Loss of SETDB1-mediated H3K9me3 in human neural progenitor cells leads to transcriptional activation of L1 retrotransposons

Ofelia Karlsson[1], Ninoslav Pandiloski[1,2], Vivien Horvath[1], Anita Adami[1], Raquel Garza[1], Pia A. Johansson[1], Jenny G Johansson[1], Christopher H. Douse[2], Johan Jakobsson[1,*]

[1]Laboratory of Molecular Neurogenetics, Department of Experimental Medical Science and Lund Stem Cell Center, BMC A11, Lund University, 221 84 Lund, Sweden

[2] Laboratory of Epigenetics and Chromatin Dynamics, Department Of Experimental Medical Science and Lund Stem Cell Center, Lund University, Lund, Sweden

*Correspondence and lead contact:

Johan Jakobsson

Lund Stem Cell Center

Department of Experimental Medical Sciences

Lund University

221 84 Lund, SWEDEN

Email: johan.jakobsson@med.lu.se

Phone: +46 46 2224225, +46709286443

Abstract

Heterochromatin is characterised by an inaccessibility to the transcriptional machinery and is associated with the histone mark H3K9me3. However, studying the functional consequences of heterochromatin loss in human cells has been challenging. Here, we used CRISPRi-mediated silencing of the histone methyltransferase SETDB1 to remove H3K9me3 heterochromatin in human neural progenitor cells. Despite a major loss of H3K9me3 peaks resulting in genome-wide reorganization of heterochromatin domains, silencing of SETDB1 had a limited effect on cell viability. Cells remained proliferative and expressed appropriate marker genes. We found that a key event following the loss of SETDB1-mediated H3K9me3 was the expression of evolutionarily young L1 retrotransposons. Derepression of L1s was associated with a loss of CpG DNA methylation at their promoters, suggesting that deposition of H3K9me3 at the L1 promoter is required to maintain DNA methylation. In conclusion, these results demonstrate that loss of H3K9me3 in human neural somatic cells transcriptionally activates evolutionary young L1 retrotransposons.

Introduction

At least 50% of the human genome consists of transposable elements (TEs) (1-4). Due to their ability to mobilize, TEs pose a potential threat to genomic integrity and are transcriptionally silenced via epigenetic mechanisms such as histone modifications and DNA methylation (5-10). Long interspersed nuclear elements–1 (L1s) are the most abundant and only autonomously mobilizing family of TEs in humans, accounting for ~17% of human DNA (1,4,11). Full-length L1s are transcribed from an internal 5′ RNA polymerase II promoter as a bicistronic mRNA encoding two proteins, ORF1p and ORF2p, which are essential for L1 mobilization (12-15). Since L1 elements have colonized the human genome throughout evolution via a copy-paste mechanism in different waves, it is possible to estimate their evolutionary age and assign them to chronologically-ordered subfamilies (16). Hundreds of thousands of L1s are primate specific, and thousands are human specific but only a small fraction of these human-specific full-length L1s are capable of transposing (17,18). Notably, the L1 promoter is bidirectional and L1 antisense transcripts can give rise to chimeric transcripts and act as alternative promoters for protein-coding genes (8,19-21).

L1 transcription is silenced in most human somatic tissues, and transcriptional repression correlates with the presence of heterochromatin marks, including H3K9me3 and CpG DNA methylation (8,9,22,23). In the cells of the human brain, it is well established that CpG DNA methylation plays a critical role in the transcriptional silencing of L1s (6,8,9,23-25). For instance, when DNMT1 is deleted using CRISPR in human neural progenitor cells (NPCs), leading to global DNA demethylation, evolutionarily young full-length L1 elements become transcriptionally active (8). However, it is unclear how DNA methylation is directed to the promoter of evolutionary young L1s in human somatic cells and the interplay between H3K9me3 and DNA methylation in the transcriptional control of L1s is poorly understood.

In this study, we used a somatic human NPC line derived from fetal brain six weeks post-conception (26). This represents a timepoint in post-implantation embryonic development when DNA methylation and other heterochromatin marks have been fully re-established (8,27). This includes the presence of H3K9me3 and DNA methylation over L1s and other TEs (8,23,28), which is established by the KRAB zinc finger (KZNF)-TRIM28 system and other epigenetic TE repressors during early development (29-31). We used CRISPRi to silence SETDB1, a histone methyltransferase associated with these repressors, to study the consequences of H3K9me3 loss at TEs in somatic cells. SETDB1 silencing in hNPCs had a limited effect on cell viability, and the cells remained proliferative and expressed the correct marker genes, despite the widespread loss of H3K9me3 domains. However, SETDB1-mediated loss of H3K9me3 resulted in the transcriptional activation of L1s. Interestingly, this activation correlated with a loss of CpG DNA methylation over the same elements. This finding links the maintenance of DNA methylation at L1s to the presence of H3K9me3. In summary, our results demonstrate the critical role of H3K9me3 in L1 silencing in human neural somatic cells.

Materials and Methods

Cell culture

A (male) embryo-derived human neural progenitor cell line (Sai2) was used throughout this manuscript (26). The Sai2 cell line is derived from a small number of cells dissected from fetal brain tissue at Carnegie stage 16, and it has the characteristics of neuroepithelial-like stem cells**.** The cells were maintained according to the established protocol (32). Cells were grown in poly-L-ornithine (15 ug/ml, Sigma) and laminin 2020 (2ug/ml, Sigma) coated wells. Cells were maintained in culture media DMEM/F12 (Thermo Fisher Scientific) supplemented with N2

(1x, Thermo Fisher Scientific), glutamine (Sigma), Penicillin/Streptomycin (1x, Gibco), B27 (0.05X, Thermo Fisher Scientific), EGF (10ng/ml, Thermo Fisher Scientific) and FGF (10ng/ml, Sigma). Media change was performed daily and cells were passaged every 2-3 days using TrypLE Express (Gibco). 10 μM Y27632 (Rock inhibitor, Miltenyi) was added to the culture media when cells were passaged.

CUT&RUN

CUT&RUN was performed as previously described by Skene and Henikoff (33). 500.000 hNPCs were washed twice with 20mM HEPES pH 7.5, 150 mM NaCl, 0.5mM spermidine and 1x Roche cOmplete Protease Inhibitor tablet. ConA-coated magnetic beads were activated using 20 mM HEPES pH 7.9, 10 mM, KCl, 1 mM $CaCl_2$ and 1 mM $MnCl_2$. Cells were attached to preactivated beads and resuspend in 20 mM HEPES pH 7.5, 0.15 M NaCl, 0.5 mM Spermidine, 1x Roche cOmplete protease inhibitors, 0.02% w/v digitonin and 2 mM EDTA together with primary antibody (rabbit anti-H3K4me3, Active Motif 39159; rabbit anti-H3K9me3, Abcam 8898; or goat anti- rabbit IgG, Abcam ab97047) at 1:50 dilution. Bead-bound cells were incubated over night with gentle shake. The following day, the bead bound cells were washed twice with digitonin buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 1x Roche cOmplete protease inhibitors, 0.02% digitonin) and incubated with pA-MNase diluted in digitonin buffer (700ng/mL) at a rotator for 10 minutes at room temperature. After, cells were washed twice with digitonin buffer, resuspended in digitonin buffer and chilled to 0°C for 5 min. To initiate genomic cleavage, 100mM $CaCl_2$ was added, and cells were incubated at 0°C for 30 min. 2x stop buffer (0.35 M NaCl, 20 mM EDTA, 4 mM EGTA, 0.02% digitonin, 50 ng/µL glycogen, 50 ng/µL RNase A) was added to stop the reaction and the cells were incubated at 37°C for 30 min to release genomic fragments. Genomic fragments were purified using PCR clean-up spin column according to manufacturer's instruction

(Macherey-Nagel). Sequencing libraries were prepared using the Hyperprep kit (KAPA) with dual index-adapters (KAPA), pooled and sequenced on a Nextseq500 instrument (Illumina).

CUT&RUN analysis

Using bowtie2 (--local --very-sensitive-local --no-mixed –no-discordant –phred33 –I –X 700), 2x75bp paired-end CUT&RUN reads were aligned to the human genome (GRCh38) and subsequently converted to BAM files using SAMtools v.1.18 (34). BAM files were then filtered by setting a mapping quality threshold (MAPQ10), which were used to create RPKM (rounds per kilobase million) normalized coverage tracks using bamCoverage (deepTools) v2.5.4 (35). All coverage tracks displayed were visualized in IGV v2.19.1. H3K9me3 peaks were called using SEACR v1.3 (36). Peaks over reference chromosomes longer than 500bp were kept for downstream analyses. Using deepTools' computeMatrix v 2.5.4 (35), heatmaps matrices were created and visualized using plotHeatmap of the same deepTools package. Finally, we used mergeBed from BedTools v2.31.0 to identify the overlaps of peaks with specific genomic regions, such as evolutionary young L1s or genes.

To identify H3K4me3 peaks that overlap with L1 promoters, we used an unbiased k-means strategy to cluster the H3K4me3 signal over all FL L1HS-L1PA3s. This resulted in three clusters. Two of the clusters contained 565 elements with high- or low-confidence H3K4me3 signal. To avoid inflating the number of L1s carrying an H3K4me3 signal, we manually curated the 548 low-confidence elements by inspecting the IGV tracks. This curation process identified 24 elements with a high-confidence H3K4me3 peak at the L1 promoter.

Lentiviral vector production

Lentivirus was produced as previously described by Zufferey et al (37). Briefly, 293T cells were transfected with third-generation packaging and envelope vectors pMDL (Addgene

#12251), psRev (Addgene #12253), and pMD2G (Addgene #12259) with Polyethyleneimine (Polysciences PN 23966) in DPBS (GIBCO). 48h post transfection, lentivirus was harvested by filtering the supernatant and centrifugation at 25,000 × g for 1.5h at 4°C. The virus was resuspended in PBS, aliquoted and stored in -80°C until further use. Lentivirus were in titers of 10^8-10^9 as determined by qRT-PCR.

CRISPRi

CRISPRi constructs were produced as previously described by Johansson et al (38). Single guide sequences were designed to target close to the transcription start site (TSS) using the GPP Portal (Broad institute). The guide sequences were cloned into a lentiviral back bone plasmid containing deadCas9-KRAB-T2A-GFP, a gift from Charles Gersback (Addgene plasmid #71237 RRID:Addgene_71237), using annealed oligos and the BsmBI cloning site. As a control we used one single guide RNA targeting LacZ, a sequence not present in the human genome. Lentivirus was produced for each guide as described above. hNPCs were transduced with MOI 5-7.5 of LacZ, SETDB1, DNMT1, MORC2 and TRIM28 targeting guide RNAs.

For the TRIM28&MORC2 double-CRISPRi we used a double transduction method. Cells were transduced with the previously mentioned deadCas9-KRAB-T2A-GFP lentivirus containing the dCas9 protein, GFP and MORC2 gRNA and simulatniously with lentivirus containing the pLV.U6BsmBI.EFS-NS.H2b-RFPW lentiviral backbone and the gRNA for *TRIM28* and mCherry as a marker. After 10 days expansion, transduced cells were detached using TrypLE Express and resuspended in cell culture media containing Y27632 (10 μM) and propidium iodide (PI, 1:500, BD Bioscience). The GFP^+ and $mCherry^+$ gates were set using untransduced hNPCs and validated via reanalysis of sorted cells (> 95% purity). Transduction efficiency was in the range of 70-95% GFP^+ cells. Single CRISPRi GFP^+/PI^- and double CRISPRi

$GFP^+/mCherry^+/PI^-$ cells were isolated and pelleted at 400 × g for 10 min. The cell pellets were snap frozen on dry ice and stored in -80°C until further use.

SETDB1 guide 1: AACTCAGGGGTCGGCCTCGA

SETDB1 guide 2: CGAAAGCGAAGGGATAAGGG

DNMT1 guide 1: TGCTGAAGCCTCCGAGATGC

MORC2 guide 1:GCTTCCAAGGACCGGATCGA

TRIM28 guide 1: CCCGCTCGCAGAAAGAGCCG

LACZ guide: TGCGAATACGCCCACGCGAT

qRT-PCR

Total RNA isolated from hNPCs (RNeasy mini kit, Qiagen) and 250 ng of the isolated RNA was reversed transcribed using the Maxima cDNA synthesis kit (Thermo Fisher Scientific). qRT-PCR was preformed using primers listed below and SYBR green master mix (Roche) on a LightCycler® 480 instrument. Sample were amplified for 40 cycles and cycle values were normalized to ß-actin and HPRT.

Primers:

ß-actin forward: CCTTGCACATGCCGGAG

ß-actin reverse: GCACAGAGCCTCCGCCTT

HPRT forward: ACCCTTTCCAAATCCTCAGC

HPRT reverse: GTTATGGCGACCCGCAG

SETDB1 forward: GACACGTCCAAATATGGGTGC

SETDB1 reverse: ACTGGCTTGAACTGGGTTCC

Western Blot

For total protein extraction, hNPCs were lysed using RIPA buffer (Sigma-Aldrich) and a complete protease inhibitor cocktail on ice for 30 min, except for the MORC2 protein, where the protein was lysed with 1% (Sigma Aldrich, 71736) instead of RIPA buffer. The lysed cells were pelleted at 17.000 × g for 20 min at 4°C. For total histone protein extraction, we used a Histone Extraction kit (Abcam) according to manufacturer's instructions. 2,5 ug of extracted total protein or 2,5 ug of extracted histone proteins were mixed with Novex LDS 4x loading dye (Thermo) and reducing agent (Thermo) before being boiled at 95°C for 5 minutes. Samples were loaded onto a 4-12% SDS/PAGE gel and run at 150V for 0.5h. The proteins were transferred to a PVDF membrane using a Transblot-Turbo transfer system (BioRad). The membrane was washed two times in TBS with 0.1% Tween (TBST) before being blocked in TBST with 5% skimmed milk for 1h. The membrane was incubated overnight at 4°C with primary antibody mouse anti-SETDB1 (Thermo Fisher MA5-15722, 1:2000), rabbit anti-H3K9me3 (Abcam 8898, 1:20.000) rabbit anti-histone 3 (Abcam 18521, 1:250.000), rabbit anti-MORC2 (A300-149A, Bethyl; 1:1000 dilution), rabbit anti-TRIM28 (Abcam, ab10484; 1:1000 dilution) diluted in TBST with 5% skimmed milk. The following day, the membrane was washed three times with TBST before being incubated with secondary antibody HRP-conjugated anti-mouse antibody (Santa Cruz Biotechnology, 1:5.000), HRP-conjugated anti-rabbit antibody (Santa Cruz Biotechnology, 1:10.000 or 1:50.000) or HRP conjugated anti-ß-actin (Sigma, A3854, 1:50.000) diluted in TBST with 5% skimmed milk for 2h in room temperature. The membrane was washed three times with TBST and once with TBS before protein was detected by chemiluminescence using ECL Prime Western blotting detection reagent (Cytiva). The signal was detected using a ChemiDoc system (BioRad).

Immunocytochemistry: SOX2 and NESTIN

hNPCs were washed three times with DPBS and fixed using 4% paraformaldehyde for 15 min at room temperature, followed by three more washed with DPBS. After incubation in blocking solution (KPBS + 0.25% Triton X-100 with 5% normal donkey serum) for 1h at room temperature, the cells were incubated with primary antibodies SOX2 (R&D Systems, AF2018, 1:100) and NESTIN (Abcam, AB176571, 1:100) diluted in blocking solution overnight at 4°C. After being washed twice with TKPBS and once with TKPBS + 5% NDS, cells were incubated with secondary antibodies Alexa647-donkey anti-rabbit (Jackson Immuno Research 711-605-152, 1:500) and Cy3-donkey anti-goat (Jackson Immuno Research 705-165-533 003, 1:500) diluted in blocking solution for 2h at room temperature following incubation with DAPI (Sigma, 1:1000) for 5 min. The cells were rinsed two times with KPBS and imaged using a Leica microscope (model DMI6000 B). Images were cropped and adjusted using ImageJ Fiji.

Immunocytochemistry: H3K9me3

hNPCs plated in poly-L-ornithine (15 ug/ml, Sigma) and laminin 2020 (2ug/ml, Sigma) coated 24 well ibidi plates were rinsed three times with DPBS before being fixed with 4% paraformaldehyde at room temperature for 15 min. The cells were washed once with KPBS before being blocked in TKPBS with 5% NDS for 1 h at room temperature. After, the cells were incubated with primary antibody rabbit anti-H3K9me3 (Abcam 8898, 1:200) diluted in TKPBS + 5% NDS overnight at 4°C. The following day, cells were washed three times with TKPBS before being incubated with secondary antibody Cy5-donkey anti-rabbit (Jackson Immuno Research,1:500) diluted in TKPBS + 5% NDS at room temperature for 1h. The cells were washed three times with KPBS and incubated with DAPI (1:1000) diluted in TKPBS at room temperature for 5 min. The cells were washed twice with KPBS before being imaged

using confocal Microscope Leica TSC SP8. Images were cropped and adjusted using ImageJ Fiji.

EdU proliferation assay

To study cell proliferation we used the Click IT™ Plus EdU Alexa Fluor™ 647 Imaging Kit (Invitrogen, C10640). GFP^+ SETDB-CRISPRi and Control-CRISPRi cells were FACS sorted 10 days after transduction and re-plated for the EdU assay. After 24h incubation with EdU (10 μM), cells were fixed in 4% paraformaldehyde. After fixing, cells were stained for Alexa647 and Hoechst following the manufacturer's instructions. Cells were imaged and analyzed using an Operetta CSL high-content analysis system.

Bulk RNA sequencing

Total RNA was isolated from hNPCs using the RNEasy mini kit (Qiagen) with on column DNAse treatment. Libraries for bulk RNA sequencing were generated using Illumina TruSeq Stranded mRNA library prep kit (with poly-A selection) optimized for long fragments. Sequencing was performed on a NovaSeq6000 (paired end 2 × 150bp).

Bulk RNA sequencing analysis

Paired-end 150bp RNAseq reads were mapped to the reference human genome (GRCh38) using STAR aligner verson 2.6.0 (39) with gencode v38 as a guide (--sjdbGTFfile). Unique mapping (described below) was performed to investigate the expression of individual transposable elements (TEs). Multi mapping (described below) was performed to investigate the expression of genes and TEs at a subfamily level.

