


RESEARCH

Open Access



# SETDB1 regulates short interspersed nuclear elements and chromatin loop organization in mouse neural precursor cells

Daijing Sun<sup>1†</sup>, Yueyan Zhu<sup>1†</sup>, Wenzhu Peng<sup>1†</sup>, Shenghui Zheng<sup>1</sup>, Jie Weng<sup>1</sup>, Shulong Dong<sup>1,2</sup>, Jiaqi Li<sup>1</sup>, Qi Chen<sup>1</sup>, Chuanhui Ge<sup>1</sup>, Liyong Liao<sup>1</sup>, Yuhao Dong<sup>1</sup>, Yun Liu<sup>1,2</sup>, Weida Meng<sup>1,2†</sup> and Yan Jiang<sup>1\*†</sup> 

<sup>†</sup>Daijing Sun, Yueyan Zhu and Wenzhu Peng contributed equally to this work.

<sup>†</sup>Weida Meng and Yan Jiang are co-senior authors.

\*Correspondence: yan\_jiang@fudan.edu.cn

<sup>1</sup> Institutes of Brain Science, State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Fudan University, Shanghai 200032, China

<sup>2</sup> MOE Key Laboratory of Metabolism and Molecular Medicine, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Fudan University, Shanghai 200032, China

## Abstract

**Background:** Transposable elements play a critical role in maintaining genome architecture during neurodevelopment. Short Interspersed Nuclear Elements (SINEs), a major subtype of transposable elements, are known to harbor binding sites for the CCCTC-binding factor (CTCF) and pivotal in orchestrating chromatin organization. However, the regulatory mechanisms controlling the activity of SINEs in the developing brain remains elusive.

**Results:** In our study, we conduct a comprehensive genome-wide epigenetic analysis in mouse neural precursor cells using ATAC-seq, ChIP-seq, whole genome bisulfite sequencing, *in situ* Hi-C, and RNA-seq. Our findings reveal that the SET domain bifurcated histone lysine methyltransferase 1 (SETDB1)-mediated H3K9me3, in conjunction with DNA methylation, restricts chromatin accessibility on a selective subset of SINEs in neural precursor cells. Mechanistically, loss of *Setdb1* increases CTCF access to these SINE elements and contributes to chromatin loop reorganization. Moreover, *de novo* loop formation contributes to differential gene expression, including the dysregulation of genes enriched in mitotic pathways. This leads to the disruptions of cell proliferation in the embryonic brain after genetic ablation of *Setdb1* both *in vitro* and *in vivo*.

**Conclusions:** In summary, our study sheds light on the epigenetic regulation of SINEs in mouse neural precursor cells, suggesting their role in maintaining chromatin organization and cell proliferation during neurodevelopment.

**Keywords:** Transposable elements, SETDB1, H3K9me3, DNA methylation, Chromatin loop, Neurodevelopment

## Background

Transposable Elements (TEs) make up a substantial portion of the mammalian genome, comprising approximately 40-50% in both human and mouse genomes [1, 2]. Their historical association with genetic diversity, evolutionary changes, and potential regulatory roles has been extensively documented [3–6]. Meanwhile, their roles in the



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

central nervous system, particularly their impact on the diversity of neurons, neuronal communication, and brain development and disorders, represent an essential area of ongoing research [7–9].

Short Interspersed Nuclear Elements (SINEs) represent a distinctive subclass among TEs due to their relatively concise size, typically spanning just a few hundred base pairs [10]. While SINEs themselves do not encode proteins, they play a role in transcribing functional RNAs [11] or acting as enhancers [11], delicately fine-tuning the intricate genetic expression vital for neurogenesis, neuronal maturation, and the establishment of neural networks [8, 12]. SINEs have been implicated in neuronal plasticity [13], actively contributing to the brain's adaptive responses in experiences, learning and memory processes [13]. For instance, the B2 subset, a distinct group within SINEs, is prevalent in the genome and widely distributed in gene-rich regions [14]. Studies have indicated that noncoding SINE\_B2 RNA can modulate gene expression under stressful conditions [15, 16]. Additionally, B2 transcription is activated in the mature hippocampus in response to novel stimuli [13]. Nevertheless, our understanding on how SINEs are regulated in the developing brain remains limited.

Epigenetic regulation of TEs plays a pivotal role in governing neurodevelopment, contributing to the establishment of a sophisticated framework for managing precise spatiotemporal gene expression patterns essential in the intricate process of neurodevelopment [12, 17, 18]. Notably, SINEs significantly contribute by serving as repositories of various epigenetic marks, including DNA methylation and histone modifications, exerting substantial influence on local chromatin configurations. DNA methylation patterns within SINEs wield significant influence over neighboring gene accessibility, modulating their expression and contributing significantly to vital developmental processes [19]. An integral repressive epigenetic mark, H3K9me3, holds particular importance in TE regulation during evolution. Although its specific role in SINE regulation within the brain remains elusive, existing reports suggest its involvement in silencing SINEs in macrophages and influencing gene expression associated with inflammatory responses [20]. Adding another layer to their epigenetic influence, SINEs maintain the capacity to act as binding sites for CCCTC-binding factor (CTCF) and participate in chromatin folding, and this process is under the regulation of chromatin remodeling complex ChAHP (CHD4/ADNP/HP1) in mouse embryonic stem cells (ESCs) [21, 22]. These features potentially impact the establishment and maintenance of the critical epigenetic landscape required for proper brain development.

SET Domain Bifurcated 1 (SETDB1), a histone methyltransferase targeting H3K9me3 in genic regions, plays a critical evolutionary role in TEs [23, 24]. Researches, including ours, underscore its critical involvement in brain development by regulating endogenous retroviruses (ERVs) [18, 25]. Analogous to the influence of SINEs on gene expression, the absence of SETDB1 in neural precursor cells (NPCs) derepresses ERVs, impacting adjacent gene transcription [25], or serving as enhancer elements [18]. Beyond its role as an H3K9me3 'writer', SETDB1 orchestrates 3D chromatin conformation within the brain [26]. Our previous study revealed its ability to maintain the topologically associating domain (TAD) conformation covering the protocadherin gene cluster, through H3K9me3 deposition, DNA methylation, and long-range repressive chromatin interactions [27]. Notably, SETDB1 loss in mature

neurons increases genome-wide CTCF binding [27], guiding our ongoing study on the epigenetic regulation of SINEs.

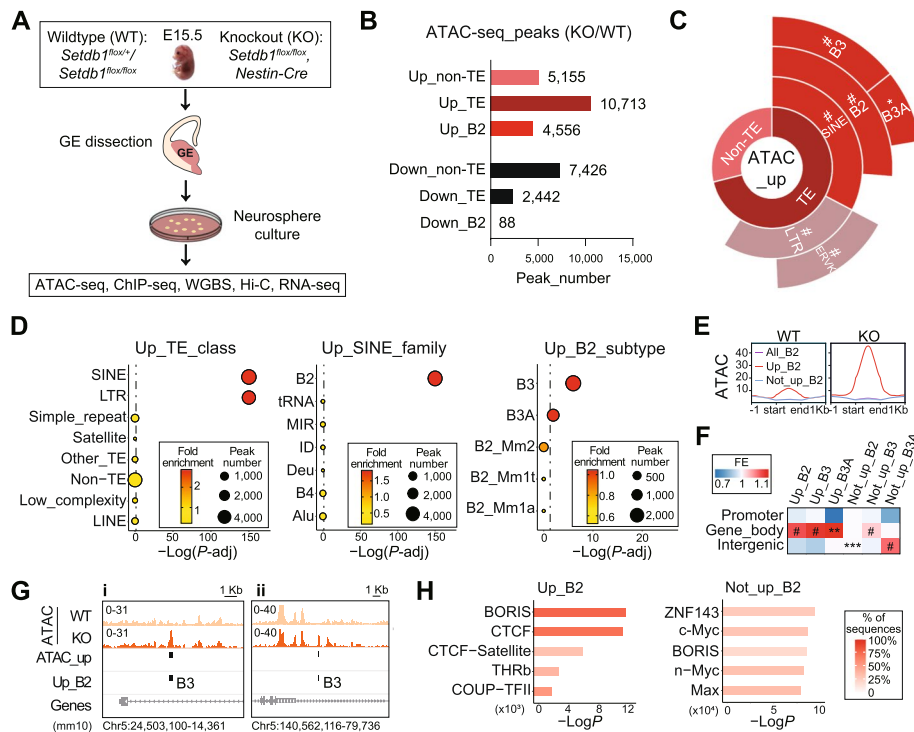
Our current study investigated the pivotal role of SETDB1-mediated H3K9me3, complemented by DNA methylation, in regulating a specific subset of SINE B2 elements. Despite their ancient origin, these elements exhibit functional conservation in evolution, particularly enriched for CTCF binding motifs. Our extensive epigenomic profiling, including ATAC-seq, anti-H3K9me3 ChIP-seq, WGBS, anti-CTCF ChIP-seq, *in situ* Hi-C, and RNA-seq analyses, deciphered intricate epigenetic regulation for these elements in mouse NPCs. Our research also sheds light on their impact on gene expression patterns through CTCF-mediated higher-order chromatin organizations, and potential influence on cell proliferation of NPCs.

## Results

### Increased chromatin accessibility on SINE\_B2 elements after *Setdb1* ablation in mouse NPCs

To investigate the regulation of TEs in the developing brain by SETDB1, we dissected out ganglionic eminences (GEs) from both wildtype (WT) and SETDB1-knockout (KO) mouse brains at embryonic day 15.5 (E15.5), and conducted neurosphere culture following an established protocol [18]. The purified NPCs were then harvested and subjected to various epigenomic profiling assays, including the evaluation of genome-wide chromatin accessibility (ATAC-seq), H3K9me3 and CTCF occupancy (ChIP-seq), DNA methylation (WGBS), 3D genome organization (*in situ* Hi-C), and gene transcription (RNA-seq) (Fig. 1A). It is worth noting that ATAC-seq and RNA-seq data were derived from our previously published work [18]. However, in this study, we undertook novel analyses of these datasets from a completely different perspective, shifting our emphasis onto the investigation of TEs instead of individual gene loci.

Differential analysis (KO/WT,  $N = 3$ ) of ATAC-seq peaks revealed genome-wide alterations in chromatin accessibility, resulting in 15,868 up-regulated and 9,868 down-regulated peaks (Fig. 1B). Notably, more than 65% (10,713) of the up-regulated ATAC-seq peaks (ATAC\_up, Additional file 1: Table S1) were located on TEs. Furthermore, nearly half of these peaks (4,556) were specifically annotated to the B2 family of SINE class (SINE\_B2). In contrast, less than 25% of the down-regulated ATAC-seq peaks (2,442) were associated with TEs, and only a limited number of 88 peaks were annotated as SINE\_B2 (Fig. 1B, Additional file 2: Fig. S1A-B). Further subtype analysis of the overlapping TEs revealed that only SINE and LTR showed significant enrichment within the ATAC\_up peaks (Fig. 1C-D). The enrichment on LTR was expected, as previous studies have reported that SETDB1 plays critical roles in repressing LTRs, particularly ERVs, in the brain and other organs [18, 24, 28–30]. In our study, within the LTR class, ERVKs were notably enriched in ATAC\_up peaks (Fig. 1C). In addition, previous work reported activation of LINEs in human acute myeloid leukemia cells after the loss of *SETDB1* [31]. However, LINEs were not enriched in ATAC\_up peaks in our current study (Fig. 1C), likely due to species specificity. Notably, the robust enrichment within the SINE class was observed. Among the SINE class, we found that B2, as opposed to Alu and other subclasses, exhibited high enrichment in the ATAC\_up peaks. Moreover, among the five B2 subtypes (B3, B3A, B2\_Mm2, B2\_Mm1t, B2\_Mm1a), B3 and B3A were the main



**Fig. 1** Increased chromatin accessibility on SINE\_B2 elements after *Setdb1* ablation in mouse neural precursor cells. **A** Experimental procedures. Ganglionic eminences (GEs) of wildtype (WT) or *Setdb1*-knockout (KO) mouse embryos were dissected at embryonic day 15.5 (E15.5) and dissociated for neurosphere culture. Neural precursor cells (NPCs) were harvested for ATAC-seq, ChIP-seq, WGBS, Hi-C and RNA-seq. WGBS, whole genome bisulfite sequencing. **B** Numbers of differential ATAC-seq peaks in NPCs annotated as non-TE, TE or B2 (overlapping  $\geq 50\%$ ). KO/WT,  $N = 3$ ,  $FDR < 0.05$ ,  $|\log_2(\text{FoldChange})| \geq 0.585$ . TE, transposable element. **C** Enrichment of up-regulated ATAC-seq peaks (ATAC\_up). **D** Enrichment of ATAC\_up on “TE\_class” (left), “SINE\_family” (middle) and “B2\_subtype” (right). Dotted lines indicate  $P\text{-adj} = 0.05$ . **E** ATAC-seq (ATAC) signal profiles on all B2 elements (All\_B2, purple), B2 with (Up\_B2, red) or without (Not\_up\_B2, blue) ATAC\_up. **F** Genomic annotation of Up\_B2, Up\_B3, Up\_B3A, Not\_up\_B2, Not\_up\_B3, and Not\_up\_B3A. FE, fold enrichment. Fisher's exact test, B-H adjusted, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , # $P < 0.0001$ . **G** IGV map tracks show ATAC signal of WT and KO on two representative regions (i and ii) on Chr5. Vertical bars indicate sites of ATAC\_up and Up\_B2. **H** Top 5 enriched Homer known motifs of Up\_B2 and Not\_up\_B2

