Loss of H3K9me3 maintenance in human neural progenitor cells leads to transcriptional activation of L1 retrotransposons

Ofelia Karlsson[1], Ninoslav Pandiloski[1,2], Anita Adami[1], Raquel Garza[1], Vivien Horvath[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
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 associated with the histone mark H3K9me3. Heterochromatin erosion is a hallmark of human ageing and H3K9me3 is lost globally in aged mammalian cells. However, the functional consequences of the loss of heterochromatin maintenance have been challenging to model. In this study, we used CRISPRi-mediated deletion of the H3K9me3 methyltransferase SETDB1 in human neural progenitor cells to model disruption of heterochromatin maintenance. Deletion of SETDB1 had a limited effect on cell viability, despite a global loss of H3K9me3 peaks resulting in genome-wide reorganization of heterochromatin domains, cells remained proliferative and expressed appropriate marker genes. We found that a key event following the loss of H3K9me3 maintenance 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, further implicating these elements in human ageing processes.

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*). Only a small fraction of these full-length L1s are capable of transposing (*16, 17*). 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, 18-20*). 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, 21, 22*).

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 (*23-27*). L1s contain strong gene regulatory sequences that can act as both transcriptional activators and repressors, affecting the expression of genes in the vicinity of their integration site (*10, 19, 20, 28, 29*). In addition, the accumulation of L1-derived nucleic acids in the cytoplasm has the potential to activate a pro-inflammatory innate immune response, leading to cellular toxicity (*24, 25, 30*). Mobilization of L1s in somatic cells has been suggested to lead to genetic mosaicism in the human brain (*9, 31-38*). However, the mechanistic basis of L1 transcriptional activation in senescence and

ageing, its link to heterochromatin loss and the ultimate functional consequences remain poorly understood.

It is well described that the pattern of H3K9me3 on TEs is established during early development and that disruption of heterochromatin establishment results in massive transcriptional activation of TEs, including L1s (*39-42*). During physiological and pathological ageing, heterochromatin and H3K9me3 are eroded, leading to changes in genome stability and alterations in gene expression (*43-46*). This process is likely to have detrimental effects on a range of cellular and physiological processes, ultimately leading to reduced cellular fitness. As such, heterochromatin erosion is considered a hallmark of ageing (*23, 47-49*). However, the molecular and functional consequences of the loss of heterochromatin maintenance in human somatic cells have been challenging to model.

In this study, we used human neural progenitor cells (hNPCs) to model disruption of H3K9me3 heterochromatin maintenance. These hNPCs represent a somatic cell type with fully established facultative heterochromatin, including the presence of H3K9me3 and DNA methylation over L1s and other TEs (*8, 22, 50*), which is established by the KRAB zinc finger (KZNF)-TRIM28 system and other epigeneitic TE repressors during early development (*51-53*). By deleting the histone methyltransferase SETDB1 associated with these repressors, we were able to study the consequences of H3K9me3 loss at TEs in somatic cells. Deletion of SETDB1 in hNPCs had a limited effect on cell viability and the cells remained proliferative and expressed appropriate marker genes despite widespread loss of H3K9me3 domains. However, loss of H3K9me3 maintenance resulted in transcriptional activation of L1s. Notably, the transcriptional activation of L1s correlated with a loss of DNA methylation over the same elements, linking the stability of DNA methylation at L1s to H3K9me3 maintenance. In summary, our results demonstrate a critical role for H3K9me3 in L1 silencing in human neural

somatic cells and imply a direct link between L1 activation and the epigenetic changes associated with physiological and pathological ageing.

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 (*54*), 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. 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. 1)**. 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)**. 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 (*55*) (**Fig. 1D)**. We found that almost all full-length evolutionarily-young elements,

including human-specific (L1HS) and hominoid-specific (L1PA2 and L1PA3) copies, were covered by H3K9me3 in hNPCs **(Fig. 1E,F)**. 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, 17*). 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.

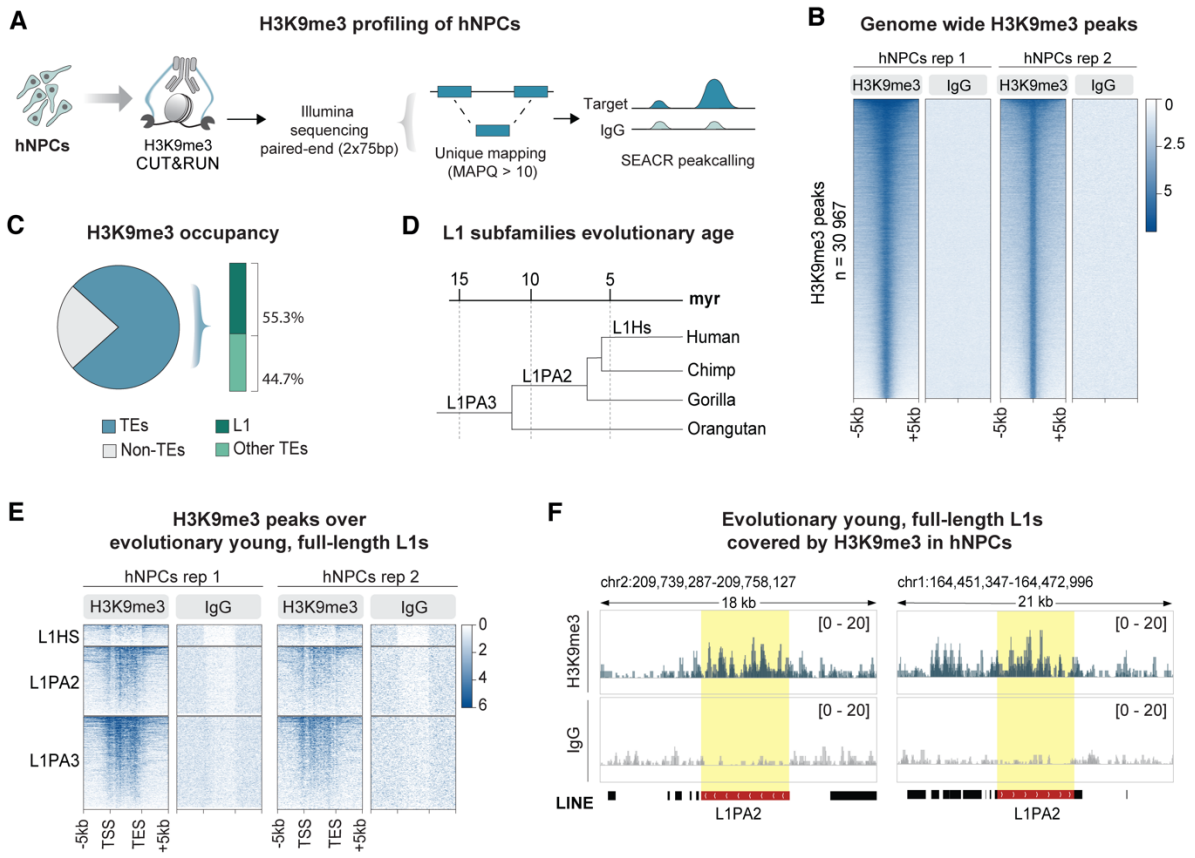


Figure 1. Evolutionary young, full-length L1s are covered by H3K9me3 in hNPCs A Schematic illustrating CUT&RUN epigenomic profiling of H3K9me3 in hNPCs. **B** RPKM normalized heat maps showing genome-wide CUT&RUN signal enrichment of H3K9me3 and non-targeting control IgG in hNPCs (*n*=2). **C** Plots 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** 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). **F** 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) (*56-58*), which is highly expressed in hNPCs and the developing and adult human brain **(Sup Fig. 2)**. To investigate the consequences of SETDB1 loss and impaired H3K9me3 maintenance, we designed a CRISPRi-based approach to 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 (*59*). 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, RNA sequencing, and Western blot (**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 (*60*). 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. 2F, Sup Fig 3E)**. Previous studies in mouse model systems have described H3K9me3 as important for maintaining the cellular identity of neural cells and silencing lineage-inappropriate genes (*61*). However, in SETDB1-CRISPRi hNPCs, proliferation and expression of NPC lineage markers remained unchanged compared to control hNPCs, and neuronal and pluripotency markers were not activated as monitored by bulk RNA sequencing, at least over the 10-day timecourse of the experiment (**Fig. 2E, Sup Fig. 3D**). Furthermore, immunofluorescence staining for the NPC markers SOX2 and NESTIN showed no obvious differences in SETDB1-CRISPRi hNPCs compared to controls (**Fig. 2F, Sup Fig. 3E**). To further investigate how loss 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 ($|LFC| > 1$, $padj < 0.05$, DESeq2) **(Sup Fig. 3F)**. 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. 3G).** Taken together, these results demonstrate that loss of SETDB1 results in transcriptional changes but that these do not appear to have a major impact on the cellular identity of hNPCs.

To further investigate the role of SETDB1 in maintaining 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. 2G, Sup Fig. 3H).** This included a profound loss of H3K9me3 over evolutionarily-young full-length L1 elements **(Fig. 2H)**. 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 **(Fig. 2H-I)**. The appearance of these ectopic regions of H3K9me3 was also visible when performing immunocytochemistry for H3K9me3 in SETDB1-CRISPRi hNPCs **(Fig. 2J, Sup Fig. 3I).** We found no changes in global H3K9me3 levels using WB analysis **(Sup Fig. 3J).** 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 (*62*) and the complete heterochromatin reorganization that occurs in SETDB1-CRISPRi hNPCs.