Unique mapping: Reads were allowed to align to a single locus (--outFilterMultimapNmax 1) with a mismatch ratio up to 0.03 (–outFilterMismatchNoverLmax 0.03). Sorted alignment files as output by STAR were indexed using SAMtools v1. 18 (34). Normalized coverage tracks (bigwig files) using RPKM were created using deeptools bamCoverage v2.5.4 (35). FeatureCounts v1.6.3 (40) was used to quantify the expression of individual TEs using a curated GTF file provided by the creators of TEtranscripts. FeatureCounts was ran on default parameters using –s2 to force the strandedness of the reads to the feature.

Differential expression analyses: DESeq2 v1.38.3 (41) with default parameters was used to perform differential expression analyses for all bulk RNAseq datasets using the read count matrix obtained with featureCounts. Mean plots were created using median-of-ratio normalized counts per condition, adding a pseudo count of 0.5, and log2 transformed values. All pvalues and p adjusted values calculated by DESeq2 are two-sided.

To obtain distribution of the expected upregulated TEs, the RepeatMasker annotation was sampled with the exact number of the observed upregulated TEs, without replacement. The number of sampled TE was then compared to the observed values and scored accordingly (+1 if expected > observed; +0 if expected ≤ observed). This then was repeated 50.000 times to calculate a double-sided p-value.

To quantify and normalise the strand-specific expression of L1s, the reads were normalised using gene-derived medians to calculate the median ratio. p-values were calculated using the Mann–Whitney U-test , a non-parametric statistical test used to compare the non-normal distributions of two independent groups. When calculating the p-value for the differences between conditions, each L1 element was considered an independent replicate. In the analysis we used the FL-L1s that were upregulated and found in at least one of the two experimental conditions (two different gRNAs). This resulted in 241 FL-L1s being included in the analysis.

Gene-ontology overrepresentation test of biological processes (enrichGO) was performed using clusterProfiler v4.8.2 over differentially expressed genes (|LFC| > 1, two-sided padj < 0.05). To measure sense and antisense transcription over evolutionary young L1s, featureCounts was used specifying the -s2 and -s1 parameter, respectively. Gene expression size factors were further used to normalize the reads by median-of-ratios normalization approach.

Oxford Nanopore DNA sequencing

High molecular weight DNA was extracted from hNPCs using the Nanobind HMW DNA extraction kit (PacBio) according to the manufacturer's instructions. DNA quality and concentration was measured using Nanodrop and the dsDNA HS assay on a Qubit (Thermo Fisher Scientific). Genomic fragments <5kb were eliminated using the Short Read Eliminator XS kit (PacBio) and remaining genomic fragments were sheered to 10kb using a g-tube kit (Covaris) following manufacturer's instructions. Libraries were prepared using the SKQ-LSK114 kit (Oxford Nanopore Technologies) with 1 µg DNA input and loaded onto R10.4.1 flow cells. Sequencing was performed on an Oxford Nanopore Technologies PromethION 2 solo for 72h following manufacturers guidelines. To obtain the data, dorado (v4.3.0)(RRID: SCR_025883; https://github.com/nanoporetech/dorado) modified basecalling was performed over the pod5 files, of the 5mCG_5hmCG super accurate basecalling model. Basecalled files were mapped to the human genome (GRCh38) using minimap2 v2.25 (42). Alignment files (BAM) sorted and indexed using SAMtools, and methylation matrices were created using the segmeth tool from MethylArtist v1.2.6 (43). Methylartist segplot and locus tools were further employed for all downstream visualization, either using methylation matrices or BAM files as input. To statistically compare the difference in DNA methylation between control and

CRISPRi conditions, Wilcoxon signed-rank test was used. Methylartist composite plots were created using L1HS and L1PA2 consensus sequences to plot FL L1HS and FL L1PA2-L1PA3 elements, respectively. By default the tool uses a maximum of 200 elements per plot, therefore if elements of interest surpassed that threshold, 200 of them were randomly selected with "shuf -n 200".

Since the modified basecalling model also provides information about 5hmC methylation, the same BAM files were used to asses the changes in 5hmC methylation over promoter regions of L1s. To visualize the data and statistically compare the differences in 5hmC methylation, identical approaches were used as above with the only exception of specifying the "-m h" flag to obtain 5hmC methylation over the chosen regions. The y-axes of the locus plots were altered using "-ymin 0 -ymax 0.2".

Results

L1s are covered by H3K9me3 in hNPCs

To investigate the distribution of H3K9me3 heterochromatin in hNPCs we performed CUT&RUN epigenomic analysis. We used a hNPC line that shares characteristics with neuroepithelial-like stem cells **(Fig. 1A)**. This cell line was derived from fetal brain at six weeks post-conception (26), a time point in post-implantation embryonic development when heterochromatin and DNA methylation patterns have been fully re-established.

Due to the high degree of sequence similarity among L1s, short read sequencing data results in a large proportion of ambiguously mapping reads. To avoid signal amplification due to multimapping artifacts over these elements, we used a strict unique mapping approach to examine individual L1 loci (**Fig. 1A**). With this bioinformatic approach, it is possible to analyze

the epigenetic status of most L1s, except for some of the evolutionarily-youngest elements, where we rely on the flanking regions **(Sup Fig 1A)**. This approach allows us to discriminate reads without ambiguity, and trace the epigenetic modification to a unique locus in the human genome.

We found that H3K9me3 is highly abundant in hNPCs and peak calling identified 30.967 regions in the human genome covered by H3K9me3 **(Fig. 1B, Sup Fig. 1B)**. More than three quarters of the peaks (76.9%) could be confidently mapped to TEs, and among these, L1s represented the most abundant TE class (55.3%) **(Fig. 1C).** We found that H3K9me3 was particularily enriched over evolutionary young L1s (**Fig. 1D-E**). Almost all full-length (>6kb) evolutionarily-young elements, including human-specific (L1HS) and hominoid-specific (L1PA2 and L1PA3) copies, were covered by H3K9me3 in hNPCs **(Fig. 1F-G)**. Notably, the majority of these full-length L1 elements contain an intact 5' promoter and can be transcribed, and some even retain the capacity to retrotranspose (11,18). These data demonstrates that hNPCs represent a useful model for investigating epigenetic mechanisms of L1 regulation, including the presence of heterochromatin, in human somatic neural cells.

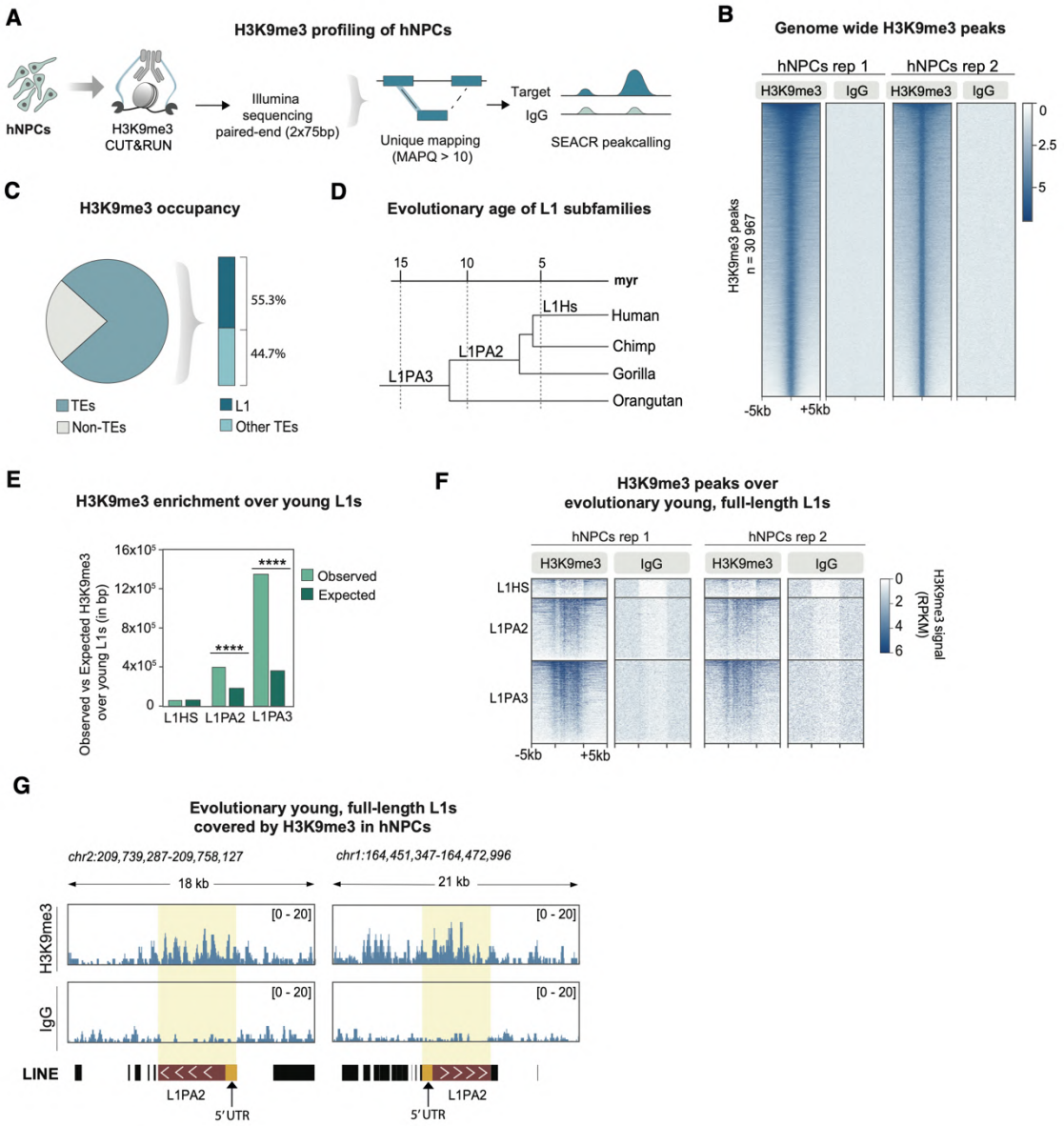


Figure 1. Evolutionary young, full-length L1s are covered by H3K9me3 in hNPCs ***A*** *Schematic illustrating CUT&RUN epigenomic profiling of H3K9me3 in hNPCs.* ***B*** *RPKM normalized heat maps showing genome-wide CUT&RUN signal enrichment of H3K9me3 and non-targeting control IgG in hNPCs (n=2).* ***C*** *Plots showing the genome-wide distribution of H3K9me3 CUT&RUN peaks in hNPCs, categorized by TEs, non-TE (pie chart); L1 elements and other TEs (bar chart).* ***D*** *Phylogenetic tree of evolutionarily young L1 subfamilies.* ***E*** *Barplot showing H3K9me3 enrichment over evolutionary young L1s* ***F*** *RPKM normalized heat maps showing CUT&RUN signal of H3K9me3 in hNPCs and their respective IgG controls, over full-length L1HS to L1PA3 elements sorted by evolutionary age (n=2).* ***G*** *Genome Browser tracks showing RPKM normalized H3K9me3 CUT&RUN data over two full-length L1PA2 elements in hNPCs.*

CRISPRi-based silencing of SETDB1 in hNPCs results in a global reorganization of H3K9me3

The deposition of H3K9me3 at TEs has been linked to the histone methyltransferase SETDB1 (also known as ESET) (44-46), which is highly expressed in hNPCs and the developing and adult human brain **(Sup Fig. 2A-C)**. To study the effects of the loss of SETDB1-mediated H3K9me3, we developed a CRISPRi-based approach to transcriptionally silence SETDB1 in hNPCs (SETDB1-CRISPRi) **(Fig. 2A)**. Two guide RNAs were designed to target near the transcription start site (TSS) of the SETDB1 gene. As a control, we used a guide RNA targeting LacZ, a sequence not found in the human genome. SETDB1 and control guide RNAs were separately cloned into lentiviral vectors expressing a catalytically inactive Cas9 (dCas9) fused to a Krüppel-associated box (KRAB) transcriptional repressor domain (38). Transduced cells were expanded for 10 days prior to FACS isolation of GFP^+ cells carrying the CRISPRi construct. Lentiviral transduction of hNPCs using the SETDB1-CRISPRi vector resulted in efficient, near-complete silencing of SETDB1 at transcript and protein level using both gRNAs, as determined by qRT-PCR, Western blot and RNA sequencing (**Fig. 2B-D, Sup Fig. 3A-C**).

Since CRISPRi works by transcriptionally silencing its target TSS through heterochromatin remodeling via KRAB repressor binding, it was surprising that SETDB1 silencing could be achieved, since the KRAB system depends on this enzyme for H3K9me3 deposition (47). To investigate the mechanism behind this, we performed CUT&RUN profiling of H3K4me3 and H3K9me3 of SETDB1-CRISPRi hNPCs. Analysis of the active promoter mark H3K4me3 revealed a complete loss of H3K4me3 over the SETDB1 TSS following SETDB1-CRISPRi, confirming the loss of transcriptional activity at the SETDB1 promoter. However, this was not accompanied by an accumulation of the repressive histone mark H3K9me3 over the SETDB1 TSS (**Fig. 2D**). These results show that the CRISPRi machinery works efficiently without H3K9me3 deposition, at least at this locus.

The SETDB1-CRISPRi hNPCs remained viable despite the absence of SETDB1 expression and no changes in morphology were observed compared to control cells **(Fig. 2E, Sup Fig 3D)**. Previous studies in mouse model systems have described H3K9me3 as important for maintaining the cellular identity of neural cells and silencing lineage-inappropriate genes (48). However, in SETDB1-CRISPRi hNPCs proliferation remained unchanged, as monitered by EdU incorporation assays **(Fig. 2F, Sup Fig3E)**. Expression of NPC lineage markers did not differ from that of control hNPCs, as determined by bulk RNA sequencing and neuronal and pluripotency markers were not activated, at least over the 10-day timecourse of the experiment (**Fig. 2G, Sup Fig. 3F**). Furthermore, immunofluorescence staining for the NPC markers SOX2 and NESTIN showed no obvious differences in SETDB1-CRISPRi hNPCs compared to controls (**Fig. 2E, Sup Fig. 3D**). To further investigate how silencing of SETDB1 affects the cellular identity of hNPCs, we performed differential gene expression analysis. We found that 748 genes were upregulated and 1202 genes were downregulated in SETDB1-CRISPRi hNPCs compared to controls ($|LFC| > 1$, two-sided padj < 0.05, DESeq2) **(Sup Fig. 3G)**. Gene-ontology overrepresentation test of differentially expressed genes ($|LFC| > 1$) in SETDB1-CRISPRi hNPCs revealed an enrichment of terms related to cell adhesion and synapse assembly among the upregulated genes, and regulation of actin filament organization among the downregulated genes **(Sup Fig. 3H).** Taken together, these results demonstrate that CRISPRi-mediated silencing of SETDB1 results in modest transcriptional changes, but that these do not appear to have a major impact on the cellular identity of hNPCs.

To investigate the role of SETDB1 in depositing H3K9me3 in human NPCs, we performed CUT&RUN profiling of H3K9me3 in SETDB1-CRISPRi and control cells. This revealed a genome-wide loss of the majority of H3K9me3 peaks. Of the 30.967 H3K9me3 peaks found in control hNPCs, 85.6 % were lost in SETDB1-CRISPRi hNPCs **(Fig. 2H).** This included a profound loss of H3K9me3 over evolutionarily-young full-length L1 elements **(Fig. 2I)**.

However, not all H3K9me3 was lost. Surprisingly, we found the appearance of large ectopic regions of H3K9me3, mainly at centromeres and telomeres, but also scattered throughout the genome including regions flanking some of the FL-L1s **(Fig. 3A,B, Sup Fig. 4A)**. The appearance of these ectopic regions of H3K9me3 was also visible when performing immunocytochemistry for H3K9me3 in SETDB1-CRISPRi hNPCs **(Fig. 3C, Sup Fig. 4B).** We found no changes in global H3K9me3 levels using WB analysis **(Fig. 3D, Sup Fig. 3C).** The underlying mechanism for the appearance of these broad ectopic H3K9me3 domains remains unclear but is likely a result of the activity of other H3K9me3 methyltransferases (49) and the complete heterochromatin reorganization that occurs in SETDB1-CRISPRi hNPCs.