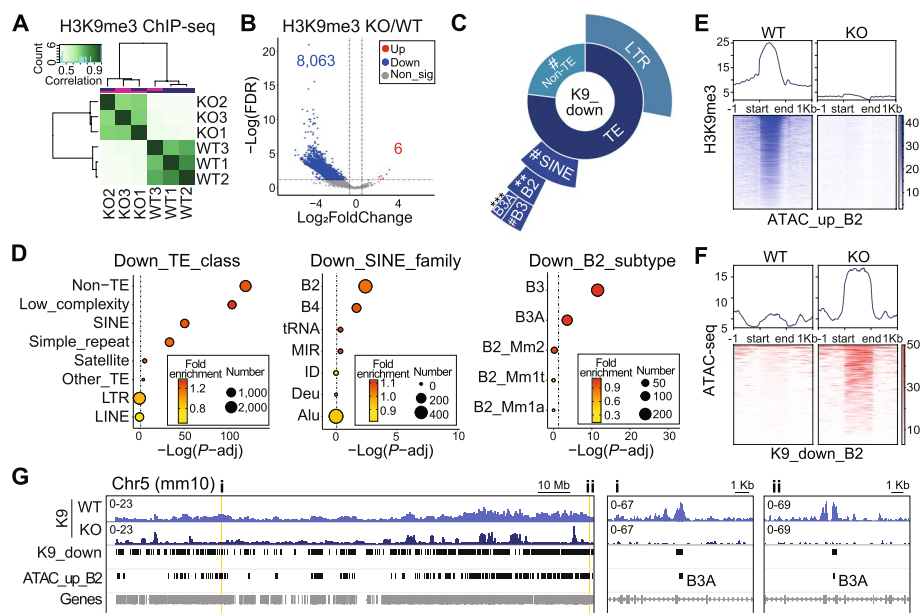
contributors to the observed effects (Fig. 1C-D). Considering the relatively short length of SINE (average length of 159 bp) compared to other types of TEs, we recalculated the enrichment by accumulative length and obtained consistent results (Additional file 2: Fig. S1C-D).

We defined the group of B2 elements that overlapped ATAC\_up peaks as “ATAC\_up\_B2”. When compared to all B2 elements in the genome (All\_B2), and the group of B2 that did not overlap with ATAC\_up peaks (ATAC\_not\_up\_B2), we observed a notable increase in chromatin accessibility for ATAC\_up\_B2, even in the WT, which was further augmented in KO (Fig. 1E). This finding suggested that this specific subset of B2 elements exhibits a baseline level of activity and potential functionality. Genomic annotation analysis revealed that ATAC\_up\_B2, as well as ATAC\_up\_B3 and ATAC\_up\_B3A, were significantly enriched in gene bodies (Fig. 1F). Interestingly, although ATAC\_up\_B2 were not enriched in promoters (Fig. 1F), they were significantly closer to transcription start sites (TSS) when compared to all B2 elements throughout genome (Additional

file 2: Fig. S1E). In contrast, ATAC\_not\_up\_B2 elements were enriched in intergenic regions (Fig. 1F), and located significantly farther from TSS (Additional file 2: Fig. S1E). Visualizations on the map tracks clearly demonstrated that ATAC\_up\_B2 elements were highly concentrated in gene-rich regions (Fig. 1G, Additional file 2: Fig. S1F). These results suggest a high potential for ATAC\_up\_B2 elements to influence gene transcription. Furthermore, we conducted Homer motif analysis and revealed that only ATAC\_up\_B2 were highly enriched with motifs for CTCE, in contrast to ATAC\_not\_up\_B2 or All\_B2 (Fig. 1H, Additional file 2: Fig. S1G), underscoring the distinct functional role of ATAC\_up\_B2 elements. In summary, these findings indicated that the chromatin accessibility of a specific population of SINE\_B2 elements is controlled by SETDB1 in NPCs.

### Reduction of H3K9me3 on SINE\_B2 after *Setdb1* ablation in NPCs

SETDB1 is one of the most important histone methyltransferases with specificity towards H3K9me3. We, therefore, investigated whether SETDB1 represses the chromatin accessibility on SINE\_B2 elements in NPC through H3K9me3 deposition. To explore this, we conducted anti-H3K9me3 ChIP-seq in NPCs from both WT and KO. The correlation analysis showed a clear distinction between WT and KO samples (Fig. 2A). Subsequently, we called peaks and performed a differential analysis ( $N = 3$ ) between KO and WT, which disclosed a substantial down-regulation of H3K9me3 peaks (K9\_down,  $N = 8,063$ ) in KO (Additional file 1: Table S2), with only 6 peaks showing up-regulation



**Fig. 2** Reduction of H3K9me3 on SINE\_B2 after *Setdb1* ablation in NPCs. **A** Correlation heatmap of H3K9me3 ChIP-seq in WT and KO NPCs. **B** Volcano plot shows differential analysis for H3K9me3 ChIP-seq. KO/WT,  $N = 3$ ,  $FDR < 0.05$ ,  $|\text{Log}_2\text{FoldChange}| \geq 0.585$ . Red, Up. Blue, Down. Grey, non-significant (Non\_sig). **C** Enrichment analysis of down-regulated H3K9me3 peaks (K9\_down) on mouse genome. **D** Enrichment of K9\_down on “TE\_class” (left), “SINE\_family” (middle), and “B2\_subtype”. Dotted lines indicate  $P\text{-adj} = 0.05$ . Fisher’s exact test, B-H adjusted,  $**P < 0.01$ ,  $***P < 0.001$ ,  $\#P < 0.0001$ . **E** H3K9me3 signal profiles and heatmaps on ATAC\_up\_B2 overlapping B2 (ATAC\_up\_B2) sites. **F** ATAC-seq signal profiles and heatmaps on K9\_down overlapping B2 (K9\_down\_B2) sites. **G** IGV map tracks show H3K9me3 (K9) signal of WT and KO on Chr5. Vertical bars indicate sites of K9\_down sites and ATAC\_up\_B2. Yellow shades indicate zoomed-in regions i and ii

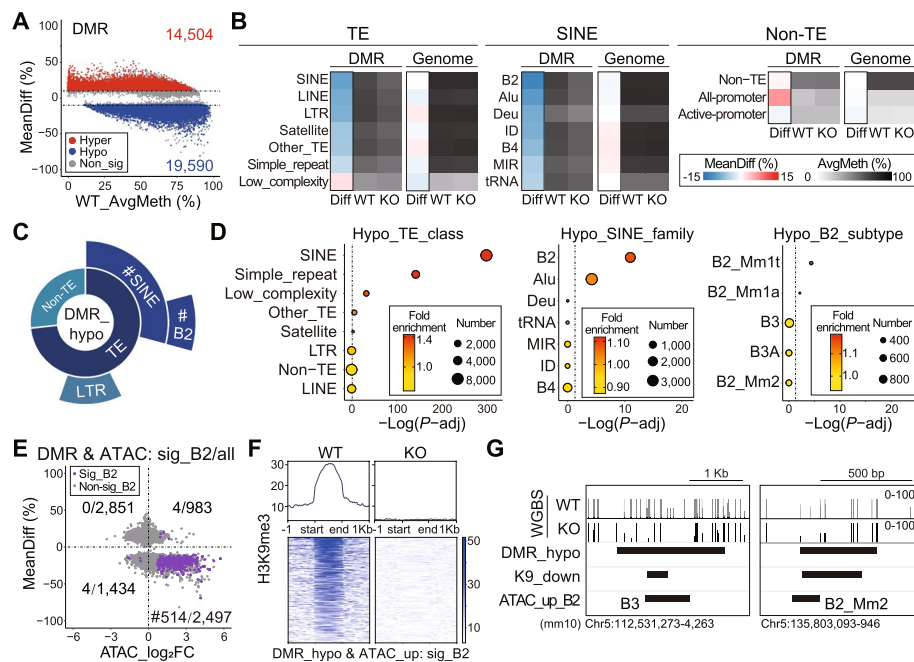
(Fig. 2B). The majority (62.1%) of these K9\_down peaks were located in genic regions (Additional file 2: Fig. S2A). Consistent with the ATAC-seq data (Fig. 1), we observed enrichments of SINE (by TE class), B2 (by SINE family) and B3/B3A (by B2 subtype) elements within the K9\_down peaks (Fig. 2C-D, Additional file 2: Fig. S2B-C). Although LTR elements constituted a substantial portion of the K9\_down peaks, they did not reach statistical significance in the enrichment tests by number (Fig. 2C-D). In addition, when considering enrichment by cumulative length, SINE\_B2 remained highly enriched in K9\_down peaks, while LTR displayed a milder level of enrichment (Additional file 2: Fig. S2C). Furthermore, CTCF motif was also enriched in the K9\_down peaks (Additional file 2: Fig. S2D).

Next, we assessed the signal correlation between differential peaks from the ATAC-seq and H3K9me3 ChIP-seq datasets. Although H3K9me3 signal was enriched on ATAC\_up\_B2 elements in WT, there was almost no signal in KO (Fig. 2E). *Vice versa*, the ATAC-seq signal was increased on K9\_down that overlap B2 elements (Fig. 2F). These observations were visually evident in the H3K9me3 map tracks, particularly on the ATAC\_up\_B2 elements (Fig. 2G). In summary, these findings suggested that the loss of H3K9me3 signal in *Setdb1* knockout NPCs may contribute to the increased chromatin accessibility observed on the SINE\_B2 elements.

#### Alterations of DNA methylation after *Setdb1* ablation in NPCs

Our previous work demonstrated the interplay between SETDB1-mediated H3K9me3 and DNA methylation, another well-established repressive epigenetic modification, on the gene cluster in neurons [27]. Given this, in the current study, we examined DNA methylation by conducting WGBS in NPCs derived from both WT and KO. The high efficiency of bisulfite conversion was confirmed using spiked-in lambda DNA (Additional file 2: Fig. S3A). The global DNA methylation level displayed only mild changes (Additional file 2: Fig. S3B), and the transcription of genes encoding DNA methyltransferases was not significantly altered in KO (Additional file 2: Fig. S3C). Nevertheless, differential analysis (KO/WT,  $N = 2$ ) revealed over 30,000 significant differentially methylated regions (DMRs) across the genome (Fig. 3A, Additional file 1: Table S3), evenly distributed on promoters, gene bodies and intergenic regions (Additional file 2: Fig. S3D). As expected, in the whole genome, most TEs exhibited high levels of DNA methylation (Fig. 3B, left), whereas non-TE regions, especially active gene promoters, displayed lower levels of methylation (Fig. 3B, right). Meanwhile, the genome-wide average methylation on SINEs and its subfamilies, including B2, remained largely unchanged (Fig. 3B, middle). However, when examining the significant DMRs, we observed prominent decrease of DNA methylation on TEs, with the exception of low-complexity regions. Among all TEs, SINE/B2 were the most hypomethylated (Fig. 3B, left and middle). Interestingly, gene promoters, but not those active promoters, exhibited increased DNA methylation in KO compared to WT (Fig. 3B, right). Moreover, enrichment analysis, both by TE number and accumulative length, revealed that significantly hypomethylated DMRs (DMR\_hypo) were enriched for SINE (by TE class) and B2 (by SINE family), but not for LTR (Fig. 3C-D, Additional file 2: Fig. S3E-F). Interestingly, within the subtypes of B2, DMR\_hypo were more enriched for Mm1t and Mm1a (Fig. 3D, Additional file 2: Fig. S3E-F), contrasting with B3/B3A that were enriched in the ATAC\_up (Fig. 1C-D)





**Fig. 3** Alterations of DNA methylation after *Setdb1* ablation in NPCs. **A** MA plot shows differential methylated regions (DMRs) derived from WGBS in WT and KO NPCs.  $N = 2$ ,  $|\text{MeanDiff}| \geq 0.1$ . Red, hypermethylated (Hyper). Blue, hypomethylated (Hypo). Grey, non-significant (Non\_sig). AvgMeth, average DNA methylation level. MeanDiff, mean DNA methylation difference between KO and WT. **B** DMR and genome-wide MeanDiff and AvgMeth of TEs, SINEs and non-TE regions. **C** Enrichment of significantly hypomethylated DMRs (DMR\_hypo) on mouse genome. **D** Enrichment of DMR\_hypo on “TE\_class” (left), “SINE\_family” (middle) and “B2\_subtype” (right). Dotted lines indicate  $P\text{-adj} = 0.05$ . Fisher’s exact test, B-H adjusted,  $^{\#}P < 0.0001$ . **E** MeanDiff and  $\text{Log}_2\text{FoldChange}$  of ATAC-seq signal (ATAC\_log<sub>2</sub>FC) for overlapping DMRs and ATAC-seq peaks (DMR & ATAC). sig\_B2/all, number of significant DMR & ATAC peaks on B2/all DMR & ATAC peaks. Purple, sig\_B2. Grey, non-sig\_B2. Note sig\_B2 is enriched in the 4<sup>th</sup> quadrant. Fisher’s exact test,  $^{\#}P < 0.0001$ . **F** Signal profiles and heatmaps of H3K9me3 on sig\_B2 from the 4<sup>th</sup> quadrant of (E). **G** IGV map tracks show DNA methylation level of WT and KO on overlapping DMR\_hypo, K9\_down, and ATAC\_up\_B2 sites

and K9\_down peaks (Fig. 2C-D). Nevertheless, the CTCF motif was also enriched in DMR\_hypo regions (Additional file 2: Fig. S3G).