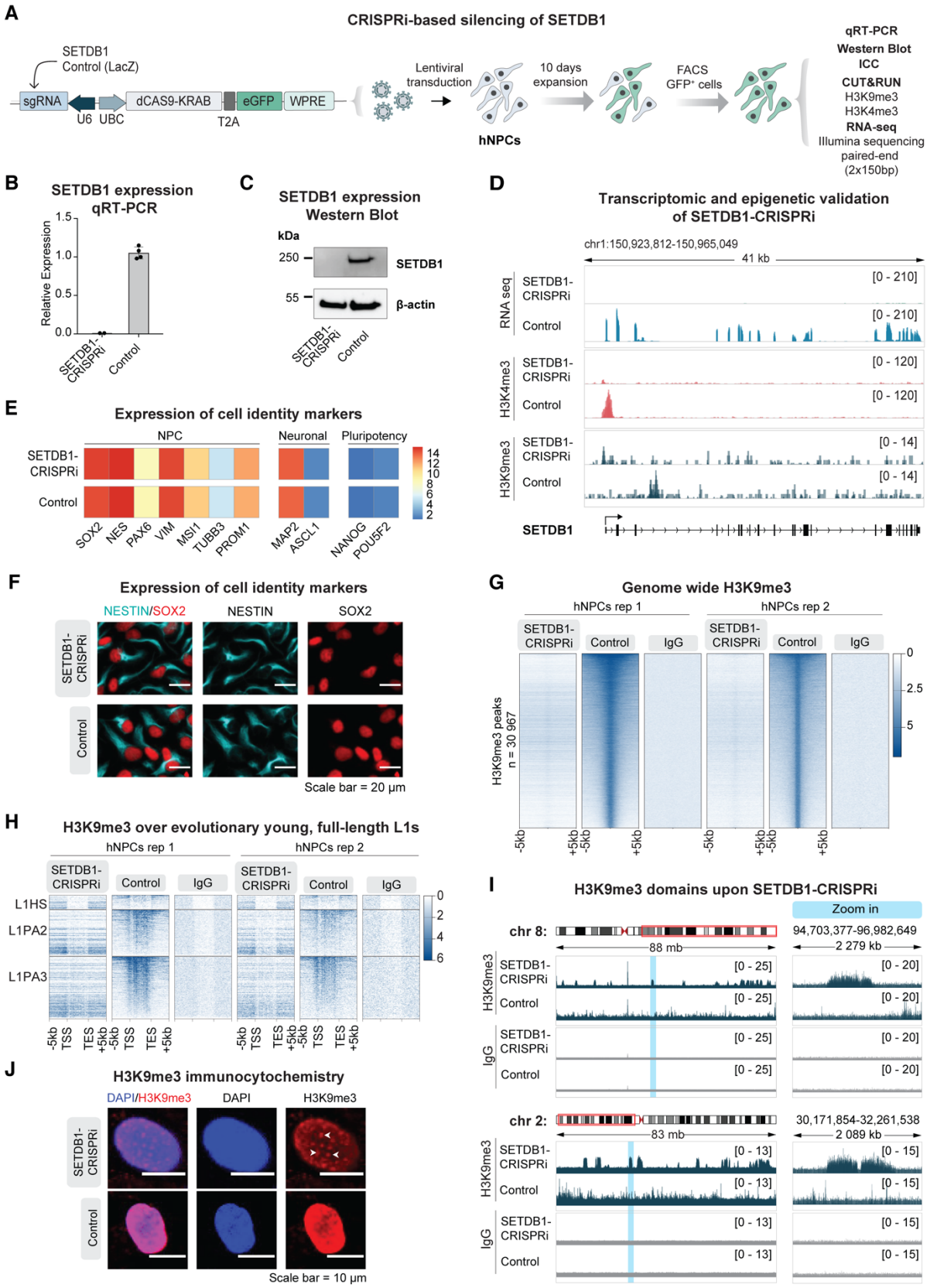


Figure 2. SETDB1-CRISPRi results in reorganization of H3K9me3 in hNPCs A Schematic of the CRISPRi construct and workflow for CRISPRi-based silencing of SETDB1 in hNPCs. **B** qRT-PCR analysis of SETDB1 expression after CRISPRi silencing in hNPCs. Results are shown as means with standard deviation (*n*=2 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** 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. **F** Immunocytochemistry of NESTIN (cyan) and SOX2 (red) in SETDB1-CRISPRi and control hNPCs. Scale bar = 20 µm. **G** 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. **H** 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. **I** Genome browser snapshot of chromosome 8 (top) and chromosome 2 (bottom) 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 blue. **J** 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.

Loss of H3K9me3 maintenance results in transcriptional activation of L1s

To investigate the transcriptional consequences on L1s and other TEs after loss of H3K9me3 maintenance 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. 3A)**. We obtained between 26 – 76 million uniquely mapped reads per sample. Sequencing data were mapped using two bioinformatic approaches. We used a stringent unique mapping approach to quantify expression of individual TEs, and we allowed for multimapping reads to quantify transposable element expression at the subfamily level using the TEtranscripts software (*20, 63*).

Using unique mapping we found 3643 individual TE loci that were upregulated in SETDB1-CRISPRi hNPCs (DESeq2 LFC >0, padj 0.05) confirming that many TEs are transcriptionally activated upon loss of H3K9me3 maintenance **(Fig. 3B-C)**. Strikingly, significantly upregulated TEs were enriched for L1s **(Fig 3B-C)**. 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. 3D)**. In total, we found 146 individual, full-length L1HS-L1PA3 elements being significantly upregulated **(Fig. 3D-E)**. We

verified the upregulation of evolutionarily young L1s using a multimapping approach in combination with TEtranscripts. The total expression level of the L1HS-L1PA3 subfamilies, as quantified by TEtranscripts, further confirmed their transcriptional activation upon loss of SETDB1 and H3K9me3 **(Fig. 3F-G)**. We found a similar upregulation of L1s when using both SETDB1 gRNAs (**Sup fig. 4**). 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 region were in sense with regards to the L1s in SETDB1-CRISPRi l hNPCs **(Fig 3H)**. This indicates that the detected L1 transcripts originate from the L1 promoter and are not a consequence of readthrough transcription. 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.)** (*50, 64-66*).

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 histone mark H3K4me3, which is associated with active promoters. Using this approach, we found several evolutionarily-young full-length L1s that lose H3K9me3 and gain H3K4me3 over their promoters, leading to their transcriptional activation upon SETDB1-CRISPRi (**Fig 3I**). Taken together, these results show that maintaince of H3K9me3 is essential to silence evolutionarily-young, full-length L1 elements in hNPCs.