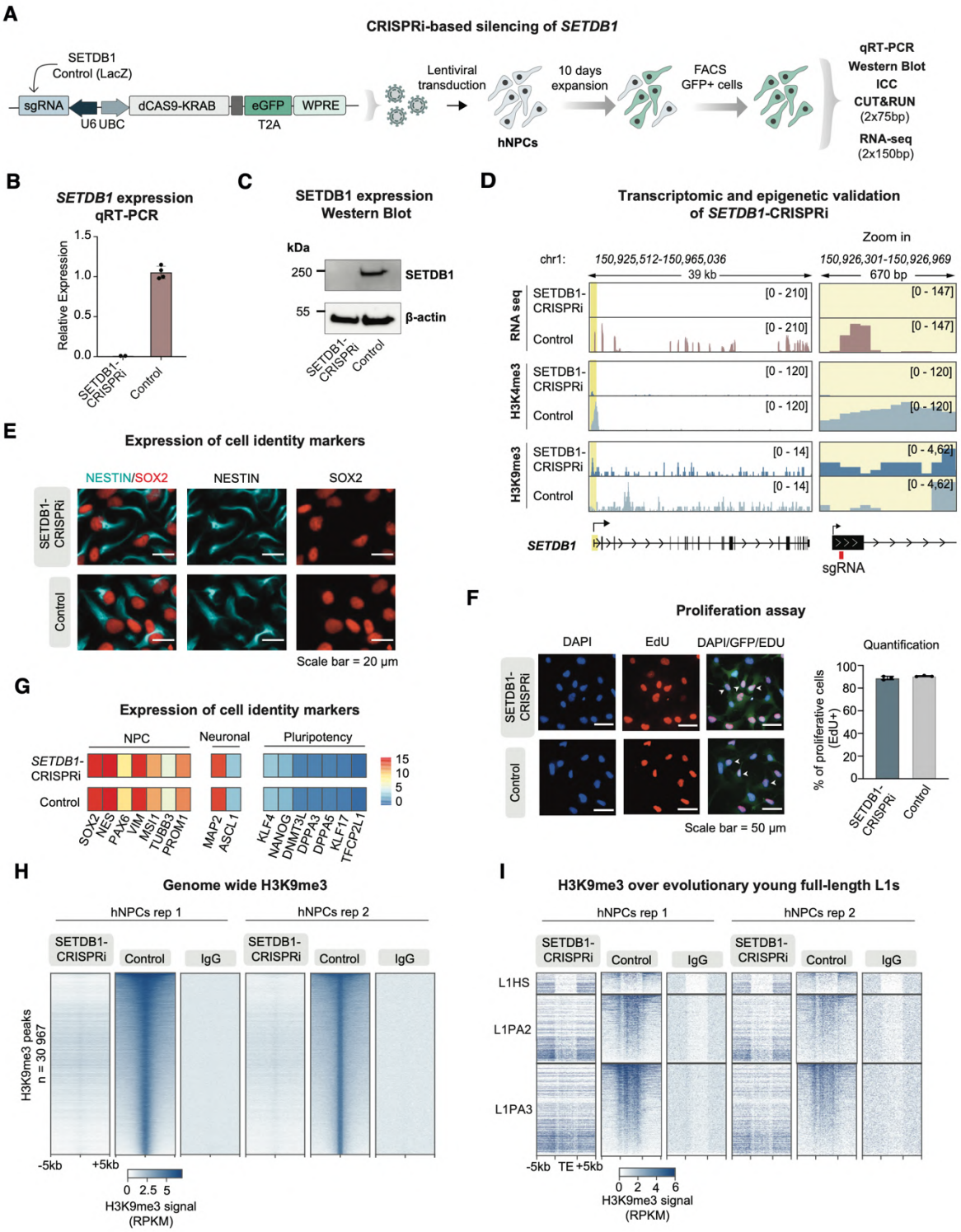


Figure 2. SETDB1-CRISPRi results in loss of H3K9me3 in hNPCs A *Schematic of the CRISPRi construct and workflow for CRISPRi-based silencing of SETDB1 in hNPCs.* ***B*** *qRT-PCR analysis of SETDB1 expression after CRISPRi silencing in hNPCs. Results are shown as means with standard deviation (n=2 and n=4 respectively).* ***C*** *Western blot analysis of SETDB1 protein levels in SETDB1-CRISPRi and control hNPCs relative to β-actin protein levels.* ***D*** *RPKM normalized genome browser tracks showing: expression of SETDB1 in SETDB1-CRISPRi*

and control hNPCs (top). Epigenetic changes as a result of the CRISPRi approach as determined by CUT&RUN profiling of H3K4me3 (middle) and H3K9me3 (bottom). ***E*** *Immunocytochemistry of NESTIN (cyan) and SOX2 (red) in SETDB1-CRISPRi and control hNPCs. Scale bar = 20 µm.* ***F*** *EdU proliferation assay showing cell proliferation in SETDB1-CRISPRi and Control hNPCs. Scale bar=50 µm.* ***G*** *Heat map showing mean expression of NPC, neuronal, and pluripotency gene markers in SETDB1-CRISPRi (n=2) and control (n=4) hNPCs as determined by bulk RNA sequencing.* ***H*** *Heat maps illustrating genome-wide RPKM normalized CUT&RUN signal of H3K9me3 and non-targeting control IgG in SETDB1-CRISPRi (n=2) and control (n=2) hNPCs.* ***I*** *Heat maps illustrating RPKM normalized CUT&RUN signal of H3K9me3 and non-targeting control IgG in SETDB1-CRISPRi (n=2) and control (n=2) hNPCs over full-length L1HS-L1PA3 elements.*

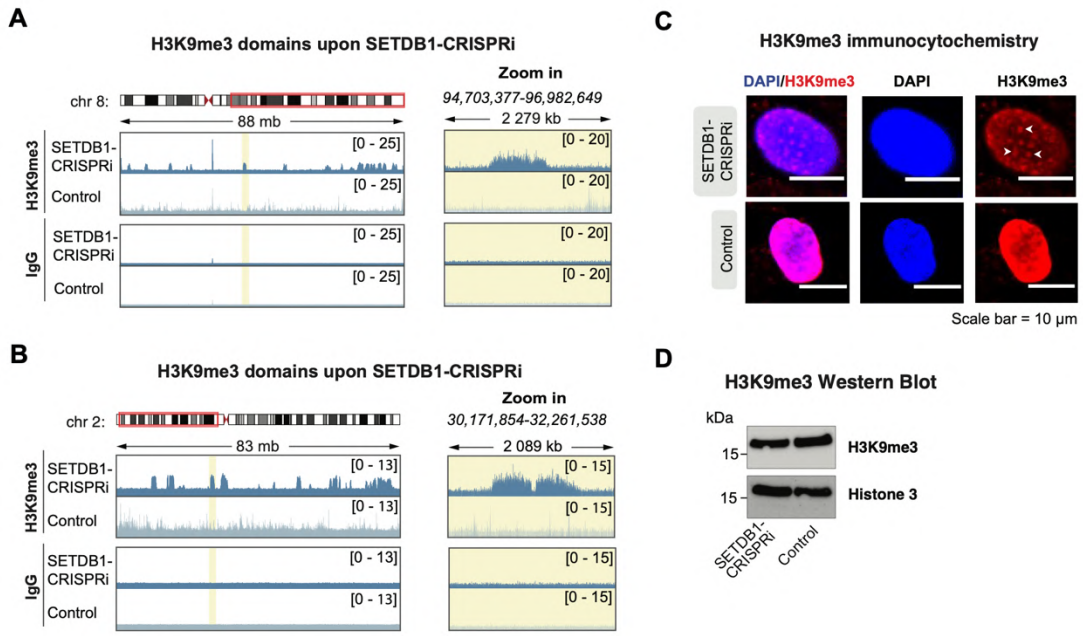


Figure 3 SETDB1-CRISPRi results in reorganisation of H3K9me3 in hNPCs ***A*** *Genome browser track of chromosome 8 showing RPKM normalized CUT&RUN signal of H3K9me3 in SETDB1-CRISPRi and control hNPCs. Right panels show zoom in of H3K9me3 dense areas highlighted in yellow.* ***B*** *Genome browser track of chromosome 2 showing RPKM normalized CUT&RUN signal of H3K9me3 in SETDB1-CRISPRi and control hNPCs. Right panels show zoom in of H3K9me3 dense areas highlighted in yellow.* ***C*** *Immunocytochemistry of H3K9me3 (red) and nuclear marker DAPI (blue) in SETDB1-CRISPRi and control hNPCs. H3K9me3 foci are indicated with white arrows. Scale bar = 10 µm.* ***D*** *Western blot analysis of H3K9me3 protein levels in SETDB1-CRISPRi and control hNPCs relative to Histone 3 protein levels.*

Loss of SETDB1-mediated H3K9me3 results in transcriptional activation of L1s

To investigate the transcriptional consequences on L1s and other TEs after loss of H3K9me3 in hNPCs, we performed bulk RNA sequencing of SETDB1-CRISPRi and control hNPCs. We used an in-house 2 x 150 bp poly(A)-enriched stranded library preparation with a reduced fragmentation step to optimize read length for L1 analysis **(Fig. 4A)**. We obtained between 26 – 76 million uniquely mapped reads per sample. We used a stringent unique mapping approach to quantify expression of individual TE loci (21).

We found 3643 individual TE loci that were upregulated in SETDB1-CRISPRi hNPCs (DESeq2 LFC > 0, two-sided padj 0.05) confirming that many TEs are transcriptionally activated upon loss of H3K9me3 **(Fig. 4B)**. Strikingly, significantly upregulated TEs were enriched for L1s **(Fig 4C)**. When analysing L1-subfamilies we found that this enrichment of upregulated elements was limited to evolutionarily young L1s, including human-specific L1HS and hominoid-specific L1PA2-L1PA3s (**Fig. 4D-E)**. In total, we found 146 individual, full-length L1HS-L1PA3 elements being significantly upregulated **(Fig. 4D-E)**. We found a similar upregulation of L1s when using both SETDB1 gRNAs (**Sup fig. 5A-C**).

When comparing the number of reads transcribed in the same orientation as the L1 (in sense) to reads in the opposite direction (in antisense) we found that most of the transcription in these regions were in sense with regards to the L1s in SETDB1-CRISPRi hNPCs **(Fig 4F, Sup fig. 5D)**. This is consistent with the idea that the detected L1 transcripts come from the L1 promoter, rather than being a result of read-through transcription. We also investigated the genomic location of the L1s that were found to be upregulated in SETDB1-CRISPRi hNPCs. The majority of the upregulated L1s were found in intergenic regions of the genome **(Fig. 4G, Sup fig. 5E)**. Of those located within genes, the majority were found to be antisense to the host genes **(Fig. 4G, Sup fig. 5E)**. These findings also suggest that the upregulated L1 expression

is a consequence of elevated L1 transcription, rather than being related to the expression of genes harbouring an L1 in their introns. In addition to L1s, we also found upregulation of other TEs, such as several ERV families and SVAs, which are know to be controlled by the KZNF-TRIM28 system that depends on SETDB1 **(Table S1.)** (28,50-52).

To further confirm that the the epigenetic status of L1 promoters is associated with an active state in SETDB1-CRISPRi hNPCs, we performed CUT&RUN analysis for the histone mark H3K4me3, which is associated with active promoters. Although this approach is limited to L1 elements where the H3K4me3 peak is broad enough to spread to the surrounding genomic context, we were able to detect 24 high-confidence H3K4me3 peaks located at the 5′ end of evolutionarily young, full-length L1s in SETDB1-CRISPRi NPCs (**Fig 4H-I, Sup fig. 5F**). Most of these loci were also transcriptionally activated and lost H3K9me3 (**Fig 4H-I, Sup fig. 5F**). Taken together, these results show that SETDB1-mediated H3K9me3 is essential to silence evolutionarily-young, full-length L1 elements in hNPCs.

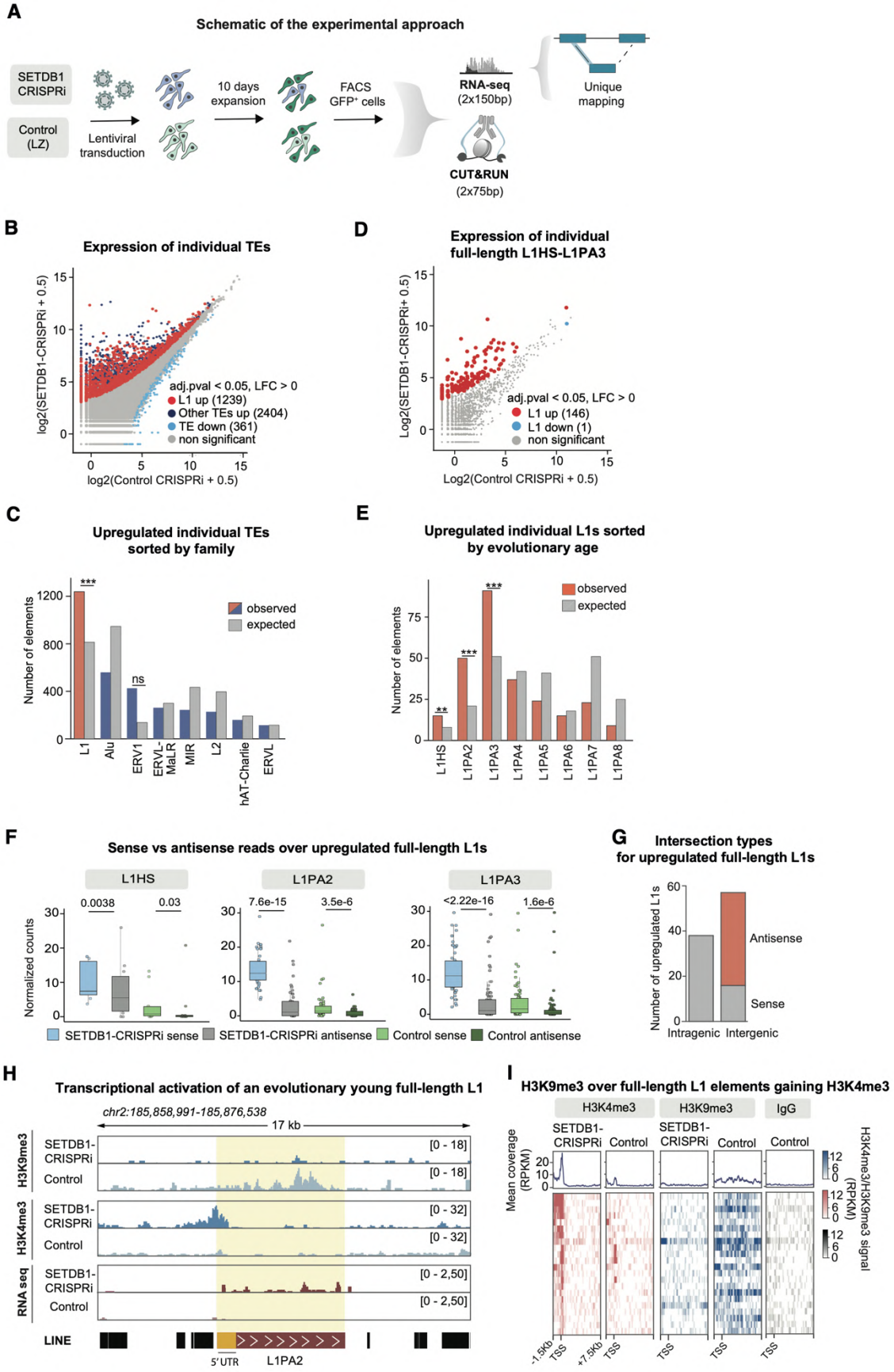

A
Schematic of the experimental approach
SETDB1 CRISPRi
Control (LZ)
Lentiviral transduction
10 days expansion
FACS GFP+ cells
RNA-seq (2x150bp)
CUT&RUN (2x75bp)
Unique mapping
B
Expression of individual TEs
log2(SETDB1-CRISPRi + 0.5)
log2(Control CRISPRi + 0.5)
adj.pval < 0.05, LFC > 0
L1 up (1239)
Other TEs up (2404)
TE down (361)
non significant
D
Expression of individual full-length L1HS-L1PA3
Log2(SETDB1-CRISPRi+ 0.5)
Log2(Control CRISPRi + 0.5)
adj.pval < 0.05, LFC > 0
L1 up (146)
L1 down (1)
non significant
C
Upregulated individual TEs sorted by family
Number of elements
observed
expected
L1
Alu
ERV1
ERVL-MaLR
MIR
L2
hAT-Charlie
ERVL
E
Upregulated individual L1s sorted by evolutionary age
Number of elements
observed
expected
L1HS
L1PA2
L1PA3
L1PA4
L1PA5
L1PA6
L1PA7
L1PA8
F
Sense vs antisense reads over upregulated full-length L1s
L1HS
L1PA2
L1PA3
Normalized counts
0.0038
0.03
7.6e-15
3.5e-6
<2.22e-16
1.6e-6
SETDB1-CRISPRi sense
SETDB1-CRISPRi antisense
Control sense
Control antisense
G
Intersection types for upregulated full-length L1s
Number of upregulated L1s
Antisense
Sense
Intragenic
Intergenic
H
Transcriptional activation of an evolutionary young full-length L1
chr2:185,858,991-185,876,538
17 kb
H3K9me3
H3K4me3
RNA seq
SETDB1-CRISPRi
Control
[0 - 18]
[0 - 32]
[0 - 2,50]
LINE
5' UTR
L1PA2
I
H3K9me3 over full-length L1 elements gaining H3K4me3
H3K4me3
H3K9me3
IgG
SETDB1-CRISPRi
Control
Mean coverage (RPKM)
H3K4me3/H3K9me3 signal (RPKM)
-1.5Kb
TSS
+7.5Kb

Figure 4. Evolutionarily-young, full-length L1s are transcriptionally activated upon loss of H3K9me3 in hNPCs A *Workflow for CUT&RUN and RNA sequencing analysis of SETDB1-CRISPRi and control hNPCs.* ***B*** *Mean plot of uniquely mapped bulk RNA sequencing reads for SETDB1-CRISPRi hNPCs (n=2) compared to control (n=2) over individual TEs categorized into L1 elements and other TEs. Log_2 fold change (LFC) >0, two-sided padj 0.05 calculated with DESeq2.* ***C*** *Observed vs expected number of individual, transcriptionally activated elements per TE family in SETDB1-CRISPRi hNPCs (n= 2) compared to control (n= 2). Pvalues were calculated using bootstrap-based empirical two-sided pvalue calculation.* ***D*** *Mean plot of uniquely mapped bulk RNA sequencing reads for SETDB1-CRISPRi hNPCs (n=4, pooled data from both gRNAs) compared to controls (n=4) over individual full-length (>6 kb) L1HS-L1PA3 elements. LFC > 0, two-sided padj < 0.05 calculated with DESeq2.* ***E*** *Observed vs expected number of transcriptionally activated L1 elements sorted by evolutionary age in SETDB1-CRISPRi (n=2) compared to control hNPCs (n=2). Pvalues were calculated using bootstrap-based empirical two-sided pvalue calculation.* ***F*** *Normalized read counts in sense and antisense for upregulated, full-length L1HS-L1PA3 elements in SETDB1-CRISPRi (n=2) vs control hNPCs (n=4). Statistical significance was calculated with Mann-Whitney U test.* ***G*** *Number, genomic location and direction of upregulated full-length (>6 kb) L1HS-L1PA3 elements in SETDB1-CRISPRi (n=2) compared to control hNPC (n=2)* ***H*** *RPKM normalized genome browser tracks showing: CUT&RUN analysis of H3K9me3 (top), H3K4me3 (middle) and bulk RNA sequencing (bottom) in SETDB1-CRISPRi and control hNPCs.* ***I*** *RPKM normalized heat maps showing CUT&RUN signal of H3K9me3 over full-length (>6 kb) L1HS-L1PA3 elements elements gaining H3K4me3 in hNPCs and their respective IgG controls.*

SETDB1 regulates the expression of evolutionarily young L1s independently from the HUSH/MORC2 corepressor and the KZNF-TRIM28 pathways