Next, we evaluated the intersection between DMRs and ATAC-seq peaks. We found that for significant overlapping DMRs and ATAC peaks on B2 elements (sig\_B2), a substantial fraction (514/2,497) was situated in the 4<sup>th</sup> quadrant (DMR\_hypo & ATAC\_up), suggesting an increase in chromatin accessibility in the DNA hypomethylation regions on B2 elements in KO (Fig. 3E). Notably, there was a near complete loss of H3K9me3 signal in this set of B2 elements (Fig. 3F-G). Altogether, these findings implied that DNA methylation also participated in SETDB1-mediated suppression of B2 elements in NPCs.

#### Increase of CTCF binding on SINE\_B2 after *Setdb1* ablation in NPCs

SINE\_B2 elements are well known for their potential to harbor CTCF binding sites (CBS) [20, 21]. In our current study, we discovered an enrichment of CTCF motifs in B2 elements with ATAC\_up, K9\_down, and DMR\_hypo events (Fig. 1H, Additional file 2: Fig. S2D, Fig. S3G). Therefore, we performed anti-CTCF ChIP-seq to characterize the CTCF binding landscape in both WT and KO NPCs. Correlation analysis showed a clear distinction between WT and KO samples (Additional file 2: Fig. S4A). Moreover,

differential analysis (KO/WT,  $N = 3$ ) revealed a substantial global increase in CTCF binding, with 8,597 significantly up-regulated and 615 down-regulated CTCF peaks in KO (Fig. 4A). Remarkably, more than 70% of these up-regulated CTCF peaks (CTCF\_up) overlapped with 12,179 SINEs (Fig. 4B, Additional file 1: Table S4, Additional file 2: Fig. S4B), and showed exclusive enrichment (both in terms of TE number and cumulative length) for SINE (by TE class), B2 (by SINE family), and B3A/ B3 (by B2 subtype) elements (Fig. 4B, Additional file 2: Fig. S4C-E). By contrast, down-regulated CTCF peaks (CTCF\_down) only overlapped with 79 SINEs (Additional file 2: Fig. S4B). Notably, despite multiple evidence indicating the activation of LTR elements in KO (Fig. 1C-D, Additional file 2: Fig. S2B-C), LTR was not enriched in the CTCF\_up peaks when calculated by TE number (Additional file 2: Fig. S4C), and showed only mild enrichment when calculated by cumulative length (Additional file 2: Fig. S4E). Furthermore, CTCF motifs were only enriched in the CTCF\_up peaks overlapping B2 elements (CTCF\_up\_B2), not in peaks that did not overlap with B2 (Fig. 4C). These findings highlight the specificity of increased CTCF binding towards SINE\_B2 elements following the loss of *Setdb1* in NPCs.

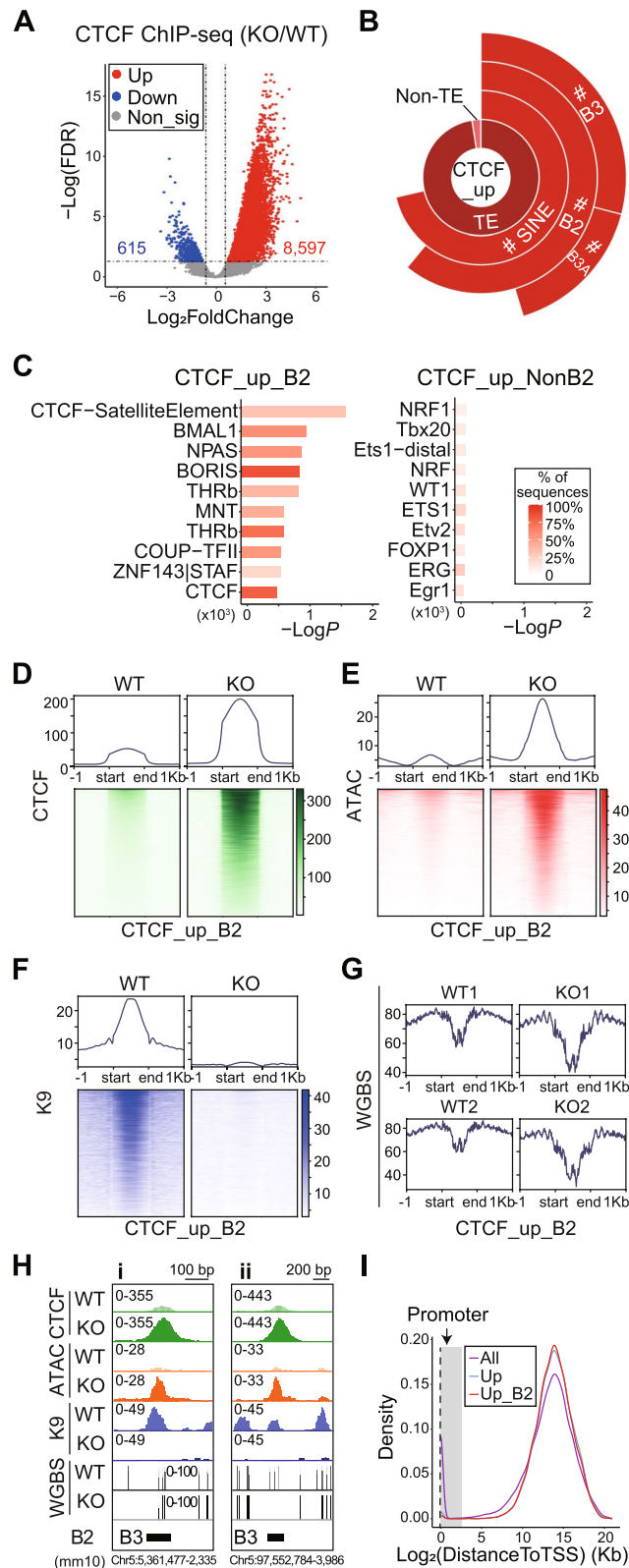
We proceeded to conduct an integrated analysis of the CTCF ChIP-seq, ATAC-seq, and WGBS datasets. Permutation tests revealed significantly more overlaps and shorter mean distance between CTCF\_up and ATAC\_up\_B2, relative to the background of all B2 loci (Additional file 2: Fig. S4F). Approximately half of the CTCF\_up peaks overlapped with ATAC\_up peaks, and over 70% (3,429/4,683) of ATAC\_up\_B2 peaks were associated with CTCF\_up peaks (Additional file 2: Fig. S4G). Furthermore, the scatter plot showed that almost all the B2 elements (99.95%) with significant changes in ATAC-seq and CTCF ChIP-seq exhibited increases of both signals in KO compared to WT (Additional file 2: Fig. S4H). Next, we investigated the correlation between differential CTCF binding and DNA methylation changes. Consistent with published literature indicating DNA methylation restricted CTCF binding [32, 33], signal profiles revealed noticeable decrease of DNA methylation on CTCF\_up peaks in KO (Additional file 2: Fig. S4I, top). Meanwhile, on CTCF\_down peaks, DNA methylation remained largely unchanged (Additional file 2: Fig. S4I, bottom). Taken together, these findings suggest that increased CTCF binding in KO was due to increased chromatin accessibility and DNA hypomethylation.

Following these observations, we looked deeper into the accumulative signal of CTCF ChIP-seq, ATAC-seq, H3K9me3 ChIP-seq, and DNA methylation on CTCF\_up\_B2 elements in both WT and KO (Fig. 4D-H). These B2 elements exhibited moderate enrichment of CTCF occupancy in WT, which increased in KO as anticipated (Fig. 4D). In

(See figure on next page.)

**Fig. 4** Increase of CTCF binding on SINE\_B2 after *Setdb1* ablation in NPCs. **A** Volcano plot shows differential CTCF ChIP-seq peaks. KO/WT,  $N = 3$ ,  $FDR < 0.05$ ,  $|\text{Log}_2\text{FoldChange}| \geq 0.585$ . Red, Up. Blue, Down. Grey, non-significant (Non\_sig). **B** Enrichment of up-regulated CTCF peaks (CTCF\_up) on mouse genome. **C** Top 10 enriched Homer known motifs of CTCF\_up peaks overlapping (CTCF\_up\_B2, left) and not overlapping B2 elements (CTCF\_up\_NonB2, right). **D-G** Signal profiles and heatmaps of CTCF ChIP-seq (**D**), ATAC-seq (**E**), H3K9me3 ChIP-seq (**F**) and WGBS (**G**) on CTCF\_up\_B2. (**H**) IGV map tracks show representative signals in (**D-G**) on CTCF\_up\_B2. (**I**) Density plot displays the distance to TSS of all CTCF peaks (All, purple), CTCF\_up (Up, blue), and CTCF\_up\_B2 (Up\_B2, red). TSS, transcription start site. Grey shade, promoter ( $\text{TSS} \pm 3 \text{ Kb}$ )





**Fig. 4** (See legend on previous page.)

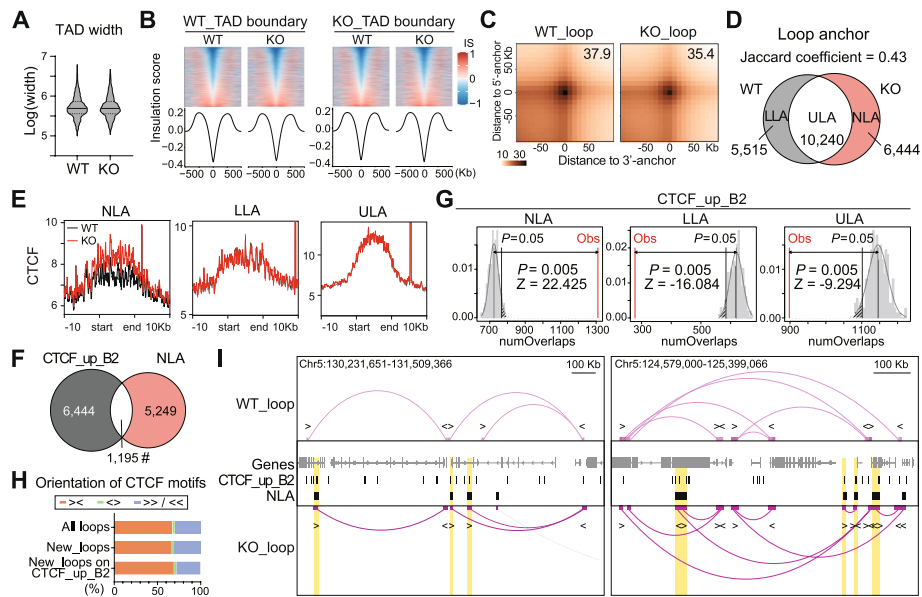
parallel, chromatin accessibility mirrored this trend (Fig. 4E). Interestingly, the repressive epigenetic markers showed a similar trend but exhibited distinct patterns. For H3K9me3, a notable enrichment (“peak”) was observed on CTCF\_up\_B2 in WT, which was almost completely diminished in KO (Fig. 4F). In contrast, DNA methylation was below the background (“valley”) on CTCF\_up\_B2 in WT, and the signal was further reduced in KO (Fig. 4G). These observations suggest that H3K9me3 and DNA methylation both contribute to the repression of CBS on SINE\_B2 elements.

CTCF is a highly conserved zinc finger protein and plays an important role in the regulation of gene transcription [34–36]. CTCF occupancy is broadly distributed across the genome. Density plot illustrated that a considerable proportion of CTCF peaks were located on gene promoters (around TSS), while the remaining peaks were concentrated around 16 Kb away from the TSS (Fig. 4I). We noticed that CTCF\_up peaks were almost exclusively absent from the promoter regions (Fig. 4I). This was in line with the fact that the observed CTCF\_up peaks were largely located on SINE\_B2 elements (Fig. 4B, Additional file 2: Fig. S4C–E), which showed a significant enrichment in non-promoter regions (Additional file 2: Fig. S4J), suggesting an indirect mechanism such as CTCF-mediated chromatin interactions in gene regulation.

#### Reorganization of chromatin loops on SINE\_B2 with increased CTCF binding after *Setdb1* ablation

In recent years, CTCF has been extensively studied for its pivotal role in maintaining 3D genome architecture through the mediation of chromatin folding [34–36]. Therefore, we explored the influences of genome-wide increased CTCF binding on chromatin organization in KO NPCs. We performed *in situ* Hi-C in biological triplicates and generated Hi-C contact maps with a resolution of < 15 Kb for individual samples, and < 5 Kb for merged samples by genotype (Additional file 2: Fig. S5A). There were no significant alterations in normalized contact maps (Additional file 2: Fig. S5B) or genome-wide relative contact probability between WT and KO (Additional file 2: Fig. S5C). Although CTCF has been recognized as an insulator in the formation of TAD boundaries [37, 38], our data did not reveal any significant changes in the number and width of TADs, the insulation score of TAD boundaries, or the interaction frequency between neighboring TADs (Fig. 5A–B, Additional file 2: Fig. S5D).