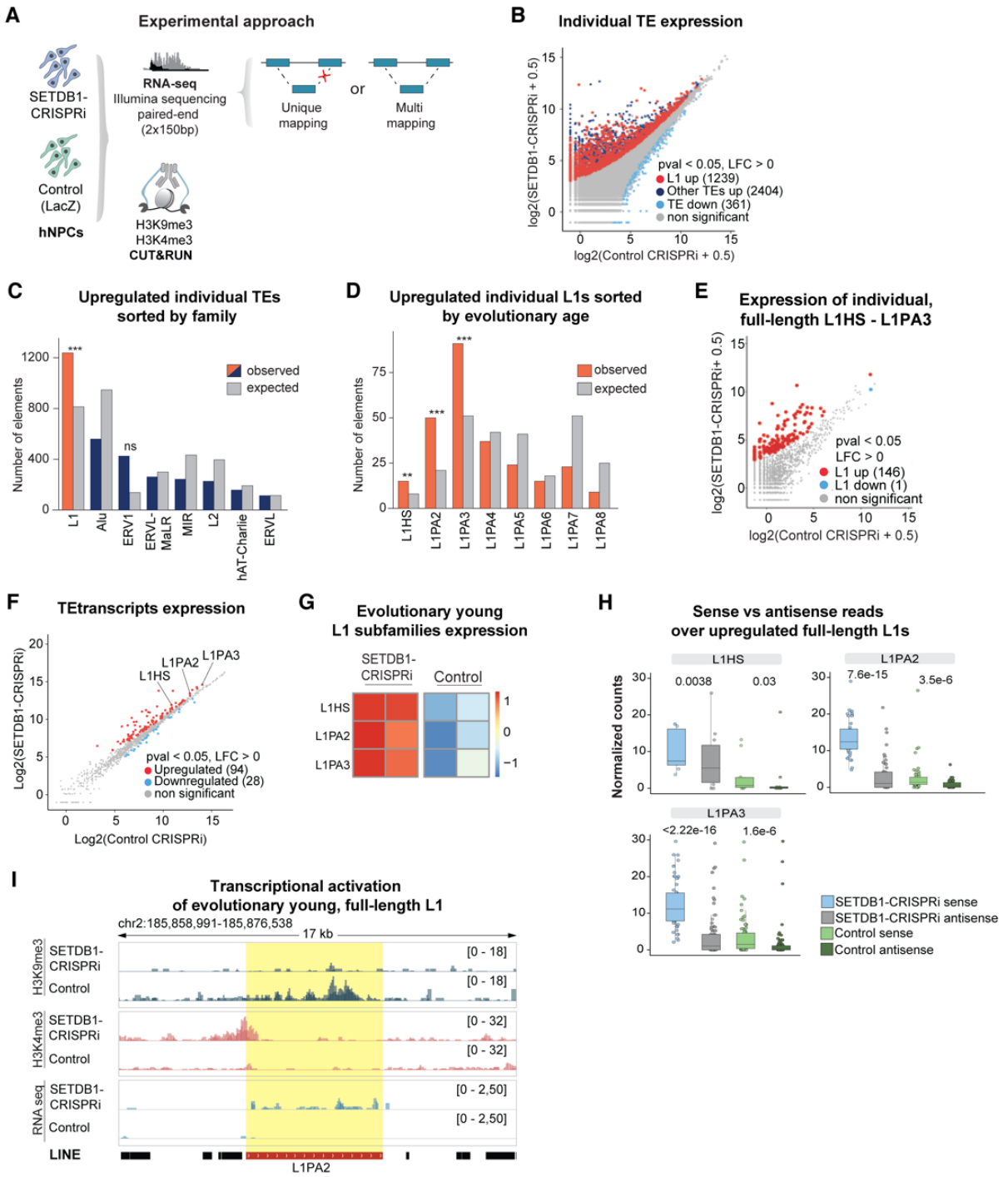


Figure 3. Evolutionarily-young, full-length L1s are transcriptionally activated upon loss of H3K9me3 maintenance in hNPCs A Workflow for CUT&RUN and RNA sequencing analysis of SETDB1-CRISPRi and control hNPCs. **B** Mean plot of uniquely mapped bulk RNA sequencing reads for SETDB1-CRISPRi hNPCs (*n*=2) compared to control (*n*=2) over individual TEs categorized into L1 elements and other TEs. Log_2 fold change (LFC) >0, 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 pvalue calculation. **D** Observed vs expected number of transcriptionally activated L1 elements sorted by subfamily in SETDB1-CRISPRi *(n=2)* compared to control hNPCs *(n=2)*. Pvalues were calculated using bootstrap-based empirical pvalue calculation. **E** Mean plot of uniquely mapped

bulk RNA sequencing reads for SETDB1-CRISPRi hNPCs (*n*=4) compared to controls (*n*=4) over individual full-length (>6 kb) L1HS-L1PA3 elements. LFC > 0, padj < 0.05 calculated with DESeq2. **F** Mean plot of TE families expression as determined by bulk RNA sequencing using TEtranscripts in SETDB1-CRISPRi (*n*=2) and control (*n*=2) hNPCs. LFC >0, padj 0.05 calculated with DESeq2. **G** LFC of L1HS-L1PA3 subfamilies as determined by bulk RNA sequencing in SETDB1-CRISPRi (*n*=2) compared to control hNPCs (*n*=2) using TEtranscripts. **H** 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 calculated with Mann-Whitney U test. **I** 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.

SETDB1 regulates the expression of evolutionarily young L1s independently from the HUSH/MORC2 corepresor 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, 22, 31*). 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 (*51, 67-70*). Previous studies in hNPCs have shown that loss of these pathways in somatic cells has only a limited effect on L1 expression (*22, 64*). To confirm this, we used our previously published H3K9me3 CUT&RUN and bulk RNA sequencing data from MORC2-CRISPRi (*22*) and TRIM28-CRISPRi hNPCs (*50*) to compare the transcriptional activation of evolutionarily-young full-length L1s with that of SETDB1-CRISPRi hNPCs **(Fig. 4A-C).** While SETDB1-CRISPRi resulted in major loss of H3K9me3 peaks over full-length L1HS-L1PA3 elements **(Fig. 2H),** MORC2-CRISPRi and TRIM28-CRISPRi had minor effects on H3K9me3 deposition over full-length L1HS-L1PA3 elements **(Fig. 4D)**. In accordance, bulk RNA sequencing analysis confirmed that loss of SETDB1 resulted in greater transcriptional activation of L1 elements, with 146 individual, full-length L1HS-L1PA3 elements being significantly upregulated **(Fig. 3E)**, compared to loss of MORC2 or TRIM28 that resulted in only 6 and 2 upregulated individual, full-length L1HS-L1PA3 elements respectively **(Fig. 4E-**

F). These results show that in addition to the TEs that are repressed by HUSH and KZNF-TRIM28 and thus dependendent on SETDB1, SETDB also acts through alternative pathways to maintain L1 promoters with H3K9me3 leading to their transcriptional repression.

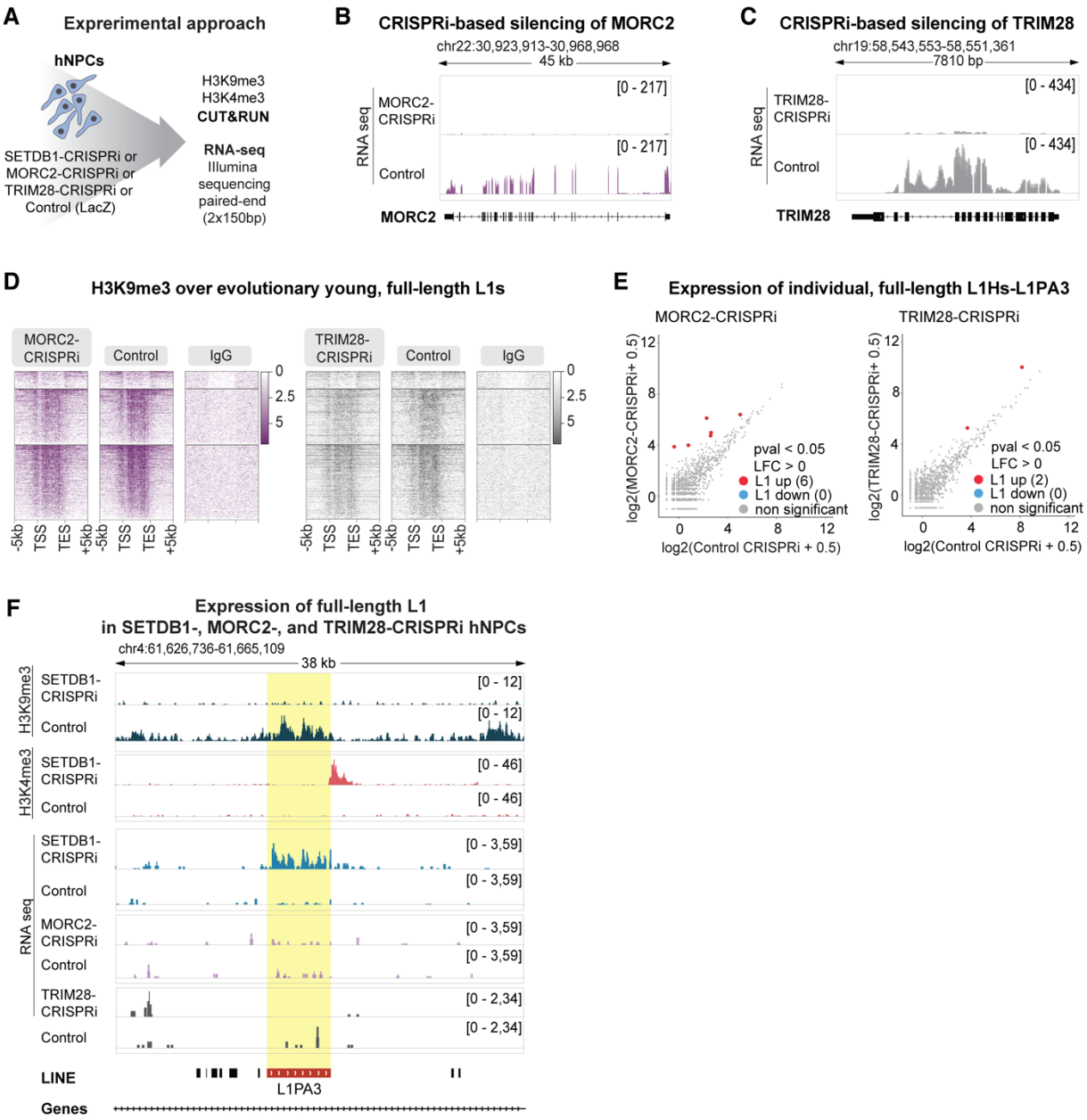


Figure 4. SETDB1 regulates evolutionary young, full-length L1 elements independently from MORC2 and TRIM28

A Workflow of H3K9me3 and H3K4me3 CUT&RUN and RNA sequencing analysis in SETDB1-CRISPRi, MORC2-CRISPRi, TRIM28-CRISPRi and control hNPCs. **B** RPKM normalized genome browser tracks showing MORC2 expression in MORC2-CRISPRi and control hNPCs. Data from Pandiloski *et al.* (2024) (*22*) **C** RPKM normalized genome browser tracks showing TRIM28 expression in TRIM28-CRISPRi and control hNPCs. Data from Horvath et al. (2024) (*50*) **D** 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) (*22*) and Horvath *et al.* (2024) (*50*) **E** Mean plot of uniquely mapped bulk RNA sequencing reads for MORC2-CRISPRi (n=4) and TRIM28-CRISPRi (n=4) hNPCs compared to controls (n=4) over full-length (>6 kb) L1HS-L1PA3 elements. LFC > 0, padj < 0.05 calculated with DESeq2. Data from Pandiloski *et al.* (2024) (*22*) and Horvath *et al.* (2024) (*50*) **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, MORC2-CRISPRi and TRIM28-CRISPRi hNPCs compared to controls (bottom).