The observation that evolutionarily young L1s are activated by loss of H3K9me3 was surprising, as previous work has indicated that DNA methylation is the key repressive mechanism of L1s in neural cells (8,9,23,24). The two major pathways that control the transcription of TEs in association with SETDB1 and H3K9me3 are the KZNF-TRIM28 system and the human silencing hub (HUSH)/MORC2 co-repressor pathway (29,53-56). Previous studies in hNPCs have shown that loss of these pathways in somatic cells has only a

limited effect on L1 expression (23,50). To confirm this, we performed lentiviral-based TRIM28-CRISPRi and MORC2-CRISPRi experiments in hNPCs, comparing the transcriptional activation of evolutionarily-young full-length L1s with that of SETDB1-CRISPRi hNPCs **(Fig. 5A)**. Ten days after lentiviral transduction, we verified the efficient silencing of MORC2 and TRIM28 using qRT-PCR, WB and RNA-seq **(Fig. 5B-D, Sup fig.6A)**. Silencing of TRIM28 or MORC2 resulted in upregulation of only 4 and 5 individual, full-length L1HS-L1PA3 elements, respectively **(Fig. 5E-F)**, which can be compared with 146 individual, full-length L1HS-L1PA3 elements being significantly upregulated after SETDB1 silencing (**Fig. 4D**). To verify that the lack of a transcriptional response of L1s upon silencing of TRIM28 or MORC2 was not due to a redundancy between these pathways, we performed a double CRISPRi-silencing experiment of these two genes (**Fig. 5A**). We verified the efficient double silencing of TRIM28 and MORC2 using qRT-PCR, WB and RNA-seq **(Fig. 5B-D, Sup fig.6A)**. Silencing of both MORC2 and TRIM28 did also not result in transcriptional activation of L1s, only 4 indivual full-length L1HS-L1PA3 elements were upregulated in this setting **(Fig. 5E-F)**. We also used our previously published H3K9me3 CUT&RUN data from TRIM28-CRISPRi and MORC2-CRISPRi hNPCs (23,28) to investigate if these genes are responsible for H3K9me3 at L1s in hNPCs**.** Gobal analysis following the inactivation of TRIM28 or MORC2 confirmed the expected loss of the H3K9me3 signal (Sup fig.6B-C). SETDB1-CRISPRi resulted in major loss of H3K9me3 peaks over full-length L1HS-L1PA3 elements **(Fig. 2I),** while MORC2-CRISPRi and TRIM28-CRISPRi had minor effects on H3K9me3 deposition over full-length L1HS-L1PA3 elements **(Fig. 5G).** Together these results show that in addition to the TEs that are repressed by HUSH and KZNF-TRIM28 and thus dependendent on SETDB1, SETDB1 also acts through alternative pathways to maintain L1 promoters with H3K9me3 leading to their transcriptional repression.

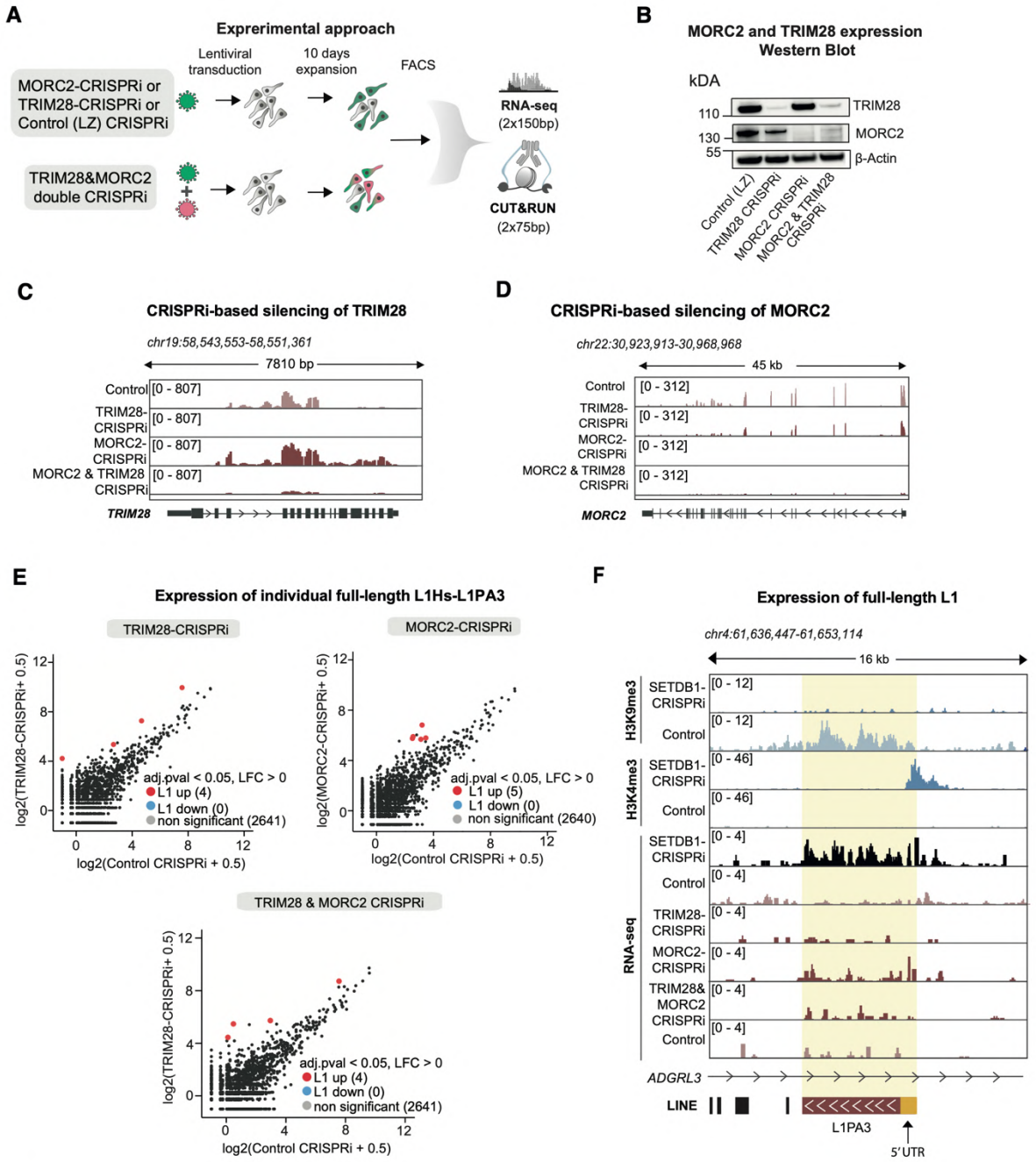

A
Exprerimental approach
Lentiviral transduction
10 days expansion
FACS
MORC2-CRISPRi or TRIM28-CRISPRi or Control (LZ) CRISPRi
TRIM28&MORC2 double CRISPRi
RNA-seq (2x150bp)
CUT&RUN (2x75bp)
B
MORC2 and TRIM28 expression Western Blot
kDA
110
130
55
TRIM28
MORC2
β-Actin
Control (LZ)
TRIM28 CRISPRi
MORC2 CRISPRi
MORC2 & TRIM28 CRISPRi
C
CRISPRi-based silencing of TRIM28
chr19:58,543,553-58,551,361
7810 bp
Control
TRIM28-CRISPRi
MORC2-CRISPRi
MORC2 & TRIM28 CRISPRi
[0 - 807]
TRIM28
D
CRISPRi-based silencing of MORC2
chr22:30,923,913-30,968,968
45 kb
[0 - 312]
MORC2
E
Expression of individual full-length L1Hs-L1PA3
TRIM28-CRISPRi
MORC2-CRISPRi
TRIM28 & MORC2 CRISPRi
log2(TRIM28-CRISPRi+ 0.5)
log2(MORC2-CRISPRi+ 0.5)
log2(Control CRISPRi + 0.5)
adj.pval < 0.05, LFC > 0
L1 up (4)
L1 down (0)
non significant (2641)
L1 up (5)
non significant (2640)
F
Expression of full-length L1
chr4:61,636,447-61,653,114
16 kb
H3K9me3
H3K4me3
RNA-seq
SETDB1-CRISPRi
Control
TRIM28-CRISPRi
MORC2-CRISPRi
TRIM28& MORC2 CRISPRi
[0 - 12]
[0 - 46]
[0 - 4]
ADGRL3
LINE
L1PA3
5'UTR
G
H3K9me3 over evolutionary young full-length L1s
TRIM28-CRISPRi
Control
IgG
MORC2-CRISPRi
L1HS
L1PA2
L1PA3
-5kb TSS TES +5kb

Figure 5. SETDB1 regulates evolutionary young, full-length L1 elements independently from MORC2 and TRIM28 A *Workflow of H3K9me3 and H3K4me3 CUT&RUN and RNA sequencing analysis in TRIM28-CRISPRi, MORC2-CRISPRi, TRIM28&MORC2 double CRISPRi and control hNPCs.* ***B*** *Western blot analysis of TRIM28 and MORC2 protein levels in TRIM28-CRISPRi, MORC2-CRISPRi, TRIM28&MORC2 double CRISPRi and control hNPCs relative to ß-actin protein levels.* ***C*** *RPKM normalized genome browser tracks showing TRIM28 expression in TRIM28-CRIPSRi, MORC2-CRISPRi, TRIM28&MORC2 double CRISPRi and control hNPCs.* ***D*** *RPKM normalized genome browser tracks showing MORC2 expression in TRIM28-CRIPSRi, MORC2-CRISPRi, TRIM28&MORC2 double CRISPRi and control hNPCs* ***E*** *Mean plot of uniquely mapped bulk RNA sequencing reads for TRIM28-CRISPRi (n=3), MORC2-CRISPRi (n=3), TRIM28&MORC2 double CRISPRi (n=3) hNPCs compared to controls (n=3) over full-length (>6 kb) L1HS-L1PA3 elements. LFC > 0, two-sided padj < 0.05 calculated with DESeq2.* ***F*** *RPKM normalized genome browser tracks showing: CUT&RUN analysis in SETDB1-CRISPRi hNPCs of H3K9me3 (top) and H3K4me3 (middle). RNA expression of an L1PA3 element in SETDB1-CRISPRi, TRIM28-CRISPRi, MORC2-CRISPRi, TRIM28&MORC2 double CRISPRi and hNPCs compared to controls (bottom).* ***G*** *RPKM normalized heat maps showing CUT&RUN signal of H3K9me3 over full-length (> 6kb) L1HS to L1PA3 elements sorted by evolutionary age in MORC2-CRISPRi and TRIM28-CRISPRi hNPCs compared to their respective controls and non-targeting control IgG. Data from Pandiloski et al. (2024) (23) and Horvath et al. (2024) (28)*

Loss of H3K9me3 results reduced CpG DNA methylation at L1 promoters

CpG DNA methylation of the L1 promoter has been described as the major mechanism for transcriptional silencing of L1 elements in somatic cells, including hNPCs (7-10,22,23). Therefore, we hypothesized that loss of H3K9me3 may have downstream effects on DNA methylation levels at L1s. To investigate the interplay between H3K9me3 and CpG DNA methylation, we performed genome-wide DNA methylation profiling using Oxford Nanopore Technologies (ONT) long-read DNA sequencing on SETDB1-CRISPRi and control hNPCs **(Fig. 6A)**. As an additional positive control, we used CRISPRi to silence DNMT1 (23,28), whose deletion in hNPCs results in a global loss of DNA methylation, including over L1s (8) (**Fig. 6A, Sup Fig 7A-B**).

We observed no major changes in global genome-wide DNA methylation between SETDB1-CRISPRi and control hNPCs, suggesting that DNA methylation patterns are generally unaffected by the loss of H3K9me3 over the 10-day timecourse of the experiment **(Fig. 6B)**. This contrasts with the silencing of DNMT1, which results in global DNA hypomethylation **(Fig. 6B)**. However, when we performed DNA methylation analysis of evolutionarily young L1s, we found that SETDB1-CRISPRi hNPCs exhibited lower levels of CpG DNA methylation compared to control hNPCs. In particular, L1s that were transcriptionally activated in hNPCs showed significantly lower levels of DNA methylation over their 5' promoter region **(Fig. 6C, Sup Fig.7C-E)**. For example, we found full-length human- and hominoid-specific L1 elements that were transcriptionally activated in SETDB1-CRISPRi hNPCs, as determined by bulk RNA sequencing and CUT&RUN profiling of the active promoter mark H3K4me3, and also lost DNA methylation **(Fig. 6D-E)**. When investigating DNA methylation over all full-length L1s and those which did not get transcriptionally activated we found that some L1s in these groups also loose DNA methylation, highlighting the challanges to detect the RNA signal for some of the evolutionary young L1s (Sup Fig.7F-I). Taken together, these results demonstrate that the loss of SETDB1-mediated H3K9me3 in human somatic cells leads to subsequent loss of DNA methylation and gain of H3K4me3 methylation, resulting in transcriptional activation of L1s.

The loss of DNA methylation could result from an active process involving ten-eleven translocation (TET) enzymes, which is characterized by the conversion of methylcytosine to hydroxymethylcytosine (57). Alternatively, it may result from a passive process associated with an inability to attract DNMT1, which is characterized by the loss of methylcytosine. Therefore, we investigated the presence of DNA hydroxymethylation in SETDB1-CRISPRi and control hNPCs, as determined from the ONT data, using DNMT1-CRISPRi as a negative control. While we did not see global changes in DNA hydroxymethylation at a genome-wide level in SETDB1-CRISPRi compared to control cells, our analysis revealed that

transcriptionally activated full-length L1s gained DNA hydroxymethylation over their promoter (Fig. 7B-D, Sup Fig.8A-F). This suggests that the loss of DNA methylation at L1s with SETDB1-CRISPRi occurs via an active process involving TET enzymes. In contrast, DNMT1-CRISPRi NPCs do not exhibit an increase in DNA hydroxymethylation at these elements (Fig. 7B). This is consistent with passive loss of DNA methylation resulting from the loss of DNMT1 activity. Thus, these results demonstrate that the loss of DNA methylation specifically over the promoter region of L1s in SETDB1-CRISPRi NPCs is an active process likely mediated by the recruitment of TET enzymes.

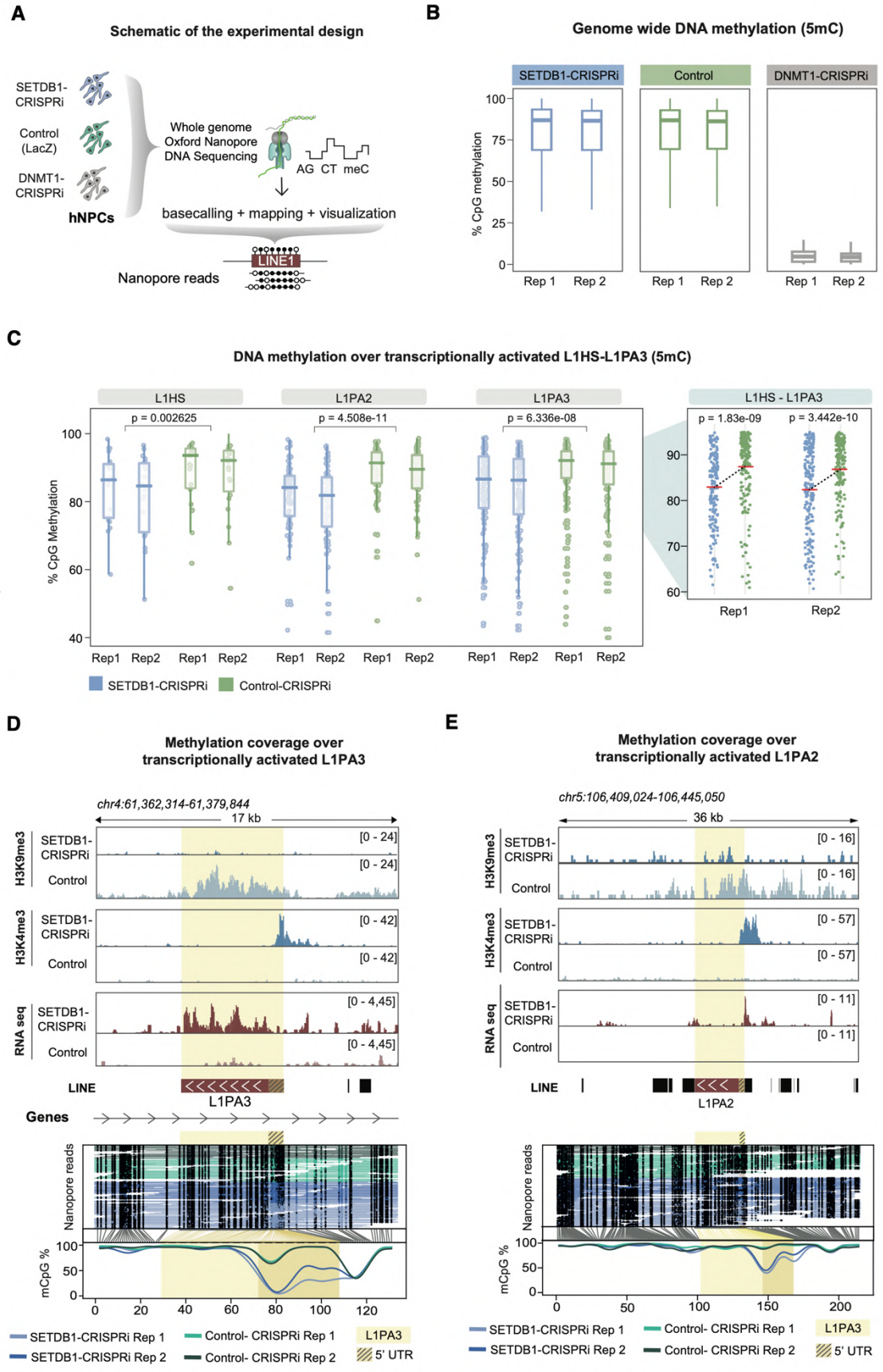

A
Schematic of the experimental design
SETDB1-CRISPRi
Control (LacZ)
DNMT1-CRISPRi
hNPCs
Whole genome Oxford Nanopore DNA Sequencing
AG CT meC
basecalling + mapping + visualization
LINE1
Nanopore reads
B
Genome wide DNA methylation (5mC)
SETDB1-CRISPRi
Control
DNMT1-CRISPRi
% CpG methylation
Rep 1
Rep 2
C
DNA methylation over transcriptionally activated L1HS-L1PA3 (5mC)
L1HS
L1PA2
L1PA3
L1HS - L1PA3
p = 0.002625
p = 4.508e-11
p = 6.336e-08
p = 1.83e-09
p = 3.442e-10
% CpG Methylation
Rep1
Rep2
SETDB1-CRISPRi
Control-CRISPRi
D
Methylation coverage over transcriptionally activated L1PA3
chr4:61,362,314-61,379,844
17 kb
H3K9me3
H3K4me3
RNA seq
SETDB1-CRISPRi
Control
LINE
L1PA3
Genes
Nanopore reads
mCpG %
SETDB1-CRISPRi Rep 1
SETDB1-CRISPRi Rep 2
Control- CRISPRi Rep 1
Control- CRISPRi Rep 2
L1PA3
5' UTR
E
Methylation coverage over transcriptionally activated L1PA2
chr5:106,409,024-106,445,050
36 kb
L1PA2

