



Published in final edited form as:

Neuron. 2019 November 06; 104(3): 471–487.e12. doi:10.1016/j.neuron.2019.09.014.

Recapitulation and reversal of schizophrenia-related phenotypes in *Setd1a*-deficient mice

Jun Mukai^{1,2,‡,*}, Enrico Cannavò^{2,*}, Gregg W. Crabtree^{1,2}, Ziyi Sun^{1,#}, Anastasia Diamantopoulou¹, Pratibha Thakur^{1,2}, Chia-Yuan Chang¹, Yifei Cai¹, Stavros Lomvardas^{2,3,4}, Atsushi Takata⁵, Bin Xu⁶, Joseph A. Gogos^{1,2,4,†}

¹Department of Physiology and Cellular Biophysics, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA

²Mortimer B. Zuckerman Mind Brain and Behavior Institute Columbia University, New York, NY 10027 USA

³Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, NY 10032 USA

⁴Department of Neuroscience, Columbia University, New York, NY 10032 USA

⁵Department of Human Genetics, Yokohama City University Graduate School of Medicine, Fukuura 3-9, Kanazawa-ku, Yokohama 236-0004, Japan

⁶Department of Psychiatry, Columbia University Medical Center, New York, NY 10032, USA

Summary

SETD1A, a lysine-methyltransferase, is a key schizophrenia susceptibility gene. Mice carrying a heterozygous loss-of-function mutation of the orthologous gene exhibit alterations in axonal branching and cortical synaptic dynamics, accompanied by working memory deficits. We show that *Setd1a* binds both promoters and enhancers with a striking overlap between *Setd1a* and *Mef2* on enhancers. *Setd1a* targets are highly expressed in pyramidal neurons and display a complex pattern of transcriptional up- and down-regulations shaped by presumed opposing functions of

Joseph A. Gogos (jag90@columbia.edu).

[‡]Present Address: Laboratory of Molecular Psychiatry and Neuroscience, Research and Development Center for Precision Medicine, University of Tsukuba, Tsukuba, Ibaraki, 305-8550, Japan

[#]Present Address: State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, and Collaborative Innovation Center for Biotherapy, Chengdu, Sichuan 610041, China

*Equal Contribution

†Lead Contact

Author Contributions. J.M., E.C., B.X., A.T., S.L., and J.A.G designed and interpreted the experiments. J.M validated the mouse model, performed and analyzed behavioral, morphological, co-IP, targeted expression and all rescue assays. E.C. performed and analyzed all ChIP-Seq and scRNA-Seq assays. G.W.C. and Z.S. performed and analyzed electrophysiology experiments. A.D performed and analyzed behavior experiments. P.T contributed to the *in vivo* axonal branching assays. C-Y.C performed and analyzed expression assays. B.X performed the bioinformatics analysis of the overlap between ChIP-Seq peaks and risk genes. A.T contributed to the validation of the mouse model and the bioinformatics analysis of the RNA-Seq and ChIP-Seq data. Y.C contributed the bioinformatics analysis. S.L. and J.A.G supervised the work. J.M., E.C., G.W.C., Z.S., A.D, A.T., B.X., S.L., J.A.G wrote the paper.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of Interests

The authors declare no competing financial interests.

Setd1a on promoters and Mef2-bound enhancers. Notably, evolutionary conserved Setd1a targets are associated with neuropsychiatric genetic risk burden. Reinstating *Setd1a* expression in adulthood rescues cognitive deficits. Finally, we identify LSD1 as a major counteracting demethylase for Setd1a, and show that its pharmacological antagonism results in a full rescue of the behavioral and morphological deficits in *Setd1a*-deficient mice. Our findings advance understanding of how *SETD1A* mutations predispose to SCZ and point to novel therapeutic interventions.

eTOC blurb

Mukai et al. find that mice with a mutation in a schizophrenia risk gene (*Setd1a*) show cognitive and circuitry deficits that are rescued by restoring Setd1a function in adulthood. Setd1a coordinates expression of genes important for neuronal development and function.

Keywords

SETD1A; MEF2; LSD1; histone methyltransferase; enhancer; loss-of-function (LoF) mutation; schizophrenia; axonal branching; working memory; pharmacological reversal

Introduction

Over the past decade, human genetic studies have revealed an important role of individually rare (but collectively common) *de novo* (DNMs) and inherited mutations in genetically complex neuropsychiatric diseases (Fromer et al., 2014; Genovese et al., 2016; Gulsuner et al., 2013; Rodriguez-Murillo et al., 2012). We have previously linked loss-of-function (LoF) mutations in *SETD1A*, a lysine methyltransferase, to risk for schizophrenia (SCZ) (Takata et al., 2016; Takata et al., 2014). This finding was subsequently confirmed by large scale exome sequencing studies (Singh et al., 2016; Takata et al., 2016; Takata et al., 2014). *SETD1A* is part of the Set/COMPASS complex, which mediates mono-, di-, and trimethylation of the lysine 4 on the histone H3 protein (H3K4) (Miller et al., 2001) (Shilatifard, 2012). In addition to *SETD1A*, mutations in subunits of Set/COMPASS have been reported in SCZ and other neurodevelopmental disorders (NDD) (Gulsuner et al., 2013; Neale et al., 2012; O’Roak et al., 2012) as well as dominant Mendelian syndromes (Fahrner and Bjornsson, 2014).

SETD1A is a confirmed SCZ risk gene and *SETD1A* mutations confer a large increase in disease risk (Singh et al., 2016), thus providing an excellent genetic substrate for disease modeling. Here, we performed a comprehensive analysis of mutant mice carrying a LoF allele in the endogenous *Setd1a* orthologue that models the various identified *SETD1A* SCZ LoF risk alleles to uncover the role of *SETD1A* in gene regulation, neuronal architecture, synaptic plasticity and behavioral paradigms dependent on learning. We also used this mouse model to examine whether malfunction of neural circuits and neurocognitive deficits emerging due to *Setd1a*-deficiency can be reversed by pharmacological interventions during adulthood, paving the way for the development of new therapeutics.

Results

Setd1a is expressed throughout the adult mouse brain, with higher expression in neocortex (Allen Brain Atlas, www.brain-map.org). Real-time Quantitative Reverse Transcription PCR (qRT-PCR) analysis on mouse prefrontal cortex (PFC) detected *Setd1a* mRNA at various developmental stages (Figure S1A). Immunostaining of the prelimbic area in medial PFC of 6-week old mice revealed that *Setd1a* positive cells are distributed in all cortical layers except L1 and merge with NeuN⁺ cells (Figure S1B). *Setd1a* protein exhibits mixtures of punctate and diffuse fibrillary nuclear staining in all NeuN⁺ neuron and parvalbumin positive (PV⁺) interneurons (INs) in the medial PFC superficial layers, but only a few small puncta in CNPase (oligodendrocytes) and GFAP (astrocytes) positive cells in the cortical plate and corpus callosum (Figure S1C). *Setd1a* localizes to euchromatin and does not overlap with heterochromatin regions in either NeuN⁺ or PV⁺ neurons (Figure S1C). We employed a mouse model carrying a *Setd1a* LoF allele (Figures S1D and S1E). Reduction of *Setd1a* RNA and protein levels to approximately half of WT levels was confirmed by qRT-PCR and immunoblots of adult frontal cortex (Figures S1F and S1G). Although *Setd1a*^{-/-} is embryonic lethal (Bledau et al., 2014), *Setd1a*^{+/-} mice were indistinguishable from WT littermates in terms of body size, weight and posture (not shown).

***Setd1a*^{+/-} mice exhibit deficits in working memory (WM) and cortical synaptic plasticity.**

Cognitive impairment is a core feature of SCZ (Barch and Ceaser, 2012). We therefore asked whether *Setd1a* deficiency in mice affects performance in a variety of cognitive tasks. First, we established that *Setd1a*^{+/-} mice show normal activity in the open-field assay (Figures S1H and S1I). We assessed spatial WM performance using a T-maze delayed non-match to place (DNMTP) task. *Setd1a*^{+/-} mice learned the task and performed almost as well as WT littermates during training (Figure 1A and 1B). In the subsequent WM test, however, *Setd1a*^{+/-} mice manifested a marked impairment (correct responses: 63.63 ± 3.03 %, *P* = 0.0038 at 10 sec; 64.58 ± 3.10 %, *P* = 0.0172 at 30 sec), compared to WT littermate controls (76.39 ± 2.68 % at 10 sec, 76.39 ± 3.38 % at 30 sec) (Figure 1C). In contrast, memory performance of *Setd1a*^{+/-} mice was intact in a series of other cognitive tasks (Figure S1J–L).