Loss of H3K9me3 maintenance on L1s results in loss of DNA methylation

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, 21, 22*). Biochemical experiments suggest that DNMT1, the enzyme that maintains DNA methylation in somatic cells, directly recognizes H3K9me3 marks to reinforce CpG methylation (*71*). 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. 5A)**. As an additional positive control, we used CRISPRi to silence DNMT1 (*22, 50*), whose deletion in hNPCs results in a global loss of DNA methylation, including over L1s (*8*) (**Fig. 5A, Sup Fig 5A-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. 5B)**. This contrasts with the loss of DNMT1, which results in global hypomethylation **(Fig. 5B)**. 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. 5C)**. For example, we found full-length human- and hominoid-specific L1 elements on chromosome 4 and 5 that were transcriptionally activated in SETDB1-CRISPRi hNPCs, as determined by bulk RNA sequencing and CUT&RUN profiling of the active promoter mark H3K4me3, that also lost DNA methylation. **(Fig. 5D-E)**. Taken together, these results demonstrate that the loss of H3K9me3 methylation maintenance in human somatic cells leads to subsequent loss of DNA methylation and gain of H3K4me3 methylation, resulting in transcriptional activation of L1s.

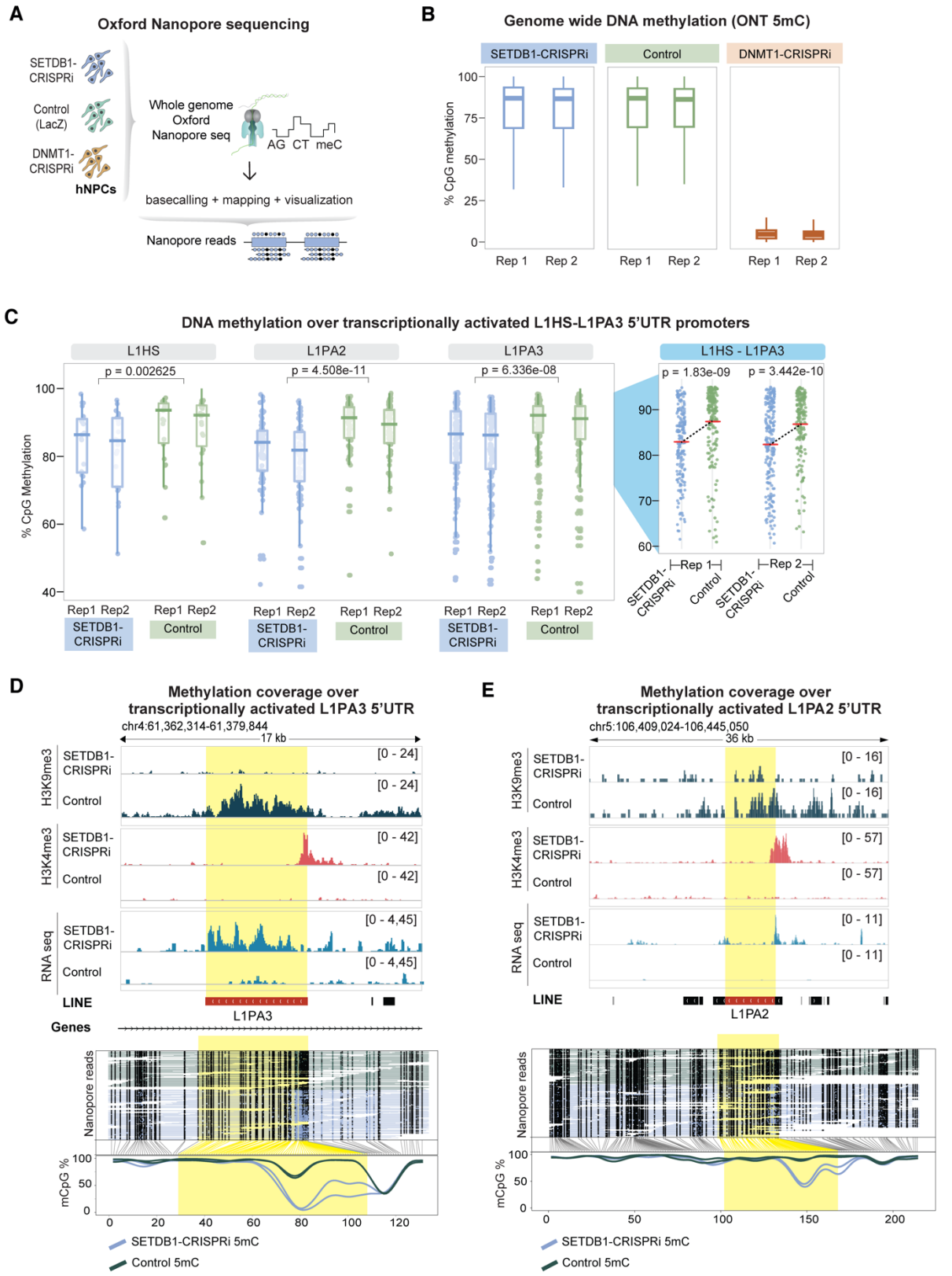


Figure 5. Loss of H3K9me3 maintenance results in loss of CpG methylation at L1 promoters in hNPCs A Schematics of whole genome ONT long-read DNA sequencing analysis in hNPCs **B** Box plots of the genome wide CpG DNA methylation status in SETDB1-CRISPRi (*n*=2), control (*n*=2) and DNMT1-CRISPRi (*n*=2) hNPCs. **C** Box plots of the methylation status over the promoter of transcriptionally activated L1HS-L1PA3 L1s 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. **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 at the bottom.

Discussion

Here we study the consequences of loss of H3K9me3 heterochromatin in human hNPCs, a cell type that allows us to model some aspects of heterochromatin maintenance, including the presence of H3K9me3 and DNA methylation at L1s (*8, 22*). Upon depletion of the H3K9me3 methyltransferase SETDB1 in hNPCs, we found a major loss of H3K9me3 peaks, including those over L1s and other TEs, as well as 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. We do not fully understand the mechanism behind this phenomenon, but previous studies have reported that a global loss of H3K9me3 leads to rearrangements of the epigenome and leads to the gain of H3K9me3 at specific loci mediated by other H3K9me3 methyltransferases such as SUV39H1/2 (*72-74*). 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 (*75*) 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* (*20, 76-80*). It is well established that CpG DNA methylation plays a critical role in the transcriptional silencing of L1 in neural cells (*6, 8, 9, 22, 31, 38*).

Upon deletion of DNMT1 in human NPCs, which leads to global DNA demethylation, evolutionarily young full-length L1 elements are transcriptionally activated (*8*). Thus, DNA methylation appears to be particularly important for silencing this family of TEs in the brain. However, it is not known how DNA methylation is directed to the promoter of evolutionary young L1s in human somatic 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 L1 promoter, 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 (*71, 81, 82*).

The best understood mechanisms for deposition of H3K9me3 on TEs are the KRAB zinc finger proteins (KZNF)-TRIM28 system and the HUSH pathway. However, a previous study found that ancestral L1PA3 elements have undergone structural sequence changes that allow them and subsequent younger L1 families to escape silencing by KZNFs (*83*). The HUSH pathway can target young L1s (*70, 84*), 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, 22*). 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 required to maintain DNA methylation at L1s in somatic cells and demonstrate that DNMT1 alone is not sufficient to maintain methylation patterns.

Loss of H3K9me3 and heterochromatin is a hallmark of ageing. The global loss and local increase of heterochromatin has been described as a hallmark for cellular senescence and has been detected in high passage cell cultures and aged human cells (*47, 74, 85, 86*). 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. This is likely to have consequences. Evolutionarily young, intact L1 elements have the potential to retrotranspose and thus can cause deleterious mutations and structural variants (*11, 87-90*). 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, 20*). 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 6A)**. 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 (*24, 25, 30, 91*). 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 6B)**.

In conclusion, our results provide a unique insight into the consequences of loss of H3K9me3 maintenance in human somatic cells. H3K9me3 maintenance is not required to maintain cellular identity and viability, but rather provides a critical mechanism to repress evolutionarily young L1 elements. These data provide a strong basis for further investigation of the role of L1s and other TEs in human ageing processes.

Methods

Cell culture

The human embryo-derived neural epithelial-like stem cell line Sai2 (*54*) was maintained according to the established protocol (*92*). 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 (*93*). 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), paired-end CUT&RUN reads were aligned to the human genome (GRCh38) and subsequently converted to BAM files using SAMtools v.1.18 (*94*). 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 (*95*). All coverage tracks displayed were visualized in IGV v2.19.1. Peaks were called using SEACR v1.3 (*96*). Peaks over reference chromosomes longer than 500bp were kept for downstream analyses. Using deepTools' computeMatrix v 2.5.4 (*95*), 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.

Lentiviral vector production

Lentivirus was produced as previously described by Zufferey et al (*97*). 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 (*59*). 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 or DNMT1 targeting guide RNAs. 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^+ gates were set using untransduced hNPCs and validated via reanalysis of sorted cells (> 95% purity). GFP^+/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
LACZ guide: TGCGAATACGCCCACGCGAT

qRT-PCR

Total RNA isolated from hNPCs (RNeasy mini kit, Qiagen) were 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. 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) or rabbit anti-histone 3 (Abcam 18521, 1:250.000) diluted in 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.

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 (*98*) 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 (*94*). Normalized coverage tracks (bigwig files) using RPKM were created using deeptools bamCoverage v2.5.4 (*95*). FeatureCounts v1.6.3 (*99*) 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.

