

SETDB1 Loss Dysregulates Translation-Related Pathways and Autism-Associated Gene Networks in Mouse Embryonic Stem Cells

Lixia Dong

Hefei Fourth People's Hospital

Yuanyuan Chen

Hefei Fourth People's Hospital

Cuizhen Zhu

zhucui zhen88@126.com

Hefei Fourth People's Hospital

Research Article

Keywords: SETDB1, H3K9 methylation, embryonic stem cells, heterochromatin, autism spectrum disorder, transcriptional regulation

Posted Date: September 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10947238/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. The authors declare that they have no conflicts of interest related to this study.

Article Title:

SETDB1 Loss Dysregulates Translation-Related Pathways and Autism-Associated Gene Networks in Mouse Embryonic Stem Cells

Authors:

Lixia Dong¹, Yuanyuan Chen², Cuizhen Zhu^{1*}

Affiliations:

¹ Brain Neurobiomedical Research Center, Hefei Fourth People's Hospital, Affiliated Psychological Hospital of Anhui Medical University, Hefei, 230022, China

² Psychiatry Department, Affiliated Psychological Hospital of Anhui Medical University, Hefei, 230022, China

* Corresponding author.

Corresponding Author:

Cuizhen Zhu

Brain Neurobiomedical Research Center, Hefei Fourth People's Hospital, Affiliated Psychological Hospital of Anhui Medical University, 316 Huangshan Road, Shushan District, Hefei, Anhui 230022, People's Republic of China

Email: zhucuizhen88@126.com

Manuscript Type: Research Article

Word Count: Abstract: 219 words; Main text (Introduction through Conclusion): approx. 3200 words

Number of: Main Figures: 6

Declarations:

Conflict of Interest: The authors declare no competing financial interests.

Funding: This work was supported by the Anhui Provincial Health Research Project (Province-Ministry Co-construction Project) Youth Project (grant No. AHWJ2024Aa30396) to L.D.; the Hefei Municipal Health Research Project for Applied Medical Science (grant No. Hwk2024zd002) to L.D.; the Scientific Research Project of Anhui Education Department (Natural Science) Youth Project (grant No. 2025AHGXZK40664) to L.D.; the National Clinical Key Specialty Construction Project of China (grant number: NO), Anhui Province Medical and Health Key Specialty Construction Project, Hefei 8th-cycle Key Clinical Specialty Funding(grant number: NO).

Declaration of Generative AI and AI-assisted technologies in the writing process: During the preparation of this work, the authors used various AI-assisted tools for language refinement, structural suggestions, and writing assistance. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the final content of the published article.

Data Availability: The RNA-seq data generated in this study have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE210606 and are publicly available as of the date of publication.

Submission Declaration: The authors declare that this manuscript is original, has not been published before, and is not currently being considered for publication elsewhere.

SETDB1 Loss Dysregulates Translation-Related Pathways and Autism-

Associated Gene Networks in Mouse Embryonic Stem Cells

Abstract

Autism spectrum disorder (ASD) has a strong genetic basis, yet the role of epigenetic regulators in its pathogenesis remains poorly understood. SETDB1, an H3K9 methyltransferase critical for heterochromatin silencing, has been implicated in neurodevelopment, but its impact on ASD-associated gene expression has not been systematically characterized. Here, we generated conditional *Setdb1* knockout mouse embryonic stem cells (mESCs) and found that SETDB1 loss causes cell death and loss of pluripotency. RNA-seq analysis revealed widespread transcriptional changes, with 1,886 upregulated and 2,170 downregulated genes ($|\log_2FC| \geq 1$, $p_{adj} < 0.05$). Notably, 266 high-confidence ASD risk genes were significantly enriched among these differentially expressed genes (hypergeometric test, $p < 0.001$), and these genes were primarily associated with synapse organization, chromatin remodeling, ion transport, and nervous system development. Gene set enrichment analysis uncovered a striking global downregulation of translation and ribosome biogenesis pathways ($NES < -1.8$, $FDR < 0.05$), suggesting that impaired protein synthesis contributes to the observed cell death. Protein-protein interaction network analysis identified a highly interconnected hub gene set (e.g., *Actb*, *Mapt*, *Syngap1*), most of which were upregulated in KO cells, whereas *Syp* was downregulated. Collectively, our results identify SETDB1 as an essential safeguard of mESC survival and a key regulator of ASD-linked gene networks, providing a valuable transcriptomic resource for understanding the epigenetic control of pathways relevant to neurodevelopmental disorders.

Keywords: SETDB1; H3K9 methylation; embryonic stem cells; heterochromatin; autism spectrum disorder; transcriptional regulation

1 Introduction

Autism spectrum disorder (ASD) is a heterogeneous group of neurodevelopmental conditions defined by persistent deficits in social communication and interaction, as well as restricted, repetitive patterns of behavior, interests, or activities¹. With a global prevalence exceeding 1%, ASD represents a major public health burden, and the demand for effective interventions continues to grow². Twin and family studies have consistently estimated its heritability at 70–90%, underscoring a strong genetic component³. Nevertheless, despite decades of genome-wide association studies and exome sequencing, known risk variants—including rare *de novo* mutations and common polymorphisms—collectively account for only a minority of cases^{4,5}. This “missing heritability” has driven a paradigm shift from single-gene causation toward the concept that convergent molecular pathways, rather than individual genes, underlie ASD pathogenesis⁶. Among the most consistently implicated pathways are synaptic transmission⁷, chromatin remodeling⁸, and transcriptional regulation⁹. Synaptic dysfunction has long been recognized as a core feature of ASD; for example, genes encoding neurexins (*NRXN1/2/3*), neuroligins (*NLGN1/2/3*)^{10,11}, and postsynaptic scaffolding proteins (*SHANK1/2/3*^{12,13}) are recurrently mutated in affected individuals, highlighting the centrality of synapse organization and plasticity. In parallel, a growing number of ASD risk genes encode proteins that regulate chromatin structure and gene expression, including *CHD8*,

EHMT1, *KDM6B*, *ADNP*, and *POGZ*^{14–17}. These observations have established epigenetic dysregulation as a key mechanistic theme in ASD, yet the precise role of many chromatin-associated factors—especially those involved in histone methylation—remains largely unexplored. Among these factors, SETDB1 (SET domain bifurcated histone lysine methyltransferase 1) stands out as a particularly compelling candidate, given its emerging links to neurodevelopment and its known function as a central H3K9 methyltransferase¹⁸.

SETDB1 is an H3K9 methyltransferase that catalyzes trimethylation of histone H3 at lysine 9 (H3K9me3). In the nervous system, SETDB1 regulates multiple aspects of neuronal development, including neural progenitor proliferation, neuronal differentiation, dendritic arborization, and synaptic plasticity^{19–21}. Conditional knockout of *Setdb1* in the mouse forebrain leads to severe cognitive deficits, abnormal social behavior, and repetitive movements, all of which are reminiscent of ASD-like phenotypes^{22,23}. Importantly, rare *de novo* missense mutations in *SETDB1* have been identified in individuals with developmental delay and autistic features, directly linking SETDB1 to neurodevelopmental disorders²⁴. Moreover, SETDB1 physically interacts with other ASD-associated chromatin factors such as ADNP and POGZ, suggesting a shared regulatory network²⁵. Despite these compelling links, the genome-wide transcriptional impact of SETDB1 loss on ASD risk genes has not been systematically studied, particularly in early stem cell states where heterochromatin is dynamically remodeled. Mouse embryonic stem cells (mESCs) provide a powerful model to investigate essential chromatin regulators because they are amenable to genetic manipulation and can be maintained in a pluripotent state, allowing the study of early events that may prime later neurodevelopmental defects²⁶.

SETDB1 is highly expressed in mESCs and is required for the repression of endogenous retroviruses and developmental genes, thereby safeguarding pluripotency^{27,28}. However, the direct consequences of SETDB1 ablation in mESCs—both for pluripotency maintenance and for the regulation of specific gene networks—have not been systematically defined. Moreover, whether SETDB1 targets in pluripotent cells intersect with ASD-associated pathways remains completely unexplored.