WM can be affected by the synaptic properties of PFC neurons (Barak and Tsodyks, 2014). The monosynaptic glutamatergic inputs from L2 neurons to the apical and basal dendrites of L5 pyramidal neurons (PNs) are critical for controlling the computations and output of L5 PNs (Quiquempoix et al., 2018) and play a key role in WM (Bastos et al., 2018; Goldman-Rakic, 1995). We performed field recordings in brain slices to examine the synaptic properties in L5 neurons in response to the stimulation of L2. The input-output relation across experiments and the fiber volley amplitude was comparable between genotypes, indicating normal basic synaptic transmission (Figures S2A, S2B, and S2C). To determine whether short-term synaptic plasticity was impaired we employed a paired-pulse ratio (PPR) protocol at various inter-stimulus intervals (ISIs) as previously described for two other genetic models (Crabtree et al., 2017; Fenelon et al., 2013). At ISIs of less than 200 ms, both genotypes exhibited a significant facilitation, but *Setd1a*^{+/-} mice showed overall significantly decreased PPR values (Figures 1G and S2J, *N* = 7, *n* = 23; *Setd1a*^{+/-}, *N* = 7, *n*

= 23; two-way repeated-measures ANOVA, $P = 0.04$) indicating an increased initial probability of presynaptic transmitter release.

During PFC-dependent WM tasks, PFC neurons in local ensembles receive trains of inputs from neighboring cells in the 20 to 60 Hz frequency range (Miller et al., 1996). We tested the effect of *Setd1a* deficiency on synaptic plasticity within this physiological frequency range by applying 40 stimuli at 5 Hz, 20 Hz and 50 Hz. In response to these stimulus trains, a rapid depression of synaptic strength was observed in both genotypes (Zucker and Regehr, 2002). Compared to WT mice, *Setd1a*^{+/-} mice show significantly greater depression of fEPSP responses at 5 Hz (Figures 1H and S2K, WT, N = 7, n = 23; *Setd1a*^{+/-}, N = 7, n = 22; two-way repeated-measures ANOVA, $P = 0.03$) and at 20 Hz (Figures 1H and S2K, WT, N = 7, n = 20; *Setd1a*^{+/-}, N = 7, n = 19; two-way repeated-measures ANOVA, $P = 0.04$) but not at 50 Hz (Figures 1H and S2K, WT, N = 7, n = 18; *Setd1a*^{+/-}, N = 7, n = 16; two-way repeated-measures ANOVA, $P = 0.95$). To determine the impact of *Setd1a* deficiency on the potentiation of synaptic efficacy of L5 cortical synapses, short-term (STP) and long-term potentiation (LTP) were induced (Crabtree et al., 2017; Fenelon et al., 2013) but no genotypic differences were found (Figures S2D and S2L, $P = 0.57$).

We also tested whether PFC neurons displayed alterations in intrinsic excitability. Neuronal excitability is known to critically regulate network information flow (Titley et al., 2017) and has previously been shown to be altered in animal models of SCZ (Benamer et al., 2018; Crabtree et al., 2017; Gruter et al., 2015; Hamm et al., 2017). We performed whole-cell current-clamp recordings in PFC L2/3 PNs – a major source of excitatory input to L5 – and found significantly increased action potential (AP) responses to current steps (two-way repeated-measures ANOVA, $P = 0.021$, Figure 1F). In contrast, we found the minimal current required for AP generation (rheobase) was unaltered (Figure S2E). Notably, neuronal input resistance, membrane time constant, and resting membrane potential were all found to be unaltered and thus did not underlie the observed enhanced neuronal excitability (Figures S2F, S2G and S2H). Both initial and terminal inter-spike intervals were significantly reduced (Figure S2M) while AP afterhyperpolarization was not altered (Figure S2N). Whole-cell voltage-clamp studies employing voltage steps revealed that *Setd1a*^{+/-} mice displayed reduced outward currents near AP threshold voltages (-40mV to -30mV, Figures S2I and S2O). These reductions in hyperpolarizing currents may in part explain the enhanced AP generation.

Laminar organization and cytoarchitecture in the cortex of *Setd1a*^{+/-} mice.

Adult 8-week-old *Setd1a*^{+/-} mice manifest normal gross brain morphology and only a modestly diminished overall density of neurons (identified by their NeuN immunoreactivity) in L2/3 and L5. Quantitative analysis of different IN subpopulations identified by their immunohistochemical markers revealed a reduced number of PV⁺ INs in L5 and Reelin⁺ INs in L2/3 and L5, in the prelimbic area of medial PFC (Figures S3A, S3B, S3C, S3E, and S3H). The numbers of VIP and SST positive INs were unchanged (Figures S3D, S3G, S3F, and S3). Further analysis (Figures S3J, S3K, S3L, and S3M) showed that alterations in the laminar organization are, at least in part, due to alterations in the proliferation of neural

progenitor cells (NPCs) during cortico-genesis (Bledau et al., 2014). Alterations in the numbers of INs may also reflect homeostatic responses to altered circuit activity.

We also investigated whether *Setd1a* deficiency affects development of cortical axons, dendrites and spines. Plasmids encoding EGFP were electroporated *in utero* into the dorsolateral ventricular zone of the mouse forebrain at E15.5, and embryos were allowed to be born and develop to P8.5 when all cells labeled with EGFP completed their migration into L2/3. As shown previously (Mukai et al., 2015), callosal axons crossed the midline, coursed beneath the gray matter of the contralateral somatosensory cortex, penetrated the gray matter, and ascended to L1–3 and 5 (Figures 2A and 2B). The total number of callosal axon-terminal branches within the contralateral cortical L1–4 was decreased by disruption of *Setd1a* (N = 10, n = 30), compared to WT brains (N = 10, n = 30, $P < 0.0001$) (Figures 2C and 2D). Notably secondary branch number was robustly decreased in *Setd1a*^{+/-} brains (N = 10, n = 30) versus WT brains (N = 10, n = 30; $P < 0.007$) (Figures 2C and 2D). Similar deficits in axonal branching were observed in cortical primary cultures (Figures 2E, 2F, 2G, 2H, and 2I).

We looked for deficits in dendrite and spine development in the basal dendrites of L5 neurons in the prelimbic area of medial PFC by intercrossing *Setd1a*^{+/-} mice with a reporter Thy1-GFP/M line (Figure S3N). *Setd1a*^{+/-} mice did not show a simplification in the dendritic arbors by Sholl analysis (Figures S3O, S3P, and S3Q). Analysis of spines in 8-week-old mice showed that mushroom spine density was modestly reduced in the basal dendrites of L5 *Setd1a*^{+/-} (N = 10, n = 20) compared to WT neurons (N = 10, n = 20, $P < 0.002$; Figures 2J and 2K). Mushroom spines showed a small decrease in the head width [N = 10, n = 605, Kolmogorov-Smirnov (K-S) test $P < 0.0001$] and length (N = 10, n = 605, K-S test $P < 0.02$) in *Setd1a*^{+/-} compared to WT neurons (N = 10, n = 644) (Figures 2L and 2M). Similar analysis at the basal dendrites of L2/3 neurons of *Setd1a*^{+/-} (N = 7, n = 21) and WT littermates (N = 7, n = 21), labeled with EGFP via AAV2-Syn-EGFP virus injection into the prelimbic area of medial PFC, revealed a modest reduction in mushroom spine density ($P < 0.035$; Figures 2N and 2O). Reduction in head width (N = 8, n = 373, K-S test $P < 0.034$), but not length ($P = 0.259$) was also observed in *Setd1a*^{+/-} L2/3 neurons compared to WT neurons (N = 9, n = 379) (Figures 2P and 2Q).

Identification of direct targets of Setd1a in the PFC of adult mice.

Given that Setd1a is a lysine demethylase that regulates transcription, we sought to identify direct genomic targets of Setd1a in the PFC by chromatin immunoprecipitation (IP) followed by sequencing (ChIP-Seq, see Methods). We performed ChIP-Seq analysis in chromatin prepared from micro-dissected PFCs of 6-week old mice for histone modifications that mark active promoters (H3K4me2, H3K4me3 and H3K27ac) and active or poised enhancers (H3K4me1 and H3K27ac) using the same chromatin preparations. Using stringent criteria [Irreproducibility Discovery Rate (IDR) < 0.02], we identified 25,658 Setd1a peaks genome-wide. Despite the fact that Setd1a mediates the trimethylation of H3K4 (Shilatifard, 2012), only a minority (34%) of Setd1a peaks are located at promoter elements as observed by the overlap of Setd1a with H3K4me3 peaks (Figures 3A, 3B, S4A, and Data Table 1). The majority of Setd1a peaks (66%) overlap with enhancer marks and not

H3K4me3 (Figures 3B, S4A, Data Table 1). Further analysis of the genomic distribution of the non-promoter Setd1a peaks, revealed that 45% reside within intergenic regions while 51% reside in intronic sites with enhancer-like chromatin properties (Figures 3B, S4A). A minority (4%) resides within exons, 3'UTRs or within 1 kb from the transcription end-point.