Figure 6. Loss of H3K9me3 results in loss of CpG DNA methylation at L1 promoters in hNPCs A *Schematics of whole genome ONT long-read DNA sequencing analysis in hNPCs* ***B*** *Box plots of the genome wide CpG DNA methylation status in SETDB1-CRISPRi (n=2), control (n=2) and DNMT1-CRISPRi (n=2) hNPCs.* ***C*** *Box plots of the methylation status over the promoter of transcriptionally activated L1HS-L1PA3 elements in SETDB1-CRISPRi (n=2) and control (n=2) hNPCs. Zoom-in panels indicating mean methylation levels (red line) per condition for L1HS-L1PA3 elements. The figure shows FL-L1s that were found to be upregulated in at least one of the two experimental conditions (two different gRNAs). This resulted in 241 FL-L1s being included in the analysis. n values: L1HS = 22; L1PA2 = 77; and L1PA3 = 142.* ***D&E*** *RPKM normalized genome browser tracks showing from top to bottom: CUT&RUN analysis of H3K9me3 and H3K4me3 in SETDB1-CRISPRi compared to control hNPCs, RNA expression of L1s in SETDB1-CRISPRi compared to controls hNPCs. ONT DNA reads across SETDB1-CRISPRi and control hNPCs. Black dots indicate methylated CpGs, and methylation coverage of the L1 element can be seen at the bottom. The L1 element is highlighted in yellow, while the 5' UTR is marked with stripes.*

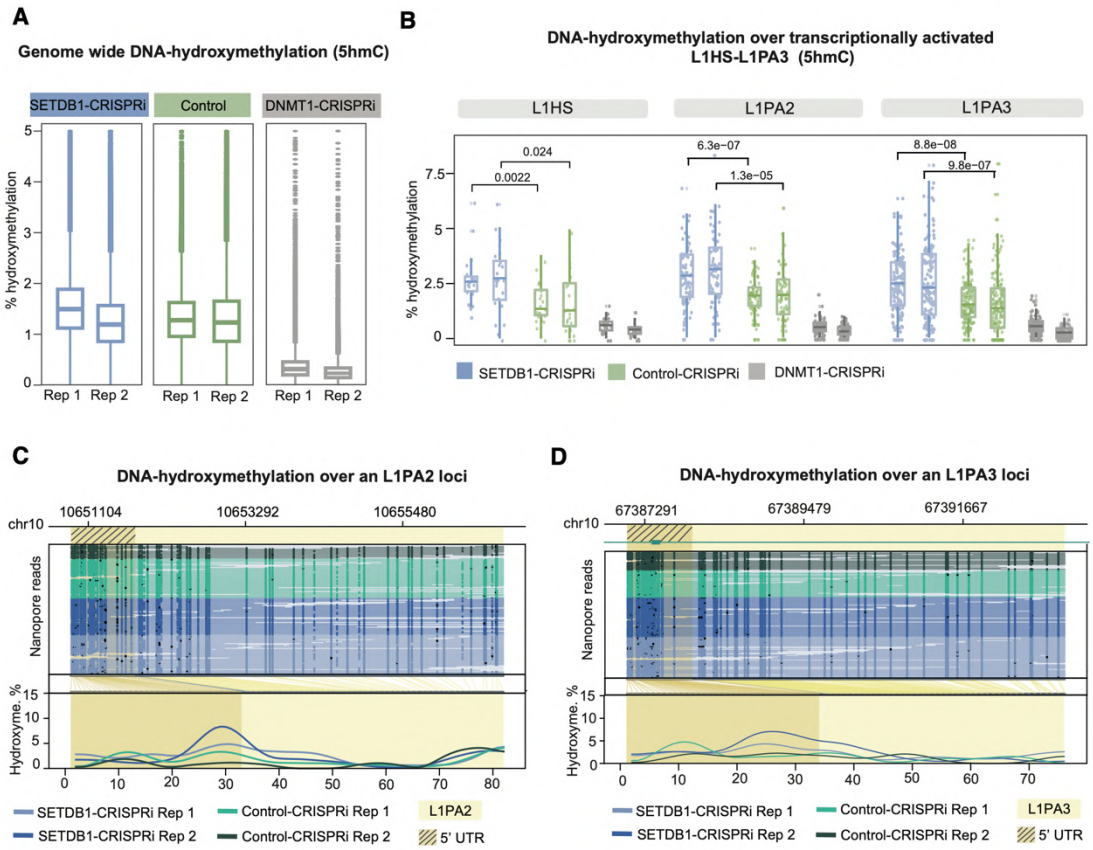


Figure 7. Loss of H3K9me3 results in gain of DNA hydroxymethylation A *Box plots of the genome wide DNA hydroxymethylation status in SETDB1-CRISPRi (n=2), control (n=2) and DNMT1-CRISPRi (n=2) hNPCs.* ***B*** *Box plots of the DNA hydroxymethylation status over the promoter of transcriptionally activated L1HS-L1PA3 elements in SETDB1-CRISPRi (n=2), control (n=2)and DNMT1-CRISPRi hNPCs. The figure shows FL-L1s that were found to be upregulated in at least one of the two experimental conditions (two different gRNAs, 241 FL-L1s being included in the analysis. N: L1HS = 22; L1PA2 = 77; and L1PA3 = 142).* **C** *Locus plot showing DNA hydroxymethylation over an upregulated L1PA2 in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. Black dots indicate methylated CpGs, and methylation coverage of the L1 element can be seen at the bottom. The L1 element is highlighted in yellow, while the 5' UTR is marked with stripes.* ***D*** *Locus plot showing DNA hydroxymethylation over an upregulated L1PA3 in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. Black dots indicate methylated CpGs, and methylation coverage of the L1 element can be seen at the bottom. The L1 element is highlighted in yellow, while the 5' UTR is marked with stripes.*

Discussion

Here we study the consequences of loss of H3K9me3 heterochromatin in hNPCs, a cell type that allows us to model some aspects of heterochromatin maintenance in a human setting, including the presence of H3K9me3 and DNA methylation at L1s (8,23). Upon CRISPRi-based silencing of the H3K9me3 methyltransferase SETDB1 in hNPCs, we found a major loss of H3K9me3 peaks, including those over L1s and other TEs. Moreover, we observed a general reorganization of H3K9me3 domains, including the appearance of new ectopically large heterochromatin domains, as detected by both CUT&RUN analysis and confocal microscopy. While we do not fully understand the mechanism behind this phenomenon, previous studies have reported that a global loss of H3K9me3 leads to rearrangements of the epigenome and the gain of H3K9me3 at specific loci mediated by other H3K9me3 methyltransferases such as SUV39H1/2 (58-60). However, despite this global heterochromatin reorganization, SETDB1-

CRISPRi hNPCs remained viable, proliferative, and express appropriate levels of NPC marker genes, suggesting that the presence of H3K9me3 at TEs is not critical for maintaining somatic cell identity and viability, at least under the experimental conditions in this study.

L1s are the only autonomously mobilizing family of TEs in humans (61) and can impact genome stability via several mechanisms, including new somatic retrotransposition events and alterations of transcriptional networks where transcriptionally active L1s may impact the genome both in *cis* and *trans* (21,62-66). L1s are therefore mostly transcriptionally silenced in somatic human cells. Our data demonstrate that H3K9me3 and SETDB1 play a critical role in this process. In the absence of SETDB1 and H3K9me3, DNA methylation is lost over the promoter of evolutionary young L1s, which correlates with their transcriptional activation. This is in line with a model where the deposition of H3K9me3 at the L1 promoter is needed to maintain DNA methylation (67-69). The best understood mechanisms for deposition of H3K9me3 on TEs are the KRAB zinc finger proteins (KZNF)-TRIM28 system and the HUSH pathway. However, the control of evolutionary young L1s by KZFNs remains poorly understood. A previous study found that ancestral L1PA3 elements underwent structural sequence changes that allowed them, and subsequent younger L1 families, to escape silencing by ZNF93, which otherwise targets L1s (70). Although other KZNFs have been reported to bind to young L1s (31), these have not been studied in detail and appear not to be active in NPCs, as evidenced by the results of our TRIM28-CRISPRi experiments. The HUSH pathway can target evolutionary young L1s (56,71), but recent data from our labs show that this is restricted to transcriptionally active elements that are not covered by DNA methylation at steady-state, as is the case for almost all L1s in hNPCs (8,9,23). Thus, the pathway by which H3K9me3 is maintained over evolutionarily-young L1s resulting in their silencing in somatic cells remains to be fully elucidated. However, our data show that the deposition of H3K9me3 at L1s is necessary to maintain DNA methylation over these elements in somatic cells. This

likely occurs by counteracting the demethylation activity of TET enzymes. Thus, our data demonstrate that DNMT1 alone is insufficient for maintaining methylation patterns at the promoter of evolutionary young L1s.

The loss of H3K9me3 and heterochromatin is a hallmark of aging (72). Experiments conducted across diverse organisms and tissues have provided evidence of age-related heterochromatin loss and there is a gradual loss of H3K9me3 during the aging process in human model systems (73-76). While our H3K9me3 deficient hNPCs have obvious limitations as a model for human brain ageing, our results show that a key event following the loss of H3K9me3 maintenance is the transcriptional activation of L1s. In line with this, a number of recent studies have shown that increased L1 expression correlates with senescence and other age-related cellular phenotypes, and suggest that aberrant transcription of L1s may be associated with ageing and related disease processes (77-81).

The transcriptional activation of L1s is likely to have consequences, as evolutionarily young, intact L1 elements have the potential to retrotranspose and thus can cause deleterious mutations and structural variants (11,82-85). In addition, several recent studies have shown that L1 transcription itself can affect multiple cellular pathways both in *cis* and in *trans*. For example, we have recently shown that when DNA methylation is experimentally removed in human NPCs, transcriptionally activated L1 elements act as alternative promoters, affecting the expression levels of many protein-coding genes (8,21). These observations suggest that the aberrant activation of L1s could lead to widespread dysregulation of gene regulatory networks. In line with this, we observed in the current study that genes located in proximity of full-length L1HS-L1PA3 are upregulated upon SETDB1-CRISPRi **(Sup Fig. 9A)**. This supports the idea that L1 elements regulated by H3K9me3 in hNPCs can have a regulatory effect on nearby gene expression. Furthermore, L1 transcripts and L1-derived peptides can be sensed by the cellular machinery to trigger an inflammatory response. For example, L1 expression has been reported

to be increased in senescent cells and to activate an interferon response (78,79,86,87). Thus, transcriptional activation by L1s in aged cells, as a consequence of H3K9me3 loss, may contribute to the ageing process via multiple mechanisms. However, we did not observe an interferon response in SETDB1-CRISPRi hNPCs, which may be explained by the cell type used in this study **(Sup Fig. 9B)**.

In conclusion, our results provide a unique insight into the consequences of loss of H3K9me3 in human somatic cells. SETDB1-mediated H3K9me3 is not required to maintain cellular identity and viability, but rather provides a critical mechanism to repress evolutionarily young L1 elements.

Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The sequencing data presented in this study has been deposited at the GEO superseries GSE292899. Reviewer access to the data: GSE294699 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE294699; token: stgpoewcllshdkz) and GSE292899 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE292899; token: srwfscggjlepdiz). Original code has been deposited at GitHub and is publicly available at git@github.com:Molecular-Neurogenetics/SETDB1_Karlsson_2025.git.

Acknowledgements

We would like to thank S. Henikoff for providing reagents and U. Jarl, for technical assistance. We acknowledge Clinical Genomics Lund, SciLifeLab, and Center for Translational Genomics (CTG) Lund University for providing expertise and service with sequencing and analysis. We are grateful to all members of the Jakobsson laboratory.

Author contributions

Design and interpretation: All authors. Conceptualization: O.K., N.P., and J.J. Experimental research: O.K., A.A., V.H., P.J. and J.G.J. Bioinformatics: N.P. and R.G. Writing—original draft: O.K., N.P. and J.J. Writing—review and editing: All authors.

Funding

This work was supported by grants from the Swedish Research Council (2022-00673 to J.J. and 2021-03494 to C.H.D.), the Swedish Brain Foundation (FO2023-0232 to J.J.), Cancerfonden (222185 to J.J.), Barncancerfonden (PR2023-0099to J.J.), the Swedish Society for Medical Research (S19-0100 to C.H.D.), Olle Engqvists stiftelse (218-0090 to J.J.) and the Swedish Government Initiative for Strategic Research Areas (MultiPark & StemTherapy). This project has been made possible in part by grant 2023-331773 to C.H.D. from the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation.

Competing interests: The authors declare that they have no competing interests.

References

1. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
2. de Koning, A.P., Gu, W., Castoe, T.A., Batzer, M.A. and Pollock, D.D. (2011) Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet*, **7**, e1002384.
3. Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A.V., Mikheenko, A., Vollger, M.R., Altemose, N., Uralsky, L., Gershman, A. *et al.* (2022) The complete sequence of a human genome. *Science*, **376**, 44–53.
4. Hoyt, S.J., Storer, J.M., Hartley, G.A., Grady, P.G.S., Gershman, A., de Lima, L.G., Limouse, C., Halabian, R., Wojenski, L., Rodriguez, M. *et al.* (2022) From telomere to

telomere: The transcriptional and epigenetic state of human repeat elements. *Science*, **376**, eabk3112.
5. Payer, L.M. and Burns, K.H. (2019) Transposable elements in human genetic disease. *Nat Rev Genet*, **20**, 760–772.
6. Deniz, O., Frost, J.M. and Branco, M.R. (2019) Regulation of transposable elements by DNA modifications. *Nat Rev Genet*, **20**, 417–431.
7. Yoder, J.A., Walsh, C.P. and Bestor, T.H. (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics*, **13**, 335–340.
8. Jonsson, M.E., Ludvik Brattas, P., Gustafsson, C., Petri, R., Yudovich, D., Pircs, K., Verschuere, S., Madsen, S., Hansson, J., Larsson, J. *et al.* (2019) Activation of neuronal genes via LINE-1 elements upon global DNA demethylation in human neural progenitors. *Nat Commun*, **10**, 3182.
9. Sanchez-Luque, F.J., Kempen, M.-J.H.C., Gerdes, P., Vargas-Landin, D.B., Richardson, S.R., Troskie, R.-L., Jesuadian, J.S., Cheetham, S.W., Carreira, P.E., Salvador-Palomeque, C. *et al.* (2019) LINE-1 Evasion of Epigenetic Repression in Humans. *Molecular Cell*, **75**, 590–604.e512.
10. Lanciano, S., Philippe, C., Sarkar, A., Pratella, D., Domrane, C., Doucet, A.J., van Essen, D., Saccani, S., Ferry, L., Defossez, P.A. *et al.* (2024) Locus-level L1 DNA methylation profiling reveals the epigenetic and transcriptional interplay between L1s and their integration sites. *Cell Genom*, **4**, 100498.
11. Beck, C.R., Garcia-Perez, J.L., Badge, R.M. and Moran, J.V. (2011) LINE-1 elements in structural variation and disease. *Annu Rev Genomics Hum Genet*, **12**, 187–215.
12. Feng, Q., Moran, J.V., Kazazian, H.H., Jr. and Boeke, J.D. (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell*, **87**, 905–916.
13. Swergold, G.D. (1990) Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol Cell Biol*, **10**, 6718–6729.
14. Mathias, S.L., Scott, A.F., Kazazian, H.H., Boeke, J.D. and Gabriel, A. (1991) Reverse Transcriptase Encoded by a Human Transposable Element. *Science*, **254**, 1808–1810.
15. Khazina, E. and Weichenrieder, O. (2018) Human LINE-1 retrotransposition requires a metastable coiled coil and a positively charged N-terminus in L1ORF1p. *eLife*, **7**, e34960.
16. Khan, H., Smit, A. and Boissinot, S. (2006) Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. *Genome Res*, **16**, 78–87.
17. Beck, C.R., Collier, P., Macfarlane, C., Malig, M., Kidd, J.M., Eichler, E.E., Badge, R.M. and Moran, J.V. (2010) LINE-1 Retrotransposition Activity in Human Genomes. *Cell*, **141**, 1159–1170.
18. Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J.V. and Kazazian, H.H., Jr. (2003) Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A*, **100**, 5280–5285.
19. Denli, A.M., Narvaiza, I., Kerman, B.E., Pena, M., Benner, C., Marchetto, M.C., Diedrich, J.K., Aslanian, A., Ma, J., Moresco, J.J. *et al.* (2015) Primate-specific ORF0 contributes to retrotransposon-mediated diversity. *Cell*, **163**, 583–593.
20. Speek, M. (2001) Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol Cell Biol*, **21**, 1973–1985.
21. Garza, R., Atacho, D.A.M., Adami, A., Gerdes, P., Vinod, M., Hsieh, P., Karlsson, O., Horvath, V., Johansson, P.A., Pandiloski, N. *et al.* (2023) LINE-1 retrotransposons drive human neuronal transcriptome complexity and functional diversification. *Sci Adv*, **9**, eadh9543.