Multi mapping: Reads were aligned to multiple loci by increasing the number of multimappers allowed to 100 (--outFilterMultimapNmax 100) and increasing the number of anchor loci to 200 (--winAnchorMultimapNmax 200). Sorted BAM output was specified, which was then indexed using SAMtools v.1.18 (*94*) and used for creating round per kilobase million (RPKM) normalized coverage tracks with bamCoverage v2.5.4 (*95*). Reads were quantified over genomic elements using featureCounts v1.6.3 (*99*) with specifying the –s2 flag, to force strandedness of each feature to be taken into consideration. Over TEsubfamilies, reads were quantified using TEtranscripts v2.0.3 (*63*) using the authors' currated GTF file using multi mode (--mode).

Differential expression analyses: DESeq2 v1.38.3 (*100*) 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.
To obtain distributions of the expected number of upregulated TEs, TE annotation was bootstrapped 50000 times without replacement. Each boostrapping attempt was directly compared to the observed values and subsequently scored (+1 if expected > observed and +0 if expected ≤ observed) to calculate a double sided p-value. Gene-ontology overrepresentation test of biological processes (enrichGO) was performed using clusterProfiler v4.8.2 over differentially expressed genes (|LFC| > 1, 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)[] 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 (*101*). Alignment files (BAM) sorted and indexed using SAMtools, and methylation matrices were created using the segmeth tool from MethylArtist v1.2.6 (*102*). 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.

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. 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.

Funding

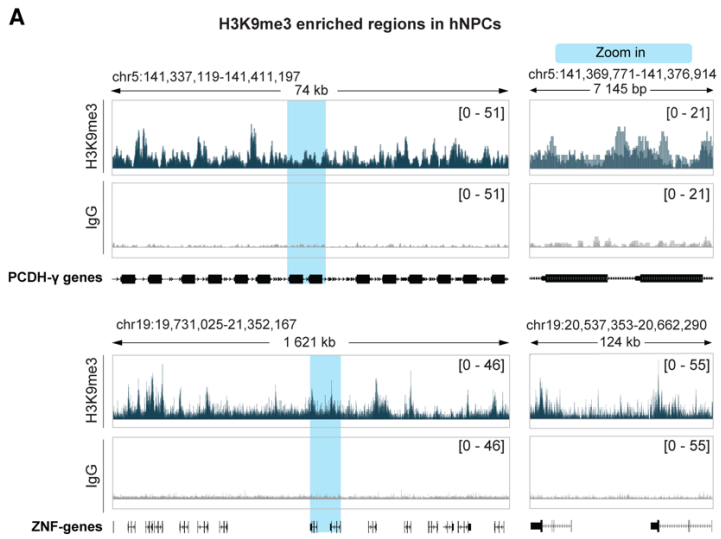
The work was supported by grants from the Swedish Research Council (2022-00673 to J.J. and 2021-03494 to C.H.D.), the Swedish Brain Foundation (FO2023-0232 to J.J.), Cancerfonden (222185 to J.J.), Barncancerfonden (PR2023-0099to J.J.), the Swedish Society for Medical Research (S19-0100 to C.H.D.), Olle 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.

Author contributions

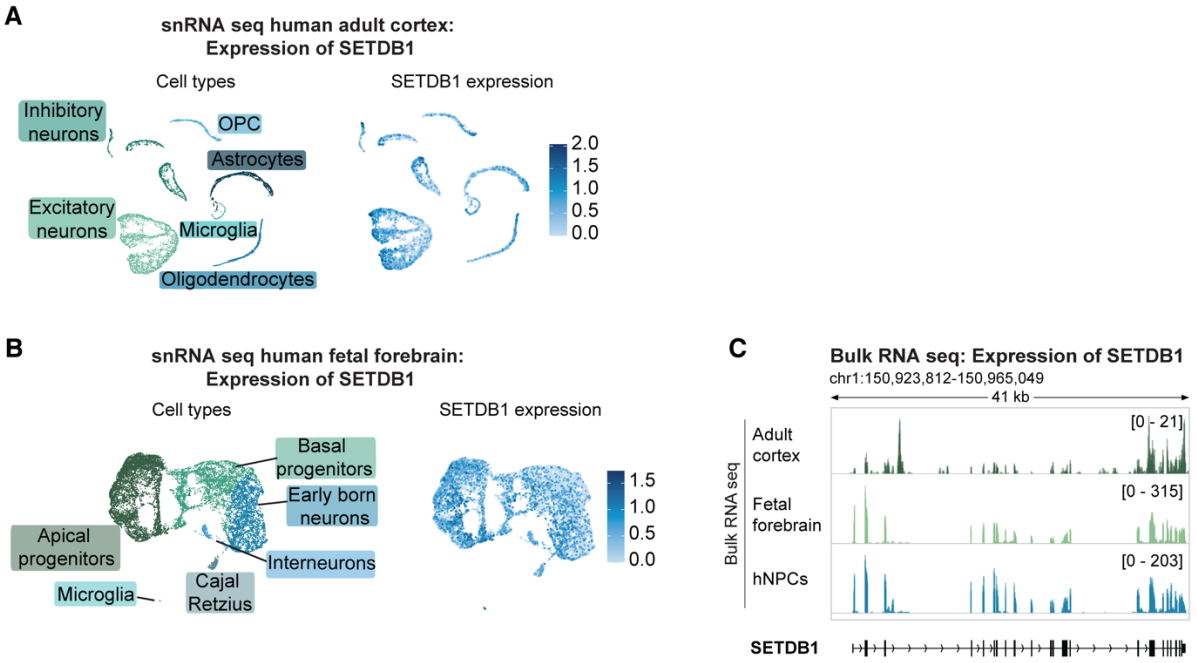
Design and interpretation: All authors. Conceptualization: O.K., N.P., and J.J. Experimental research: O.K., A.A., V.H., P.J. and J.G.J. Bioinformatics: N.P. and R.G. Writing—original draft: O.K., N.P. and J.J. Writing—review and editing: All authors.

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

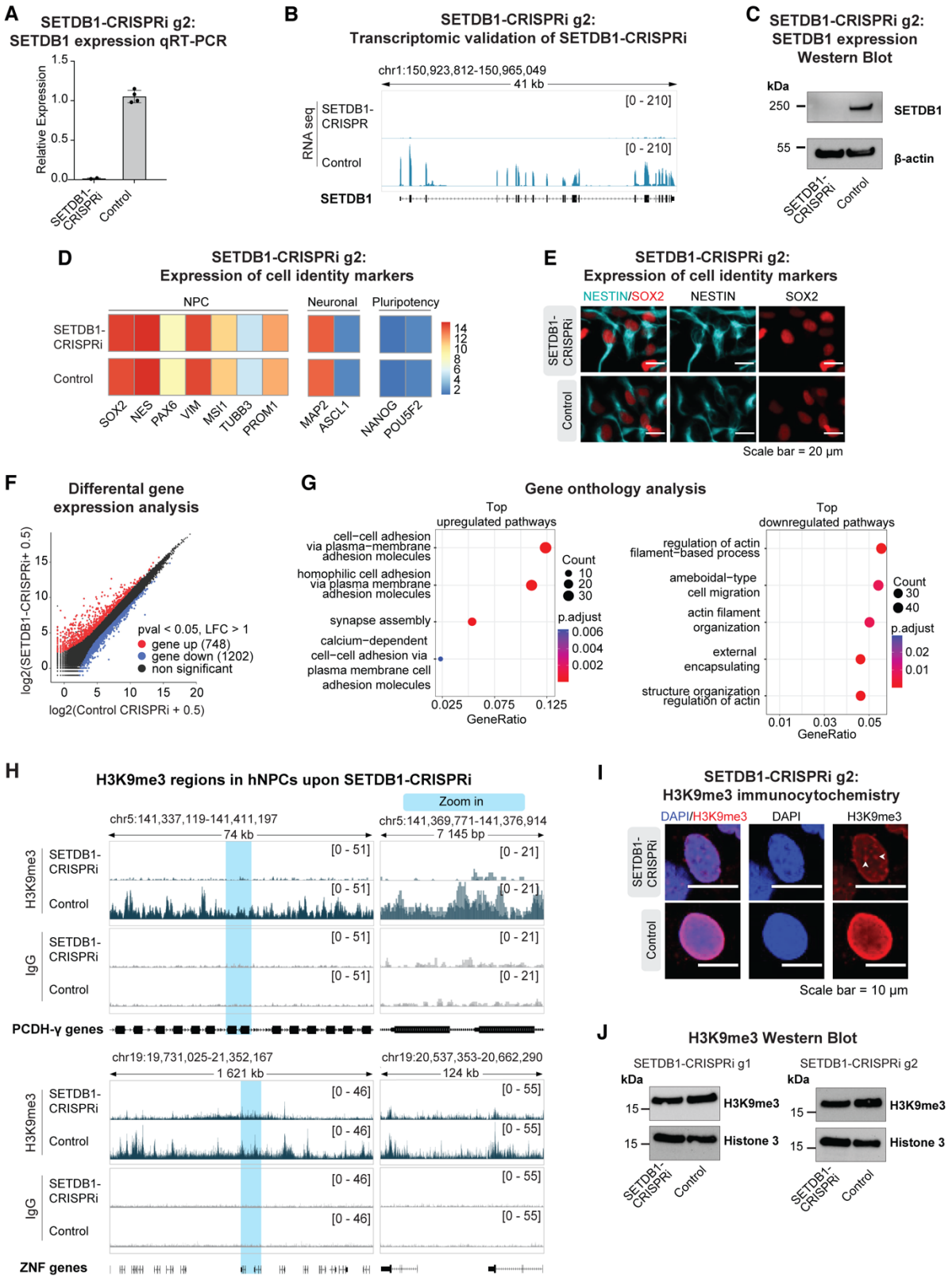
Supplementary Figures



Supplementary Figure 1. H3K9me3 in hNPCs localizes to regions with well established H3K9me3-enrichment A CUT&RUN analysis of H3K9me3 in hNPCs over genomic regions with well-established H3K9me3-enrichment; PCDH-ɣ cluster (top); KRAB-ZNF cluster (bottom).