We then turned our focus to chromatin loops. While the number and span of loops remained largely unchanged (Additional file 2: Fig. S5E–F), we observed notable change in the aggregate contact intensity of loops, which was 37.9 in WT and 35.4 in KO (Fig. 5C). Additionally, we detected redistribution of loop anchors (Jaccard coefficient = 0.43), resulting in 6,444 new loop anchors (NLA), 5,515 lost loop anchors (LLA) and 10,240 unchanged loop anchors (ULA) in KO compared with WT (Fig. 5D). As CTCF is well-known to function as loop anchors [36], we investigated whether the redistribution of loop anchors arose from altered CTCF binding. Indeed, CTCF binding signal was increased on NLA (Fig. 5E, left), but remained unchanged on LLA (Fig. 5E, middle) and ULA (Fig. 5E, right). In addition, the average number of CTCF peaks was also increased on NLA (Additional file 2: Fig. S5G, left), but not on LLA (Additional file 2: Fig. S5G, middle) or ULA (Additional file 2: Fig. S5G, right). Notably, a significant portion (1,195 out of 6,444) of NLAs ( $P < 0.0001$ , fold enrichment = 1.85) overlapped with



**Fig. 5** Reorganization of chromatin loops on SINE\_B2 with increased CTCF binding after *Setdb1* ablation. **A** Violin plot displays width of topologically associating domains (TAD,  $N = 3,733$  WT/ $3,734$  KO) in WT and KO NPCs. **B** Insulation score heatmaps and profiles on WT (left) or KO (right) TAD boundaries. IS, insulation score. **C** Heatmaps show aggregate contact intensity of loops in WT and KO. WT\_loop,  $N = 9,286$ , left. KO\_loop,  $N = 9,955$ , right. Color scale indicates contact intensity. Numbers indicate contact intensity in the center. **D** Venn diagram displays overlap between WT and KO loop anchors. Anchors were categorized as “new\_loop\_anchor” (NLA, red), “lost\_loop\_anchor” (LLA, grey), and “unchanged\_loop\_anchor” (ULA, white). **E** Signal profiles of CTCF ChIP-seq on NLA (left), LLA (middle) and ULA (right). Black, WT; red, KO. **F** Venn diagram displays overlap between CTCF\_up\_B2 and NLA. Fisher’s exact test,  $^{\#}P < 0.0001$ . Fold enrichment = 1.85. **G** Permutation tests on the number of overlaps (numOverlaps) between CTCF\_up\_B2 and NLA (left), LLA (middle), and ULA (right). Black,  $P = 0.05$ . Red, observed (Obs). **H** Bar plot displays percentage of loops with convergent (><, orange), divergent (<>, green), or tandem (>> or <<, blue) CTCF motif orientation at their anchors. New\_loops, new loops in KO. **I** IGV map tracks show chromatin loops in WT and KO NPCs. Yellow shades highlight NLAs overlapping CTCF\_up\_B2. Arrows indicate the orientation of CTCF binding motifs

CTCF\_up\_B2 peaks (Fig. 5F, Additional file 1: Table S5), and this overlap exceeded expectations (Fig. 5G, left). Meanwhile, the overlaps between CTCF\_up\_B2 and LLA (Fig. 5G, middle) or ULA (Fig. 5G, right) significantly fell below anticipations. CTCF motif orientation is known to be critical for anchoring chromatin loops [39, 40]. We, therefore, checked the direction of CTCF motifs on the loops mediated by NLA (New\_loops,  $N = 7,020$ ). Consistent with the literature [39, 40], all loops, New\_loops, and those on CTCF\_up\_B2, were mainly (~70%) anchored on convergent (><) CTCF motifs (Fig. 5H). Visualization of the chromatin loops from WT and KO showed the overlap of NLA and CTCF\_up\_B2 peaks and the orientation of CTCF motifs, which were largely convergent on loop anchors. (Fig. 5I).

In conclusion, these findings suggest that although increased CTCF binding did not interfere with the conformation of TADs, it could contribute to the reorganization of chromatin loops due to the loss of *Setdb1* in NPCs.

#### Differential gene expression associated with loop reorganization after *Setdb1* ablation

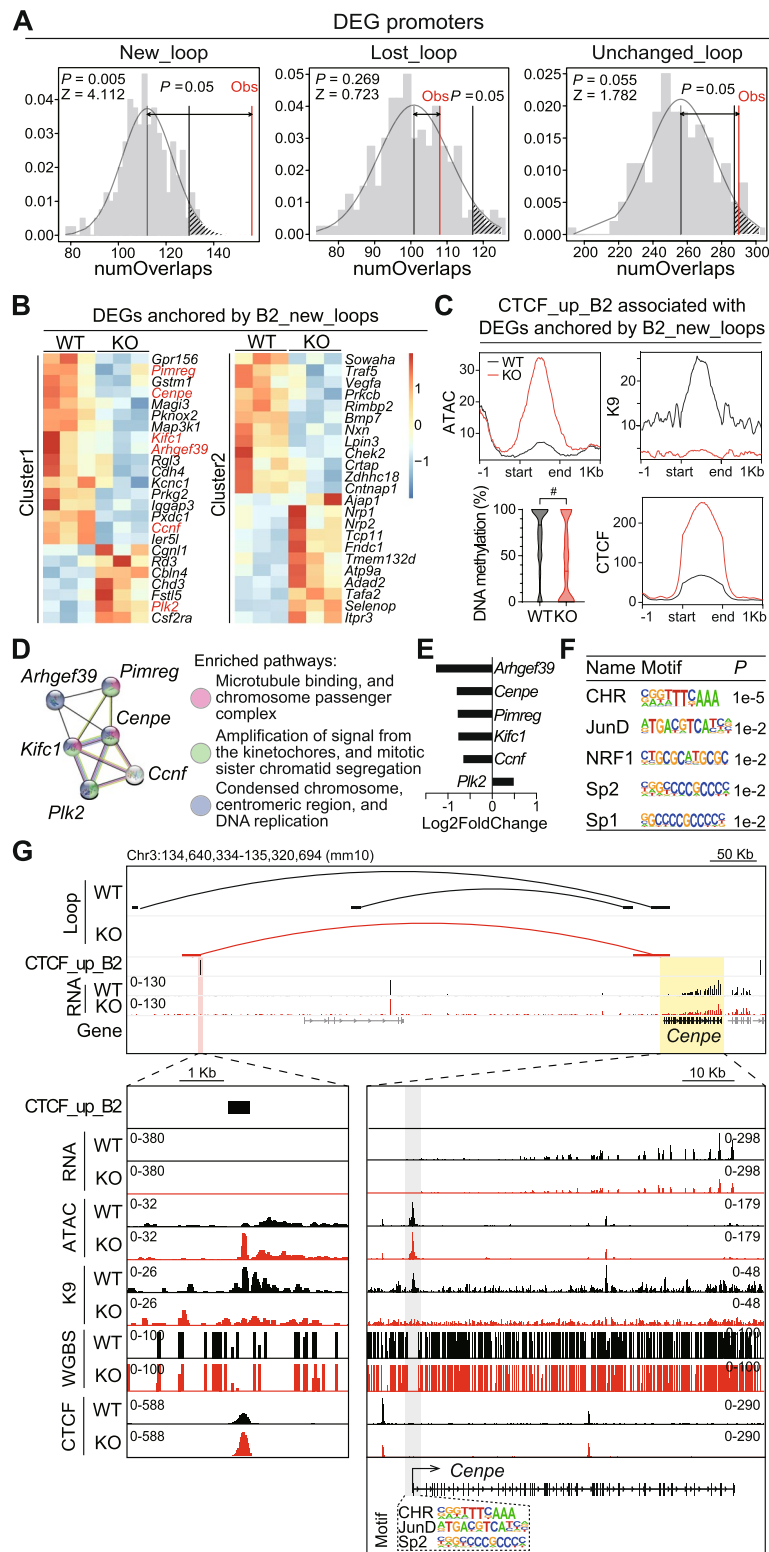
Chromatin loops, which bring *cis*-regulatory elements into close proximity with gene promoters, are known to regulate gene transcription [35]. To understand the transcriptional implications of reorganized chromatin loops following *Setdb1* loss in NPCs, we

reanalyzed our recently published RNA-seq data [18]. We noticed that *Setdb1* ablation in NPCs led to transcriptional dysregulation, including 120 differentially expressed genes (DEGs) that were enriched for cell cycle and mitosis pathways [18]. Interestingly, in our current study, we found that the promoters of DEGs significantly overlapped with the anchors of New\_loops (loops with at least one NLA), but not with Lost\_loops (loops with at least one LLA) or Unchanged\_loops in KO (Fig. 6A), suggesting these newly formed chromatin loops in KO might influence gene transcription. Next, we turned our focus to the DEGs anchored by New\_loops on CTCF\_up\_B2 elements (B2\_new\_loops) (Fig. 6B, Additional file 1: Table S6). Consistent with Fig. 4D-G, this population of CTCF\_up\_B2 elements exhibited increased chromatin accessibility, loss of H3K9me3, significant DNA hypomethylation, and increased CTCF bindings in KO (Fig. 6C). Furthermore, STRING analysis of these DEGs identified enriched network and pathways associated with mitotic chromatid segregation (Fig. 6D-E), and their promoters were enriched with CHR (the homology region for cell cycle gene promoters that recruited regulatory DREAM complex), and motifs of JunD, NRF1 and Sp transcription factors (TFs) (Fig. 6F). Interestingly, these TFs were reported as CTCF binding partners or co-regulatory factors [41–44], suggesting protein-protein interactions between CTCF and these TFs could contribute to the observed loop formation and regulation of mitosis genes in our study. Among these mitosis genes, *Cenpe* encodes a critical kinesin motor that contributes to microtubule-kinetochore interactions and spindle assembly checkpoint in replicating cells including NPCs [45–49]. As illustrated, new loops in KO connected the promoters and CTCF\_up\_B2 with altered epigenetic signatures and CTCF binding on *Cenpe* (Fig. 6G), as well as on other mitosis-related genes (Additional file 2: Fig. S6). These findings suggest that the newly formed chromatin loops in KO, resulting from excessive CTCF binding at B2 elements, contributed to the dysregulation of mitotic genes in *Setdb1*-knockout NPCs.

SETDB1 was reported in the regulation of cell division and proliferation in multiple systems, including germ cells [50], skeletal system [51] and cancer cells [52, 53]. In our current study, the dysregulated mitosis genes in KO suggested that SETDB1 may also regulate cell division in NPCs. Therefore, we assessed cell proliferation in both WT and KO NPCs. Our *in vitro* analysis of primarily cultured neurospheres showed a significant reduction in diameter for KO compared to WT (Fig. 7A). Cell cycle analysis with

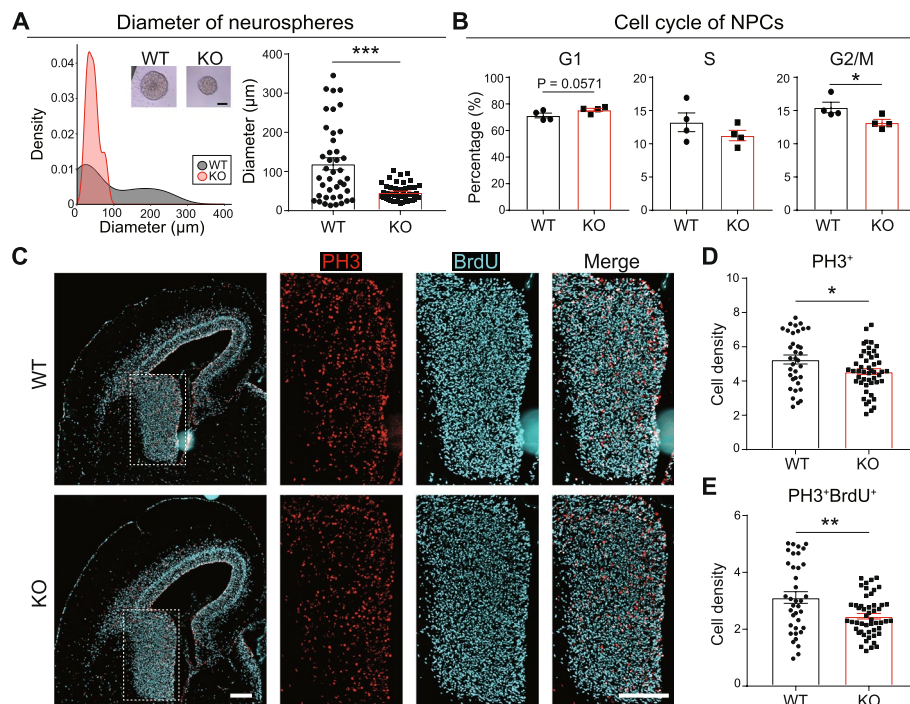
(See figure on next page.)