To address these gaps, we hypothesized that SETDB1 is essential for maintaining proper expression of specific gene networks in mESCs, and that its disruption could reveal pathways relevant to both stem cell biology and neurodevelopmental disorders. To test this hypothesis, we generated conditional *Setdb1* knockout mESCs using a FLP-FRT/Cre-loxP strategy, and performed high-throughput RNA-sequencing followed by comprehensive bioinformatics analyses, including differential expression, functional enrichment, gene set enrichment analysis (GSEA), protein-protein interaction (PPI) network construction, and hub gene identification. Our study provides a systematic transcriptomic resource for understanding SETDB1-dependent gene regulation in pluripotent stem cells and offers a foundation for exploring its potential intersection with ASD-associated pathways.

2 Materials and Methods

2.1 Cell culture

Mouse embryonic stem cells (mESCs) were cultured in high-glucose Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco), 2 mM L-glutamine (Gibco), 1%

non-essential amino acids (Gibco), 0.1 mM β -mercaptoethanol (Gibco), 100 U/mL penicillin-streptomycin (Gibco), and 1000 U/mL leukemia inhibitory factor (LIF). Cells were maintained at 37 °C in a humidified atmosphere containing 5% CO₂ and routinely tested for mycoplasma contamination²⁹.

2.2 Generation of conditional SETDB1 knockout mESCs

To delete exons 11–13 of the mouse *Setdb1* gene, we designed a conditional knockout strategy. A targeting vector containing loxP sites flanking exons 11–13 and an FRT-flanked neomycin resistance (Neo) cassette was constructed. The linearized vector was transfected into mESCs using Lipofectamine 2000 (Invitrogen), and homologous recombinants were selected with G418. Correct targeting was confirmed by long-range PCR and Southern blot. To excise the Neo cassette, we transiently expressed FLP recombinase by transfecting a pCAG-FLPe plasmid using Lipofectamine 2000, followed by negative selection. The resulting floxed (FF) allele carries exons 11–13 flanked by loxP sites. To generate the null (KO) allele, a Cre recombinase expression vector (pCAG-Cre) was transiently transfected using Lipofectamine 2000, and deletion of the floxed region was verified by PCR and Sanger sequencing.

All oligonucleotides were synthesized by GeneScript (Nanjing, China). The sgRNA sequences used for CRISPR/Cas9 targeting were TGGCATGCACAATCACCGCC (upstream) and AGTTATTATGAAGCCTAGGC (downstream). For genomic PCR genotyping, the forward primer was TTGCACTCACAGAAATCCAC and the reverse primer was GCTTAAACTGGATTCCAGTA. For RT-PCR validation of *Setdb1* expression, the forward primer was GATGACAAAAGATGTGAGTG and the reverse primer was

GTGGGCAGACACTCTTCTAG.

2.3 Morphological analysis and alkaline phosphatase staining

WT, FF, and KO mESCs were seeded on gelatin-coated 6-well plates at a density of 5×10^4 cells per well. After 48 h, bright-field images were captured using an Olympus IX73 inverted microscope. Alkaline phosphatase (AP) staining was performed using an Alkaline Phosphatase Detection Kit (Sigma-Aldrich) according to the manufacturer's instructions. Briefly, cells were fixed with 4% paraformaldehyde for 2 min and then incubated with AP staining solution for 15 min at room temperature. Stained colonies were imaged using a dissection microscope.

2.4 RNA extraction, library preparation, and sequencing

Total RNA was extracted from three independent WT and SETDB1 KO mESC cultures using TRIzol reagent (Invitrogen). RNA concentration and purity were measured using a NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific), and RNA integrity was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies); all samples had RNA Integrity Numbers (RIN) > 9.0. cDNA libraries were constructed and sequenced by Shanghai Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China). In brief, poly(A) mRNA was enriched from 1 μ g of total RNA using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB), and strand-specific RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit (NEB) following the manufacturer's instructions. The constructed libraries were sequenced on an Illumina NovaSeq 6000 platform with paired-end 150 bp (PE150) reads, generating approximately 30 million raw reads per sample.

2.5 Read alignment and differential expression analysis

Raw sequencing reads were filtered to remove adapter sequences and low-quality bases using fastp (v0.23.4). Clean reads were aligned to the mouse reference genome (GRCm39, GENCODE annotation M30) using STAR (v2.7.10a) with default parameters. Gene-level read counts were obtained using featureCounts (v2.0.3). Differential expression analysis was performed using PyDESeq2 (v0.4.0) with the Wald test; genes with $|\log_2 \text{ fold change}| \geq 1$ and Benjamini-Hochberg adjusted p-value (padj) < 0.05 were considered significantly differentially expressed. Principal component analysis (PCA) was performed on the $\log_2(\text{TPM}+1)$ -transformed expression values of the top 2,000 most variable genes (by variance), after standardization.

2.6 ASD risk gene list and overlap analysis

High-confidence ASD risk genes were retrieved from the Simons Foundation Autism Research Initiative (SFARI) Gene database (<https://gene.sfari.org/>). The overlap between SETDB1-regulated DEGs and ASD risk genes was assessed using a hypergeometric test, and the resulting overlapping gene set was used for subsequent functional and network analyses.

2.7 Gene Ontology and pathway enrichment analysis

Gene Ontology (GO) enrichment analysis (Biological Process, Cellular Component, Molecular Function) was performed using Metascape (v3.5) (<http://metascape.org>) with default parameters. Significantly enriched terms ($\text{padj} < 0.05$) were visualized as horizontal bar plots and bubble charts. In addition, gene set enrichment analysis (GSEA) was conducted using the gseapy Python package (v1.1.0) in prerank mode. The ranking metric was

defined as $\text{sign}(\log_2\text{FC}) \times (-\log_{10}(\text{pvalue}))$. Gene sets were obtained from the MSigDB Hallmark (H) and GO Biological Process (C5) collections. Statistical significance was set at $\text{FDR} < 0.05$.

2.8 Protein-protein interaction network construction and hub gene identification

Protein-protein interaction (PPI) networks were constructed using the STRING database (v12.0; <https://string-db.org/>) with a confidence score threshold of 0.4³⁰. Networks were imported into Cytoscape (v3.10.0) for visualization and analysis. Node degree (number of connections) was calculated using the NetworkAnalyzer plugin, and the top 10 genes ranked by degree were designated as hub genes. Expression changes ($\log_2\text{FC}$) were mapped to node colors (red: upregulation; blue: downregulation).

2.9 Statistical analysis and data visualization

All statistical analyses were performed using Python (v3.9) or R (v4.5). Comparisons between two groups were conducted using two-tailed Student's t-test for normally distributed data (assessed by Shapiro-Wilk test) or the Mann-Whitney U test for non-normally distributed data. Multiple testing correction was performed using the Benjamini-Hochberg method. Data are presented as mean $\pm$ standard deviation (SD) unless otherwise stated. Graphs were generated using matplotlib (v3.7) and seaborn (v0.12) in Python. All code is available upon request.

3 Results

3.1 Generation of conditional SETDB1 knockout mESCs

To investigate the biological function of SETDB1, we generated conditional *Setdb1* knockout mESCs using a FLP-FRT/Cre-loxP strategy (Fig. 1A). In the targeting construct, loxP sites were inserted flanking exons 11–13 of the mouse *Setdb1* locus, together with an FRT-flanked neomycin resistance cassette for positive selection. After homologous recombination, the Neo cassette was removed by transient FLP recombinase expression, yielding the floxed (FF) allele. Subsequent Cre-mediated recombination excised exons 11–13, generating a null (KO) allele predicted to cause a frameshift and premature translation termination.