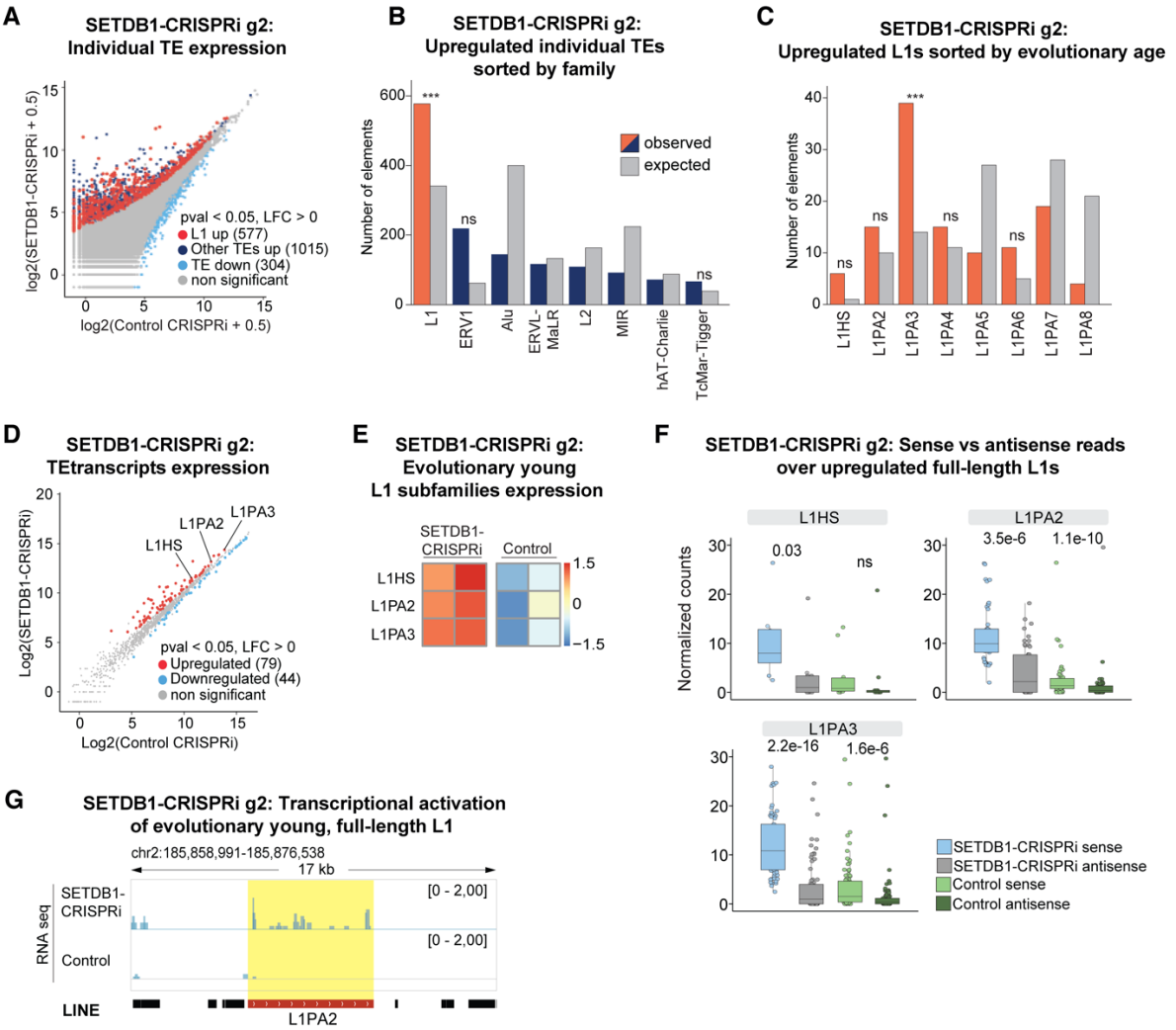


Supplementary Figure 2. SETDB1 is expressed in the adult and developing human brain and hNPCs A-B UMAP of snRNA-seq data showing SETDB1 expression in (A) human adult cortical tissue and (B) human fetal forebrain tissue. **C** Genome browser tracks showing bulk RNA sequencing data of SETDB1 expression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

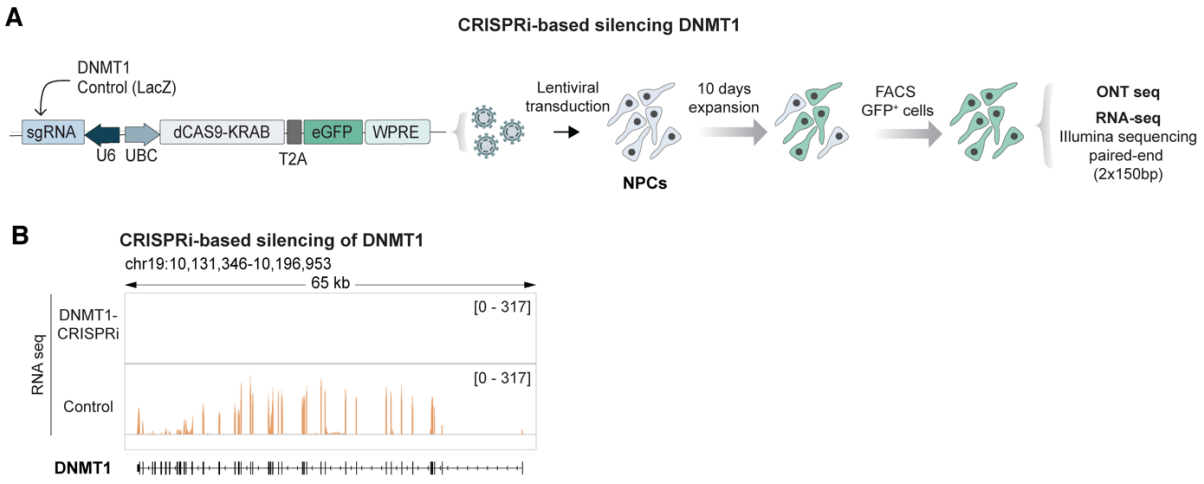


Supplementary Figure 3. Validation of SETDB1-CRISPRi using an alternative guide A qRT-PCR analysis of SETDB1 expression after CRISPRi g2-based silencing in hNPCs. Results are shown as means with standard deviation (*n*=2 and *n*=4 respectively). **B** RPKM normalized genome browser tracks showing expression of SETDB1 in SETDB1-CRISPRi g2 and control hNPCs as determined by bulk

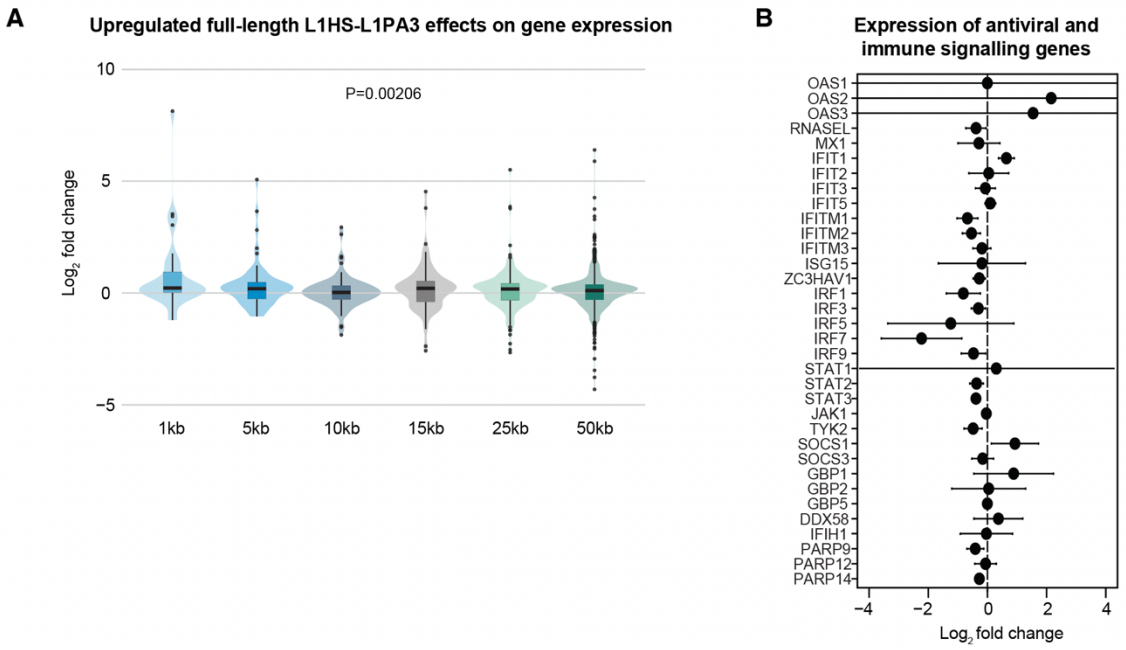
RNA sequencing **C** Western blot analysis of SETDB1 protein levels relative to β-actin protein levels in SETDB1-CRISPRi g2 and control hNPCs. **D** Heat map of NPC, neuronal and pluripotency gene marker mean expression in SETDB1-CRISPRi g2 (*n*=2) and control (*n*=4) hNPCs as determined by bulk RNA sequencing. **E** Immunocytochemistry of NESTIN (cyan) and SOX2 (red) in SETDB1-CRISPRi g2 and control hNPCs. Scalebar = 20 µm. **F** 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. **G** 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. **H** CUT&RUN analysis of H3K9me3 over genomic regions with well-established H3K9me3-enrichment in SEDTB1-CRISPRi and control hNPCs **I** 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. **J** Western blot analysis of H3K9me3 levels relative to Histone 3 in SETDB1-CRISPRi g1 hNPCs compared to control (left) and SETDB1-CRISPRi g2 hNPCs compared to control (right).