**Fig. 6** Differential gene expression associated with loop reorganization after *Setdb1* ablation. **A** Permutation tests on the number of overlaps (numOverlaps) between loops and promoters of differentially expressed genes (DEGs) from RNA-seq analysis. KO/WT.  $N = 3$ .  $P < 0.05$ . Loops were categorized as New\_loop (new in KO), Lost\_loop (lost in KO), and Unchanged\_loop (unchanged in KO). Black,  $P = 0.05$ . Red, observed (Obs). **B** Heatmaps display normalized expression and enriched clusters of DEGs anchored by New\_loops on B2 elements (B2\_new\_loops). Color scale indicates  $\log_{10}(\text{TPM}+1)$ . **C** Signal profiles of ATAC-seq (ATAC, upper left), H3K9me3 ChIP-seq (K9, upper right), and CTCF ChIP-seq (CTCF, lower right), as well as the level of DNA methylation (lower left) on CTCF\_up\_B2 elements that were associated with B2\_new\_loops and DEGs. Black, WT; red, KO. Two-tailed Mann-Whitney  $U$  test.  $^{\#}P < 0.0001$ . **D** STRING enriched functional network and pathways. **E** Bar graph shows fold changes of gene expression in KO vs. WT. **F** Top 5 enriched Homer known motifs on promoters of DEGs in (D-E). **G** Representative images show transcriptional and epigenetic signatures on *Cenpe* gene and its associated B2\_new\_loop. Red shade, CTCF\_up\_B2 element overlapping B2\_new\_loop in KO. Yellow shade, *Cenpe* gene. Grey shade, promoter of *Cenpe*. Dotted box, motifs on *Cenpe* promoter. RNA, RNA-seq; ATAC, ATAC-seq; K9, H3K9me3 ChIP-seq; CTCF, CTCF ChIP-seq



**Fig. 6** (See legend on previous page.)





**Fig. 7** Compromised NPC proliferation after *Setdb1* ablation. **A** Density plot (left) and bar graph (right) show diameter of NPC neurospheres.  $N = 41$  WT (black)/45 KO (red). Mean  $\pm$  SEM. Two-tailed unpaired *t*-test.  $***P < 0.001$ . Inset, representative images of WT and KO NPC neurospheres. Scale bar, 50  $\mu$ m. **B** Flow cytometry analysis of cell cycle for NPCs. Bar plots display the percentages of WT (black) and KO (red) NPC in G1 (left), S (middle), and G2/M (right) phases. Mean  $\pm$  SEM. Two-tailed Mann-Whitney *U* test.  $*P < 0.05$ . **C** Representative images for immunofluorescence staining of PH3 (red) and BrdU (cyan) in E15.5 brain slices derived from WT and KO mice. Scale bars, 200  $\mu$ m. **D-E** Bar graphs show density of PH3<sup>+</sup> (**D**) and PH3<sup>+</sup>BrdU<sup>+</sup> (**E**) cells in ganglionic eminence (white dotted box in **C**) from WT and KO.  $N = 36$  WT (black)/50 KO (red). Mean  $\pm$  SEM. Two-tailed Mann-Whitney *U* test.  $*P < 0.05$ ,  $**P < 0.01$

flow cytometry further revealed significantly decreased proportion of KO cells in G2/M phase, as well as a trend ( $P = 0.0571$ ) of increased proportion in G1 phase (Fig. 7B, Additional file 2: Fig. S7). For *in vivo* studies, we administered intraperitoneal injections of BrdU to pregnant mice at E15.5 and sacrificed them at E16.5. We examined the fetal brains for BrdU labeling (marker for S phase) and phospho-histone H3 (PH3) expression (marker for M phase) using immunofluorescence staining. We observed significant decreases in the number of PH3<sup>+</sup> and PH3<sup>+</sup>BrdU<sup>+</sup> cells in the GE of KO brains (Fig. 7C-E). Together, our findings from both *in vitro* and *in vivo* studies suggest slowing of cell cycle and NPC proliferation due to the loss of *Setdb1*.

## Discussion

In this study, we have elucidated that in mouse NPCs, SETDB1 represses the chromatin accessibility on a selective subset of SINE\_B2 elements, particularly the B3 and B3A subtypes. In the absence of SETDB1, increase of chromatin accessibility, loss of H3K9me3, and DNA hypomethylation happened on these B2 elements in NPCs. Consequently, CTCF occupancy was markedly increased, due to the increased availability of CBS harbored by B2 elements. Excessive CTCF binding led to the reorganization of chromatin loops and contributed to transcriptional dysregulation of mitosis genes, which

resulted in a prolonged cell cycle and compromised proliferation of NPCs in the *Setdb1* knockouts.

SETDB1 is a well-established epigenetic repressor that directly catalyzes H3K9me3 modification with its pre-SET, SET, and post-SET domains [54, 55]. On the other hand, SETDB1 also contains triple Tudor domains and MBD domain, through which it interacts with KAP1/KRAB-Zfp and MBD1/ATF7IP protein complex, respectively [55, 56]. Truncated splicing isoform of SETDB1, devoid of catalytic SET domains, was detected in mouse brain tissue [57]. Accordingly, non-enzymatic roles of SETDB1 in H3K9me3-independent pathways have been reported in controlling mouse ESC pluripotency and cell fate decisions. This includes its regulation of the activity of the Polycomb Repressive Complex 2 (PRC2) and the associated H3K27me3 modification [58], as well as its direct binding to cohesin [59], thereby contributing to gene regulation and genome organization. In our study, we ablated *Setdb1* in NPCs with conditional knockout mice, in which Cre recombinase driven by *Nes* promoter deleted the exon 3 of *Setdb1*, leading to frameshift and a premature stop codon that caused loss of Tudor, MBD and SET domains [27]. Therefore, in our KO model, we disrupted both enzymatic and non-enzymatic functions of SETDB1. Our current study focused on SETDB1-mediated H3K9me3 repression on SINE\_B2 activity; however, whether the non-enzymatic functions of SETDB1 also contribute to the regulation on SINEs would be important for future investigation.

The interplay between H3K9me3 and DNA methylation, another crucial repressive epigenetic modification, has been well studied. A complex interdependency exists between these two mechanisms; the loss of a critical methyltransferase in either one can impair the stability of the other [60, 61]. DNA methylation is often perceived as more stable and enduring, while H3K9me3 is considered more dynamic. Previous studies have demonstrated that these two markers can function either synergistically or sequentially to suppress ERVs during development [62, 63]. Conversely, these two marks have been reported to regulate different ERVs in ESCs [64]. Our study further elucidates the cooperation between DNA methylation and H3K9me3 in the developing brain. We found that SETDB1-mediated H3K9me3 is critical for maintaining genome-wide DNA methylation homeostasis in NPCs, and both contribute to the repression of SINE\_B2. However, this cooperative "double-lock" model may be specific to NPCs, as research in macrophages has ruled out the involvement of DNA methylation in the regulation of SINE\_B2 [20]. In addition, DNA methylation may be dispensable for the maintenance of H3K9me3 on SINE\_B2, as recent study in ESCs did not observe loss of H3K9me3 or increase of CTCF binding on SINE\_B2 after *Dnmt* triple ablation [65].

SETDB1 is well recognized as a master regulator of the LTR class of TEs, particularly as a suppressor of ERV transcription in ESCs [66, 67], and in differentiated brain cells, including NPCs and mature neurons [18, 24, 25]. However, less is known about the regulatory roles of SETDB1 on other classes of TEs. One study has uncovered the regulation of SETDB1 on SINEs in bone marrow-derived macrophages [20]. Our study extended its scope into the field of neurodevelopment. We found that SETDB1 is vital for repressing the non-LTR SINE class, particularly the B2 family of SINE. Interestingly, the affected B2 elements, compared to unaffected ones, were enriched with ATAC-seq and CTCF ChIP-seq signals and exhibited lower DNA methylation

levels in WT. This suggests that this population of B2 elements exhibits regulatory activity in NPCs at the baseline level. Furthermore, in our NPC system, the loss of SETDB1 only triggered an increase in chromatin accessibility, while the transcription of SINE\_B2 remained silenced. These findings suggest that this population of B2 elements plays a role in regulating NPCs via *cis*-regulation pathways, rather than through *trans*-regulation avenues.

SINE\_B2 elements are classified into five subtypes: B2\_Mm1a, B2\_Mm1t, B2\_Mm2, B3, and B3A, listed from evolutionarily younger to older [21]. An interesting finding in our study was the divergence in the regulation of SINE\_B2 subtypes by DNA methylation and H3K9me3. We discovered that the down-regulation of H3K9me3 was prevalent in the evolutionarily older B3/B3A, while DNA hypomethylation loci were more common in the evolutionarily younger B2 elements, B2\_Mm1a and B2\_Mm1t. This was unexpected, as DNA methylation, one of the most stable repressive epigenetic mechanisms, was anticipated to be primarily in charge of relatively ancient TEs during evolution. On the other hand, our findings support the notion that the B3/B3A loci, which are repressed by SETDB1-mediated H3K9me3, a reversible histone modification, are more open to regulation. Indeed, our study provided evidence that the loss of SETDB1 selectively opened the B3/B3A loci and possibly enabled CTCF binding. This aligns with previous research indicating that CTCF motifs were enriched in the evolutionarily older B3/B3A, rather than their younger counterparts [68]. Furthermore, a recent research confirmed that B3/B3A are also enriched for CBS that mediate conserved chromatin loops between mouse and human [69]. This suggests that these loci are not merely "ancient rubbish DNA" that must be suppressed from activation. Instead, they might be functionally critical, having survived natural selection, evolved into *bona fide* CBS, and contributed to the regulatory flexibility of the genome.

SETDB1 is conventionally recognized as a transcriptional repressor. However, our previous study has indicated that its function extends beyond merely facilitating H3K9me3 in gene regulation. In mature neurons, we have demonstrated that SETDB1 influences the transcription of clustered protocadherin genes by preserving the integrity of the chromatin domain at this locus [27]. Our current investigation further substantiates this role, providing additional insights into the involvement of SETDB1 in regulating higher-order chromatin conformation. We discovered that the absence of SETDB1 in NPCs leads to the emergence of newly accessible SINE\_B2 sites. These sites allow for CTCF binding, serving as chromatin loop anchors, and possibly triggering a reorganization of chromatin folding. Furthermore, our results suggest that the formation of chromatin loops primarily contributes to the adaptability of gene expression, rather than directly affecting transcriptional activity. It appears that the ultimate control over changes of gene expression is determined by TFs [70]. As neural differentiation progresses, SETDB1 expression declines, which coincides with NPCs exiting the cell cycle and ceasing proliferation [25]. Our study indicates that certain newly formed chromatin loops could enhance the regulatory flexibility of genes associated with mitotic cell division, thereby offering more opportunities for repressive TFs to inhibit their expression.

Despite the aforementioned insights, limitations exist within our current research. We still lack evidence for causal effects between regulatory elements and gene

expression, as our emphasis was laid on genome-wide effects, rather than select single gene or locus. On the other hand, more evidence is needed to clarify the mechanistic conservation between mice and human beings.

## Conclusions

Collectively, our data unravel the epigenetic signatures of a select population of SINE\_B2 elements in mouse NPCs. These SINE\_B2 elements, under the control of SETDB1-mediated H3K9me3, together with DNA methylation, contribute to the CTCF binding and chromatin loop formations, and regulate mitosis genes critical for NPC proliferation during mouse brain development.

## Methods

### Animals

All mice were housed 5 per cage, at 21–26 °C under 12/12-hour day/night cycle, and were allowed sterile food and water *ad libitum*.

### Neurosphere culture

Generation of knockout mice and neurosphere culture of neural precursor cells (NPCs) were conducted according to established protocol [18]. In brief, female *Setdb1*<sup>2lox/2lox</sup> mice were crossed with male *Setdb1*<sup>2lox/+</sup>, *Nestin-Cre* mice and checked for vaginal plug. At gestational day 15.5, mouse fetuses were dissected, genotyped and classified into wildtype (WT, *Setdb1*<sup>2lox/+</sup> or *Setdb1*<sup>2lox/2lox</sup>) and knockout (KO, *Setdb1*<sup>2lox/2lox</sup>, *Nestin-Cre*). Ganglionic eminences of fetal brains were harvested, dissociated with Accutase solution (Sigma A6964), and cultured in ultra-low attachment plates (Corning 3471) with advanced DMEM/F-12 (Gibco 12634-010) supplemented with N-2 (Gibco 17502048), B-27 (Gibco 12587010), GlutaMAX (Gibco 35050061), P/S (Sigma V900929), heparin (5 ug/ml, Sinopharm 63007101), EGF (20 ng/ml, Gibco PHG0314) and FGF (20 ng/ml, Gibco PHG0264). Fresh medium was added every 3 days, and neurospheres were dissociated with Accutase and passed every 5 days.