WT, FF, and KO clonal lines were successfully obtained and verified by PCR genotyping and Sanger sequencing. Morphologically, WT and FF cells formed compact, domed colonies with smooth borders, characteristic of undifferentiated mESCs (Fig. 1B, top). In contrast, KO clones exhibited flattened, spread-out morphology, reduced cell density, and increased numbers of floating dead cells, suggesting compromised viability. Alkaline phosphatase (AP) staining, a well-established marker of pluripotency, was strongly positive in WT and FF colonies but markedly diminished in KO clones (Fig. 1B, bottom). Collectively, these results demonstrate that SETDB1 is essential for maintaining normal mESC colony morphology, pluripotency, and cell survival.

3.2 SETDB1 KO induces global transcriptomic changes

To characterize the molecular consequences of SETDB1 loss, we performed RNA-seq on three independent WT and KO mESC cultures. Principal component analysis (PCA) based on the 2,000 most variable genes ($\log_2(\text{TPM}+1)$ -transformed) revealed that PC1 accounted for 97.3% of the

total variance, clearly separating WT from KO samples (Fig. 2A). A kernel density plot of global expression showed a leftward shift in KO cells (Fig. 2B), indicating a genome-wide reduction in transcript abundance.

Differential expression analysis identified 4,056 significantly altered genes ($|\log_2FC| \geq 1$, $p_{adj} < 0.05$), of which 1,886 were upregulated and 2,170 were downregulated in KO cells (Fig. 2C). Gene Ontology (GO) enrichment analysis of all differentially expressed genes (DEGs) revealed significant enrichment of terms related to nervous system development, cell morphogenesis, postsynapse, axon development, chromatin remodeling, and regulation of membrane potential (Fig. 2D). Notably, many of these categories are associated with neurodevelopment and synaptic function, suggesting that SETDB1 may help maintain a poised transcriptional state for neuronal genes even in undifferentiated pluripotent stem cells.

3.3 GSEA reveals global downregulation of translation and ribosome biogenesis

Given the observed global reduction in transcript abundance and the compromised cell viability in KO cells, we performed gene set enrichment analysis (GSEA) using the ranking metric $\text{sign}(\log_2FC) \times (-\log_{10}(pvalue))$. Strikingly, multiple significantly negatively enriched gene sets were centered around protein synthesis and ribonucleoprotein complex assembly (Fig. 3). Fundamental biological processes such as *translation* (NES = -2.52, FDR < 0.001), *cytoplasmic translation* (NES = -2.46, FDR < 0.001), *ribosomal small subunit biogenesis* (NES = -2.02, FDR < 0.001), and *ribonucleoprotein complex biogenesis* (NES = -2.04, FDR < 0.001) were all profoundly suppressed in KO cells. Additionally, broader categories including

macromolecule biosynthetic process and peptide biosynthetic process also exhibited strong negative enrichment. Collectively, these results suggest that SETDB1 loss leads to a marked inhibition of the protein synthesis machinery, which may contribute to the observed cell death and loss of pluripotency.

3.4 SETDB1 KO significantly dysregulates ASD risk genes

We next examined whether SETDB1-regulated genes overlap with known ASD risk genes. From the SFARI Gene database, we retrieved 1,268 high-confidence ASD risk genes. Among the 4,056 DEGs, 266 were identified as ASD risk genes, representing a highly significant enrichment (hypergeometric test, $p < 0.001$; Fig. 4A). GSEA of this gene set did not reach statistical significance (NES = -1.100 , FDR = 0.228 ; Fig. 4B), although the overall trend toward downregulation was consistent with the global expression shift observed in Fig. 2B.

To identify individual ASD genes most affected by SETDB1 loss, we generated a volcano plot highlighting the 266 overlapping genes (Fig. 4C). Notably, the most upregulated genes included the potassium channel genes *Kcnq2* ($\log_2\text{FC} = 2.4$, $\text{padj} = 1.2 \times 10^{-12}$) and *Kcnq3* ($\log_2\text{FC} = 1.9$, $\text{padj} = 3.4 \times 10^{-10}$), as well as *Cntnap2* ($\log_2\text{FC} = 1.7$, $\text{padj} = 2.1 \times 10^{-8}$) and *Scn8a* ($\log_2\text{FC} = 1.5$, $\text{padj} = 5.6 \times 10^{-7}$). Conversely, the most downregulated ASD genes included *Foxp1* ($\log_2\text{FC} = -1.6$, $\text{padj} = 1.1 \times 10^{-6}$) and *Syp* ($\log_2\text{FC} = -1.2$, $\text{padj} = 3.3 \times 10^{-5}$). These data demonstrate that SETDB1 loss causes bidirectional expression changes in ASD-associated genes, with a prominent upregulation of ion channel genes. A heatmap of the 266 overlapping genes ($\log_2(\text{TPM}+1)$, row-wise Z-score) clearly separated WT from KO samples and revealed distinct subclusters of co-regulated genes (Fig. 4D), further confirming the widespread regulatory impact of SETDB1.

Finally, module-based GO analysis grouped the enriched terms into three functional modules: Synaptic Function (e.g., “postsynapse”, “regulation of synaptic plasticity”), Neuronal Development (e.g., “axon guidance”, “brain development”), and Chromatin Regulation (e.g., “chromatin remodeling”, “histone modification”) (Fig. 4E). Collectively, these results indicate that SETDB1-sensitive ASD genes are primarily involved in synapse organization, neuronal differentiation, and epigenetic regulation.

3.5 PPI network and hub genes associated with SETDB1 KO

To identify core genes that may mediate the downstream effects of SETDB1 loss, we constructed a protein-protein interaction (PPI) network using the 266 ASD-related DEGs and STRING database (confidence score ≥ 0.4). The network comprised 234 nodes and 1,527 edges. Node degree (number of interacting partners) was highly variable, and we selected the top 50 hub genes based on degree (Fig. 5A). The network displayed several densely interconnected clusters, with major hubs including *Actb* (degree = 29), *Mapt* (degree = 26), *Syngap1* (degree = 24), *Dlg2* (degree = 22), *Nrxn3* (degree = 20), *Cntnap2* (degree = 19), *Kcnq2* (degree = 18), and *Scn8a* (degree = 17). Node colors were mapped to $\log_2FC$, revealing that the majority of hub genes were upregulated (red) in KO cells.

Ranking all genes by degree identified the top 10 hub genes: *Actb*, *Syp*, *Nrxn3*, *Cntnap2*, *Camk2a*, *Dlg2*, *Mapt*, *Syngap1*, *Kcnq2*, and *Scn8a* (Fig. 5B). All except *Syp* (downregulated) were upregulated. To visualize these expression changes, we extracted the TPM values for these genes and plotted their WT and KO levels ($\log_2(TPM+1)$, mean $\pm$ SD; Fig. 5C). Statistical testing confirmed that the upregulation of *Actb*, *Nrxn3*, *Cntnap2*, *Camk2a*, *Dlg2*,

Mapt, *Syngap1*, *Kcnq2*, and *Scn8a* was highly significant ($p < 0.001$), whereas the downregulation of *Syp* was also significant ($p = 0.003$). These results establish that SETDB1 loss causes robust changes in a core set of ASD-related hub genes that are central to synapse function and neuronal excitability.

Based on these findings, we summarize a mechanistic model of the dual consequences of SETDB1 loss (Fig. 6). In wild-type mESCs, SETDB1 maintains translation and ribosome biogenesis while keeping ASD-associated genes (e.g., *Kcnq2*, *Cntnap2*, *Nrxn3*) at basal levels, supporting cell survival and pluripotency. Upon SETDB1 knockout, translation and ribosome biogenesis are globally suppressed, and ASD-related genes become markedly upregulated, contributing to cell death and loss of pluripotency.

4 Discussion

In this study, we provide the first comprehensive transcriptomic analysis of SETDB1 knockout mESCs and its impact on ASD-associated gene networks. Our findings demonstrate that SETDB1 is essential for mESC survival and pluripotency, and that its loss leads to widespread transcriptional alterations, including global suppression of translation and ribosome biogenesis, alongside significant upregulation of ASD-related ion channel and synaptic genes. Through PPI network analysis, we identified a core hub gene set—including *Actb*, *Nrxn3*, *Cntnap2*, and *Kcnq2*—that is highly interconnected and predominantly upregulated in KO cells. These results establish SETDB1 as a critical regulator of mESC transcriptome homeostasis and suggest a potential link between SETDB1-dependent heterochromatin silencing and gene networks implicated in neurodevelopmental disorders.