22. Deniz, Ö., Frost, J.M. and Branco, M.R. (2019) Regulation of transposable elements by DNA modifications. *Nature Reviews Genetics*, **20**, 417–431.
23. Pandiloski, N., Horváth, V., Karlsson, O., Koutounidou, S., Dorazehi, F., Christoforidou, G., Matas-Fuentes, J., Gerdes, P., Garza, R., Jönsson, M.E. *et al.* (2024) DNA methylation governs the sensitivity of repeats to restriction by the HUSH-MORC2 corepressor. *Nature Communications*, **15**, 7534.
24. Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V. and Gage, F.H. (2009) L1 retrotransposition in human neural progenitor cells. *Nature*, **460**, 1127–1131.
25. Macia, A., Widmann, T.J., Heras, S.R., Ayllon, V., Sanchez, L., Benkaddour-Boumzaouad, M., Muñoz-Lopez, M., Rubio, A., Amador-Cubero, S., Blanco-Jimenez, E. *et al.* (2017) Engineered LINE-1 retrotransposition in nondividing human neurons. *Genome Res*, **27**, 335–348.
26. Tailor, J., Kittappa, R., Leto, K., Gates, M., Borel, M., Paulsen, O., Spitzer, S., Karadottir, R.T., Rossi, F., Falk, A. *et al.* (2013) Stem Cells Expanded from the Human Embryonic Hindbrain Stably Retain Regional Specification and High Neurogenic Potency. *The Journal of Neuroscience*, **33**, 12407.
27. Greenberg, M.V.C. and Bourc'his, D. (2019) The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*, **20**, 590–607.
28. Horváth, V., Garza, R., Jönsson, M.E., Johansson, P.A., Adami, A., Christoforidou, G., Karlsson, O., Castilla Vallmanya, L., Koutounidou, S., Gerdes, P. *et al.* (2024) Mini-heterochromatin domains constrain the cis-regulatory impact of SVA transposons in human brain development and disease. *Nature Structural & Molecular Biology*.
29. Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J. *et al.* (2010) KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature*, **463**, 237–240.
30. Ecco, G., Imbeault, M. and Trono, D. (2017) KRAB zinc finger proteins. *Development*, **144**, 2719–2729.
31. Imbeault, M., Helleboid, P.-Y. and Trono, D. (2017) KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature*, **543**, 550–554.
32. Falk, A., Koch, P., Kesavan, J., Takashima, Y., Ladewig, J., Alexander, M., Wiskow, O., Tailor, J., Trotter, M., Pollard, S. *et al.* (2012) Capture of neuroepithelial-like stem cells from pluripotent stem cells provides a versatile system for in vitro production of human neurons. *PLoS One*, **7**, e29597.
33. Skene, P.J. and Henikoff, S. (2017) An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife*, **6**.
34. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
35. 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**, W187–191.
36. Meers, M.P., Tenenbaum, D. and Henikoff, S. (2019) Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. *Epigenetics Chromatin*, **12**, 42.
37. Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L. and Trono, D. (1997) Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat Biotechnol*, **15**, 871–875.

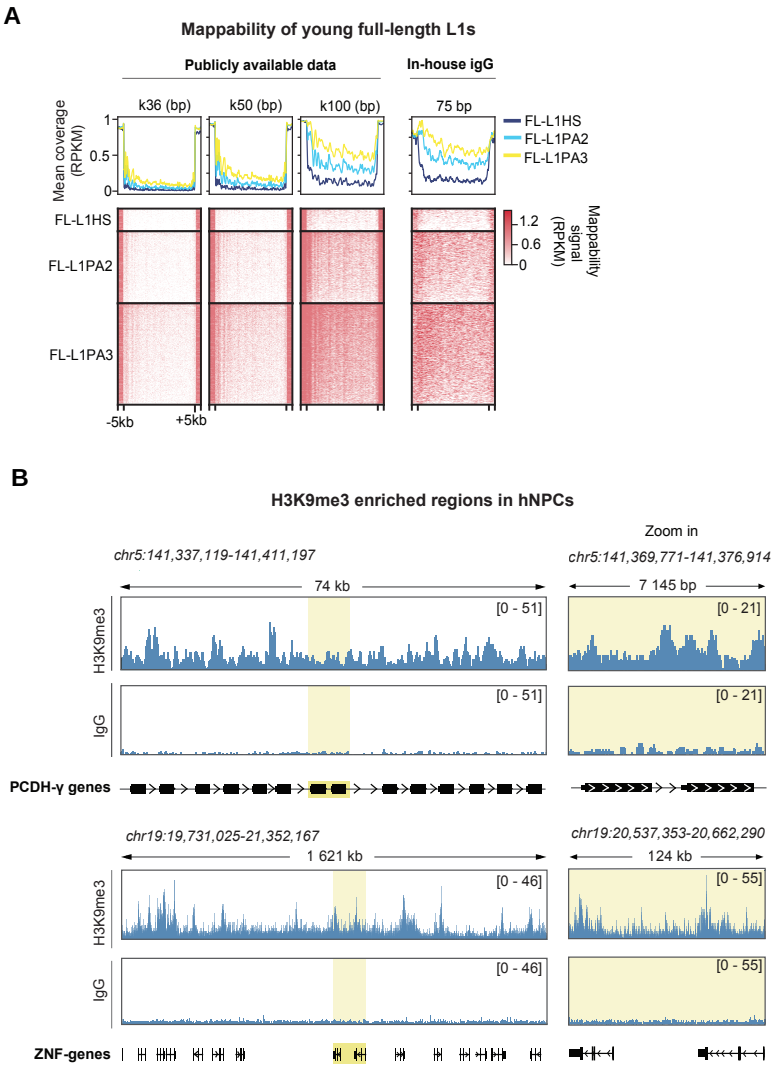
38. Johansson, P.A., Adami, A. and Jakobsson, J. (2022) CRISPRi-mediated transcriptional silencing in iPSCs for the study of human brain development. *STAR Protoc*, **3**, 101285.
39. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
40. 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.
41. 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.
42. Li, H. (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, **34**, 3094–3100.
43. Cheetham, S.W., Kindlova, M. and Ewing, A.D. (2022) Methylartist: tools for visualizing modified bases from nanopore sequence data. *Bioinformatics*, **38**, 3109–3112.
44. Kato, M., Takemoto, K. and Shinkai, Y. (2018) A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing. *Nat Commun*, **9**, 1683.
45. Sun, D., Zhu, Y., Peng, W., Zheng, S., Weng, J., Dong, S., Li, J., Chen, Q., Ge, C., Liao, L. *et al.* (2024) SETDB1 regulates short interspersed nuclear elements and chromatin loop organization in mouse neural precursor cells. *Genome Biol*, **25**, 175.
46. Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G. and Rauscher, F.J., 3rd. (2002) SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev*, **16**, 919–932.
47. Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E. and Gersbach, C.A. (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nature Methods*, **12**, 1143–1149.
48. Tan, S.L., Nishi, M., Ohtsuka, T., Matsui, T., Takemoto, K., Kamio-Miura, A., Aburatani, H., Shinkai, Y. and Kageyama, R. (2012) Essential roles of the histone methyltransferase ESET in the epigenetic control of neural progenitor cells during development. *Development*, **139**, 3806–3816.
49. Padeken, J., Methot, S.P. and Gasser, S.M. (2022) Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. *Nature Reviews Molecular Cell Biology*, **23**, 623–640.
50. Brattas, P.L., Jonsson, M.E., Fasching, L., Nelander Wahlestedt, J., Shahsavani, M., Falk, R., Falk, A., Jern, P., Parmar, M. and Jakobsson, J. (2017) TRIM28 Controls a Gene Regulatory Network Based on Endogenous Retroviruses in Human Neural Progenitor Cells. *Cell Rep*, **18**, 1–11.
51. Turelli, P., Playfoot, C., Grun, D., Raclot, C., Pontis, J., Coudray, A., Thorball, C., Duc, J., Pankevich, E.V., Deplancke, B. *et al.* (2020) Primate-restricted KRAB zinc finger proteins and target retrotransposons control gene expression in human neurons. *Sci Adv*, **6**, eaba3200.
52. Turelli, P., Castro-Diaz, N., Marzetta, F., Kapopoulou, A., Raclot, C., Duc, J., Tieng, V., Quenneville, S. and Trono, D. (2014) Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. *Genome Res*, **24**, 1260–1270.
53. Ecco, G., Cassano, M., Kauzlaric, A., Duc, J., Coluccio, A., Offner, S., Imbeault, M., Rowe, Helen M., Turelli, P. and Trono, D. (2016) Transposable Elements and Their

KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Developmental Cell*, **36**, 611–623.
54. Douse, C.H., Tchasovnikarova, I.A., Timms, R.T., Protasio, A.V., Seczynska, M., Prigozhin, D.M., Albecka, A., Wagstaff, J., Williamson, J.C., Freund, S.M.V. *et al.* (2020) TASOR is a pseudo-PARP that directs HUSH complex assembly and epigenetic transposon control. *Nat Commun*, **11**, 4940.
55. Liu, N., Lee, C.H., Swigut, T., Grow, E., Gu, B., Bassik, M.C. and Wysocka, J. (2018) Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature*, **553**, 228–232.
56. Robbez-Masson, L., Tie, C.H.C., Conde, L., Tunbak, H., Husovsky, C., Tchasovnikarova, I.A., Timms, R.T., Herrero, J., Lehner, P.J. and Rowe, H.M. (2018) The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. *Genome Res*, **28**, 836–845.
57. Wu, X. and Zhang, Y. (2017) TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet*, **18**, 517–534.
58. Montavon, T., Shukeir, N., Erikson, G., Engist, B., Onishi-Seebacher, M., Ryan, D., Musa, Y., Mittler, G., Meyer, A.G., Genoud, C. *et al.* (2021) Complete loss of H3K9 methylation dissolves mouse heterochromatin organization. *Nat Commun*, **12**, 4359.
59. Park, J.-W., Kim, J.J. and Bae, Y.-S. (2018) CK2 downregulation induces senescence-associated heterochromatic foci formation through activating SUV39h1 and inactivating G9a. *Biochemical and Biophysical Research Communications*, **505**, 67–73.
60. Tsurumi, A. and Li, W.X. (2012) Global heterochromatin loss: a unifying theory of aging? *Epigenetics*, **7**, 680–688.
61. Mills, R.E., Bennett, E.A., Iskow, R.C. and Devine, S.E. (2007) Which transposable elements are active in the human genome? *Trends in Genetics*, **23**, 183–191.
62. Jonsson, M.E., Garza, R., Johansson, P.A. and Jakobsson, J. (2020) Transposable Elements: A Common Feature of Neurodevelopmental and Neurodegenerative Disorders. *Trends Genet*, **36**, 610–623.
63. Percharde, M., Lin, C.-J., Yin, Y., Guan, J., Peixoto, G.A., Bulut-Karslioglu, A., Biechele, S., Huang, B., Shen, X. and Ramalho-Santos, M. (2018) A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. *Cell*, **174**, 391–405.e319.
64. Jachowicz, J.W., Bing, X., Pontabry, J., Bošković, A., Rando, O.J. and Torres-Padilla, M.-E. (2017) LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nature Genetics*, **49**, 1502–1510.
65. Toda, T., Bedrosian, T.A., Schafer, S.T., Cuoco, M.S., Linker, S.B., Ghassemzadeh, S., Mitchell, L., Whiteley, J.T., Novaresi, N., McDonald, A.H. *et al.* (2024) Long interspersed nuclear elements safeguard neural progenitors from precocious differentiation. *Cell Rep*, **43**, 113774.
66. Mangoni, D., Simi, A., Lau, P., Armaos, A., Ansaloni, F., Codino, A., Damiani, D., Floreani, L., Di Carlo, V., Vozzi, D. *et al.* (2023) LINE-1 regulates cortical development by acting as long non-coding RNAs. *Nature Communications*, **14**, 4974.
67. Ren, W., Fan, H., Grimm, S.A., Guo, Y., Kim, J.J., Yin, J., Li, L., Petell, C.J., Tan, X.F., Zhang, Z.M. *et al.* (2020) Direct readout of heterochromatic H3K9me3 regulates DNMT1-mediated maintenance DNA methylation. *Proc Natl Acad Sci U S A*, **117**, 18439–18447.
68. Du, J., Johnson, L.M., Jacobsen, S.E. and Patel, D.J. (2015) DNA methylation pathways and their crosstalk with histone methylation. *Nature Reviews Molecular Cell Biology*, **16**, 519–532.

69. Liu, X., Gao, Q., Li, P., Zhao, Q., Zhang, J., Li, J., Koseki, H. and Wong, J. (2013) UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nature Communications*, **4**, 1563.
70. Jacobs, F.M., Greenberg, D., Nguyen, N., Haeussler, M., Ewing, A.D., Katzman, S., Paten, B., Salama, S.R. and Haussler, D. (2014) An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature*, **516**, 242–245.
71. Müller, I., Moroni, A.S., Shlyueva, D., Sahadevan, S., Schoof, E.M., Radzisheuskaya, A., Højfeldt, J.W., Tatar, T., Koche, R.P., Huang, C. *et al.* (2021) MPP8 is essential for sustaining self-renewal of ground-state pluripotent stem cells. *Nature Communications*, **12**, 3034.
72. Lopez-Otin, C., Blasco, M.A., Partridge, L., Serrano, M. and Kroemer, G. (2023) Hallmarks of aging: An expanding universe. *Cell*, **186**, 243–278.
73. Lee, J.H., Kim, E.W., Croteau, D.L. and Bohr, V.A. (2020) Heterochromatin: an epigenetic point of view in aging. *Exp Mol Med*, **52**, 1466–1474.
74. Liu, G.H., Barkho, B.Z., Ruiz, S., Diep, D., Qu, J., Yang, S.L., Panopoulos, A.D., Suzuki, K., Kurian, L., Walsh, C. *et al.* (2011) Recapitulation of premature ageing with iPSCs from Hutchinson-Gilford progeria syndrome. *Nature*, **472**, 221–225.
75. Miller, J.D., Ganat, Y.M., Kishinevsky, S., Bowman, R.L., Liu, B., Tu, E.Y., Mandal, P.K., Vera, E., Shim, J.W., Kriks, S. *et al.* (2013) Human iPSC-based modeling of late-onset disease via progerin-induced aging. *Cell Stem Cell*, **13**, 691–705.
76. Zhang, W., Li, J., Suzuki, K., Qu, J., Wang, P., Zhou, J., Liu, X., Ren, R., Xu, X., Ocampo, A. *et al.* (2015) Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science*, **348**, 1160–1163.
77. Della Valle, F., Reddy, P., Yamamoto, M., Liu, P., Saera-Vila, A., Bensaddek, D., Zhang, H., Prieto Martinez, J., Abassi, L., Celii, M. *et al.* (2022) LINE-1 RNA causes heterochromatin erosion and is a target for amelioration of senescent phenotypes in progeroid syndromes. *Sci Transl Med*, **14**, eabl6057.
78. De Cecco, M., Ito, T., Petrashen, A.P., Elias, A.E., Skvir, N.J., Criscione, S.W., Caligiana, A., Brocculi, G., Adney, E.M., Boeke, J.D. *et al.* (2019) L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature*, **566**, 73–78.
79. Simon, M., Van Meter, M., Ablaeva, J., Ke, Z., Gonzalez, R.S., Taguchi, T., De Cecco, M., Leonova, K.I., Kogan, V., Helfand, S.L. *et al.* (2019) LINE1 Derepression in Aged Wild-Type and SIRT6-Deficient Mice Drives Inflammation. *Cell Metabolism*, **29**, 871–885.e875.
80. Tang, H., Yang, J., Xu, J., Zhang, W., Geng, A., Jiang, Y. and Mao, Z. (2024) The transcription factor PAX5 activates human LINE1 retrotransposons to induce cellular senescence. *EMBO Rep*, **25**, 3263–3275.
81. Ardeljan, D., Steranka, J.P., Liu, C., Li, Z., Taylor, M.S., Payer, L.M., Gorbounov, M., Sarnecki, J.S., Deshpande, V., Hruban, R.H. *et al.* (2020) Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication. *Nature Structural & Molecular Biology*, **27**, 168–178.
82. Ewing, A.D. and Kazazian, H.H., Jr. (2010) High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res*, **20**, 1262–1270.
83. Gilbert, N., Lutz-Prigge, S. and Moran, J.V. (2002) Genomic Deletions Created upon LINE-1 Retrotransposition. *Cell*, **110**, 315–325.
84. Gilbert, N., Lutz, S., Morrish, T.A. and Moran, J.V. (2005) Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol Cell Biol*, **25**, 7780–7795.