To assess if Setd1a-bound regulatory elements have pan-neuronal or cortical-specific functions we compared the H3K27ac ChIP-Seq from the PFC with ENCODE H3K27ac ChIP-Seq from the olfactory bulb (OB) and cerebellum (Shen et al., 2012). Setd1a bound PFC promoters, as a whole, have H3K27ac in the OB and less so in the cerebellum (Figure 3C). In contrast the majority of the enhancer peaks (74%) are enriched for H3K27ac specifically in the PFC (Figures 3C) suggesting that cortex-specific transcription factors (TFs) recruit Setd1a at enhancers whereas more generally expressed TFs recruit Setd1a to promoters.

Setd1a targets are extensively enriched for Mef2 binding sites.

De novo motif analysis of the Setd1a peaks using HOMER revealed Mef2 consensus motifs as the most enriched DNA motif within the enhancer peaks (Figure 3D, Data Table 1) and Sp1 consensus motifs as the most enriched in the promoters (Figure 3D, Data Table 1). ChIP-Seq experiments with an antibody that recognizes PFC-expressed Mef2a/c/d isoforms showed a striking overlap between Setd1a and Mef2 targets specifically at enhancers, with 11,688 of the 16,889 Setd1a peaks being occupied by Mef2 (Figure 3E S4B). Co-IP experiments with nuclear extracts prepared from the cortical plate, revealed an interaction between Setd1a and Mef2a/c (Figures 3F). To determine if this regulatory relationship is conserved across evolution, we examined the PhastCons conservation score in the Euarchontoglires clade of Setd1a bound enhancers. This analysis revealed both conserved (~46%) and non-conserved (~54%) Setd1a targets. Among the former, we detected conservation of Mef2 binding sites, providing additional support for a role of Mef2 in Setd1a recruitment (Figure S4C and S4D). To investigate if the interaction between Setd1a and Mef2 is conserved in humans in disease-relevant brain areas, we mapped Setd1a peaks from the mouse genome (mm10) to the corresponding positions in the human genome (hg38) focusing in particular to PFC-specific bound regions. We mapped 3,813 mouse genomic regions (out of 12,499) to the human genome (Data Table 1). To enrich for functional elements, we explored the chromatin accessibility of these peaks in the human brain. We analyzed a previously published ATAC-Seq dataset from 14 regions of human postmortem brains (Fullard et al., 2018) focusing specifically on neuronal and non-neuronal cells of dorsal and ventral PFC, hippocampus, medial dorsal thalamus and putamen. Of the 3,813 conserved human peaks, 1,765 encompass significant ATAC-Seq peaks in at least one of the considered brain areas. Interestingly, these sites are accessible in neuronal cell types and not glia. Moreover, we found that a minority of peaks is accessible across all five regions of the brain (14%) while the majority is mostly accessible in cortical regions and hippocampus (86%) (Figure S4E). To investigate the functionality of these sites in human neurons, we performed motif analysis on the accessible conserved Setd1a regions with HOMER and identified MEF2 motifs as one of the most enriched motifs ($P = 1e-41$; Data Table 1). These data suggest that the role of Mef2 and Setd1a in transcriptional regulation is

conserved between mouse and human and that intergenic Setd1a peaks likely originate from neuronal cortical populations (see also below).

Setd1a targets are enriched for orthologues of disease-risk genes.

To elucidate whether Setd1a-dependent genome regulation in mice is linked to neuropsychiatric etiology, we determined more directly if conserved Setd1a genomic binding sites are enriched for genes and genomic intervals implicated in SCZ. We first performed an interval-based analysis to assess enrichment of Setd1a peaks of the 108 loci identified by SCZ-GWAS (Schizophrenia Working Group of the Psychiatric Genomics, 2014). We focused only on PFC-specific enhancer peaks which are conserved and accessible in the human brain (see above). We applied a permutation test ($N=20,000$ permutations) implemented in the RegioneR program and found a significant overlap ($P=0.0055$ $Z=2.9$) between these Setd1a peak intervals and the 108 SCZ-GWAS associated intervals (Schizophrenia Working Group of the Psychiatric Genomics, 2014) (Figures 4A and Data Table 2). This was true specifically for intronic Setd1a peaks (see Figure S5A–S5D). To confirm the specificity of this finding, we conducted an identical analysis on data from other GWAS of CNS and non-CNS diseases, with similar sizes of genome-wide significant intervals (Welter et al., 2014), but we did not observe any significant overlap (Tables S1 and S2).

Using a complementary gene-based analysis we directly assessed whether the observed enrichment of Setd1a regulatory elements within SCZ-GWAS loci extends to Setd1a target gene content. We used the same set of human enhancer peaks and restricted our analysis to targets genes with Setd1a peaks that represent well-annotated, high-confidence targets (all peaks with an annotation by PAVIS, $N=708$ genes) as opposed to multiple ambiguous target genes abutting each intergenic peak (Data Table 2). We assessed enrichment of these high-confidence Setd1a targets within the SCZ-GWAS loci, using INRICH, and found a significant overlap (FDR corrected $P=0.008$, Table S3). The list of genes spans diverse neuronal functions ranging from neurogenesis to synaptic function and neuronal excitability and includes several candidate risk-genes (Table S4). Identical analysis of data from other GWAS of CNS and non-CNS diseases did not show significant enrichment except for Autism (Tables S3 and S5). Notably, in keeping with the hypothesis that Mef2 and Setd1a have a coordinated role in SCZ-related gene regulation we observed an excess within SCZ-GWAS intervals of target genes with peaks co-bound by Setd1a and Mef2 but less significant with Mef2-only peaks ($P=0.003$ and 0.03 respectively) (Table S6).

We also investigated whether the high-confidence Setd1a target genes ($N=708$) are significantly overrepresented among genes implicated in SCZ and other NDDs by studies of rare LoF DNMs. Using DNENRICH, we observed that there is significant enrichment of the high-confidence Setd1a target genes (Figure 4B, $P=1.6 \times 10^{-5}$, $FC=1.5$) among the genes hit by LoF DNMs in the combined group of NDD (SCZ, ID/DD and ASD). This was especially true when restricting to the LoF-intolerant genes ($pLI > 0.9$ or $LoF-FDR < 0.01$) hit by LoF DNMs (Figure 4B, $P=1.0 \times 10^{-6}$, $FC=2.1$ with pLI , $P < 1.0 \times 10^{-6}$, $FC=2.1$ with $LoF-FDR$). Significant enrichment was not observed among the genes hit by LoF DNMs, irrespective of their selective pressure, in unaffected subjects (Figure 4B, smallest P

= 0.42) or congenital heart disease patients (Figure 4B, smallest $P=0.24$). Stratification of NDD by disease, revealed an enrichment in ID/DD, ASD and SCZ with variable significance probably due to the number of input DNMs (Figure 4B).

We further searched for disease-related functional and expression properties of *Setd1a* target genes within the conserved and chromatin accessible genomic regions used in our analysis. GO analysis of high-confidence *Setd1a* target genes revealed a significant enrichment in terms relevant to “synaptic function”, “ion channel activity” and “neurogenesis”, as well as significant enrichment for the “schizophrenia” term as defined in the DisGeNET database (umls:C0036341) (Figure 4C). Brain regions and developmental stage analysis using the CSEA tool showed a significant enrichment primarily in the cortex at early fetal/early mid fetal stages (FDR corrected pSI $P=0.003$ and $P=0.002$, respectively) and to a lesser extent at the young adulthood stage (pSI $P<0.05$) (Figures 4D). Notably removal of *Setd1a* target genes overlapping with genes carrying LoF DNMs in NDDs (which are prenatally-biased in their expression, Fig S5E), resulted in stronger enrichment at young adulthood/adolescence stages (FDR corrected pSI $P=2.544e-07$ and $7.600e-06$, respectively) (Figures 4E), suggesting that *Setd1a* targets genes with distinct roles in both prenatal and postnatal neurodevelopment.