The morphological changes and loss of alkaline phosphatase activity in SETDB1 KO mESCs indicate that SETDB1 is required for maintaining the undifferentiated state. This finding aligns with previous reports that SETDB1 safeguards embryonic stem cell (ESC) identity by silencing developmental genes and transposable elements^{29,31,32}. However, the rapid cell death observed in our KO cells suggests that SETDB1 also regulates genes essential for cell viability beyond pluripotency maintenance. The strong downregulation of translation and ribosome biogenesis pathways (Fig. 3) provides a plausible mechanistic explanation: insufficient protein synthesis would be expected to lead to cell cycle arrest and apoptosis. Indeed, similar phenotypes have been reported for other chromatin regulators such as EHMT1 and G9a, where impaired translation or ribosome function triggers cell death^{33,34}. Collectively, these observations suggest that SETDB1 likely coordinates both pluripotency and survival through distinct downstream programs.

The PPI network analysis identified a set of hub genes that are highly interconnected and largely upregulated in KO cells. Among them, *Cntnap2* and *Nrxn3* are well-established ASD risk genes encoding neurexins and their ligands, which lead to autism-like phenotypes³⁵⁻³⁷. Their upregulation in KO cells is noteworthy, as loss-of-function mutations in these genes generally reduce synapse formation^{38,39}. Our data suggest that SETDB1 normally represses these genes; its absence leads to overexpression, which could disrupt synaptic stoichiometry. Similarly, *Kcnq2* and *Scn8a* are voltage-gated channel genes linked to epileptic encephalopathies and ASD^{40,41}. Their overexpression may contribute to aberrant membrane excitability⁴². The downregulation of *Syp* (synaptophysin) in KO cells further supports the

notion that SETDB1 is required for maintaining balanced expression of synaptic proteins⁴³.

Several limitations of this study should be acknowledged. First, our experiments were performed in mESCs, which do not fully recapitulate the complexity of mature neurons. While the enrichment of neuronal genes in our GO analysis indicates that SETDB1 regulates a neuronal program even in pluripotent cells, validation in neural progenitors and differentiated neurons will be essential to confirm the relevance of our findings to neurodevelopment. Second, we did not perform ChIP-seq for H3K9me3 or SETDB1 binding, so we cannot distinguish direct from indirect transcriptional targets. Future genome-wide chromatin occupancy studies would help clarify the mechanisms by which SETDB1 regulates these genes. Third, the GSEA of the ASD gene set did not reach statistical significance (FDR = 0.228), which may reflect the heterogeneity of ASD genes, as some are induced while others are repressed in KO cells. Finally, our study does not include functional rescue experiments (e.g., overexpression or knockdown of *Kcnq2* in KO cells) to establish causality. Despite these limitations, our results provide a strong foundation for future mechanistic studies on SETDB1-mediated gene regulation in pluripotent stem cells and its potential relevance to neurodevelopmental disorders.

Conclusions

In summary, our study demonstrates that SETDB1 is essential for mESC survival and for maintaining proper expression of a large cohort of ASD-associated genes, particularly those involved in synaptic function, ion transport, and chromatin regulation. The global suppression of translation and the prominent upregulation of ion channel genes suggest a dual consequence

of SETDB1 loss—impaired protein synthesis and dysregulated expression of neurodevelopmentally relevant gene networks—that together may contribute to cell death and, by extension, to potential neuronal dysfunction upon differentiation. Our findings highlight SETDB1 as a critical regulator of gene networks linked to ASD in pluripotent stem cells and provide a valuable transcriptomic resource for understanding the epigenetic control of pathways relevant to neurodevelopmental disorders.

5 Author contributions

Lixia Dong: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Software, Visualization, Writing – original draft, Writing – review & editing, Funding acquisition, Project administration. Yuanyuan Chen: Investigation, Formal analysis, Data curation, Methodology, Validation, Visualization. Cuizhen Zhu: Conceptualization, Supervision, Resources, Writing – original draft, Writing – review & editing, Funding acquisition, Project administration. All authors have read and approved the final manuscript and agree to be accountable for all aspects of the work.

6 Acknowledgements

We are grateful to all members of the laboratory. This work was supported by the Anhui Provincial Health Research Project (Province-Ministry Co-construction Project) Youth Project (grant No. AHWJ2024Aa30396) to L.D.; the Hefei Municipal Health Research Project for Applied Medical Science (grant No. Hwk2024zd002) to L.D.; the Scientific Research Project of Anhui Education Department (Natural Science) Youth Project (grant No. 2025AHGXZK40664) to L.D.; the National Clinical Key Specialty

Construction Project of China (grant number: NO), Anhui Province Medical and Health Key Specialty Construction Project, Hefei 8th-cycle Key Clinical Specialty Funding(grant number: NO).

7 Conflict of interest

The authors declare no competing financial interests.

8 Data availability

The RNA-seq data generated in this study have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE210606 and are publicly available as of the date of publication. This paper does not report original code. All other data are available from the corresponding author upon reasonable request.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used AI-assisted tools for language refinement, structural suggestions, and writing assistance. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the final content of the published article.

Reference

1. Pruitt A, Gupta AR, Hoffman EJ. Molecular and Genetic Mechanisms in Autism Spectrum Disorder. *Ann Neurol.* 2025;98(6):1163-1177. doi:10.1002/ana.70013
2. Salari N, Rasoulpoor S, Rasoulpoor S, et al. The global prevalence of autism spectrum disorder: a comprehensive systematic review and meta-analysis. *Ital J Pediatr.* 2022;48(1):112. doi:10.1186/s13052-022-01310-w

3. Tick B, Bolton P, Happé F, Rutter M, Rijdsdijk F. Heritability of autism spectrum disorders: a meta-analysis of twin studies. *J Child Psychol Psychiatry*. 2016;57(5):585-595. doi:10.1111/jcpp.12499
4. Pugsley K, Scherer SW, Bellgrove MA, Hawi Z. Environmental exposures associated with elevated risk for autism spectrum disorder may augment the burden of deleterious de novo mutations among probands. *Mol Psychiatry*. 2022;27(1):710-730. doi:10.1038/s41380-021-01142-w
5. Wang T, Guo H, Xiong B, et al. De novo genic mutations among a Chinese autism spectrum disorder cohort. *Nat Commun*. 2016;7:13316. doi:10.1038/ncomms13316
6. Ruzzo EK, Pérez-Cano L, Jung JY, et al. Inherited and De Novo Genetic Risk for Autism Impacts Shared Networks. *Cell*. 2019;178(4):850-866.e26. doi:10.1016/j.cell.2019.07.015
7. Vilela J, Pereira AC, Violante IR, et al. Neurometabolic profiles of autism spectrum disorder patients with genetic variants in specific neurotransmission and synaptic genes. *Sci Rep*. 2025;15(1):36318. doi:10.1038/s41598-025-20090-x
8. Wang X, Zeng C, Wu Z, et al. Chromatin remodeling factor BAF155 coordinates oligodendroglial-neuronal communications linked to regional myelination and autism-like behavioral deficits in mice. *Nat Commun*. 2025;17(1):1165. doi:10.1038/s41467-025-67930-y
9. Fazel Darbandi S, An JY, Lim K, et al. Five autism-associated transcriptional regulators target shared loci proximal to brain-expressed genes. *Cell Rep*. 2024;43(6):114329. doi:10.1016/j.celrep.2024.114329
10. Gerik-Celebi HB, Bolat H, Unsel-Bolat G. Rare heterozygous genetic variants of NRXN and NLGN gene families involved in synaptic function and their association with neurodevelopmental disorders. *Dev Neurobiol*. 2024;84(3):158-168. doi:10.1002/dneu.22941
11. Roddick KM, Habib EB, Brown RE, Balci F. Interval Timing Is Altered in Male *Nrxn1*^{+/-} Mice: A Model of Autism Spectrum Disorder. *Autism Res*. 2026;19(5):e70226. doi:10.1002/aur.70226
12. Jung S, Park M. Shank postsynaptic scaffolding proteins in autism spectrum disorder: Mouse models and their dysfunctions in behaviors, synapses, and molecules. *Pharmacol Res*. 2022;182:106340. doi:10.1016/j.phrs.2022.106340