### ChIP-seq

All ChIP experiments were conducted in biological triplicates (3 KO vs. 3 WT) with published protocol [18, 27], with minor modifications. ~ 3 million NPCs were harvested, washed in PBS and lysed with NP40 lysis buffer (0.32 M sucrose, 5 mM CaCl<sub>2</sub>, 3 mM Mg(Ace)<sub>2</sub>, 0.1 mM EDTA, 10 mM Tris pH8, 0.1% NP40). For native ChIP (nChIP, for H3K9me3), nuclei were collected through centrifugation, resuspended in douncing buffer (10 mM Tris pH8, 4 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>) and chromatin was fragmented with micrococcal nuclease (Sigma N3755). Hypotonic solution (0.2 mM EDTA, cComplete inhibitor (Roche 11697498001), 0.1% NP40) was added for nuclei lysis and chromatin release. For fixed ChIP (xChIP, for CTCF), nuclei were collected through centrifugation, fixed in 1% formaldehyde (Sigma 252549) at room temperature (RT) for 5 minutes with rotation, and quenched in 0.125 M glycine at RT for 10 minutes with rotation. After collected with centrifugation, the fixed nuclei were resuspended in FSB solution (5 mM EDTA, 20 mM Tris pH8, 500 mM NaCl) supplemented with 0.1% SDS and cComplete inhibitor, and sonicated

with Covaris E220 ultrasonicator system to 300~500 bp. Next, for both nChIP and xChIP, samples were incubated with 4 µg antibody (anti-H3K9me3: Abcam ab8898, RRID: AB\_306848; anti-CTCF: Millipore 07-729, RRID: AB\_441965) in FSB solution (5 mM EDTA, 20 mM Tris pH8, 500 mM NaCl) supplemented with 0.1% NP40 and cOmplete inhibitor at 4 °C with overnight rotation. Protein A/G magnetic beads (Thermo 88803) were added to capture antibody-conjugated chromatin, collected on magnetic rack, washed sequentially in low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH8, 150mM NaCl), high salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH8, 500mM NaCl), lithium chloride buffer (1% NP40, 1% deoxycholic acid, 1 mM EDTA, 10 mM Tris pH8, 0.25 M LiCl) , twice in TE buffer (1 mM EDTA, 10 mM Tris pH8), and were eluted in 0.1 M NaHCO<sub>3</sub> and 1% SDS. The eluate was treated with 1 µl of 10 mg/ml RNase A (Sigma R6513) and 2 µl of 10 mg/ml proteinase K (Sigma P2308), and ChIP DNA was finally collected by ethanol precipitation. For library preparation, ChIP DNA was blunted with End-it DNA Repair Kit (Epicentre ER0720), A-tailed with Klenow Exo-minus (Epicentre KL06041K), ligated with adapter (Vazyme N802-01) using Fast-Link kit (Epicentre LK11025) and amplified with VAHTS HiFi Universal Amplification Mix for Illumina (Vazyme N618-02). The library was size-selected with SPRIselect beads (Beckman B23318) and sent for sequencing on Illumina NovaSeq 6000 (paired-end, 150 bp).

#### **Whole-genome bisulfite sequencing (WGBS)**

WGBS was conducted in biological duplicates (2 KO vs. 2 WT) according to published protocol [71], with minor modifications. ~ 2 million cultured cells were harvested, washed in PBS and lysed with NP40 lysis buffer (0.32 M sucrose, 5 mM CaCl<sub>2</sub>, 3 mM Mg(Ace)<sub>2</sub>, 0.1 mM EDTA, 10 mM Tris pH8, 0.1% NP40). Nuclei were collected with centrifugation, incubated with 5 µl of 10 mg/ml RNase A (Sigma R6513) in 500 µl PK lysis buffer (1 M Tris pH8, 50 mM EDTA, 2% SDS, 2 M NaCl) at 37 °C for 15 minutes, and then with 10 µl of 20 mg/ml proteinase K (Sigma P2308) at 52 °C overnight with rotation. The samples were then added with phenol/chloroform (1:1) (Solarbio P1021), mixed with vortex, and centrifugated with 15000 g at RT for 10 minutes. The supernatant was taken and precipitated with isopropanol. The precipitated genomic DNA (gDNA) was washed in 70% ethanol, air dried and dissolved in buffer EB (Qiagen 19086). The gDNA was then spiked in with 1% of unmethylated lambda DNA (Promega D1521) for efficiency validation of bisulfite conversion. 500 ng of gDNA was taken and sonicated to ~ 250 bp with Covaris M220 ultrasonicator. The DNA fragments were end-repaired and A-tailed with NEBNext Ultra II End Repair/dA-Tailing Module (NEB E7546), and then ligated with NEBNext Multiplex Oligos for Illumina (Methylated Adaptor, Index Primers Set 1, NEB E7535). After size-selected to 300~400 bp with Ampure XP beads (Beckman A63881), DNA was bisulfite converted with EZ DNA methylation kit (Zymo Research D5001), PCR amplified with KAPA HiFi HotStart Uracil+ ReadyMix (Kapa Biosystems KK2801), and the library was finally size-selected to 300~500 bp with Ampure XP beads. After analysis with Agilent TapeStation 4200, the library was sent for deep sequencing on Illumina NovaSeq 6000 (paired-end, 150 bp).



### Hi-C

*In situ* Hi-C was conducted in biological triplicates (3 KO vs. 3 WT) according to published protocol [27], with minor modifications. ~ 3 million cells were lysed in NP40 lysis buffer (0.32 M sucrose, 5 mM CaCl<sub>2</sub>, 3 mM Mg(Ace)<sub>2</sub>, 0.1 mM EDTA, 10 mM Tris pH8, 0.1% NP40), fixed in 1% formaldehyde (Sigma 252549) at RT for 5 minutes and quenched with 125 mM glycine at RT for 10 minutes. Nuclei were incubated in 0.5% SDS at 62 °C for 10 minutes, and quenched with 1.25% Triton X-100 (Sigma 93443) at 37 °C for 15 minutes. Chromatin was digested with 100 U of MboI restriction enzyme (NEB R0147) in 1X NEBuffer2 (NEB B7002S) at 37 °C overnight with rotation. Enzyme digestion was inactivated with incubation at 65 °C for 20 minutes, and nuclei were collected with centrifugation. Restriction fragment overhangs were filled and DNA ends were marked with biotin using End-it DNA Repair Kit (Epicentre ER0720) supplemented with 0.256 mM biotin-14-dATP (Life Technologies 19524-016), 0.4 mM dCTP, 0.4 mM dGTP and 0.4 mM dTTP (NEB N0446S). For proximity ligation, DNA fragments were ligated with 2000 U of T4 DNA ligase (NEB M0202) in 1200 µl of ligation master mix (T4 DNA ligase buffer (NEB B0202), 0.83% Triton X-100, 0.1 mg/ml bovine serum albumin (NEB B9200S)) at RT for 4 hours with slow rotation. After ligase inactivation at 65 °C for 5 minutes, nuclei were collected with centrifugation, and the chromatin was reverse-crosslinked in PK lysis buffer (1 M Tris pH8, 50 mM EDTA, 2% SDS, 2 M NaCl) at 65 °C for 4 hours, followed by RNase A (Sigma R6513) and proteinase K (Sigma P2308) digestion. DNA was further purified by phenol/chloroform (1:1) and ethanol precipitation, dissolved in buffer EB (Qiagen 19086) and sonicated to ~ 300 base pairs by Covaris E220 ultrasonicator system. For biotin pull-down, 100 µl per sample of 10 mg/ml Dynabeads MyOne Streptavidin T1 beads (Life technologies 65602) was washed with 400 µl per sample of 1 X TWB (5 mM Tris-HCl pH7.5, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween-20), resuspended in 100 µl per sample of 2X binding buffer (10 mM Tris-HCl pH7.5, 1 mM EDTA, 2 M NaCl) and incubated with samples at RT for 15 minutes with rotation. The beads were separated on magnetic stand, washed three times with 1 X TWB (pre-warmed to 55 °C) and once in 100 µl buffer EB (Qiagen 19086). For library preparation, the beads were collected for DNA end repair with End-it DNA Repair Kit, A-tailing with Klenow Exo-minus (Epicentre KL06041K), adapter (Vazyme N802-01) ligation with Fast-Link kit (Epicentre LK11025) and PCR amplification with VAHTS HiFi Universal Amplification Mix for Illumina (Vazyme N618-02). Supernatant was taken as the library, which was further size-selected with SPRIselect beads (Beckman B23318) and sent for deep sequencing on Illumina NovaSeq 6000 (paired-end, 150 bp).

### Cell cycle analysis

Cell cycle analysis was conducted in 4 biological replicates (4 KO vs. 4 WT) according to published protocol [72], with minor modifications. In brief, ~ 1 million NPCs were dissociated with Accutase solution (Sigma A6964), collected by centrifugation and resuspended in 0.5 mL of PBS (Biosharp BL302A), followed by gentle pipetting. For fixation, the resuspended cells were transferred into 4.5 mL of 70% ethanol and fixed overnight. After centrifugation, ethanol was decanted, and fixed cells were washed for 30 seconds in 5 mL of PBS, followed by centrifugation. Fixed cell pellet was resuspended in 1 mL of PI staining buffer (0.1% Triton X-100 (Sigma 93443), 10 µg/ml propidium iodide

(Sigma S7109), 100 µg/ml DNase-Free RNaseA (Sigma R6513) in PBS), incubated for 30 minutes at room temperature in darkness and transferred to flow cytometer (Beckman Cytoflex). PI fluorescence was excited by 488 nm laser, and emission was measured with 600 nm filter.

#### **BrdU labeling & immunofluorescence staining**

BrdU labeling and immunofluorescence staining was conducted in biological replicates (6 KO vs. 5 WT) according to published protocol [18, 73], with minor modifications. Female *Setdb1*<sup>2lox/2lox</sup> mice were crossed with male *Setdb1*<sup>2lox/+</sup>, *Nestin-Cre* mice and checked for vaginal plug. At gestational day 15.5, dams were injected with 50 mg/kg BrdU (Sigma B5002) intraperitoneally and sacrificed 24 hours post-injection. Fetus mice were dissected, genotyped, and classified into WT and KO. Fetus mice brains were fixed in 4% paraformaldehyde (Sigma P6148) at 4 °C for 6 hours, dehydrated in 30% sucrose (Sinopharm 10021463) at 4 °C overnight and embedded in optimal cutting temperature compound (Sakura 4583). Brains were sliced for 15 µm on cryostat (Leica CM1950), dried and preserved at -80 °C. For immunofluorescence staining, brain slices were dried at RT for 30 minutes, washed in PBS (Solarbio P1022) for 3 X 5 minutes, treated with 2 M HCl at RT for 30 minutes, washed in 0.1 M boric acid buffer (80 mM boric acid (Sinopharm 10004818) and 20 mM sodium borate (Sangon A500832)) for 3 X 5 minutes and in PBS for 5 minutes. For blocking, brain slices were incubated in blocking solution (5% goat serum (Thermo 16210064) + 0.5% Triton X-100 (Sigma 93443) in PBS) at RT for 30 minutes. For first antibody conjugation, brain slices were incubated in Anti-BrdU (Abcam ab6326, RRID: AB\_305426, 1:1000 dilution) and Anti-PH3 (Millipore 06-570, RRID: AB\_310177, 1:2000 dilution) diluted in blocking solution at 4 °C overnight, and then washed in PBS for 3 X 5 minutes. For second antibody conjugation, brain slices were incubated in Goat Anti-Rat 488 (Jackson Immunoresearch 112-545-003, RRID: AB\_2338351, 1:1000 dilution) or Goat Anti-Rabbit 594 (Jackson Immunoresearch 111-585-003, RRID: AB\_2338059, 1:1000 dilution) diluted in PBS at RT for 1 hour, and then washed in PBS for 3 X 5 minutes in darkness. Finally, brain slices were covered in mounting medium (Sigma F4680), and images were captured with Olympus VS120 virtual slide microscope. Manual cell counting was conducted with Imaris 9.5.

#### **Data analysis**

Statistical plots were generated with Graphpad Prism 9 and R package “ggplot2” [74] (version 3.4.2). Representative images were generated with IGV [75] (version 2.14.0).

#### **ATAC-seq & ChIP-seq**

ATAC-seq and ChIP-seq were analyzed according to published protocol [18], with minor modifications. ATAC-seq data were previously published (GSE186806) [18]. Quality control of ChIP-seq reads were conducted with FastQC [76] (version 0.11.9). Reads were trimmed with Trim Galore [77] (version 0.6.7, --quality 20 --phred33 --stringency 1 --length 20 --fastqc --paired) and aligned to mm10 with Bowtie2 [78] (version 2.4.1) in default settings. Reads were further binarized, sorted, deduplicated and indexed with SAMtools [79] (version 1.6). Bigwig files were generated with “bam-Coverage” function of Deeptools [80] (version 3.5.1, --binSize 10 --normalizeUsing

RPKM --effectiveGenomeSize 2652783500 --ignoreForNormalization chrX chrM). Narrow peaks were called with MACS2 [81] (version 2.2.7.1) (ATAC-seq: --shift -100 --extsize 200 --nomodel; ChIP-seq: default settings). Differential analysis ( $N = 3$ ) was conducted with R package “DiffBind” [82] (version 3.4.11). Peaks were counted with `summits=FALSE`, `bUseSummarizeOverlaps=TRUE`. Difference was reported with `method=DBA_DESEQ2`, `contrast=1`, `th=1`. Differential peaks were defined as  $FDR < 0.05$  &  $\text{Log}_2\text{FoldChange} \geq 0.585$  (up) or  $\leq -0.585$  (down). Genomic annotation of peaks was conducted with R package “ChIPseeker” [83] and “clusterProfiler” [84], with promoters defined as transcription start site (TSS)  $\pm 3$  Kb. For TE analysis, Bedtools [85] (version 2.30.0) was used to intersect differential peaks and RepeatMasker mm10 reference [86] (`-f 0.5 -F 0.5 -e`). For enrichment analysis of peak annotation, all detected peaks in respective datasets were used as background. For enrichment analysis of B2 genomic annotation, all SINE\_B2 elements were used as background. Profiles and heatmaps of ATAC-seq and ChIP-seq signals were generated with “computeMatrix scale-regions” (`-a 1000 -b 1000 --skipZeros`) and “plotHeatmap” (default settings) functions of Deeptools [80] (version 3.5.1). Motif analysis was conducted with “findMotifsGenome.pl” function of Homer [87] (version 4.11.1), with all WT peaks in respective datasets as background (`-bg`). Permutation tests were conducted with “permTest” function (`ntimes = 200`, `randomize.function = resampleRegions`, `universe = “all_loop_anchors”`, `evaluate.function = numOverlap` or `MeanDistance`) of R package “regionR” [88] (version 1.26.1).