***Setd1a* targets display a complex pattern of transcriptional dysregulation.**

We used RNA-Seq analysis to profile transcriptomes of micro-dissected PFCs from four 6-week-old *Setd1a*^{+/-} mice and four WT littermates. We identified 342 genes that significantly change their expression in *Setd1a*^{+/-} mice (FDR<5%) (Figure 5A, Data Table 1). Transcriptional alterations were modest with most of the genes (>70%) displaying a fold change in the range of 0.7–1.3. Interestingly, the majority of the differentially expressed genes (80%) in *Setd1a*^{+/-} PFC overlap with direct targets of *Setd1a*, likely representing primary rather than secondary transcriptional alterations. To dissect the effect of the *Setd1a* mutation on gene expression, we partitioned differentially expressed genes based on the binding regions of *Setd1a* in the PFC. Genes were classified as up- or down-regulated and carrying *Setd1a* binding sites at promoters or enhancers (Figure 5A and 5B). As anticipated from the known function of *Setd1a* as activator (Shilatifard, 2012), the majority (78%) of genes with *Setd1a*-bound promoters were down-regulated in *Setd1a*^{+/-} mice (Figure 5B). In contrast, genes that have *Setd1a*-bound intergenic enhancers can be either up- (40%) or down-regulated (60%) (Figure 5B). Although the complex pattern of transcriptional changes in the PFC of *Setd1a*^{+/-} mice could be explained by a combination of direct and indirect effects, we explored the possibility that *Setd1a* may have opposing functions and molecular targets in promoters and enhancers. Because *Setd1a* binding shows an almost complete overlap with *Mef2* binding on enhancers, and because *Mef2* activity can be inhibited by lysine methylation (Choi et al., 2014), we reasoned that *Setd1a* may act likewise as a *Mef2* repressor on enhancers restricting its transactivation activity (while it acts as activator on promoters via H3K4 trimethylation). Consistently, *Setd1a* inhibits *Mef2c*-mediated transcriptional activation in transient transfection assays (that remove the influence of histone methylation, Figure 5C), providing an explanation for the observed transcriptional up-regulation of genes with *Setd1a*-bound enhancers upon deletion of one *Setd1a* allele.

In addition to the presumed opposing functions of Setd1a on promoters and Mef2-bound enhancers, cellular heterogeneity of the PFC may also contribute to the transcriptional effects observed at the PFC of *Setd1a*^{+/-} mice. To test this, we performed single cell RNA-Seq (scRNA-Seq) on dissociated cells from the PFC of WT and *Setd1a*^{+/-} mice. We obtained expression profiles for 5,944 cells from WT mice and 6,031 cells from *Setd1a*^{+/-} mice (with a mean of 2262 genes per cells). We identified 8 main clusters that represent the major neuronal [PNs (*Camk2a*) and INs (*Igf1*)] and non-neuronal [mature and immature oligodendrocytes (*Olig2* and *Mbp*), microglia (*Aif1*), astrocytes (*Aldoc*) and blood brain barrier cells (*Cldn5*] populations of PFC (Figures S6A, S6B, S6C and S6D). As an indicator for the quality of our scRNA-Seq data, we compared expression, as a whole, of genes that we detected as significantly changing in bulk PFC RNA-Seq and confirmed the same trends of expression changes between bulk and scRNA-Seq (Figures S6E and S6F). Differential expression analysis failed to reveal significant changes in the expression of individual genes. While this finding may be in part due to well-known limitations in sensitivity and technical noise of scRNA-Seq (Liu and Trapnell, 2016), it strongly suggests that reduction of Setd1a levels by half leads to low magnitude changes in the expression of downstream targets even at the cellular level.

We used the scRNA-Seq data to dissect the cellular origins of the expression alterations identified by bulk RNA-Seq and highlight potential targets underlying the phenotypic alterations observed in cortical PNs. Analysis of cellular distribution of all up- and down-regulated genes in *Setd1a*^{+/-} mice revealed a heterogeneous peak expression pattern spanning several neuronal as well as disease-relevant (Chen et al., 2018; Volk, 2017; Windrem et al., 2017) non-neuronal cell-types, such as oligodendrocytes, astrocytes and microglia (Figure 5C, S6G and S6H). However, pyramidal cells display the highest fraction of differentially expressed genes compared to any other cell type identified by scRNA-Seq, consistent with the observation that genes that have Setd1a peaks at their intergenic enhancers have their highest expression in PNs (Figure S6I). The transcriptional dysregulation of some of these PN-specific genes, such as *Slitrk4* (Proenca et al., 2011) (an up-regulated gene involved in the repression of neurite growth) as well as *Synpo* (Kos et al., 2016) and *Unc5a* (Rajasekharan and Kennedy, 2009) (downregulated genes involved in dendritic spine formation and maintenance) (Figure 5D) could contribute to the cellular and synaptic alterations in *Setd1a*^{+/-} mutant mice.

The correlation between PN-specific expression and proximity to Mef2/Setd1a-bound intergenic peaks immediately suggests that many of these genomic regions might represent pyramidal-neuron specific regulatory elements. To test this, we FAC-sorted PNs and INs from double transgenic mice (from either *Emx1-ires-Cre* or *Parvalbumin-ires-Cre* lines crossed to *tdTomato* reporter mice) and analyzed them by ATAC-Seq. This approach confirmed that Setd1a-bound intergenic peaks are accessible and occupied by TFs specifically at PNs (Figures 5E). As an example, the *Tnem4* gene locus is shown in more detail in Figure 5F. This gene is expressed in PFC PNs and has been found to be up-regulated in *Setd1a*^{+/-} mice, with Setd1a bound at its enhancers but not at its promoter. We identified one enhancer element (regions marked by H3K27ac but not H3K4me3) bound by both Setd1a and Mef2c specifically in PNs and not in INs, as shown by cell-type specific ATAC-Seq.

Reinstating *Setd1a* expression in adulthood rescues WM deficits.

The lack of gross structural or neurodegenerative abnormalities in *Setd1a*-deficient brains prompted us to investigate the possibility of reversing cognitive deficits in adulthood. We took advantage of the conditional nature of the targeted allele and adopted a genetic method that allows for inducible reactivation of *Setd1a* in *Setd1a*^{+/-} adult mice. To achieve temporal control, we crossed a *Setd1a*^{+/-} mice to an inducible *R26^{FlpoER}* mouse line (*R26FlpoER*) that activates global Flpo function upon tamoxifen (TAM) treatment (Figures 6A). The *Flp/Frt*-dependent genetic switch strategy enables the conditional manipulation of *Setd1a* at its endogenous locus, preserving physiological expression. When the *Setd1a*^{+/-}:*R26FlpoER*^{+/-} mice reached 8 weeks, we used oral gavage to deliver TAM and implement Flpo-mediated gene restoration (Figures 6A and 6B).

Behavioral analysis was performed on the following 4 groups: *Setd1a*^{+/+}:*R26FlpoER*^{+/-} mice treated with TAM (WT + TAM), *Setd1a*^{+/-}:*R26FlpoER*^{+/-} mice treated with TAM (*Setd1a*^{+/-} + TAM), *Setd1a*^{+/+}:*R26FlpoER*^{+/-} mice treated with corn oil vehicle (WT + oil), and *Setd1a*^{+/-}:*R26FlpoER*^{+/-} mice treated with vehicle (*Setd1a*^{+/-} + oil). Investigation of the efficiency of *Setd1a* reactivation upon Flpo-mediated deletion of the *LacZ* cassette confirmed that upon TAM treatment *Setd1a* mRNA and protein levels were reinstated in the adult brain of *Setd1a*^{+/-} + TAM mice to ~75% of WT levels (Figures 6C, 6D, S7A, S7B, and S7C). The effect of elevation of *Setd1a* levels on spatial WM performance was assessed in a T-maze DNMT task. As expected, *Setd1a*^{+/-} + oil mice showed impaired WM performance compared to WT + oil control littermates, which was partially rescued upon TAM treatment (Figures 6E, S7D, and S7E). Rescue was particularly evident at 10s in *Setd1a*^{+/-} + TAM mice (N = 24) compared to WT + TAM littermates (N = 24), or *Setd1a*^{+/-} + oil (N = 21) ($P = 0.35$ versus WT + TAM; $P = 0.0012$ versus *Setd1a*^{+/-} + oil).

Partial restoration of cognitive deficits might have resulted from the incomplete reinstatement of *Setd1a* expression or could reflect a partial developmental contribution requiring additional earlier interventions. These limitations notwithstanding and consistent with continued plasticity in the adult brain (Mei et al., 2016), our findings indicate that adult *Setd1a* function reinstatement has substantial effects on ameliorating cognitive deficits highlighting a broad window of therapeutic opportunity (Neale et al., 2012). We explored this finding further employing the more homogeneous biodistribution of systemically administered pharmacological compounds.

Pharmacological Inhibition of LSD1 demethylase activity in adulthood counteracts the effects of *Setd1a* deficiency.