13. Lu X, Ni P, Suarez-Meade P, et al. Transcriptional determinism and stochasticity contribute to the complexity of autism-associated SHANK family genes. *Cell Rep.* 2024;43(7):114376. doi:10.1016/j.celrep.2024.114376
14. Brauer B, Ancatén-González C, Ahumada-Marchant C, et al. Impact of KDM6B mosaic brain knockout on synaptic function and behavior. *Sci Rep.* 2024;14(1):20416. doi:10.1038/s41598-024-70728-5
15. Teter OM, McQuade A, Hagan V, et al. CRISPRi-based screen of autism spectrum disorder risk genes in microglia uncovers roles of ADNP in microglia endocytosis and synaptic pruning. *Mol Psychiatry.* 2025;30(9):4176-4193. doi:10.1038/s41380-025-02997-z
16. Conrow-Graham M, Williams JB, Martin J, et al. A convergent mechanism of high risk factors ADNP and POGZ in neurodevelopmental disorders. *Brain.* 2022;145(9):3250-3263. doi:10.1093/brain/awac152
17. Wang ZJ, Zhong P, Ma K, et al. Amelioration of autism-like social deficits by targeting histone methyltransferases EHMT1/2 in Shank3-deficient mice. *Mol Psychiatry.* 2020;25(10):2517-2533. doi:10.1038/s41380-019-0351-2
18. Huang Y, Zhu Y, Zhu Y, Jiang Y, Xie Y. SETDB1 ensures the continuity of embryonic to adult neural stem cells through metabolic alterations in the dentate gyrus. *Proc Natl Acad Sci U S A.* 2025;122(30):e2424315122. doi:10.1073/pnas.2424315122
19. Tan SL, Nishi M, Ohtsuka T, et al. Essential roles of the histone methyltransferase ESET in the epigenetic control of neural progenitor cells during development. *Development.* 2012;139(20):3806-3816. doi:10.1242/dev.082198
20. Rodriguez JD, Reeves MN, Chavez SR, et al. Ectopic transcription due to inherited histone methylation may interfere with the ongoing function of differentiated neurons. *Proc Natl Acad Sci U S A.* 2025;122(39):e2513137122. doi:10.1073/pnas.2513137122
21. Chen S, Zhang B, Qin T, et al. Endogenous retrovirus-derived RNA-DNA hybrids induce microglial synaptic pruning in autism models. *Neuron.* 2026;114(10):1765-1781.e7. doi:10.1016/j.neuron.2026.01.011
22. Chen S, Zhang B, Qin T, et al. Endogenous retrovirus-derived RNA-DNA hybrids induce microglial synaptic pruning in autism models. *Neuron.* 2026;114(10):1765-1781.e7. doi:10.1016/j.neuron.2026.01.011

23. Chen S, Zhang B, Qin T, et al. Endogenous retrovirus-derived RNA-DNA hybrids induce microglial synaptic pruning in autism models. *Neuron*. 2026;114(10):1765-1781.e7. doi:10.1016/j.neuron.2026.01.011
24. Tovo PA, Davico C, Marcotulli D, et al. Enhanced Expression of Human Endogenous Retroviruses, TRIM28 and SETDB1 in Autism Spectrum Disorder. *Int J Mol Sci*. 2022;23(11):5964. doi:10.3390/ijms23115964
25. Li F, Elbakry A, Zhou FY, et al. CHAMP1 complex promotes heterochromatin assembly and reduces replication stress. *Proc Natl Acad Sci U S A*. 2026;123(1):e2525144122. doi:10.1073/pnas.2525144122
26. Chen CY, Cheng YY, Yen CYT, Hsieh PCH. Mechanisms of pluripotency maintenance in mouse embryonic stem cells. *Cell Mol Life Sci*. 2017;74(10):1805-1817. doi:10.1007/s00018-016-2438-0
27. Huang Y, Zhu Y, Zhu Y, Jiang Y, Xie Y. SETDB1 ensures the continuity of embryonic to adult neural stem cells through metabolic alterations in the dentate gyrus. *Proc Natl Acad Sci U S A*. 2025;122(30):e2424315122. doi:10.1073/pnas.2424315122
28. Jin L, He J, Feng H, et al. Transposable elements activation triggers necroptosis in mouse embryonic stem cells. *Cell Death Dis*. 2023;14(3):184. doi:10.1038/s41419-023-05705-3
29. Dong L, Liao H, Zhao L, et al. A functional crosstalk between the H3K9 methylation writers and their reader HP1 in safeguarding embryonic stem cell identity. *Stem Cell Reports*. 2023;18(9):1775-1792. doi:10.1016/j.stemcr.2023.08.004
30. Szklarczyk D, Kirsch R, Koutrouli M, 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-D646. doi:10.1093/nar/gkac1000
31. Canat A, Veillet A, Batrin R, et al. DAXX safeguards heterochromatin formation in embryonic stem cells. *J Cell Sci*. 2023;136(19):jcs261092. doi:10.1242/jcs.261092
32. Sun X, Zhang T, Tong B, Cheng L, Jiang W, Sun Y. POGZ suppresses 2C transcriptional program and retrotransposable elements. *Cell Rep*. 2023;42(8):112867. doi:10.1016/j.celrep.2023.112867
33. Meng TG, Zhou Q, Ma XS, et al. PRC2 and EHMT1 regulate H3K27me2 and H3K27me3 establishment across the zygote genome. *Nat Commun*. 2020;11(1):6354. doi:10.1038/s41467-020-20242-9

34. Zhang Z, Wang R, Zhou H, et al. Inhibition of EHMT1/2 rescues synaptic damage and motor impairment in a PD mouse model. *Cell Mol Life Sci*. 2024;81(1):128. doi:10.1007/s00018-024-05176-5
35. Zhang QM, Chen YF, Xing YY, et al. Anterior insular cortex regulates depression-like and ASD-like behaviors via the differential contribution of two subsets of microglia. *Mol Psychiatry*. 2026;31(1):173-191. doi:10.1038/s41380-025-03139-1
36. Zhang Q, Xing M, Bao Z, et al. Contactin-associated protein-like 2 (CNTNAP2) mutations impair the essential α -secretase cleavages, leading to autism-like phenotypes. *Signal Transduct Target Ther*. 2024;9(1):51. doi:10.1038/s41392-024-01768-6
37. Tromp A, Mowry B, Giacomotto J. Neurexins in autism and schizophrenia-a review of patient mutations, mouse models and potential future directions. *Mol Psychiatry*. 2021;26(3):747-760. doi:10.1038/s41380-020-00944-8
38. Hauser D, Behr K, Konno K, et al. Targeted proteoform mapping uncovers specific Neurexin-3 variants required for dendritic inhibition. *Neuron*. 2022;110(13):2094-2109.e10. doi:10.1016/j.neuron.2022.04.017
39. Antoine MW, Langberg T, Schnepel P, Feldman DE. Increased Excitation-Inhibition Ratio Stabilizes Synapse and Circuit Excitability in Four Autism Mouse Models. *Neuron*. 2019;101(4):648-661.e4. doi:10.1016/j.neuron.2018.12.026
40. Robinson M, Wettschurack K, Halurkar MS, et al. Precision medicine for sodium channelopathy-related autism and epilepsy. *Trends Mol Med*. 2026;32(4):338-353. doi:10.1016/j.molmed.2025.09.007
41. Oh H, Lee S, Oh Y, et al. Kv7/KCNQ potassium channels in cortical hyperexcitability and juvenile seizure-related death in Ank2-mutant mice. *Nat Commun*. 2023;14(1):3547. doi:10.1038/s41467-023-39203-z
42. Aiba I, Ning Y, Noebels JL. Persistent Na⁺ current couples spreading depolarization to seizures in Scn8a gain-of-function mice. *Brain*. 2025;148(9):3325-3339. doi:10.1093/brain/awaf120
43. Wang C, Jiang W, Leitz J, et al. Structure and topography of the synaptic V-ATPase-synaptophysin complex. *Nature*. 2024;631(8022):899-904. doi:10.1038/s41586-024-07610-x