### WGBS

WGBS reads were analyzed according to published protocol [71], with minor modifications. Briefly, WGBS reads were filtered and trimmed with “fastp” [89] (version 0.23.4, `-q 20 -u 50 -n 5`), and the quality control was conducted by FastQC [76] (version 0.11.9). High quality clean reads were aligned to mouse genome mm10 and lambda phage genome with Bowtie2 [78] (version 2.2.3) in default settings, and then deduplicated with “deduplicate\_bismark” function of Bismark [90] (version 0.22.1). Methylation measurement was extracted with “bismark\_methylation\_extractor” function of Bismark (`--no_overlap --ignore_r2 2 --comprehensive`), and the bisulfite conversion rate was calculated according to the methylation level of spiked-in lambda DNA. R package “bsseq” [91] (version 1.36.0) was used to identify differential methylated regions (DMRs) ( $N = 2$ ,  $ns=70$ ,  $h=1000$ ), and only CpGs covered by at least 2 reads in at least 1 sample per genotype group were retained. The final definition of DMRs is as follows: (1) t-statistic with a qcutoff between (0.01,0.990); (2) `maxGap = 300`; (3) contain at least 3 CpGs; (4) absolute mean methylation difference  $\geq 0.1$ ; (5) located on autosome. For enrichment analysis of hypomethylated DMRs, all DMRs were used as background. Genomic annotation, signal profile generation and motif analysis were conducted in the same methods as in ATAC-seq and ChIP-seq.

### Hi-C

Hi-C data were analyzed according to published protocol [27], with minor modifications. Briefly, quality control was conducted by FastQC (version 0.11.9) [76], and reads were trimmed with Trim Galore [77] (version 0.6.7, `--quality 20 --phred33 --stringency 1 --length 20 --fastqc --paired`). Restriction sites of MboI on mm10 were generated

with “hiccup\_digester” function of HiCUP [92] (version 0.8.0). High quality clean reads were aligned to mm10 with Bowtie2 program [78] (version 2.4.1) in HiCUP (max di-tag length: 700, min di-tag length: 50). “.ValidPairs” were generated with the “mapped\_2hic\_fragments.py” script of HiC-Pro [93] (version 2.11.4) in default settings. “.allValidPairs” were merged with “.ValidPairs” of samples with the same genotype (WT or KO), sorted, deduplicated and down-sampled to 1.038 billion for downstream comparison. “.hic” files were generated with “hicpro2juicebox.sh” utility of HiC-Pro and JuicerTools [94] (version 1.22.01). Contact maps were generated with Juicebox [95] (version 1.11.08). Raw and ICE-normalized matrices (iced\_matrix) were built with HiC-Pro in default settings (bin\_size: 10000 20000 40000 100000). Relative contact probability was calculated in bin\_size of 40 Kb with “RCP” function of R package “GENOVA” [96] (version 1.0.0.9000). Iced\_matrices were converted into “.h5” and “.cool” files with “hicConvertFormat” tool of HiCEXplorer [97] (version 3.7.1). Compartments were called in bin\_size of 100 Kb by “hicPCA” tool of HiCEXplorer in default settings, using bigwig file of WT ATAC-seq [18] to decide whether the eigenvector needs a sign flip. Insulation scores were calculated in bin\_size of 10 Kb with “insulation\_score” function (window = 32) of GENOVA, and TADs were called with “call\_TAD\_insulation” function of GENOVA in default settings. Insulation score heatmaps and profiles were generated with “tornado\_insulation” function (bed\_pos = ‘center’) of GENOVA. TAD+N analysis was conducted in bin\_size of 10 Kb with “intra\_inter\_TAD” function of GENOVA. Loops were called in bin\_size of 10 Kb and 20 Kb with “hicDetectLoops” tool of HiCEXplorer in default settings, and were merged with “hicMergeLoops” of HiCEXplorer. Aggregate contact intensity of loops was calculated with “APA” function of GENOVA in default settings. Profiles of CTCF ChIP-seq signal on loop anchors were generated with “computeMatrix scale-regions” (-a 10000 -b 10000 --regionBodyLength 10000 --skipZeros) and “plotHeatmap” (default settings) functions of DeepTools (version 3.5.1). To calculate the enrichment of CTCF ChIP-seq peak on loop anchors, each anchor was divided into 100 bins with “makewindow” function (-n 100 -i winnum) of Bedtools, and each bin was intersected with CTCF peaks in default settings with “intersect” function of Bedtools. Permutation tests were conducted with “permTest” function (ntimes = 200, randomize.function = resampleRegions, universe = “all\_loop\_anchors”, evaluate.function = numOverlap) of R package “regionR” [88] (version 1.26.1). Bedtools was used to intersect CTCF\_up\_B2 and NLA in default settings. CTCF motif orientation was defined according to CTCF.H12CORE.0.P.B of HOCOMOCO [98] by STORM program (--maxscore) of CREAD (version 0.84) [99].

### RNA-seq

RNA-seq data was analyzed according to published protocol [18], with minor modifications. Raw data were previously published (GSE186806) [18]. Quality control was conducted by FastQC [76] (version 0.11.9), and reads were trimmed with Trim Galore [77] (version 0.6.7, --quality 20 --phred33 --stringency 1 --length 20 --fastqc --paired). High quality clean reads were aligned to mm10 with HISAT2 [100] (version 2.2.1) in default settings. Reads were further binarized, sorted, and indexed with SAMtools (version 1.6). Reads on exons were counted with featureCounts [101] (version 2.0.3) (-p -t exon -g gene\_id -a gencode.vM20.annotation.gtf). Differential gene expression

analysis was conducted with DESeq2 [102] (version 1.34.0) in default settings, and differentially-expressed genes (DEGs) were defined as  $P < 0.05$ . Gene annotations were conducted with R package “biomaRt” [103] (version 2.50.3) (biomart = “ensembl”, version = 102, dataset = “mmusculus\_gene\_ensembl”), and promoters were defined as TSS  $\pm$  3 Kb. Permutation tests were conducted with “permTest” function (ntimes = 200, randomize.function = resampleRegions, universe = “all\_gene\_promoters”, evaluate.function = numOverlap) of R package “regioner” (version 1.26.1). Bedtools (version 2.30.0) was used to intersect DEG promoters and B2\_new\_loop in default settings. Protein-protein network analysis was conducted by STRING [104] (version 11.5) with  $k$ -means clustering ( $k = 2$ ).

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-024-03327-2>.

Additional file 1: Table S1. List of up-regulated NPC ATAC-seq peaks. Table S2. List of down-regulated NPC H3K9me3 ChIP-seq peaks. Table S3. List of DMRs in NPC WGBS. Table S4. List of up-regulated NPC CTCF ChIP-seq peaks. Table S5. List of chromatin loops with new loop anchors overlapping CTCF\_up\_B2. Table S6. List of differentially-expressed genes associated with B2\_new\_loop.

Additional file 2: Fig. S1. Increased chromatin accessibility on SINE\_B2 elements after *Setdb1* ablation in mouse neural precursor cells. Fig. S2. Reduction of H3K9me3 on SINE\_B2 after *Setdb1* ablation in NPCs. Fig. S3. Alterations of DNA methylation after *Setdb1* ablation in NPCs. Fig. S4. Increase of CTCF binding on SINE\_B2 after *Setdb1* ablation in NPCs. Fig. S5. Reorganization of chromatin loops on SINE\_B2 with increased CTCF binding after *Setdb1* ablation in NPCs. Fig. S6. Differential gene expression associated with loop reorganization after *Setdb1* ablation. Fig. S7. Compromised NPC proliferation after *Setdb1* ablation.

Additional file 3. Review history.

### Acknowledgements

We would like to thank Dr. Ying Zhu from Fudan University, and Dr. Ningyi Shao from University of Macau for their helpful advice and guidance in bioinformatic analysis.

### Peer review information

Tatiana Kutateladze and Wenjing She were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

### Review history

The review history is available as Additional file 3.

### Authors' contributions

Y.J. and W.M. conceived the ideas and supervised the research; D.S. and Y.Z. designed and performed the experiments; W.P. conducted most of the bioinformatic analysis; all the other authors contributed to experiment performance and data analysis; all authors contributed to scientific discussion and manuscript preparation.

### Funding

This work was supported by STI2030-Major Projects (2021ZD0203000) (Y.J.), and the National Natural Science Foundation of China (No.32170601, No. 81971272) (Y.J.).

### Availability of data and materials

The datasets generated and analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository under the accession number GSE247620 [105]. We used the previously published dataset (GSE186806) [18, 106]. No other scripts and software were used other than those mentioned in the Methods section.

### Declarations

#### Ethics approval and consent to participate

Mouse usage and experiments were authorized by Animal Care and Use Committee of Fudan University (FE21154).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.



Received: 8 March 2024 Accepted: 28 June 2024

Published online: 03 July 2024

**References**

- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860–921.
- Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, et al. Initial sequencing and comparative analysis of the mouse genome. *Nature*. 2002;420(6915):520–62.
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, et al. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet*. 2007;8(12):973–82.
- Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet*. 2017;18(2):71–86.
- Anania C. Transposable element evolution in mammals. *Nat Genet*. 2023;55(6):904.
- Osmanski AB, Paulat NS, Korstian J, Grimshaw JR, Halsey M, Sullivan KAM, et al. Insights into mammalian TE diversity through the curation of 248 genome assemblies. *Science*. 2023;380(6643):eabn1430.
- Reilly MT, Faulkner GJ, Dubnau J, Ponomarev I, Gage FH. The role of transposable elements in health and diseases of the central nervous system. *J Neurosci*. 2013;33(45):17577–86.
- Chesnokova E, Beletskiy A, Kolosov P. The role of transposable elements of the human genome in neuronal function and pathology. *Int J Mol Sci*. 2022;23(10).
- Jonsson ME, Garza R, Johansson PA, Jakobsson J. Transposable elements: a common feature of neurodevelopmental and neurodegenerative disorders. *Trends Genet*. 2020;36(8):610–23.
- Kramerov DA, Vassetzky NS. SINEs. *Wiley Interdiscip Rev RNA*. 2011;2(6):772–86.
- Elbarbary RA, Lucas BA, Maquat LE. Retrotransposons as regulators of gene expression. *Science*. 2016;351(6274):aac7247.
- Ahmadi A, De Toma I, Vilor-Tejedor N, Eftekhariyan Ghamsari MR, Sadeghi I. Transposable elements in brain health and disease. *Ageing Res Rev*. 2020;64:101153.
- Linker SB, Randolph-Moore L, Kottlilil K, Qiu F, Jaeger BN, Barron J, et al. Identification of bona fide B2 SINE retrotransposon transcription through single-nucleus RNA-seq of the mouse hippocampus. *Genome Res*. 2020;30(11):1643–54.
- Ferrigno O, Virolle T, Djabari Z, Ortonne JP, White RJ, Aberdam D. Transposable B2 SINE elements can provide mobile RNA polymerase II promoters. *Nat Genet*. 2001;28(1):77–81.
- Bartlett AA, Guffanti G, Hunter RG. B2 SINE RNA as a novel regulator of glucocorticoid receptor transcriptional activity. *Neurobiol Stress*. 2023;23:100522.
- Zovoilis A, Cifuentes-Rojas C, Chu H-P, Hernandez AJ, Lee JT. Destabilization of B2 RNA by EZH2 activates the stress response. *Cell*. 2016;167(7):1788–802.
- Luo X, Liu Y, Dang D, Hu T, Hou Y, Meng X, et al. 3D Genome of macaque fetal brain reveals evolutionary innovations during primate corticogenesis. *Cell*. 2021;184(3):723–40.
- Li J, Zheng S, Dong Y, Xu H, Zhu Y, Weng J, et al. Histone methyltransferase SETDB1 regulates the development of cortical Htr3a-positive interneurons and mood behaviors. *Biol Psychiatry*. 2023;93(3):279–90.
- Estéicio MRH, Gallegos J, Dekmezian M, Lu Y, Liang S, Issa JPJ. SINE retrotransposons cause epigenetic reprogramming of adjacent gene promoters. *Mol Cancer Res*. 2012;10(10):1332–42.
- Gualdrini F, Polletti S, Simonatto M, Prosperini E, Pileri F, Natoli G. H3K9 trimethylation in active chromatin restricts the usage of functional CTCF sites in SINE B2 repeats. *Genes Dev*. 2022;36(7–8):414–32.
- Kaaij LJT, Mohn F, van der Weide RH, de Wit E, Bühler M. The ChAHP complex counteracts chromatin looping at CTCF sites that emerged from SINE expansions in mouse. *Cell*. 2019;178(6):1437–51.
- Ostapczuk V, Mohn F, Carl SH, Basters A, Hess D, Iesmantavicius V, et al. Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature*. 2018;557(7707):739–43.
- Zhu Y, Sun D, Jakovcevski M, Jiang Y. Epigenetic mechanism of SETDB1 in brain: implications for neuropsychiatric disorders. *Transl Psychiatry*. 2020;10(1):115.
- Chandrasekaran S, Espeso-Gil S, Loh YHE, Javidfar B, Kassim B, Zhu Y, et al. Neuron-specific chromosomal megadomain organization is adaptive to recent retrotransposon expansions. *Nat Commun*. 2021;12(1):7243.
- Tan S-L, Nishi M, Ohtsuka T, Matsui T, Takemoto K, Kamio-Miura A, et al. Essential roles of the histone methyltransferase ESET in the epigenetic control of neural progenitor cells during development. *Development*. 2012;139(20):3806–16.
- Sun D, Weng J, Dong Y, Jiang Y. 3D genome organization in the central nervous system, implications for neuropsychological disorders. *J Genet Genomics*. 2021;48(12):1045–56.
- Jiang Y, Loh YE, Rajarajan P, Hirayama T, Liao W, Kassim BS, et al. The methyltransferase SETDB1 regulates a large neuron-specific topological chromatin domain. *Nat Genet*. 2017;49(8):1239–50.
- Wang Z, Fan R, Russo A, Cernilogar FM, Nuber A, Schirge S, et al. Dominant role of DNA methylation over H3K9me3 for IAP silencing in endoderm. *Nat Commun*. 2022;13(1):5447.
- Kato M, Takemoto K, Shinkai Y. A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing. *Nat Commun*. 2018;9(1):1683.
- Adoue V, Binet B, Malbec A, Fourquet J, Romagnoli P, van Meerwijk JPM, et al. The Histone Methyltransferase SETDB1 Controls T Helper Cell Lineage Integrity by Repressing Endogenous Retroviruses. *Immunity*. 2019;50(3).
- Cuellar TL, Herzner AM, Zhang X, Goyal Y, Watanabe C, Friedman BA, et al. Silencing of retrotransposons by SETDB1 inhibits the interferon response in acute myeloid leukemia. *J Cell Biol*. 2017;216(11):3535–49.
- Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, et al. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res*. 2012;22(9):1680–8.