Within the past few years, a number of histone demethylases have been discovered, as well as an ever-increasing number of potent and selective histone demethylases inhibitors with good brain penetration. We surmised that these compounds would most likely generate robust methylation changes, which could counteract effects of *Setd1a* deficiency. *Setd1a*-counteracting demethylase(s) have yet to be described. We hypothesized that LSD1 may be one such demethylase since both our whole tissue and scRNA-Seq analysis identified LSD1 as the most abundant demethylase in the frontal cortex. To test whether LSD1 binds *in vivo* the same regulatory elements bound by *Setd1a* we performed LSD1 ChIP-Seq from the PFC

of 6 weeks old mice, identifying 43,917 LSD1 bound regions (IDR<0.05). Strikingly, as for Mef2, we observed a compelling overlap: 22,532 out of 25,658 Setd1a-bound regions are co-bound by LSD1, both at promoter and enhancer elements (Figures 7A, 7B, and 7C, the *Neurod6* locus is shown as an example in Figure 7A). Notably, in agreement with previous observations that Mef2 activity can be inhibited by lysine methylation (Choi et al., 2014), we showed that LSD1 enhances Mef2c-mediated transcriptional activation in transient transfection assays (Figure 7D), an effect opposite to the one observed for Setd1a (Figure 5C), supporting the view that Setd1a and LSD1 can modulate Mef2 activity in opposite directions independently of their effects on histones.

To confirm the ChIP-Seq findings we undertook parallel phenotypic screens on primary cultures of *Setd1a*^{+/-} cortical neurons using well-characterized demethylase inhibitors (Figures 7E and 7F). *Setd1a*^{+/-} neurons were treated with inhibitors of LSD1/KDM1a [RN-1, tranylcypromine (TCP), ORY-1001, and SP2509], KDM5B (PBIT) and the pan-KDM inhibitor IOX-1 from DIV3 to DIV5, transfected with plasmids expressing EGFP at DIV4 and fixed at DIV5. TCP, ORY-1001, SP2509, and RN-1 (all LSD1 inhibitors), but not PBIT or IOX-1 rescued the axon branching deficits in *Setd1a*^{+/-} neurons, confirming an important role of LSD1 antagonism in counteracting the effects of *Setd1a* deficiency (Figures 7E and 7F).

We followed up the LSD1 findings by examining the effect of inhibition of LSD1 on cognitive deficits following systemic TCP administration (Figures 7I, S7F, S7G, S7H, S7I, and S7J). *Setd1a*^{+/-} mice and WT littermates were injected with 3 mg/kg/day TCP or vehicle for 14 consecutive days followed by the non-match to sample T-maze task. During the task, mice were administered daily with the same dose of TCP or vehicle. Assessment of spatial WM performance revealed modest genotypic effects during training (Figures S7F and S7G), but a robust impairment in the performance of vehicle-treated mutant mice over vehicle treated WT control littermates during testing with increasing time delays (Figure 7I). This deficit was completely reversed by TCP treatment (Fig. 7I). TCP had no effect on WT mice. Follow-up dose-response studies indicated that the effective TCP dose range was rather narrow: a lower dose (1 mg/kg/day, i.p.) was ineffective in improving WM performance (Figures S7H, S7I and S7J) whereas a higher dose of TCP (10 mg/kg/day, i.p.) resulted in occasional lethality precluding further assessment.

We asked whether *in vivo* post-developmental LSD1 antagonism by TCP can repair axonal terminal branching parallel to the rescue of the cognitive deficits. To address this question, we injected adeno-associated virus rAAV2/hSynapsin/EGFP into the frontal cortex of 12–14-week-old mice, following the T-maze test, to label cortical neurons with EGFP in TCP-or vehicle-treated mutant and WT mice. Mice were sacrificed 2 weeks later. We analyzed EGFP intensity in contralateral axon target regions as a quantitative index of the complexity of terminal axon branching. As expected mutant mice treated with vehicle showed significantly reduced callosal axon-terminal branching within the contralateral cortical L2/3 (Figures 7G and 7H), which was restored to near WT level by TCP treatment (Figures 7G and 7H).

While TCP is a prototype irreversible LSD1 inhibitor, it also inhibits two major isoforms of monoamine oxidases MAO-A and MAO-B. To eliminate confounding effects of MAO inhibition on WM we also evaluated the effect of ORY-1001, a TCP-derivative that represents the most potent and selective LSD1 inhibitor described to date (Hosseini and Minucci, 2017; Maes et al., 2015; Maes et al., 2018). The compound is an irreversible inhibitor and has an IC50 of 18 nM and an over 1000-fold selectivity for LSD1 compared to MAOs (Maes et al., 2015; Maes et al., 2018). Indeed, in our in vitro axon branching rescue assays ORY-1001 was 1000-fold more potent than TCP and demonstrated sub-nanomolar cellular efficacy (Figure 7F). Mutant mice and WT littermates were administered with ORY-1001, or vehicle (PBS) for 14 consecutive days prior to the T-maze task and daily throughout the duration of the task. As expected, we observed a robust impairment in the performance of vehicle-treated mutant mice over vehicle treated WT control littermates during testing with increasing time delays. This WM deficit was reversed by 10 µg/kg/day ORY-1001 treatment of *Setd1a*^{+/-} mice (Figure 7J). Days to criterion and percentage of correct choices during training phase were also improved in both ORY-1001-treated *Setd1a*^{+/-} and WT mice (Figures S7K and S7L). Notably, administration of lower of ORY-1001 (3 µg/kg/day) proved ineffective in improving the WM performance (Figures S7M, S7N, and S7O).

Discussion

The discovery of highly penetrant genetic risk variants for SCZ has provided unprecedented insights into disease pathophysiology (Arguello and Gogos, 2012; Del Pino et al., 2018). A major remaining challenge is how to begin to translate these insights into effective therapies. In that context, this study describes a mouse model of *SETD1A* LoF SCZ-risk mutations and evaluates *how* a bona fide pathogenic mutation, which disturbs a relatively ubiquitous molecular process, affects the structure and function of neural circuits and whether these effects are reversible in the adult brain. We describe a profile of robust but circumscribed synaptic and cognitive alterations due to *Setd1a*-haploinsufficiency. Overall, *Setd1a*^{+/-} phenotypes relate well to established evidence that SCZ patients have reduced WM capacity (Arguello and Gogos, 2012) as well as to cognitive and synaptic alterations described in other models of highly penetrant SCZ-risk mutations (Crabtree et al., 2017; Fenelon et al., 2013; Koike et al., 2006; Mukai et al., 2008; Stark et al., 2008). Comparison among strains suggests a pattern of altered connectivity and coherence within and between brain regions, shaped by alterations in rapid synaptic dynamics, altered intrinsic excitability and sparser connectivity, which may underlie the emergence of behavioral alterations (Mi et al., 2017; Hamm et al., 2017; Tamura et al., 2016; (Gogos et al., 2019).

Our ChIP-Seq analysis, revealed a large number of *Setd1a* direct chromatin targets, distributed over enhancers and promoters. *Setd1a*-bound enhancers, in particular, appear to have a cortex- and neuronal-specific function both in mouse and human, suggesting that the molecular mechanisms that lead to the recruitment of *Setd1a* to enhancers are conserved. This finding strongly implies the presence of auxiliary factors that may mediate the brain region, cell-type-specific and possibly developmental effects of this relatively ubiquitously expressed regulator. Indeed, our ChIP-Seq and IP analysis revealed a key role for Mef2 TFs (Harrington et al., 2016; Mitchell et al., 2017) in mediating the effects of *Setd1a* enhancer

occupancy, inviting to speculate that Setd1a and Mef2, when co-expressed, may be working in concert to regulate gene expression. Notably, genes with Setd1a-bound enhancers show a highly specific overlap with lists of SCZ and NDD risk genes suggesting that *Setd1a* deficiency in humans may lead to physiological deficits in part via coordinated mis-regulation of a network of disease-risk genes.

Reduction of Setd1a levels by half may lead to low magnitude changes in the expression of many downstream targets with likely cumulative effects on cellular structure and function, especially in PNs where Setd1a enhancer-bound targets show the highest expression. We highlighted the transcriptional dysregulation of some of these PN-specific genes as potentially contributing to the phenotypic alterations in *Setd1a*^{+/-} mice. Due to technical limitations, a more definitive identification of culprit target genes as well as a better correlation between chromatin modifications and target gene expression will likely require use of the conditional allele in our Setd1a mouse line to implement spatially-restricted homozygous inactivation of *Setd1a*. One unexpected finding from our analysis, is that reduction of Setd1a levels by half results to a large extent in upregulation of direct target genes in addition to predicted downregulation. It is possible that Setd1a may have other targets as a methyl-transferase, in addition to catalyzing methylation of H3K4. For example, recent data suggest that Mef2 can be regulated by G9a-dependent lysine methylation, which inhibits its activity (Choi et al., 2014; Ow et al., 2016). In that vein, Setd1a may also methylate Mef2, restraining its activity in some Mef2/Setd1a-cobound enhancers. Consistently, we provide evidence that Setd1a inhibits Mef2c-mediated transcriptional activation in transient transfection assays, offering a possible explanation for the transcriptional upregulation of genes with Setd1a-bound enhancers upon deletion of one Setd1a allele.