Figure

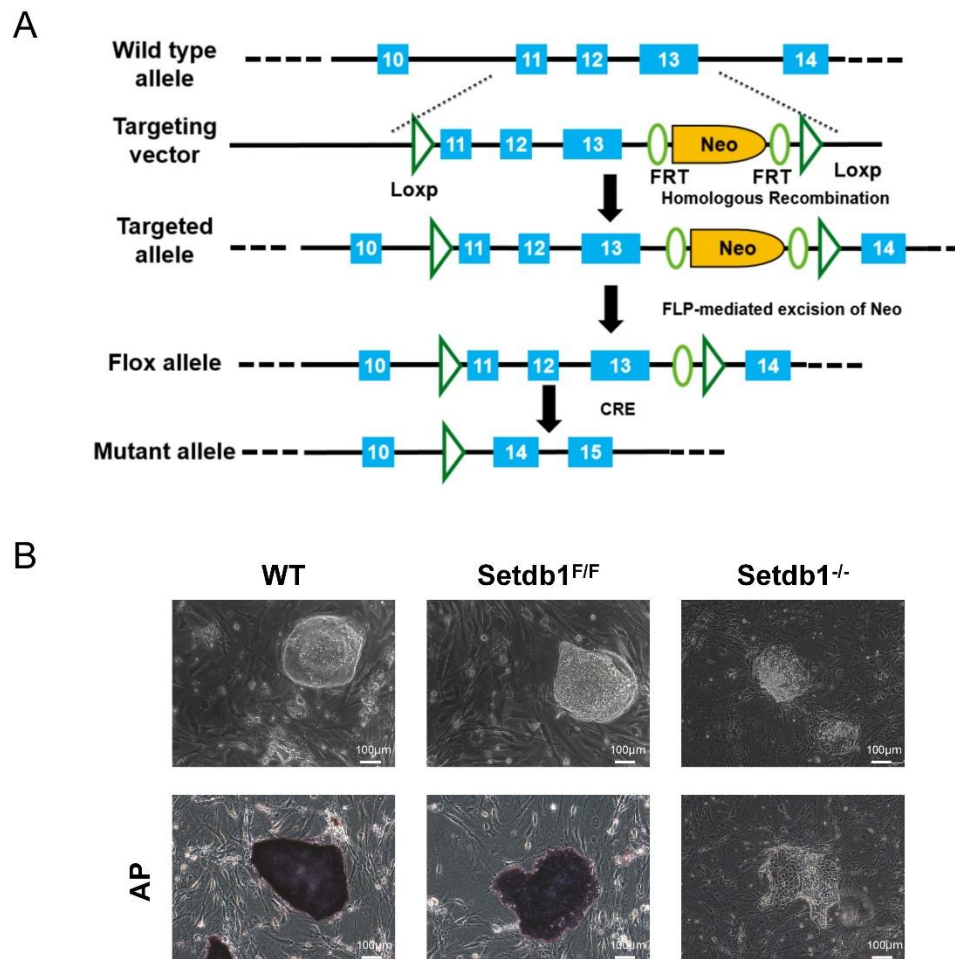


Figure 1. Generation of conditional SETDB1 knockout mESCs and validation of knockout clones.

(A) Schematic of the targeting strategy. LoxP sites (red triangles) flank exons 11–13; an FRT-flanked Neo cassette (green diamonds) is inserted. FLP recombinase removes Neo via FRT sites to generate the floxed (FF) allele; Cre recombinase then deletes exons 11–13 to produce the KO allele. (B) Representative bright-field images (top panels) and alkaline phosphatase (AP)

staining (bottom panels) of wild-type (WT), floxed (FF), and KO clones. KO clones exhibit flattened morphology, reduced cell density, and loss of AP activity. Scale bar = 100 μm .

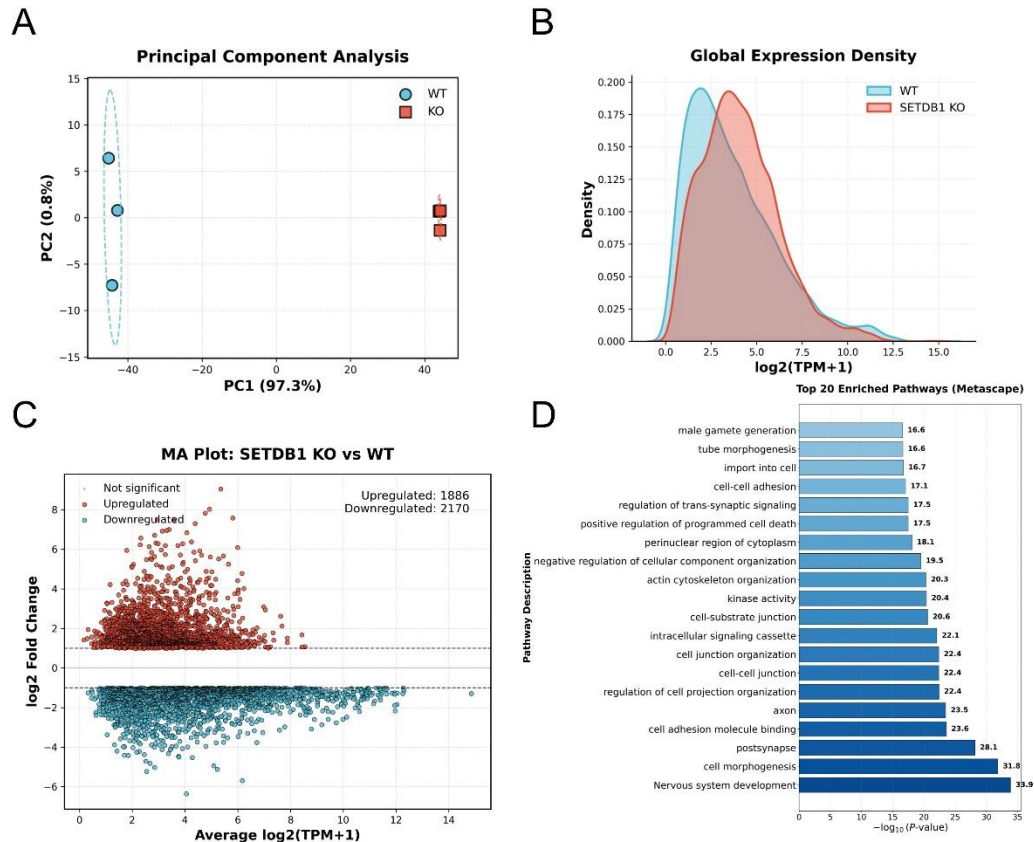


Figure 2. Transcriptomic profiling and functional enrichment of SETDB1 KO mESCs.

(A) Principal component analysis (PCA) of WT (blue) and KO (red) samples ($n = 3$ each). PC1 explains 97.3% of the variance. (B) Kernel density plot of global gene expression ($\log_2(\text{TPM}+1)$). (C) Volcano plot of differentially expressed genes (DEGs; $|\log_2\text{FC}| \geq 1$, $p_{\text{adj}} < 0.05$). Upregulated genes (1,886)

are shown in red, downregulated (2,170) in blue, and non-significant in gray.
(D) Horizontal bar plot of the top 20 enriched GO Biological Process terms among all DEGs.

