,Z., Reik,W. and Hemberger,M. (2008) Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nat. Cell Biol.*, **10**, 1280–1290.
17. Ye,J., Coulouris,G., Zaretskaya,I., Cutcutache,I., Rozen,S. and Madden,T.L. (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, **13**, 134.
18. Zhang,W., Chen,F., Chen,R., Xie,D., Yang,J., Zhao,X., Guo,R., Zhang,Y., Shen,Y., Göke,J. *et al.* (2019) Zscan4c activates endogenous retrovirus MERVL and cleavage embryo genes. *Nucleic Acids Res.*, **47**, 8485–8501.
19. Kim,D., Paggi,J.M., Park,C., Bennett,C. and Salzberg,S.L. (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.*, **37**, 907–915.
20. Frankish,A., Diekhans,M., Jungreis,I., Lagarde,J., Loveland,J.E., Mudge,J.M., Sisu,C., Wright,J.C., Armstrong,J., Barnes,I. *et al.* (2021) Gencode 2021. *Nucleic Acids Res.*, **49**, D916–D923.
21. Liao,Y., Smyth,G.K. and Shi,W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
22. Jin,Y., Tam,O.H., Paniagua,E. and Hammell,M. (2015) TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics*, **31**, 3593–3599.
23. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
24. Wu,T., Hu,E., Xu,S., Chen,M., Guo,P., Dai,Z., Feng,T., Zhou,L., Tang,W., Zhan,L. *et al.* (2021) clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (N Y)*, **2**, 100141.
25. Yu,G. and He,Q.Y. (2016) ReactomePA: an R/bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.*, **12**, 477–479.
26. Pertea,M., Kim,D., Pertea,G.M., Leek,J.T. and Salzberg,S.L. (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.*, **11**, 1650–1667.
27. Robinson,J.T., Thorvaldsdottir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and Mesirov,J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.
28. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
29. Chung,D., Kuan,P.F., Li,B., Sanalkumar,R., Liang,K., Bresnick,E.H., Dewey,C. and Keles,S. (2011) Discovering transcription factor binding sites in highly repetitive regions of genomes with multi-read analysis of ChIP-Seq data. *PLoS Comput. Biol.*, **7**, e1002111.
30. Ramirez,F., Dundar,F., Diehl,S., Gruning,B.A. and Manke,T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.*, **42**, W187–W191.

31. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nussbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-seq (MACS). *Genome Biol.*, **9**, R137.
32. Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.
33. Yu,G.C., Wang,L.G. and He,Q.Y. (2015) ChIPseeker: an R/bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, **31**, 2382–2383.
34. Fu,X.D., Djekidel,M.N. and Zhang,Y. (2020) A transcriptional roadmap for 2C-like-to-pluripotent state transition. *Sci. Adv.*, **6**, eaay5181.
35. Fu,X.D., Wu,X.J., Djekidel,M.N. and Zhang,Y. (2019) Myc and Dnmt1 impede the pluripotent to totipotent state transition in embryonic stem cells. *Nature Cell Biology*, **21**, 835–844.
36. Yang,M., Yu,H., Yu,X., Liang,S., Hu,Y., Luo,Y., Izsvak,Z., Sun,C. and Wang,J. (2022) Chemical-induced chromatin remodeling reprograms mouse escs to totipotent-like stem cells. *Cell Stem Cell*, **29**, 400–418.
37. Shen,H., Yang,M., Li,S., Zhang,J., Peng,B., Wang,C., Chang,Z., Ong,J. and Du,P. (2021) Mouse totipotent stem cells captured and maintained through spliceosomal repression. *Cell*, **184**, 2843–2859.
38. Bitoun,E., Oliver,P.L. and Davies,K.E. (2007) The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Hum. Mol. Genet.*, **16**, 92–106.
39. Park,G., Gong,Z., Chen,J. and Kim,J.E. (2010) Characterization of the DOT1L network: implications of diverse roles for DOT1L. *Protein J.*, **29**, 213–223.
40. Holmberg Olausson,K., Nister,M. and Lindstrom,M.S. (2014) Loss of nucleolar histone chaperone NPM1 triggers rearrangement of heterochromatin and synergizes with a deficiency in DNA methyltransferase DNMT3A to drive ribosomal DNA transcription. *J. Biol. Chem.*, **289**, 34601–34619.
41. Gadad,S.S., Senapati,P., Syed,S.H., Rajan,R.E., Shandilya,J., Swaminathan,V., Chatterjee,S., Colombo,E., Dimitrov,S., Pelicci,P.G. *et al.* (2011) The multifunctional protein nucleophosmin (NPM1) is a human linker histone H1 chaperone. *Biochemistry*, **50**, 2780–2789.
42. Hogan,A.K. and Foltz,D.R. (2021) Reduce, retain, recycle: mechanisms for promoting histone protein degradation versus stability and retention. *Mol. Cell. Biol.*, **41**, e0000721.