Supplementary Figure 4. Validation of transcriptional activation of L1s upon SETDB1-CRISPRi using an alternative guide 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** Number of individual, transcriptionally activated elements per TE family in SETDB1-CRISPRi g2 hNPCs (*n*= 2) compared to control (*n*= 2) versus shuffled element (50.000 bootstraps). Pvalues were calculated using bootstrap-based empirical pvalue calculation. **C** Observed vs expected number of transcriptionally activated L1 elements sorted by subfamily in SETDB1-CRISPRi g2 *(n=2)* compared to control hNPCs *(n=2)*. Pvalues were calculated using bootstrap-based empirical pvalue calculation. **D** Mean plot of TE families expression as determined by bulk RNA sequencing using TEtranscripts in SETDB1-CRISPRi g2 (*n*=2) and control (*n*=2) hNPCs. LFC >0, padj 0.05 calculated with DESeq2. **E** LFC of L1HS-L1PA3 subfamilies as determined by bulk RNA sequencing in SETDB1-CRISPRi g2 (*n*=2) compared to control hNPCs (*n*=2) using TEtranscripts. **F** 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. **G** RPKM normalized genome browser tracks showing RNA-seq in SETDB1-CRISPRi g2 and control hNPCs.



Supplementary Figure 5. CRISPRi-based silencing of DNMT1 A Workflow for CRISPRi-based silencing of DNMT1 in hNPCs and downstream analysis. **B** Genome browser tracks showing DNMT1 expression in DNMT1-CRISPRi and control hNPCs.



Supplementary Figure 6. Upregulated evolutionary young, full-length L1s affect nearby gene expression and expression of antiviral and immune signalling genes 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 signalling genes in SETDB1-CRISPRi hNPCs (*n*=4) compared to control (*n*=4). LFC values are shown ± LFC standard error calculated by DESeq2.

MER34B:ERV1:LTR	3.16247631399295	1.45501726782391e-63
HERV9NC-int:ERV1:LTR	3.10846928418097	2.30189174112048e-120
LTR10B2:ERV1:LTR	2.99836567884871	4.63835750363592e-34
MER39B:ERV1:LTR	2.60437244532883	2.23726839487073e-18
LTR12C:ERV1:LTR	2.60213211962628	8.3661822665309e-75
MER9a1:ERVK:LTR	2.58575378129185	2.24949446236682e-69
MER51E:ERV1:LTR	2.40367369822561	7.1781250462725e-53
LTR10C:ERV1:LTR	2.24230947433876	9.28390220924681e-75
LTR12D:ERV1:LTR	2.19577480151062	1.12998346538734e-35
HERVIP10B3-int:ERV1:LTR	2.09020258899439	2.78960713570136e-36
HERVI-int:ERV1:LTR	1.82575524679422	5.04817352736761e-08
LTR14C:ERVK:LTR	1.72611419969613	0.0341443925193485
HERVKC4-int:ERVK:LTR	1.72217603055252	0.0335341896160611
MER57E2:ERV1:LTR	1.71386136440529	0.0382707811830221
LTR10B:ERV1:LTR	1.57278041059048	0.0282268093627355
L1PA15-16:L1:LINE	1.12359075637236	1.59974804769826e-15
L1PA7:L1:LINE	0.9149628266938	4.21134905876528e-18
SVA_E:SVA:Retroposon	0.909851230811479	1.94385818840227e-06
HERVE_a-int:ERV1:LTR	0.893227064562727	4.39000414895656e-10
SVA_D:SVA:Retroposon	0.771243471030396	1.77998370887488e-08

Supplementary Table 1. Various TE families are upregulated upon SETDB1-CRISPRi Table showing selected differentially expressed TEs families as determined by bulk RNA sequencing using TEtranscripts in SETDB1-CRISPRi (*n*=4) and control (*n*=4) hNPCs. LFC and padj calculated with DESeq2.

References

1. E. S. Lander *et al.*, Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
2. A. P. de Koning, W. Gu, T. A. Castoe, M. A. Batzer, D. D. Pollock, Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* **7**, e1002384 (2011).
3. S. Nurk *et al.*, The complete sequence of a human genome. *Science* **376**, 44-53 (2022).
4. S. J. Hoyt *et al.*, From telomere to telomere: The transcriptional and epigenetic state of human repeat elements. *Science* **376**, eabk3112 (2022).
5. L. M. Payer, K. H. Burns, Transposable elements in human genetic disease. *Nat Rev Genet* **20**, 760-772 (2019).
6. O. Deniz, J. M. Frost, M. R. Branco, Regulation of transposable elements by DNA modifications. *Nat Rev Genet* **20**, 417-431 (2019).
7. J. A. Yoder, C. P. Walsh, T. H. Bestor, Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics* **13**, 335-340 (1997).
8. M. E. Jönsson *et al.*, Activation of neuronal genes via LINE-1 elements upon global DNA demethylation in human neural progenitors. *Nat Commun* **10**, 3182 (2019).
9. F. J. Sanchez-Luque *et al.*, LINE-1 Evasion of Epigenetic Repression in Humans. *Molecular Cell* **75**, 590-604.e512 (2019).
10. S. Lanciano *et al.*, Locus-level L1 DNA methylation profiling reveals the epigenetic and transcriptional interplay between L1s and their integration sites. *Cell Genom* **4**, 100498 (2024).
11. C. R. Beck, J. L. Garcia-Perez, R. M. Badge, J. V. Moran, LINE-1 elements in structural variation and disease. *Annu Rev Genomics Hum Genet* **12**, 187-215 (2011).
12. Q. Feng, J. V. Moran, H. H. Kazazian, Jr., J. D. Boeke, Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905-916 (1996).
13. G. D. Swergold, Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol Cell Biol* **10**, 6718-6729 (1990).
14. S. L. Mathias, A. F. Scott, H. H. Kazazian, J. D. Boeke, A. Gabriel, Reverse Transcriptase Encoded by a Human Transposable Element. *Science* **254**, 1808-1810 (1991).
15. E. Khazina, O. Weichenrieder, Human LINE-1 retrotransposition requires a metastable coiled coil and a positively charged N-terminus in L1ORF1p. *eLife* **7**, e34960 (2018).
16. C. R. Beck *et al.*, LINE-1 Retrotransposition Activity in Human Genomes. *Cell* **141**, 1159-1170 (2010).
17. B. Brouha *et al.*, Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A* **100**, 5280-5285 (2003).
18. A. M. Denli *et al.*, Primate-specific ORF0 contributes to retrotransposon-mediated diversity. *Cell* **163**, 583-593 (2015).
19. M. Speek, Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol Cell Biol* **21**, 1973-1985 (2001).
20. R. Garza *et al.*, LINE-1 retrotransposons drive human neuronal transcriptome complexity and functional diversification. *Sci Adv* **9**, eadh9543 (2023).
21. Ö. Deniz, J. M. Frost, M. R. Branco, Regulation of transposable elements by DNA modifications. *Nature Reviews Genetics* **20**, 417-431 (2019).
22. N. Pandiloski *et al.*, DNA methylation governs the sensitivity of repeats to restriction by the HUSH-MORC2 corepressor. *Nature Communications* **15**, 7534 (2024).
23. F. Della Valle *et al.*, LINE-1 RNA causes heterochromatin erosion and is a target for amelioration of senescent phenotypes in progeroid syndromes. *Sci Transl Med* **14**, eabl6057 (2022).