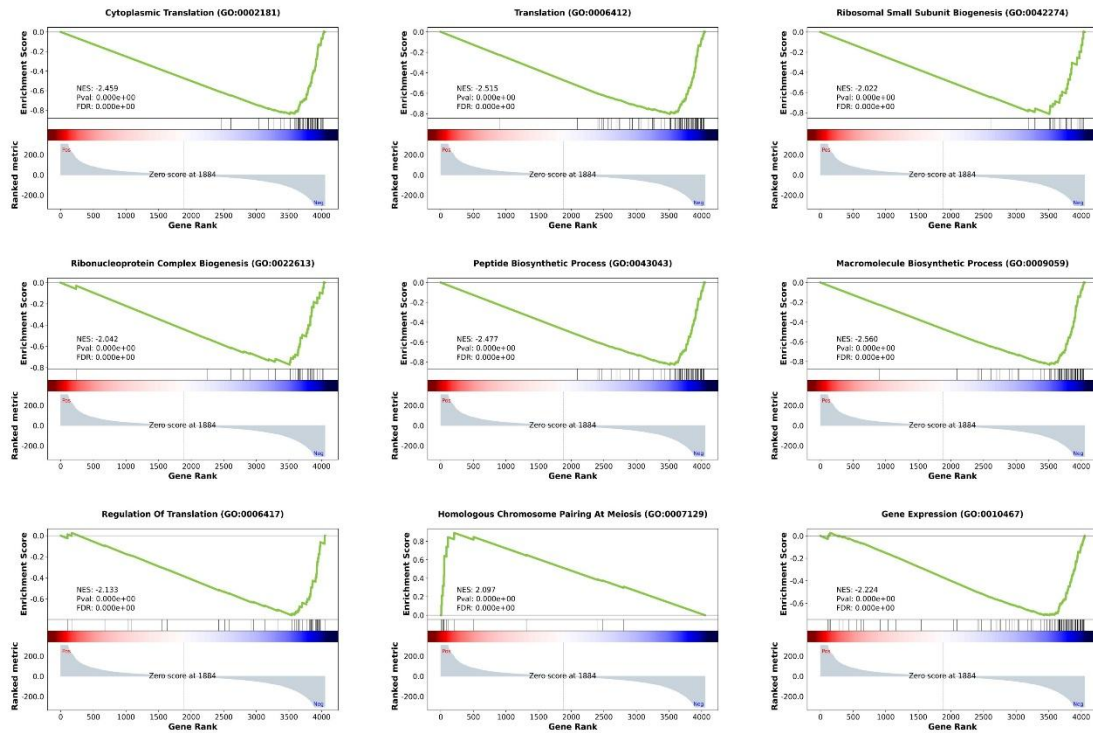
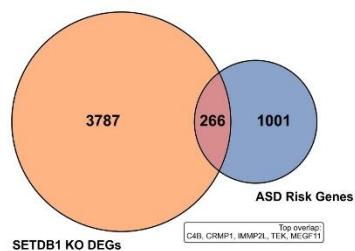


Figure 3. Gene set enrichment analysis (GSEA) of SETDB1 KO mESCs.

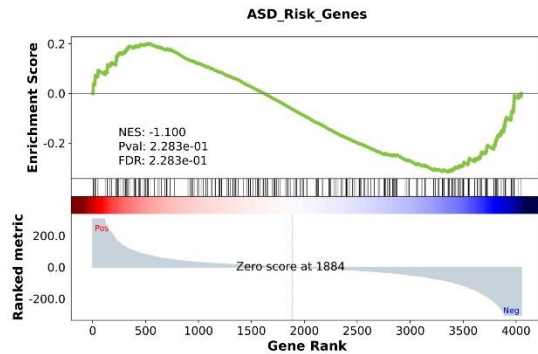
Representative enrichment plots for significantly downregulated pathways (FDR < 0.05). Green curve: running enrichment score; black vertical lines: positions of gene set members in the ranked list; NES: normalized enrichment score.

A

Overlap of SETDB1 KO Differentially Expressed Genes with ASD Risk Genes

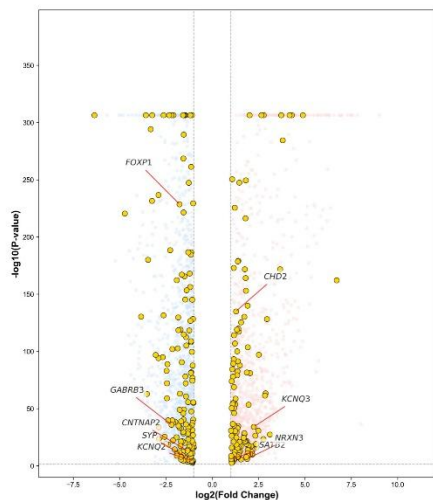


B

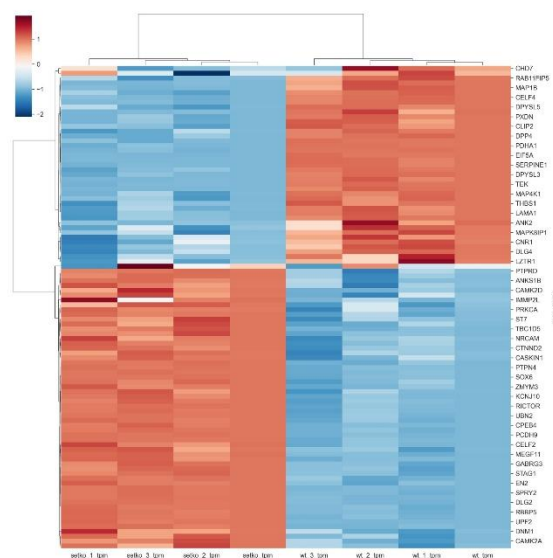


C

Dysregulated ASD Risk Genes in SETDB1 KO Cells



D



E

Enriched Pathways in SETDB1 KO ASD-Associated Genes

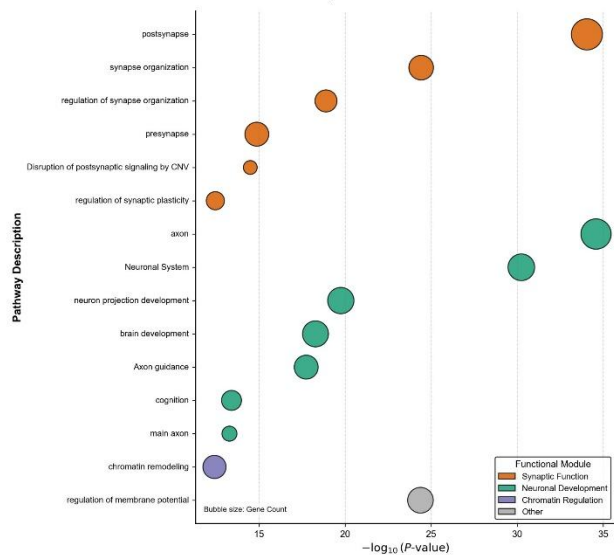


Figure 4. Comprehensive analysis of SETDB1 KO-sensitive ASD risk genes.

(A) Venn diagram showing the overlap between SETDB1 DEGs and SFARI ASD risk genes (266 overlapping genes, hypergeometric test $p < 0.001$). (B) GSEA of the ASD risk gene set (NES = -1.100 , FDR = 0.228). (C) Volcano plot highlighting the 266 overlapping ASD genes with selected genes labeled. (D) Heatmap of the 266 overlapping genes ($\log_2(\text{TPM}+1)$, row Z-score). Hierarchical clustering separates WT from KO samples. (E) Module-based bubble plot of enriched GO terms among the 266 genes. Three functional modules are distinguished: Synaptic Function (orange), Neuronal Development (green), and Chromatin Regulation (blue). Bubble size corresponds to gene count, and color indicates $-\log_{10}(p)$.