43. Sheban,D., Shani,T., Maor,R., Aguilera-Castrejon,A., Mor,N., Oldak,B., Shmueli,M.D., Eisenberg-Lerner,A., Bayerl,J., Hebert,J. *et al.* (2022) SUMOylation of linker histone H1 drives chromatin condensation and restriction of embryonic cell fate identity. *Mol. Cell*, **82**, 106–122.
44. Rodriguez-Terrones,D., Gaume,X., Ishiuchi,T., Weiss,A., Kopp,A., Kruse,K., Penning,A., Vaquerizas,J.M., Brino,L. and Torres-Padilla,M.E. (2018) A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat. Genet.*, **50**, 106–119.
45. He,J., Fu,X., Zhang,M., He,F., Li,W., Abdul,M.M., Zhou,J., Sun,L., Chang,C., Li,Y. *et al.* (2019) Transposable elements are regulated by context-specific patterns of chromatin marks in mouse embryonic stem cells. *Nat. Commun.*, **10**, 34.
46. Zhang,M., Zhao,X., Feng,X., Hu,X., Lu,W. and Lu,X. (2022) Histone chaperone HIRA complex regulates retrotransposons in embryonic stem cells. *Stem Cell Res. Ther.*, **13**, 137.
47. Chen,F., Zhang,W., Xie,D., Gao,T., Dong,Z. and Lu,X. (2020) Histone chaperone FACT represses retrotransposon MERVL and MERVL-derived cryptic promoters. *Nucleic Acids Res.*, **48**, 10211–10225.
48. Li,P., Wang,L., Bennett,B.D., Wang,J., Li,J., Qin,Y., Takaku,M., Wade,P.A., Wong,J. and Hu,G. (2017) Rif1 promotes a repressive chromatin state to safeguard against endogenous retrovirus activation. *Nucleic Acids Res.*, **45**, 12723–12738.
49. Yang,B.X., El Farran,C.A., Guo,H.C., Yu,T., Fang,H.T., Wang,H.F., Schlesinger,S., Seah,Y.F., Goh,G.Y., Neo,S.P. *et al.* (2015) Systematic identification of factors for provirus silencing in embryonic stem cells. *Cell*, **163**, 230–245.
50. Lu,F., Liu,Y., Jiang,L., Yamaguchi,S. and Zhang,Y. (2014) Role of tet proteins in enhancer activity and telomere elongation. *Genes Dev.*, **28**, 2103–2119.
51. Zuo,F., Jiang,J., Fu,H., Yan,K., Liefke,R., Zhang,J., Hong,Y., Chang,Z., Liu,N., Wang,Z. *et al.* (2022) A TRIM66/DAX1/dux axis suppresses the totipotent 2-cell-like state in murine embryonic stem cells. *Cell Stem Cell*, **29**, 948–961.
52. Zhao,X., Shen,J., Zhang,M., Feng,X., Zhang,W. and Lu,X. (2022) PIM3-AMPK-HDAC4/5 axis restricts MuERVL-marked 2-cell-like state in embryonic stem cells. *Stem Cell Reports*, **17**, 2256–2271.
53. Soucy,T.A., Smith,P.G., Milhollen,M.A., Berger,A.J., Gavin,J.M., Adhikari,S., Brownell,J.E., Burke,K.E., Cardin,D.P., Critchley,S. *et al.* (2009) An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature*, **458**, 732–736.
54. Hollmuller,E., Geigges,S., Niedermeier,M.L., Kammer,K.M., Kienle,S.M., Rosner,D., Scheffner,M., Marx,A. and Stengel,F. (2021) Site-specific ubiquitylation acts as a regulator of linker histone H1. *Nat. Commun.*, **12**, 3497.
55. Chang,L., Shen,L., Zhou,H., Gao,J., Pan,H., Zheng,L., Armstrong,B., Peng,Y., Peng,G., Zhou,B.P. *et al.* (2019) ITCH nuclear translocation and H1.2 polyubiquitination negatively regulate the DNA damage response. *Nucleic Acids Research*, **47**, 824–842.
56. Malla,A.B., Yu,H., Farris,D., Kadimi,S., Lam,T.T., Cox,A.L., Smith,Z.D. and Lesch,B.J. (2023) DOT1L bridges transcription and heterochromatin formation at mammalian pericentromeres. *EMBO Rep.*, **24**, e56492.
57. Li,Y., Wen,H., Xi,Y., Tanaka,K., Wang,H., Peng,D., Ren,Y., Jin,Q., Dent,S.Y., Li,W. *et al.* (2014) AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell*, **159**, 558–571.
58. Mohan,M., Herz,H.M., Takahashi,Y.H., Lin,C., Lai,K.C., Zhang,Y., Washburn,M.P., Florens,L. and Shilatifard,A. (2010) Linking H3K79 trimethylation to wnt signaling through a novel Dot1-containing complex (DotCom). *Genes Dev.*, **24**, 574–589.