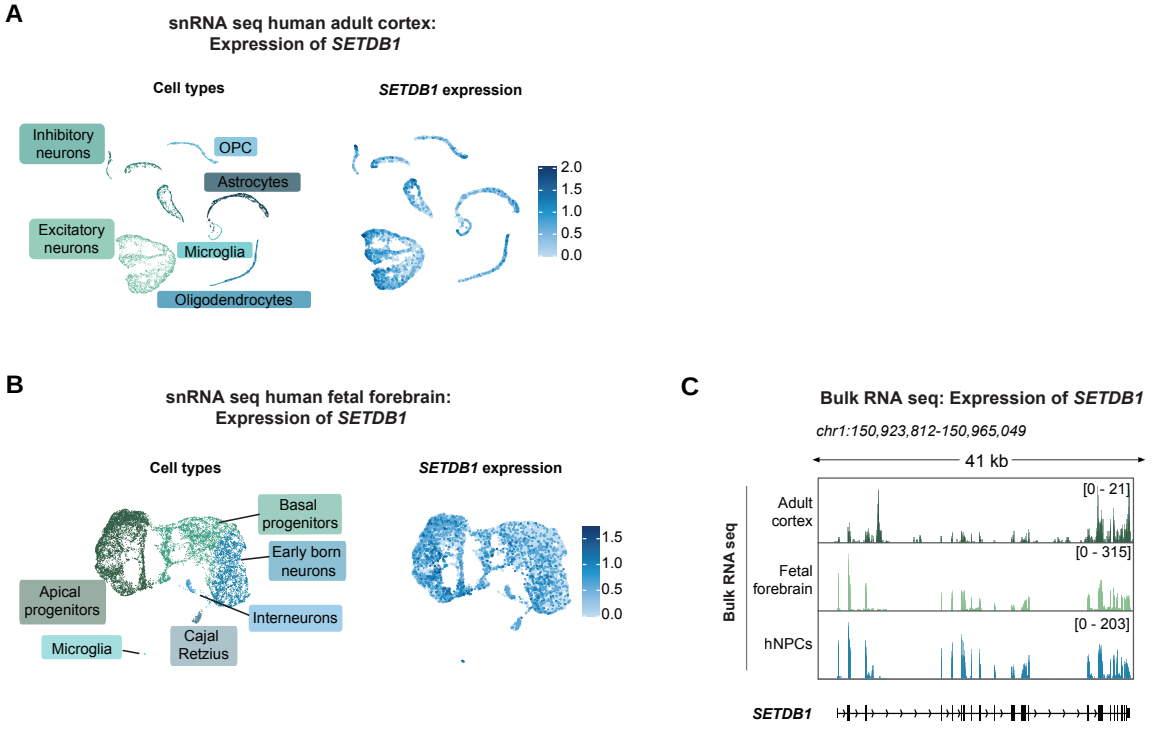
85. Symer, D.E., Connelly, C., Szak, S.T., Caputo, E.M., Cost, G.J., Parmigiani, G. and Boeke, J.D. (2002) Human L1 Retrotransposition Is Associated with Genetic Instability In Vivo. *Cell*, **110**, 327–338.
86. Tunbak, H., Enriquez-Gasca, R., Tie, C.H.C., Gould, P.A., Mlcochova, P., Gupta, R.K., Fernandes, L., Holt, J., van der Veen, A.G., Giampazolias, E. *et al.* (2020) The HUSH complex is a gatekeeper of type I interferon through epigenetic regulation of LINE-1s. *Nature Communications*, **11**, 5387.
87. Buttler, C.A., Ramirez, D., Dowell, R.D. and Chuong, E.B. (2023) An intronic LINE-1 regulates IFNAR1 expression in human immune cells. *Mob DNA*, **14**, 20.

Supplementary Figures

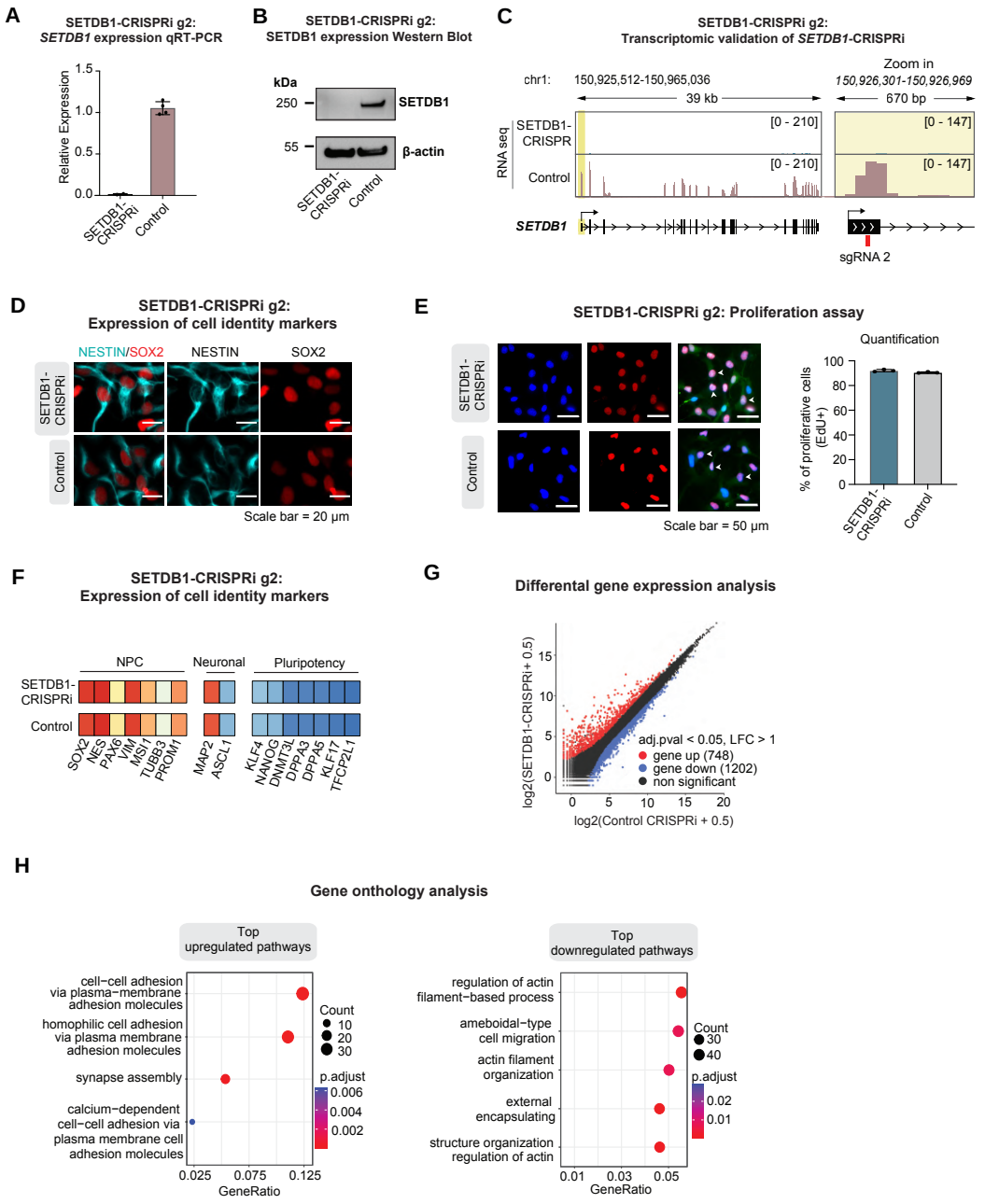


Supplementary Figure 1. H3K9me3 in hNPCs localizes to regions with well-established H3K9me3-enrichment A Mappability analysis of young full-length L1HS to L1PA3 using 36bp, 50bp and 100bp uniquely mappable kmers obtained from publicly available dataset (10.1093/nar/gky677), and an example of an in-house IgG experiment. **B** CUT&RUN analysis of H3K9me3 in hNPCs over

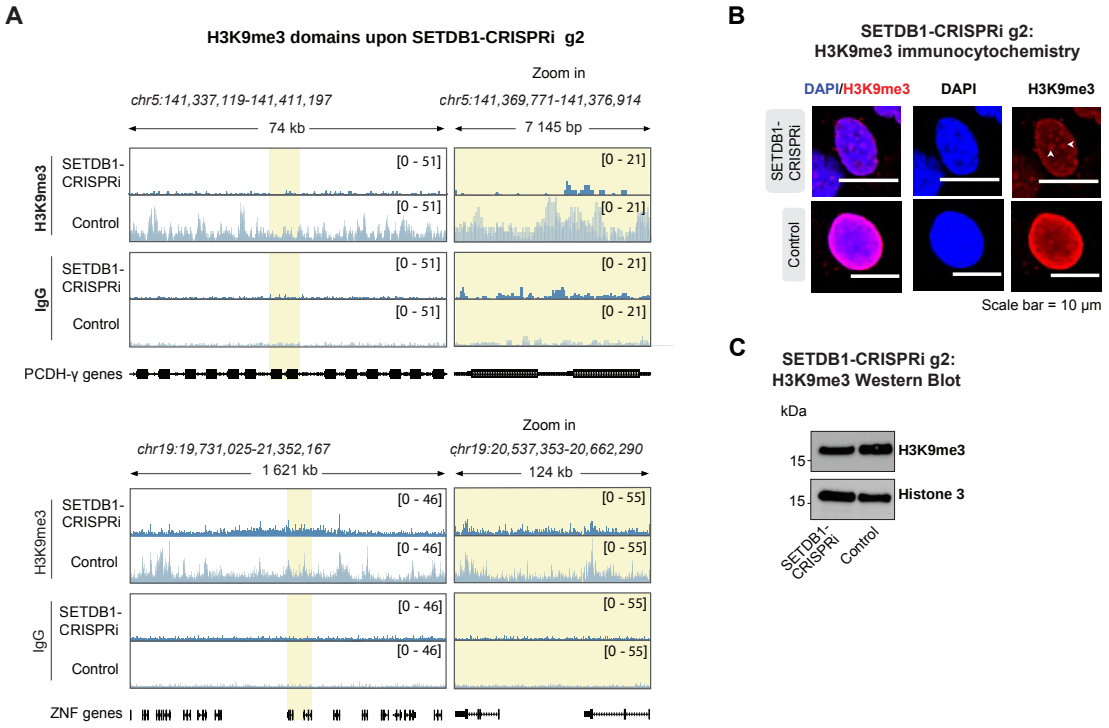
genomic regions with well-established H3K9me3-enrichment; PCDH-ɣ cluster (top); KRAB-ZNF cluster (bottom).



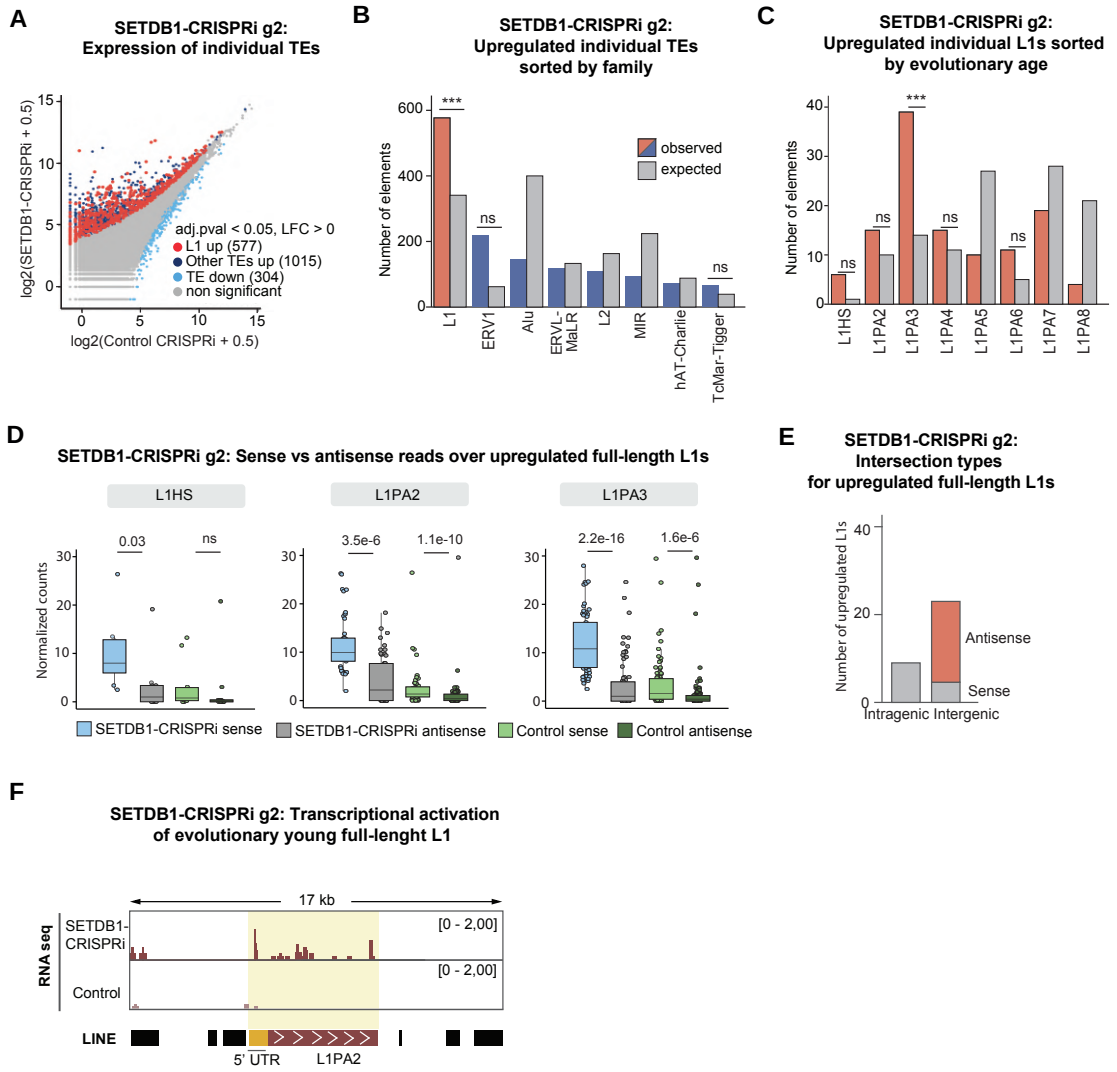
Supplementary Figure 2. SETDB1 is expressed in the adult and developing human brain and hNPCs A-B UMAP of snRNA-seq data showing SETDB1 expression in (A) human adult cortical tissue and (B) human fetal forebrain tissue. **C** Genome browser tracks showing bulk RNA sequencing data of SETDB1 expression in the human adult cortex, human fetal forebrain, and hNPCs. Data from Garza *et al.* (2023) (20).



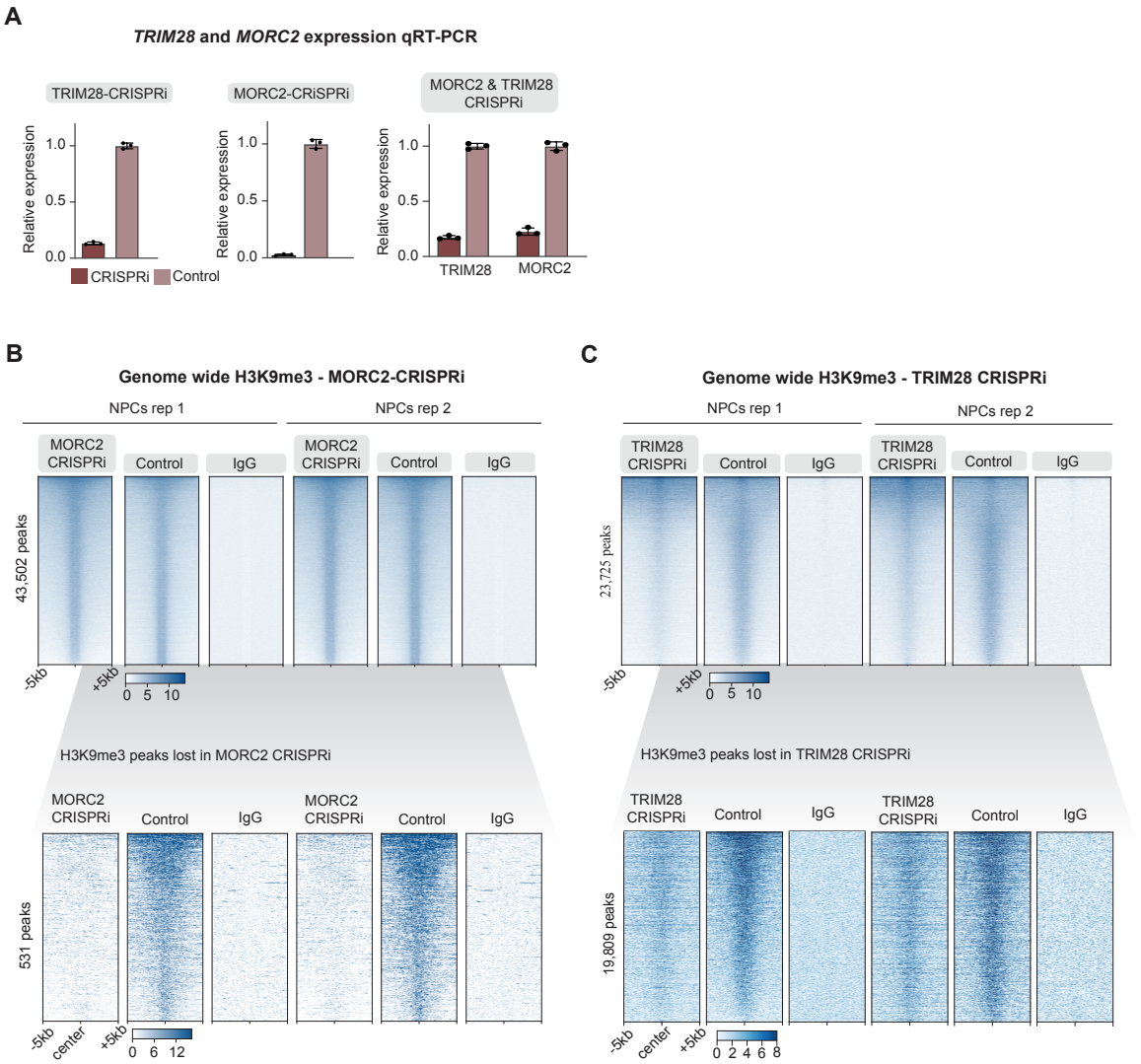
Supplementary Figure 3. Validation of SETDB1-CRISPRi using an alternative guide RNA A qRT-PCR analysis of SETDB1 expression after CRISPRi g2-based silencing in hNPCs. Results are shown as means with standard deviation (*n*=2 and *n*=4 respectively). **B** Western blot analysis of SETDB1 protein levels relative to β-actin protein levels in SETDB1-CRISPRi g2 and control hNPCs. **C** RPKM normalized genome browser tracks showing expression of *SETDB1* in SETDB1-CRISPRi g2 and control hNPCs as determined by bulk RNA sequencing **D** Immunocytochemistry of NESTIN (cyan) and SOX2 (red) in SETDB1-CRISPRi g2 and control hNPCs. Scalebar = 20 µm. **E** EdU proliferation assay showing cell proliferation in SETDB1-CRISPRi g2 and Control hNPCs. Scale bar=50 µm. **F** Heat map of NPC, neuronal and pluripotency gene marker mean expression in SETDB1-CRISPRi g2 (*n*=2) and control (*n*=4) hNPCs as determined by bulk RNA sequencing. **G** Mean plot showing differential gene expression in SETDB1-CRISPRi (*n*=4) compared to control (*n*=4) hNPCs. LFC >1, padj < 0.05 calculated with DESeq2. **H** Gene-ontology overrepresentation test of significantly differentially expressed genes (|LFC| > 1) associated with the selected top upregulated terms (left) and the selected top downregulated terms (right) in SETDB1-CRISPRi compared to control hNPCs.