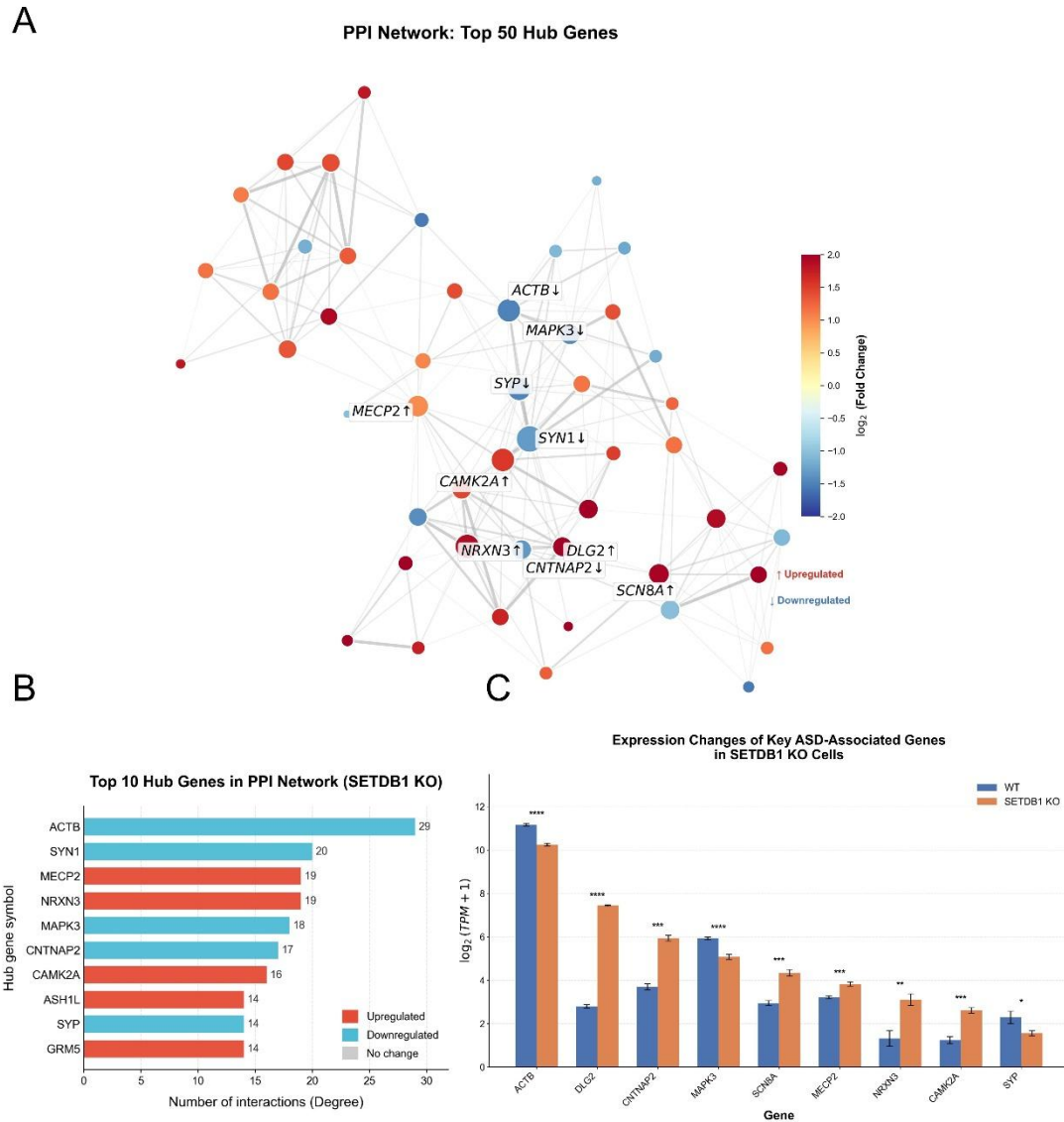


Figure 5. Protein-protein interaction network and validation of hub genes.

(A) PPI network of the top 50 ASD-related DEGs (STRING confidence ≥ 0.4). Nodes are colored by $\log_2FC$ (red: upregulated, blue: downregulated), and node size is proportional to degree (number of interactions). (B) Horizontal bar plot of the top 10 hub genes ranked by degree. Bar colors indicate expression direction (red: upregulated, blue: downregulated). (C) Bar plot of expression levels ($\log_2(TPM+1)$, mean $\pm$ SD) of the top 10 hub genes in WT and KO cells. Statistical significance was determined by two-tailed t -test (ns:

not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

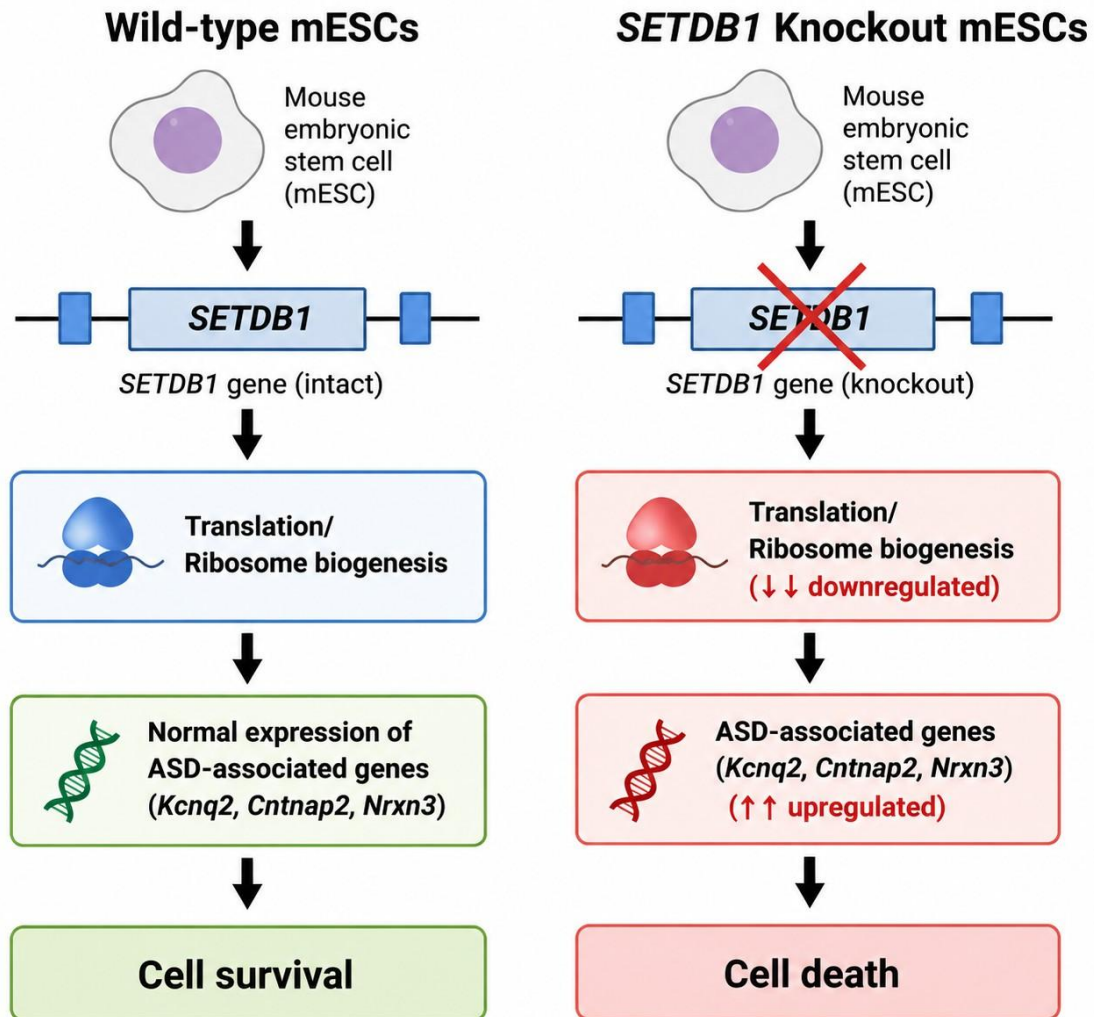


Figure 6. Proposed model of SETDB1 function in mESCs.

In wild-type cells, SETDB1 maintains normal translation and ribosome biogenesis while keeping ASD-associated genes at basal levels, supporting cell survival and pluripotency. Upon SETDB1 knockout, translation and ribosome biogenesis are globally suppressed, and ASD-related genes become markedly upregulated, leading to cell death and loss of pluripotency.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Highlights.docx](#)