59. Mueller,D., Bach,C., Zeisig,D., Garcia-Cuellar,M.P., Monroe,S., Sreekumar,A., Zhou,R., Nesvizhskii,A., Chinnaiyan,A., Hess,J.L. *et al.* (2007) A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood*, **110**, 4445–4454.
60. Kuntimaddi,A., Achille,N.J., Thorpe,J., Lokken,A.A., Singh,R., Hemenway,C.S., Adli,M., Zeleznik-Le,N.J. and Bushweller,J.H. (2015) Degree of recruitment of DOT1L to MLL-AF9 defines level of H3K79 di- and tri-methylation on target genes and transformation potential. *Cell Rep.*, **11**, 808–820.
61. Kim,S.K., Jung,I., Lee,H., Kang,K., Kim,M., Jeong,K., Kwon,C.S., Han,Y.M., Kim,Y.S., Kim,D. *et al.* (2012) Human histone H3K79 methyltransferase DOT1L protein [corrected] binds actively transcribing RNA polymerase II to regulate gene expression. *J. Biol. Chem.*, **287**, 39698–39709.
62. Yi,D., Nguyen,H.P., Dinh,J., Viscarra,J.A., Xie,Y., Lin,F., Zhu,M., Dempersmier,J.M., Wang,Y. and Sul,H.S. (2020) Dot1l interacts with Zc3h10 to activate Ucp1 and other thermogenic genes. *Elife*, **9**, e59990.
63. Lee,S., Oh,S., Jeong,K., Jo,H., Choi,Y., Seo,H.D., Kim,M., Choe,J., Kwon,C.S. and Lee,D. (2018) Dot1 regulates nucleosome dynamics by its inherent histone chaperone activity in yeast. *Nat. Commun.*, **9**, 240.
64. Yang,L., Lin,C., Jin,C., Yang,J.C., Tanasa,B., Li,W., Merkurjev,D., Ohgi,K.A., Meng,D., Zhang,J. *et al.* (2013) lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature*, **500**, 598–602.
65. Yao,T., Jing,W., Hu,Z., Tan,M., Cao,M., Wang,Q., Li,Y., Yuan,G., Lei,M. and Huang,J. (2019) Structural basis of the crosstalk between histone H2B monoubiquitination and H3 lysine 79 methylation on nucleosome. *Cell Res.*, **29**, 330–333.
66. Okuda,H., Stanojevic,B., Kanai,A., Kawamura,T., Takahashi,S., Matsui,H., Takaori-Kondo,A. and Yokoyama,A. (2017) Cooperative gene activation by AF4 and DOT1L drives MLL-rearranged leukemia. *J. Clin. Invest.*, **127**, 1918–1931.
67. Nguyen,A.T., Taranova,O., He,J. and Zhang,Y. (2011) DOT1L, the H3K79 methyltransferase, is required for MLL-AF9-mediated leukemogenesis. *Blood*, **117**, 6912–6922.
68. Kuhn,M.W., Song,E., Feng,Z., Sinha,A., Chen,C.W., Deshpande,A.J., Cusan,M., Farnoud,N., Mupo,A., Grove,C. *et al.*

(2016) Targeting chromatin regulators inhibits leukemogenic gene expression in NPM1 mutant leukemia. *Cancer Discov.*, **6**, 1166–1181.
69. Morganti,C., Ito,K., Yanase,C., Verma,A., Teruya-Feldstein,J. and Ito,K. (2022) NPM1 ablation induces HSC aging and inflammation to develop myelodysplastic syndrome exacerbated by p53 loss. *EMBO Rep.*, **23**, e54262.
70. Leon,K.E., Buj,R., Lesko,E., Dahl,E.S., Chen,C.W., Tangudu,N.K., Imamura-Kawasawa,Y., Kossenkov,A.V., Hobbs,R.P. and Aird,K.M. (2021) DOT1L modulates the senescence-associated secretory phenotype through epigenetic regulation of IL1A. *J. Cell Biol.*, **220**, e202008101.
71. Fan,X., Zhang,X., Wu,X., Guo,H., Hu,Y., Tang,F. and Huang,Y. (2015) Single-cell RNA-seq transcriptome analysis of linear and circular rnas in mouse preimplantation embryos. *Genome Biol.*, **16**, 148.
72. Qiao,Y., Ren,C., Huang,S., Yuan,J., Liu,X., Fan,J., Lin,J., Wu,S., Chen,Q., Bo,X. *et al.* (2020) High-resolution annotation of the mouse preimplantation embryo transcriptome using long-read sequencing. *Nat. Commun.*, **11**, 2653.
73. Huang,T.C., Wang,Y.F., Vazquez-Ferrer,E., Theofel,I., Requena,C.E., Hanna,C.W., Kelsey,G. and Hajkova,P. (2021) Sex-specific chromatin remodelling safeguards transcription in germ cells. *Nature*, **600**, 737–742.