One important aspect of the current study was the demonstration that specific cognitive and neural circuitry impairments in the Setd1a-deficient brain can be overcome by genetic or pharmacological interventions aiming at delayed restoration of Setd1a function. Similar analysis in other models of monogenic brain disorders has provided mixed results (Barnes et al., 2015; Rotaru et al., 2018; Tillotson et al., 2017). In particular, a series of pharmacological interventions that directly targeted downstream demethylation processes showed that LSD1 antagonism in *Setd1a*-deficient mice results in a full rescue of the behavioral abnormalities and axonal arborization deficits. While it will be important to determine the specific methylation sites underlying the beneficial effects of LSD1 antagonism, our findings indicate that reactivating Setd1a function or counteracting downstream effects of *Setd1a* deficiency in the adult brain may be sufficient to ameliorate the manifestation of some neurocognitive phenotypes. Therapies targeted to SETD1A may have wider implications as suggested by the fact that pathways related to histone methylation are strongly enriched among biological pathways in GWAS data of SCZ (Network and Pathway Analysis Subgroup of Psychiatric Genomics, 2015).

STAR Methods

LEAD CONTACT AND MATERIALS AVAILABILITY

All unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement. Information and requests for resources and reagents should be directed to, and will be fulfilled by the Lead Contact Joseph A. Gogos (jag90@columbia.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Setd1a^{tm1a(EUCOMM)Wtsi} mice (referred to as *Setd1a^{+/-}*) were obtained from EMMA (<https://www.infrafrontier.eu/search>). *Setd1a^{+/-}* mice were backcrossed for 10–11 generations in the C57BL/6J (The Jackson Laboratory, Bar Harbor, ME) background. B6N.129S6(Cg)-*Gt(ROSA)26Sor^{tm3(CAG-flpo/ERT2)Alj/J}* (referred to as *R26^{FlpoER}*), *Emx1tm1(cre)Krl/J*, *Palvtm1(cre)Arbr/J*, and *Gt(ROSA)26Sortm14(CAG-tdTomato)Hze/J* were purchased from Jackson laboratory. To achieve temporal control of *Setd1a* expression for the *Setd1a* rescue experiments, we crossed the *Setd1a^{+/-}* mice to an inducible *R26^{FlpoER+/+}* (Lao et al., 2012) mouse line that activates global *Flp* function after TAM treatment. B6.Cg-*Gt(ROSA)^{26Sortm1.3(CAG-tdTomato,-EGFP)Pjen/J}* (referred to as *RC::FLTG*) (Plummer et al., 2015) was purchased from Jackson Labs. To evaluate the TAM-inducible *Flp* recombination in neurons, we crossed the *RC::FLTG* mice to an inducible *R26^{FlpoER+/+}* (Lao et al., 2012) mouse line that activates global *Flp* function after TAM treatment. For ATAC-Seq experiment *Emx-Cre* mice (Jackson lab) and *Parv-Cre* mice (Jackson lab) were crossed to CRE inducible *tdTomato* mice (Jackson lab) so that PNs or parvalbumin expressing INs expressed the fluorescent protein *tdTomato*.

Animals were housed in groups of five per cage, under a 12-h light/12-h dark cycle, with controlled room temperature and humidity, and then housed singly for behavioral experiments. Food and water were provided *ad libitum*. All behavioral experiments were performed on adult male mice aged 8–14 weeks during the light cycle. All animal procedures were carried out in accordance with and approved by the Columbia University Institutional Animal Care and Use Committee.

METHOD DETAILS

Co-IP.—Nuclear extract was prepared from fresh cortex of 8-wk old mice using Nuclear Complex Co-IP Kit (Active Motif. Inc.) according to the manufacturer's instructions. For co-immunoprecipitation analysis, 600 µg each of nuclear extracts were incubated with 10 µg of rabbit anti-Setd1a antibody (Bethyl) or 10 µg of normal rabbit IgG (Jackson ImmunoResearch Inc.) at 4°C, overnight. The immuno-complex was adsorbed onto protein A-agarose (Sigma) beads by incubating one hour at 4 °C. For immunoblot analysis we used rabbit anti-Mef2a/c antibody (1:1000, Abcam) or anti-Setd1a antibody (1:500) as a primary antibody and mouse monoclonal anti-rabbit IgG light chain (HRP) (1:5000 and 1:8000, clone SB62a, ab99697, Abcam) as a secondary antibody.

Immunocytochemistry.—Primary neurons were fixed in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS) (pH 7.4) for 10 min at room temperature (RT), washed 3

times for 5 min in 1×PBS, permeabilized in 0.1% Triton X-100 in PBS for 5 min and washed again in PBS (3 x 5 min). Coverslips were blocked for 30 min in 3% bovine serum albumin supplemented with normalized serum and incubated for 1 hr in primary antibodies diluted into blocking buffer and washed in PBS (3 x 5 min). Prior to application of secondary antibodies, cells were blocked again in blocking buffer supplemented with 0.5M NaCl for 30 min at RT. Secondary antibodies were applied in blocking/0.5M NaCl for 30 min at RT. Following another round of washing, the coverslips were mounted onto SuperfrostPlus glass slides using Prolong Gold (Invitrogen).

Immunohistochemistry.—Experiments utilized brains from P8.5 or 6-weeks old males, perfused with PBS and 4% PFA, and post-fixed in PFA overnight. A vibratome was used to cut 60 µm coronal sections. Free-floating sections were blocked and permeabilized overnight at 4°C in PBTX (PBS/Triton X-100) containing 5% normal goat serum (PBTX / 5% NGS). Primary antibodies were diluted in PBTX / 5% NGS and incubated with sections for 24 hrs. at 4°C. Sections were washed overnight at 4°C in 1× PBS. Sections were incubated with secondary antibody (1:500, diluted in PBTX / 5% NGS) for 4 hrs. at RT. Following another round of washing in 1× PBS, sections were incubated with TO-PRO3 (1:2500 in PBS) for 30 min, washed in 1× PBS (4 x 15 min), and mounted on slides using Prolong Gold (Invitrogen).

Primary neuronal cultures and transfection.—Primary cultures were prepared from cortices dissected bilaterally from E17 embryos. Cortices were individually dissociated using fire-polished Pasteur pipettes, counted, and plated in DMEM supplemented with 0.5% penicillin/streptomycin and 10% FBS. Neurons for immunocytochemical analyses were plated on poly D-lysine coated 12 mm coverslips (2 x 10⁴ cells/coverslip). Following 1.5 hr incubation, the modified DMEM was replaced with neurobasal media supplemented with B27 and Glutamax (Invitrogen) in which neurons were cultured for 5 days. Post-hoc genotyping was performed on DNA extracted from tissue taken from each embryo. Transient transfections were carried out with Lipofectamine 2000 (Invitrogen) at DIV4 using the protocol suggested for neuronal transfection. The original media used to culture cells prior to transfection were replaced in each well/dish following 2 hr of incubation at 37°C with transfection reagents.

Administration of Kdm inhibitors into primary neuronal cultures.: KDM inhibitors were dissolved at 10 mM in 100% DMSO. The compounds were diluted to 0.1 nM - to 30 µM in fresh neurobasal media supplemented with B27 and Glutamax. Primary neuronal cultures were placed in neurobasal media with compounds at DIV3 and cultured for 48 hours. Transient transfections of pCAG-GFP were carried out with Lipofectamine 2000 at DIV4. Media with compounds used to culture cells prior to transfection were replaced in each well/dish following 2 hr of incubation at 37°C with transfection reagents .

Electroporation *in utero*.—Pregnant mothers with E15.5 embryos were anaesthetized, and a 1 inch incision was made and one uterine horn removed. A total of 0.1–0.5 ml DNA (1–2 mg/ml) was injected into one ventricle. Four to five 40–45 V pulses were applied, each

pulse lasting 50 ms, with a 1 s pause between each pulse using an ECM-830 square wave electroporator (BTX) as previously described (Mukai et al., 2015; Saito, 2006).

Virus injection.—rAAV2/hSynapsin/EGFP was obtained from Addgene. Viral titer was 3×10^{12} particles per ml. Mice (8–12 week old) were deeply anesthetized under isoflurane and placed in a stereotaxic frame. For analysis of contralateral axon terminal branching *in vivo*, Bregma and lambda were used for leveling and zeroing, and bilateral cranial windows were excised over S1 cortex, centered at 0.97 mm posterior, 2.9 mm lateral. Virus was administered to layer 2/3 by an injection pump at a rate of 100 nL/min via an elongated 10–40 μ m-diameter glass micropipette. 200 nl were infused at single injection site. For morphological analysis of spines in L2/3 neurons, Bregma and lambda were used for leveling and zeroing, and bilateral cranial windows were drilled over PFC, centered at 1.93 mm anterior, 0.25 mm lateral. Virus was administered to L2/3 by an injection pump at a rate of 100 nL/min via an elongated 5–10 μ m-diameter glass micropipette. 500 nl were infused at single injection site. Acrylic dental cement was used to cover the craniotomy after infusion.