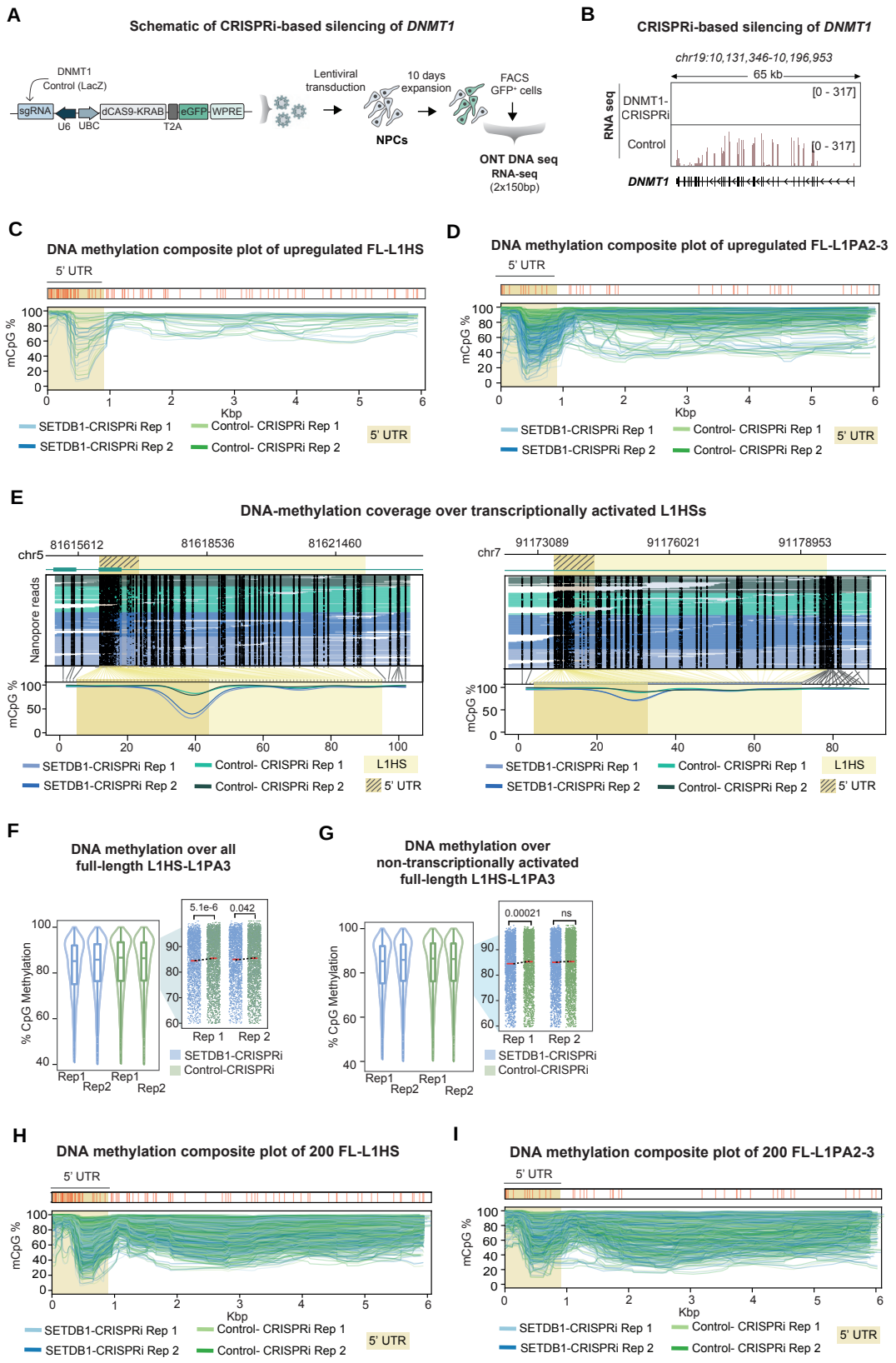
Supplementary Figure 4. H3K9me3 enrichment over genomic regions. **A** CUT&RUN analysis of H3K9me3 over genomic regions with well-established H3K9me3-enrichment in SEDTB1-CRISPRi and control hNPCs **B** Immunocytochemistry of H3K9me3 (red) and nuclear marker DAPI (blue) in SETDB1-CRISPRi g2 and control hNPCs. H3K9me3 foci are indicated with white arrows. Scalebar = 10 µm. **C** Western blot analysis of H3K9me3 levels relative to Histone 3 in SETDB1-CRISPRi g2 hNPCs compared to control.



Supplementary Figure 5. Validation of transcriptional activation of L1s upon SETDB1-CRISPRi using an alternative guide RNA A Mean plot of uniquely mapped bulk RNA sequencing reads for SETDB1-CRISPRi g2 hNPCs (*n*=2) compared to control (*n*=2) over individual TEs categorized into L1 elements and other TEs. Log_2 fold change (LFC) >0, padj 0.05 calculated with DESeq2. **B** Observed vs expected number of individual transcriptionally activated elements per TE family in SETDB1-CRISPRi g2 hNPCs (n= 2) compared to control (n= 2). P-values were calculated using bootstrap-based empirical two-sided p-value calculation. **C** Observed vs expected number of transcriptionally activated L1 elements sorted by evolutionary age in SETDB1-CRISPRi g2 *(n=2)* compared to control hNPCs *(n=2)*. Pvalues were calculated using bootstrap-based empirical pvalue calculation. **D** Normalized read counts in sense and antisense for upregulated, full-length L1HS-L1PA3 elements in SETDB1-CRISPRi g2 (*n*=2) vs control hNPCs (*n*=4). Statistical significance calculated with Mann-Whitney U test. **E** Number, genomic location and direction of upregulated full-length (>6 kb) L1HS-L1PA3 elements in SETDB1-CRISPRi g2 (n=2) compared to control hNPC (n=2) ***F*** RPKM normalized genome browser tracks showing RNA-seq in SETDB1-CRISPRi g2 and control hNPCs.

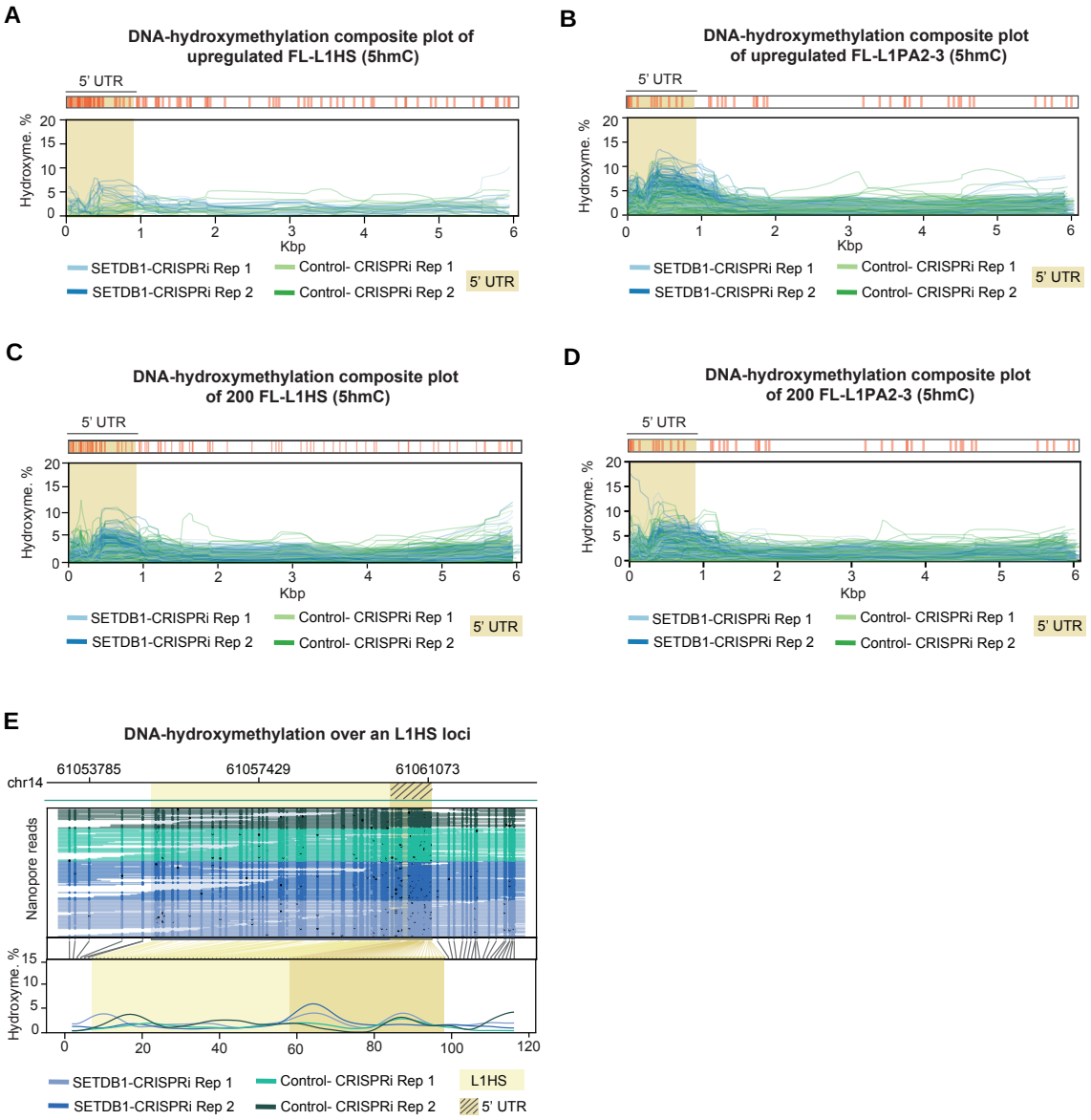


Supplementary Figure 6. SETDB1 regulates evolutionary young, full-length L1 elements independently from MORC2 and TRIM28. **A** qRT-PCR analysis of TRIM28 and MORC2 expression after CRISPRi based silencing in hNPCs. Results are shown as means with standard deviation (*n*=3). **B** RPKM normalized heat maps showing H3K9me3 CUT&RUN signal in MORC2-CRISPRi hNPCs genome-wide (top) and over genomic regions where H3K9me3 peaks were lost (bottom) compared to their respective controls and non-targeting control IgG. Data from Pandiloski et al. (2024) (23)**.** **C** RPKM normalized heat maps showing H3K9me3 CUT&RUN signal in TRIM28-CRISPRi hNPCs genome-wide (top) and over genomic regions where H3K9me3 peaks were lost (bottom) compared to their respective controls and non-targeting control IgG. Data from Horvath et al. (2024) (26).

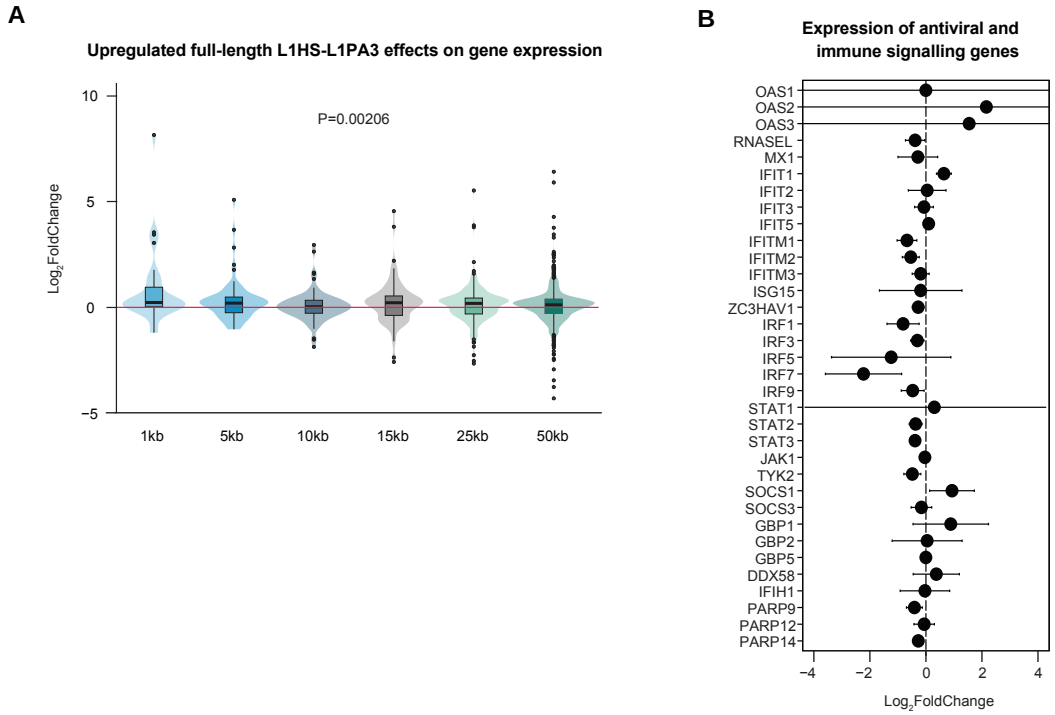


Supplementary Figure 7. DNA methylation analysis over evolutionary young L1s A Workflow for CRISPRi-based silencing of *DNMT1* in hNPCs and downstream analysis. **B** RPKM normalized genome browser tracks showing *DNMT1* expression in DNMT1-CRISPRi and control hNPCs. **C** Composite DNA methylation profiles of upregulated full-length L1HS in SETDB1-CRISPRi (blue, n=2) and

control (green, n=2) hNPCs. The 5' UTR is highlighted in yellow. **D** Composite DNA methylation profiles of upregulated full-length L1PA2 and L1PA3 in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. The 5' UTR is highlighted in yellow. **E** Locus plot showing DNA methylation over upregulated L1HS in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. Black dots indicate methylated CpGs, and methylation coverage of the L1 element can be seen at the bottom. The L1 element is highlighted in yellow, while the 5' UTR is marked with stripes. **F** Box plots of the methylation status over the promoter of all full-length L1HS, L1PA2 and L1PA3 elements in SETDB1-CRISPRi (n=2) and control (n=2) hNPCs. Zoom-in panels indicating mean methylation levels (red line) per condition for L1HS-L1PA3 elements. n values: L1HS = 306; L1PA2 = 972; L1PA3 = 1367. **G** Box plots of the methylation status over the promoter of full-length non-transcriptionally activated L1HS, L1PA2 and L1PA3 elements in SETDB1-CRISPRi (n=2) and control (n=2) hNPCs. Zoom-in panels indicating mean methylation levels (red line) per condition for L1HS-L1PA3 elements. n values: L1HS = 284; L1PA2 = 895; L1PA3 = 1225. **H** Composite DNA methylation profiles of 200 randomly picked full-length L1HS elements in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. The 5' UTR is highlighted in yellow. **I** Composite DNA methylation profiles of 200 randomly picked full-length L1PA2-L1PA3 elements in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. The 5' UTR is highlighted in yellow.



Supplementary Figure 8. DNA-hydroxymethylation over evolutionary young L1s A Composite DNA hydroxymethylation profiles of upregulated full-length L1HS in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. The 5' UTR is highlighted in yellow. **B** Composite DNA hydroxymethylation profiles of upregulated full-length L1PA2 and L1PA3 in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. The 5' UTR is highlighted in yellow. **C** Composite DNA hydroxymethylation profiles of 200 randomly picked full-length L1HS elements in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. The 5' UTR is highlighted in yellow. **D** Composite DNA hydroxymethylation profiles of 200 randomly picked full-length L1PA2-L1PA3 elements in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. The 5' UTR is highlighted in yellow. **E** Locus plot showing DNA hydroxymethylation over upregulated L1HS in SETDB1-CRISPRi (blue, n=2) and control (green, n=2) hNPCs. Black dots indicate methylated CpGs, and methylation coverage of the L1 element can be seen at the bottom. The L1 element is highlighted in yellow, while the 5' UTR is marked with stripes.



Supplementary Figure 9. The effect of upregulated evolutionary young, full-length L1s on nearby gene expression. A Violin plots showing the effect of upregulated full-length L1HS-L1PA3 elements on nearby gene expression (2-50kb) upon SETDB1-CRISPRi in hNPCs. pvalue calculated with one-way ANOVA. **B** Differential expression analysis of selected antiviral and immune signaling genes in SETDB1-CRISPRi hNPCs (*n*=4) compared to control (*n*=4). LFC values are shown ± LFC standard error calculated by DESeq2.

Individual TE	Log2FoldChange	p-adjusted value
MER39B_dup439	12.5660730750503	1.8427529832473e-26
L2c_dup54301	12.2209237438216	4.09869177917656e-25
MER34B_dup283	11.7388139351731	2.93429232239475e-24
MLT1B_dup6701	11.1359864402149	4.05191990872606e-19
MIRb_dup150062	10.3913791660113	9.6211096384187e-19
MIRb_dup83332	10.3742843350223	4.39773387345283e-17
AluSx1_dup13234	10.3463773877304	2.18345277280823e-20
L2c_dup54302	10.0803146120823	2.89355744423237e-16
SVA_D_dup667	10.0527688112821	8.80024954039596e-19
MIR_dup66266	9.91282533692722	1.32675089321779e-16
Tigger1_dup11477	9.86311883738296	2.10590058343799e-18
LTR10C_dup335	9.64204668384122	1.93395417652645e-17
SVA_E_dup244	9.57541896859938	2.85991551525974e-12
LTR12C_dup2419	9.34014514735954	8.07950807168578e-14
THE1C_dup7048	9.17288181094644	1.27595527171481e-15
LTR10C_dup29	9.03749224093489	6.18977818132119e-15
L2c_dup96897	8.97118041487722	9.18595524144081e-15
THE1B_dup16084	8.73842550859459	5.00823323103967e-14
LTR79_dup800	8.51355207135647	7.37068570999639e-13
THE1C_dup7049	8.49481998710542	5.63699471444779e-13

Supplementary Table 1. Top 20 upregulated individual TEs upon SETDB1-CRISPRi Table showing top 20 differentially expressed individual TEs as determined by bulk RNA sequencing in SETDB1-CRISPRi (*n*=4) and control (*n*=4) hNPCs. LFC and padj calculated with DESeq2.