33. Hashimoto H, Wang D, Horton JR, Zhang X, Corces VG, Cheng X. Structural basis for the versatile and methylation-dependent binding of CTCF to DNA. *Mol Cell*. 2017;66(5):711–20.e3.
34. Zheng H, Xie W. The role of 3D genome organization in development and cell differentiation. *Nat Rev Mol Cell Biol*. 2019;20(9):535–50.
35. Zuin J, Roth G, Zhan Y, Cramard J, Redolfi J, Piskadlo E, et al. Nonlinear control of transcription through enhancer-promoter interactions. *Nature*. 2022;604(7906):571–7.
36. Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, et al. CTCF-Mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell*. 2015;163(7):1611–27.
37. Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*. 2017;169(5):930–44.
38. Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007;129(4):823–37.
39. Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, et al. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell*. 2015;162(4):900–10.
40. Lu Y, Shou J, Jia Z, Wu Y, Li J, Guo Y, et al. Genetic evidence for asymmetric blocking of higher-order chromatin structure by CTCF/cohesin. *Protein Cell*. 2019;10(12):914–20.
41. Bohla D, Herold M, Panzer I, Buxa MK, Ali T, Demmers J, et al. A functional insulator screen identifies NURF and dREAM components to be required for enhancer-blocking. *PLoS One*. 2014;9(9):e107765.
42. Xu B, Wang H, Wright S, Hyle J, Zhang Y, Shao Y, et al. Acute depletion of CTCF rewires genome-wide chromatin accessibility. *Genome Biol*. 2021;22(1):244.
43. Akıncılar SC, Chua JYH, Ng QF, Chan CHT, Eslami SZ, Chen K, et al. Identification of mechanism of cancer-cell-specific reactivation of hTERT offers therapeutic opportunities for blocking telomerase specifically in human colorectal cancer. *Nucleic Acids Res*. 2023;51(1):1–16.
44. Kang Y, Hong JA, Chen GA, Nguyen DM, Schrupp DS. Dynamic transcriptional regulatory complexes including BORIS, CTCF and Sp1 modulate NY-ESO-1 expression in lung cancer cells. *Oncogene*. 2007;26(30):4394–403.
45. Yen TJ, Li G, Schaar BT, Szilak I, Cleveland DW. CENP-E is a putative kinetochore motor that accumulates just before mitosis. *Nature*. 1992;359(6395):536–9.
46. Wood KW, Sakowicz R, Goldstein LS, Cleveland DW. CENP-E is a plus end-directed kinetochore motor required for metaphase chromosome alignment. *Cell*. 1997;91(3):357–66.
47. Abrieu A, Kahana JA, Wood KW, Cleveland DW. CENP-E as an essential component of the mitotic checkpoint in vitro. *Cell*. 2000;102(6):817–26.
48. Yao X, Abrieu A, Zheng Y, Sullivan KF, Cleveland DW. CENP-E forms a link between attachment of spindle microtubules to kinetochores and the mitotic checkpoint. *Nat Cell Biol*. 2000;2(8):484–91.
49. Chen YJ, Friedman BA, Ha C, Durinck S, Liu J, Rubenstein JL, et al. Single-cell RNA sequencing identifies distinct mouse medial ganglionic eminence cell types. *Sci Rep*. 2017;7:45656.
50. Eymery A, Liu Z, Ozonov EA, Stadler MB, Peters AH. The methyltransferase Setdb1 is essential for meiosis and mitosis in mouse oocytes and early embryos. *Development*. 2016;143(15):2767–79.
51. Zhang L, Xu L, Zhang X, Wang K, Tan Y, Li G, et al. Methyltransferase Setdb1 Promotes Osteoblast Proliferation by Epigenetically Silencing Macrod2 with the Assistance of Atf7ip. *Cells*. 2022;11(16):2580.
52. Wang G, Long J, Gao Y, Zhang W, Han F, Xu C, et al. SETDB1-mediated methylation of Akt promotes its K63-linked ubiquitination and activation leading to tumorigenesis. *Nat Cell Biol*. 2019;21(2):214–25.
53. Wong CM, Wei L, Law CT, Ho DW, Tsang FH, Au SL, et al. Up-regulation of histone methyltransferase SETDB1 by multiple mechanisms in hepatocellular carcinoma promotes cancer metastasis. *Hepatology*. 2016;63(2):474–87.
54. Yang L, Xia L, Wu DY, Wang H, Chansky HA, Schubach WH, et al. Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. *Oncogene*. 2002;21(1):148–52.
55. Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ 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*. 2002;16(8):919–32.
56. Ichimura T, Watanabe S, Sakamoto Y, Aoto T, Fujita N, Nakao M. Transcriptional Repression and Heterochromatin Formation by MBD1 and MCAF/AM Family Proteins\*. *J Biol Chem*. 2005;280(14):13928–35.
57. Blackburn ML, Chansky HA, Zielinska-Kwiatkowska A, Matsui Y, Yang L. Genomic structure and expression of the mouse ESET gene encoding an ERG-associated histone methyltransferase with a SET domain. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*. 2003;1629(1):8–14.
58. Fei Q, Yang X, Jiang H, Wang Q, Yu Y, Yu Y, et al. SETDB1 modulates PRC2 activity at developmental genes independently of H3K9 trimethylation in mouse ES cells. *Genome Res*. 2015;25(9):1325–35.
59. Warrier T, El Farran C, Zeng Y, Ho BSQ, Bao Q, Zheng ZH, et al. SETDB1 acts as a topological accessory to Cohesin via an H3K9me3-independent, genomic shunt for regulating cell fates. *Nucleic Acids Res*. 2022;50(13):7326–49.
60. Sharif J, Endo TA, Nakayama M, Karimi MM, Shimada M, Katsuyama K, et al. Activation of Endogenous Retroviruses in Dnmt1(-/-) ESCs Involves Disruption of SETDB1-Mediated Repression by NP95 Binding to Hemimethylated DNA. *Cell Stem Cell*. 2016;19(1):81–94.
61. Leung D, Du T, Wagner U, Xie W, Lee AY, Goyal P, et al. Regulation of DNA methylation turnover at LTR retrotransposons and imprinted loci by the histone methyltransferase Setdb1. *Proc Natl Acad Sci U S A*. 2014;111(18):6690–5.
62. Mochizuki K, Sharif J, Shirane K, Uranishi K, Bogutz AB, Janssen SM, et al. Repression of germline genes by PRC1.6 and SETDB1 in the early embryo precedes DNA methylation-mediated silencing. *Nat Commun*. 2021;12(1):7020.
63. Liu S, Brind'Amour J, Karimi MM, Shirane K, Bogutz A, Lefebvre L, et al. Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. *Genes Dev*. 2014;28(18):2041–55.
64. Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, et al. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell*. 2011;8(6):676–87.

65. Tam PLF, Cheung MF, Chan LY, Leung D. Cell-type differential targeting of SETDB1 prevents aberrant CTCF binding, chromatin looping, and cis-regulatory interactions. *Nat Commun.* 2024;15(1):15.
66. Deniz Ö, de la Rica L, Cheng KCL, Spensberger D, Branco MR. SETDB1 prevents TET2-dependent activation of IAP retroelements in naïve embryonic stem cells. *Genome Biol.* 2018;19(1):6.
67. Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T, et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature.* 2010;463(7278):237–40.
68. Bourque G, Leong B, Vega VB, Chen X, Lee YL, Srinivasan KG, et al. Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res.* 2008;18(11):1752–62.
69. Choudhary MNK, Quaid K, Xing X, Schmidt H, Wang T. Widespread contribution of transposable elements to the rewiring of mammalian 3D genomes. *Nat Commun.* 2023;14(1):634.
70. Stadhouders R, Filion GJ, Graf T. Transcription factors and 3D genome conformation in cell-fate decisions. *Nature.* 2019;569(7756):345–54.
71. Liu P, Zhang J, Du D, Zhang D, Jin Z, Qiu W, et al. Altered DNA methylation pattern reveals epigenetic regulation of Hox genes in thoracic aortic dissection and serves as a biomarker in disease diagnosis. *Clin Epigenet.* 2021;13(1):124.
72. Hang H, Fox MH. Analysis of the mammalian cell cycle by flow cytometry. *Methods Mol Biol.* 2004;241:23–35.
73. Ligasová A, Frydrych I, Koberna K. Basic Methods of Cell Cycle Analysis. *Int J Mol Sci.* 2023;24(4):3674.
74. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*: Springer-Verlag New York; 2016.
75. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011;29(1):24–6.
76. Andrews S. FastQC. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 7 March 2024.
77. Krueger F. Trim Galore. <https://github.com/FelixKrueger/TrimGalore>. Accessed 7 March 2024.
78. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–9.
79. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience.* 2021;10(2):giab008.
80. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 2016;44(W1):W160–5.
81. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9(9):R137.
82. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature.* 2012;481(7381):389–93.
83. Yu G, Wang L-G, He Q-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics.* 2015;31(14):2382–3.
84. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284–7.
85. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26(6):841–2.
86. He X, Zhang XM, Wu J, Fu J, Mou L, Lu DH, et al. Olfactory experience modulates immature neuron development in postnatal and adult guinea pig piriform cortex. *Neuroscience.* 2014;259:101–12.
87. Benner C. HOMER. <http://homer.ucsd.edu/homer/>. Accessed 7 March 2024.
88. Gel B, Diez-Villanueva A, Serra E, Buschbeck M, Peinado MA, Malinverni R. regionR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics.* 2016;32(2):289–91.
89. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34(17):i884–90.
90. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics.* 2011;27(11):1571–2.
91. Hansen KD, Langmead B, Irizarry RA. BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biol.* 2012;13(10):R83.
92. Wingett S, Ewels P, Furlan-Magaril M, Nagano T, Schoenfelder S, Fraser P, et al. HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res.* 2015;4:1310.
93. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* 2015;16:259.
94. Robinson J DN, Aiden EL. JuicerTools. <https://github.com/aidenlab/juicertools>. Accessed 7 March 2024.
95. Robinson JT, Turner D, Durand NC, Thorvaldsdóttir H, Mesirov JP, Aiden EL. Juicebox.js provides a cloud-based visualization system for Hi-C data. *Cell Syst.* 2018;6(2):256–8.e1.
96. van der Weide RH, van den Brand T, Haarhuis JHI, Teunissen H, Rowland BD, de Wit E. Hi-C analyses with GENOVA: a case study with cohesin variants. *NAR Genom Bioinform.* 2021;3(2):lqab040.
97. Wolff J, Backofen R, Grüning B. Loop detection using Hi-C data with HiCExplorer. *Gigascience.* 2022;11:giac061.
98. Vorontsov IE, Eliseeva IA, Zinkevich A, Nikonov M, Abramov S, Boytsov A, et al. HOCOMOCO in 2024: a rebuild of the curated collection of binding models for human and mouse transcription factors. *Nucleic Acids Res.* 2024;52(D1):D154–63.
99. Schones DE, Smith AD, Zhang MQ. Statistical significance of cis-regulatory modules. *BMC Bioinformatics.* 2007;8:19.
100. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37(8):907–15.
101. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923–30.
102. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
103. Kinsella RJ, Kähäri A, Haider S, Zamora J, Proctor G, Spudich G, et al. Ensembl BioMart: a hub for data retrieval across taxonomic space. *Database (Oxford).* 2011;2011:bar030.

104. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 2023;51(D1):D638-d46.
105. Sun D, Zhu Y, Peng W, Zheng S, Weng J, Dong S, et al. Histone methyltransferase SETDB1 silenced SINE B2 retrotransposons and maintained mitotic cell cycle of neural progenitor cells. GSE247620. *Gene Expression Omnibus.* 2024. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE247620>.
106. Li J, Zheng S, Dong Y, Xu H, Zhu Y, Weng J, et al. Histone methyltransferase SETDB1 selectively regulates cortical HTR3A interneuron development and anxiety-like behavior in mice. GSE186806. *Gene Expression Omnibus.* 2022. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE186806>.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.