Electrophysiology in medial PFC slice preparations.—Field recordings were performed on 8- to 10-week-old male *Setd1a*^{+/-} mutant mice and their WT littermates. Isoflurane was used to anesthetize mice that were then decapitated. After a skull incision, the brain was removed and placed in ice-cold dissecting solution, which contained 195 mM sucrose, 10 mM NaCl, 2.5 mM KCl, 1 mM NaH₂PO₄, 25 mM NaHCO₃, 10 mM glucose, 4 mM MgSO₄, and 0.5 mM CaCl₂ and was bubbled with 95% O₂/5% CO₂ mixture. Coronal brain slices (300 μ m) containing the prelimbic and infralimbic PFC were cut using a vibratome (VT1200S; Leica, Nussloch, Germany). The freshly cut medial PFC slices were immediately transferred to an interface chamber and allowed to recover for at least 2 h at 34–36°C. During all recordings, the slices were continuously perfused with artificial cerebrospinal fluid (aCSF) (bubbled with 5% CO₂/95% O₂) that had the following composition: 124 mM NaCl, 2.5 mM KCl, 1 mM NaH₂PO₄, 25 mM NaHCO₃, 10 mM glucose, 1 mM MgSO₄, and 2 mM CaCl₂. The aCSF was maintained at 34–36°C and fed by gravity at a rate of 2–3 ml/min. Field EPSPs (fEPSPs) were recorded via a glass microelectrode (3–5 M Ω) filled with aCSF and placed in L5 of the medial PFC (600–700 μ m from midline). The stimulation site was always aligned ~200 μ m away from the recording site along the axis perpendicular to the pial surface. Basic synaptic transmission was characterized at 0.033 Hz, with stimulation intensities of 1–24 V (pulse duration, 0.1 ms). Subsequent experiments were performed at the stimulus intensity that generated a fEPSP one-third of the maximum fEPSP obtained at 24 V. Short-term synaptic facilitation was induced using a paired-pulse protocol with ISIs of 20, 50, 100, 200, 400, and 800 ms. To assess STD, fEPSPs were evoked by using a 40-pulse train at 5, 20 and 50 Hz (pulse duration, 0.1 ms). LTP was induced by one 800-ms (40 pulses), 50-Hz train after a stable 10 min baseline and monitored during 15 min. Then, 15 min after the first tetanus, four additional 50 Hz trains (separated by 10 s) were applied. The fEPSPs were then monitored for 40 min. Fiber volley was quantified by measuring the amplitude of the first peak negativity of the field responses, and the fEPSPs were quantified by measuring the initial slope of the second peak negativity of the responses. Electrophysiological signals were

acquired using an extracellular amplifier (Cygnus Technologies) and pClamp 10 (Molecular Devices)

Whole cell recordings were performed on 8- to 10-week-old male *Setd1a*^{+/-} mutant mice and their WT littermates. Following brain slicing medial PFC slices were immediately transferred to a recovery chamber and incubated at room temperature in recording solution (ACSF) for a minimum of 1h before recording. At the time of recording, slices were transferred to a submerged recording chamber and continuously perfused with ACSF with the temperature maintained between 30°C-34°C. Whole-cell patch clamp recordings were made using borosilicate glass pipettes (initial resistance 3.0–5.5 MΩ). Series resistance (5–20 MΩ) was monitored throughout each experiment and cells exhibiting a >20% change in series resistance were discarded. The internal solution used contained (in mM): KMeSO₄ 145, HEPES 10, NaCl 10, CaCl₂ 1, MgCl₂ 1, EGTA 10, MgATP 5, Na₂GTP 0.5, pH 7.2 with KOH, adjusted to 290 mOsm with sucrose. Layer 2/3 PNs in medial PFC prelimbic and infralimbic regions were identified by morphology and only neurons exhibiting spike frequency accommodation within action potential trains were included in analysis (Madison and Nicoll, 1984). All data were acquired with a Multiclamp 700B amplifier (Molecular Device) in conjunction with pClamp 10 software at 10kHz and low-pass filtered (10 kHz Bessel).

For recordings assessing neuronal excitability, action potential firing was assessed in current-clamp mode in response to incremental (20 pA steps), depolarizing current injections of 500 ms duration. Bridge balance of series resistance was employed and recordings with series resistance >20 MΩ were rejected. The resting membrane potential was determined within the first 5 s of achieving whole-cell access. For current-step experiments, the resting membrane potential of all cells was adjusted to approximately -70 mV by injection of a small standing current. Input resistance was calculated in current-clamp mode at -70 mV from the quasi steady-state (after 450–500 ms) voltage response to a 40 pA hyperpolarizing current-step. The membrane time constant ($\tau_{\text{memb.}}$) was calculated by fitting a single exponential to the voltage response to hyperpolarizing current steps (for 20–60 pA steps, using the voltage response during the first 200 ms) employing the built-in Chebyshev routine in Clampfit. For action potential afterhyperpolarization (AHP) analysis, the 1st current step sweep producing at least two APs was used. As the interval between the 1st and 2nd APs was often very brief and complex (with the frequent presence of a “contaminating” after-depolarization waveform), the AHP after the 2nd AP was used for analysis. The AHP was quantified as the difference between the AP threshold voltage and the minimal trough value of the AHP. The current-step trace used for analysis was overlaid with its own dV/dt plot to enable more confident assignment of AP threshold voltages. For recordings assessing voltage-gated channel activation in response to voltage steps, cells were held at -70mV, 10mV steps were applied, and series resistance-related errors were partially corrected by using 70% prediction and 70% series resistance compensation. Voltage-activated currents are reported as the quasi steady-state currents at the end of a 100 ms voltage step.

Behavior analysis.

Open field: Mouse activity was monitored in an illuminated chamber equipped with infrared sensors to automatically record horizontal and vertical activity (Colbourn Instruments). Each mouse was placed in the center of the chamber and its activity was recorded and collected in 1-min bins for 1 h during day 1 and for a 30-min re-exposure 24 h later (day 2).

T-maze: 6 weeks old animals were trained on a delayed non-match to place (DNMTP) T-maze task. The behavior protocol began with 2 days of habituation during which animals freely explored the maze for 10 minutes, followed by 2 days of 28 shaping when animals were required to alternate between goal arms of the maze to receive food rewards. After shaping was completed, training and testing sessions were conducted. Each trial of the task consisted of a sample and choice phase. In the sample phase, a mouse ran down the center arm of the maze and was directed into one of the goal arms. The mouse returned to the start box where it remained for a delay of 10 seconds. In the choice phase, the mouse was required to select the arm opposite to that visited in the sample phase to receive a reward. Animals were given daily training of 10 trials until they reached criterion performance, defined as performance of at least 70% correct per day for 3 consecutive days. After criterion was reached, animals completed daily testing sessions composed of 20–25 trials for delays of 10 s, 30 s, and 60 s. In the case of rescue experiments following treatment with TAM, TCP, or ORY-1001, animals completed daily testing sessions composed of 12 trials for 10 s, 20 s, and 40 s delays.

Social memory-Direct interaction with juveniles: All mice were housed two to five in each cage and given *ad libitum* access to food and water. They were kept on a 12-h (6:00 to 18:00) light–dark cycle with the room temperature regulated between 21–23 °C. Behavioral tests were performed during the light cycle in a testing room adjunct to the mouse housing room, which minimizes any transportation stress. The social memory/direct interaction test with juveniles was adapted from previously described (Kogan et al., 2000). Immediately prior to the experimental sessions, 10–12-week-old *Setd1a*^{+/-} and WT littermates were transferred to the testing room and placed into individual cages, identical to the ones used for housing, where they were allowed to habituate to the new environment for 15 min. Male juvenile mice (C57BL/6J, 4–5-week-old) were also placed in the testing room in their home cages and allowed to habituate a similar amount of time. Testing began when a novel juvenile mouse was introduced to a cage with one of the adult experimental mice. Activity was monitored for 5 min (trial 1) and scored online by a trained observer blind to the genotype of the test mice for social behavior (anogenital and nose-to-nose sniffing, close following and allogrooming) initiated by the experimental subject, as described in Hitti et al. (Hitti and Siegelbaum, 2014). After an inter-trial interval of 1 h, the experimental mice were introduced with either the previously encountered mouse or a novel mouse again for 5 min (Trial 2). The time spent in social interaction during trial 1 was subtracted from the social interaction time during trial 2 to obtain the difference score.