24. M. De Cecco *et al.*, L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature* **566**, 73-78 (2019).
25. M. Simon *et al.*, LINE1 Derepression in Aged Wild-Type and SIRT6-Deficient Mice Drives Inflammation. *Cell Metabolism* **29**, 871-885.e875 (2019).
26. H. Tang *et al.*, The transcription factor PAX5 activates human LINE1 retrotransposons to induce cellular senescence. *EMBO Rep* **25**, 3263-3275 (2024).
27. D. Ardeljan *et al.*, Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication. *Nature Structural & Molecular Biology* **27**, 168-178 (2020).
28. G. J. Faulkner *et al.*, The regulated retrotransposon transcriptome of mammalian cells. *Nature Genetics* **41**, 563-571 (2009).
29. M. E. Oomen *et al.*, An atlas of transcription initiation reveals regulatory principles of gene and transposable element expression in early mammalian development. *Cell* **188**, 1156-1174.e1120 (2025).
30. H. Tunbak *et al.*, The HUSH complex is a gatekeeper of type I interferon through epigenetic regulation of LINE-1s. *Nature Communications* **11**, 5387 (2020).
31. N. G. Coufal *et al.*, L1 retrotransposition in human neural progenitor cells. *Nature* **460**, 1127-1131 (2009).
32. J. A. Erwin *et al.*, L1-associated genomic regions are deleted in somatic cells of the healthy human brain. *Nature Neuroscience* **19**, 1583-1591 (2016).
33. G. D. Evrony *et al.*, Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* **151**, 483-496 (2012).
34. A. R. Muotri *et al.*, Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* **435**, 903-910 (2005).
35. K. R. Upton *et al.*, Ubiquitous L1 mosaicism in hippocampal neurons. *Cell* **161**, 228-239 (2015).
36. J. K. Baillie *et al.*, Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* **479**, 534-537 (2011).
37. Gilad D. Evrony *et al.*, Cell Lineage Analysis in Human Brain Using Endogenous Retroelements. *Neuron* **85**, 49-59 (2015).
38. A. Macia *et al.*, Engineered LINE-1 retrotransposition in nondividing human neurons. *Genome Res* **27**, 335-348 (2017).
39. T. Matsui *et al.*, Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**, 927-931 (2010).
40. N. Castro-Diaz *et al.*, Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev* **28**, 1397-1409 (2014).
41. H. M. Rowe *et al.*, TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. *Genome Res* **23**, 452-461 (2013).
42. A. Bulut-Karslioglu *et al.*, Suv39h-Dependent H3K9me3 Marks Intact Retrotransposons and Silences LINE Elements in Mouse Embryonic Stem Cells. *Molecular Cell* **55**, 277-290 (2014).
43. J. G. Wood, S. L. Helfand, Chromatin structure and transposable elements in organismal aging. *Front Genet* **4**, 274 (2013).
44. M. De Cecco *et al.*, Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging (Albany NY)* **5**, 867-883 (2013).
45. C. J. Driver, S. W. McKechnie, Transposable elements as a factor in the aging of Drosophila melanogaster. *Ann N Y Acad Sci* **673**, 83-91 (1992).
46. C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer, Hallmarks of aging: An expanding universe. *Cell* **186**, 243-278 (2023).

47. M. De Cecco *et al.*, Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. *Aging Cell* **12**, 247-256 (2013).
48. J. G. Wood *et al.*, Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in Drosophila. *Proceedings of the National Academy of Sciences* **113**, 11277-11282 (2016).
49. W. Li *et al.*, Activation of transposable elements during aging and neuronal decline in Drosophila. *Nature Neuroscience* **16**, 529-531 (2013).
50. V. Horváth *et al.*, Mini-heterochromatin domains constrain the cis-regulatory impact of SVA transposons in human brain development and disease. *Nature Structural & Molecular Biology*, (2024).
51. H. M. Rowe *et al.*, KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237-240 (2010).
52. G. Ecco, M. Imbeault, D. Trono, KRAB zinc finger proteins. *Development* **144**, 2719-2729 (2017).
53. M. Imbeault, P.-Y. Helleboid, D. Trono, KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* **543**, 550-554 (2017).
54. J. Tailor *et al.*, Stem Cells Expanded from the Human Embryonic Hindbrain Stably Retain Regional Specification and High Neurogenic Potency. *The Journal of Neuroscience* **33**, 12407 (2013).
55. H. Khan, A. Smit, S. Boissinot, Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. *Genome Res* **16**, 78-87 (2006).
56. M. Kato, K. Takemoto, Y. Shinkai, A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing. *Nat Commun* **9**, 1683 (2018).
57. D. Sun *et al.*, SETDB1 regulates short interspersed nuclear elements and chromatin loop organization in mouse neural precursor cells. *Genome Biol* **25**, 175 (2024).
58. D. C. Schultz, K. Ayyanathan, D. Negorev, G. G. Maul, F. J. Rauscher, 3rd, 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 (2002).
59. P. A. Johansson, A. Adami, J. Jakobsson, CRISPRi-mediated transcriptional silencing in iPSCs for the study of human brain development. *STAR Protoc* **3**, 101285 (2022).
60. P. I. Thakore *et al.*, Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nature Methods* **12**, 1143-1149 (2015).
61. S. L. Tan *et al.*, Essential roles of the histone methyltransferase ESET in the epigenetic control of neural progenitor cells during development. *Development* **139**, 3806-3816 (2012).
62. J. Padeken, S. P. Methot, S. M. Gasser, Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. *Nature Reviews Molecular Cell Biology* **23**, 623-640 (2022).
63. Y. Jin, O. H. Tam, E. Paniagua, M. Hammell, TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**, 3593-3599 (2015).
64. P. L. Brattås *et al.*, TRIM28 Controls a Gene Regulatory Network Based on Endogenous Retroviruses in Human Neural Progenitor Cells. *Cell Rep* **18**, 1-11 (2017).
65. P. Turelli *et al.*, Primate-restricted KRAB zinc finger proteins and target retrotransposons control gene expression in human neurons. *Sci Adv* **6**, eaba3200 (2020).

66. P. Turelli *et al.*, Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. *Genome Res* **24**, 1260-1270 (2014).
67. G. Ecco *et al.*, Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Developmental Cell* **36**, 611-623 (2016).
68. C. H. Douse *et al.*, TASOR is a pseudo-PARP that directs HUSH complex assembly and epigenetic transposon control. *Nat Commun* **11**, 4940 (2020).
69. N. Liu *et al.*, Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature* **553**, 228-232 (2018).
70. L. Robbez-Masson *et al.*, The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. *Genome Res* **28**, 836-845 (2018).
71. W. Ren *et al.*, Direct readout of heterochromatic H3K9me3 regulates DNMT1-mediated maintenance DNA methylation. *Proc Natl Acad Sci U S A* **117**, 18439-18447 (2020).
72. T. Montavon *et al.*, Complete loss of H3K9 methylation dissolves mouse heterochromatin organization. *Nat Commun* **12**, 4359 (2021).
73. J.-W. Park, J. J. Kim, Y.-S. Bae, CK2 downregulation induces senescence-associated heterochromatic foci formation through activating SUV39h1 and inactivating G9a. *Biochemical and Biophysical Research Communications* **505**, 67-73 (2018).
74. A. Tsurumi, W. X. Li, Global heterochromatin loss: a unifying theory of aging? *Epigenetics* **7**, 680-688 (2012).
75. R. E. Mills, E. A. Bennett, R. C. Iskow, S. E. Devine, Which transposable elements are active in the human genome? *Trends in Genetics* **23**, 183-191 (2007).
76. M. E. Jönsson, R. Garza, P. A. Johansson, J. Jakobsson, Transposable Elements: A Common Feature of Neurodevelopmental and Neurodegenerative Disorders. *Trends Genet* **36**, 610-623 (2020).
77. M. Percharde *et al.*, A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. *Cell* **174**, 391-405.e319 (2018).
78. J. W. Jachowicz *et al.*, LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nature Genetics* **49**, 1502-1510 (2017).
79. T. Toda *et al.*, Long interspersed nuclear elements safeguard neural progenitors from precocious differentiation. *Cell Rep* **43**, 113774 (2024).
80. D. Mangoni *et al.*, LINE-1 regulates cortical development by acting as long non-coding RNAs. *Nature Communications* **14**, 4974 (2023).
81. J. Du, L. M. Johnson, S. E. Jacobsen, D. J. Patel, DNA methylation pathways and their crosstalk with histone methylation. *Nature Reviews Molecular Cell Biology* **16**, 519-532 (2015).
82. X. Liu *et al.*, UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nature Communications* **4**, 1563 (2013).
83. F. M. Jacobs *et al.*, An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* **516**, 242-245 (2014).
84. I. Müller *et al.*, MPP8 is essential for sustaining self-renewal of ground-state pluripotent stem cells. *Nature Communications* **12**, 3034 (2021).
85. P. Scaffidi, T. Misteli, Lamin A-dependent nuclear defects in human aging. *Science* **312**, 1059-1063 (2006).
86. D. K. Shumaker *et al.*, Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc Natl Acad Sci U S A* **103**, 8703-8708 (2006).

87. A. D. Ewing, H. H. Kazazian, Jr., High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res* **20**, 1262-1270 (2010).
88. N. Gilbert, S. Lutz-Prigge, J. V. Moran, Genomic Deletions Created upon LINE-1 Retrotransposition. *Cell* **110**, 315-325 (2002).
89. N. Gilbert, S. Lutz, T. A. Morrish, J. V. Moran, Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol Cell Biol* **25**, 7780-7795 (2005).
90. D. E. Symer *et al.*, Human L1 Retrotransposition Is Associated with Genetic Instability In Vivo. *Cell* **110**, 327-338 (2002).
91. C. A. Buttler, D. Ramirez, R. D. Dowell, E. B. Chuong, An intronic LINE-1 regulates IFNAR1 expression in human immune cells. *Mob DNA* **14**, 20 (2023).
92. A. Falk *et al.*, 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 (2012).
93. P. J. Skene, S. Henikoff, An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**, (2017).
94. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
95. F. Ramírez, F. Dündar, S. Diehl, B. A. Grüning, T. Manke, deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187-191 (2014).
96. M. P. Meers, D. Tenenbaum, S. Henikoff, Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. *Epigenetics Chromatin* **12**, 42 (2019).
97. R. Zufferey, D. Nagy, R. J. Mandel, L. Naldini, D. Trono, Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat Biotechnol* **15**, 871-875 (1997).
98. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
99. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
100. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
101. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).
102. S. W. Cheetham, M. Kindlova, A. D. Ewing, Methylartist: tools for visualizing modified bases from nanopore sequence data. *Bioinformatics* **38**, 3109-3112 (2022).