Fear Conditioning: Experimentally naive mice were used for fear conditioning in plastic, sound-attenuating cabinets containing a house light. On the 1st day, mice were placed in chambers and received 2 pairings of a tone (30 s, 82 db) and a co-terminating shock (2 s, 0.7

mA). On the 2nd day, initially, mice were placed in the chambers for 6 min, in the absence of tone and shock, and freezing was scored for the whole session, corresponding to hippocampal-dependent contextual fear memory. 2 h later, the context and handling of the mice was changed to assess conditioned fear of the tone alone (cue fear memory). Freezing was assessed before and during the tone presentation to distinguish fear generalization and specific cue fear memory for cue.

Novel Object Recognition: Mice were given a single 10 min trial of habituation to the room, the testing arena and to the procedure over three consecutive days to reduce stress and prevent a neophobic response and consequently to promote the exploratory activity of mice toward the objects (days 1–3). On the two following days mice were given a 5 min trial of exposure to two identical objects (days 4, 5). A memory test trial was performed 1 h later where one of the familiar objects was replaced with a novel object. 24 h later mice were placed back in the arena and were given another 5 min memory test in which one of the identical objects they explored during the training trial was replaced with a novel one. The location of the novel object was randomized between trials and genotypes for both the 24 h and the 1 h novel object recognition test. A minimal exploration time for both objects during the test phase (~20 s) was used to ensure a similar exploration time of the two objects and between animals, independently of their individual exploratory activity. All trials were video recorded and analyzed blindly by two investigators. Results were expressed as a discrimination index, calculated as the amount of time each mouse spends exploring the novel object divided by the total exploration time for both objects.

Tamoxifen preparation and feeding.—We used a modified version of a TAM feeding protocol previously described in Mei, et al (Mei et al., 2016). TAM (Sigma, T5648) was dissolved in corn oil (Sigma, C8267) at 20 mg/ml by vortexing. Freshly prepared TAM was protected from light by aluminum foil. Animal feeding needles (Harvard Apparatus, 52–4025) were used for oral gavage. To avoid toxicity of TAM, the following dosages were used for adult animals (8 weeks): mice at 17–21 g body weight were fed 5 mg/day; mice at 22–25 g body weight were fed 6 mg/day; mice at 26–29 g body weight were fed 7mg/day; mice at 30–35 g body weight were fed 8mg/day. Adult animals were fed for 5 consecutive days followed by 2 days of rest. Animals were then fed for 5 more consecutive days followed by 3 weeks of rest. Corn oil was used as a control. Mice fed with TAM and with corn oil were housed separately to avoid contamination.

TCP preparation and injection.—Tranlycypromine (TCP, Sigma, P8511) was dissolved in PBS at 0.75 mg/ml by vortexing. Mice at 17–20 g body weight were injected 2.7 mg/kg/day; mice at 20–28g body weight were injected 3.0 mg/kg/day. Adult animals were injected for 14 consecutive days followed by a non-match to sample T-maze task. During the behavior task, mice were injected with the same dose at 6 pm daily.

ORY-1001 preparation and injection.—ORY-1001 (Cayman Chemical, 19136) was dissolved in PBS at 1mg/ml by vortexing. Mice at 17–20 g body weight were injected 9.0 µg/kg/day; mice at 20–28g body weight were injected 10.0 µg/kg/day. Adult animals were

injected for 14 consecutive days followed by a non-match to sample T-maze task. During the behavior task, mice were injected with the same dose at 6 pm daily.

RNA sequencing.—Total RNA was extracted from PFC of 6-week-old mice using RNeasy Mini Kit (Qiagen). We analyzed each four male *Setd1a*^{+/-} mice and their wild-type (WT) male littermates. Quality of RNA samples was assessed by Bioanalyzer (Agilent). Mean (\pm SD) RNA integrity number (RIN) was 9.3 (\pm 0.5) in the mutant mice and 9.7 (\pm 0.2) in WT. An RNA-seq library was prepared from ~400 ng of total RNA by using TruSeq RNA prep kit (Illumina). Messenger RNAs were enriched by poly-A selection. Libraries were sequenced on HiSeq2500/HiSeq4000 (Illumina) with single-end 100bp reads at Columbia Genome Center.

ChIP-Seq analysis.—Whole PFC was dissected from 6-week old wild-type (WT) mice. Dissected tissue was thoroughly minced on ice using razor blades, then fixed with EGS for 30 min and/or for 5 min at room temperature with methanol-free formaldehyde (Pierce, Thermo Fisher Scientific) diluted to 1% in PBS. Fixation was quenched by adding 1/10th volume of 1.25 M glycine. Fixed tissue was collected by centrifugation at 200 g for 5 min at 4°C, then washed twice with cold PBS. Washed tissue was flash-frozen with liquid nitrogen then cryofractured using a Covaris CryoPrep Impactor on power setting 6. Nuclei and chromatin immunoprecipitation were performed as in Markenscoff-Papadimitriou *et al.* (Markenscoff-Papadimitriou *et al.*, 2014): briefly, the cryofractured tissue was lysed in ChIP Lysis Buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% NP-40, 0.25% Sodium Deoxycholate, 0.1% SDS) in rotation for 40 min at 4°C. Nuclei were centrifuged (1700x g, 5 min, 4°C), and the pellet was resuspended in shearing buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8, 0.25% SDS), then sheared on a Covaris S2 sonicator (12 or 16 min, 2% Duty Cycle, Intensity 3, 200 cycles per burst, frequency sweeping). Sheared chromatin was centrifuged (10,000 g for 10 min at 4°C) to remove insoluble material. Sheared chromatin was diluted 5-fold with ChIP Dilution Buffer (CDB: 16.7 mM Tris-HCl pH 8.1, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS) and pre-cleared for two hours at 4°C with protein G dynabeads (Thermo Fisher Scientific). Each ChIP was set up with the appropriate antibody [10 μ g of anti-Setd1a (hSET1, Abcam, ab70378), 1 μ g of anti-H3K27ac, -H3K4me2 and -H3K4me3, 3 μ g of anti-H3K4me1, 5 μ g of anti-Mef2] and 10–20 μ g of cleared chromatin, incubating then overnight at 4°C. Protein G dynabeads were blocked overnight with 2 mg/ml yeast tRNA (Thermo Fisher Scientific), then added to antibody bound chromatin and rotated for 3 hr at 4°C. Bead-bound chromatin was washed 5 times with LiCl Wash Buffer (100 mM Tris-HCl pH 7.5, 500 mM LiCl, 1% NP-40, 1% Sodium Deoxycholate) and once with TE (pH7.5). DNA was eluted from beads by incubating at 65°C for 30 min with 25 μ L ChIP Elution Buffer (1% SDS, 0.1 M Sodium Bicarbonate). This elution was repeated and the combined elution fractions were incubated overnight at 65°C with proteinase K (Thermo Fisher Scientific) and RNaseA (Thermo Fisher Scientific). ChIP DNA was purified 1.8X of AMPure XP beads and eluted in 30 μ L of nuclease free water. Libraries were prepared with Ovation Ultralow System V2 1–16 (NuGEN technologies) according to manufacturers' recommendation. All samples were multiplexed (in batches of 6) and sequenced using Illumina HiSeq as 50 bp or 75 bp single

end. Two biological replicates for each CHIP-Seq for each TF or histone mark were sequenced. As control, 1% chromatin input was also sequenced.

Dissociation of PFC to single cells and scRNA-Seq analysis.—Cells were dissociated using the papain dissociation kit (Worthington) following the manufacturers' recommendation. In brief, the PFC of 2 WT and 2 *Setd1a*^{+/-} mice was dissected and transferred to ice-cold PBS. The PFC was cut in small pieces using a razor blade and dissociated with a papain dissociation system (Worthington Biochemical). The tissue was incubated for 1h at 37°C on a rocking platform with papain-EBSS (1 PFC/mL). The tissue was triturated 20 times and cells were pelleted (400 g, 5 min, RT). Remaining papain was inhibited by resuspending the cell pellet with Ovomuroid protease inhibitor solution diluted 1:10 in EBSS. An extra 5 ml of Ovomuroid protease inhibitor solution was slowly layered at the bottom of the cell suspension to create a discontinuous gradient and cells were pelleted (100 g, 6 min, RT). To enrich for living cells, we removed dead cells from single cell suspension using the MACS® Dead Cell Removal Kit (Miltenyi Biotec) with the MS Columns following manufacturer's recommendation. In brief, 200 µl per PFC of Dead Cell Removal MicroBeads were added the pelleted cells. Cell were resuspended and incubated for 15 min at RT. The MS columns were washed once with 0.5 ml of Binding buffer. After incubation is complete, the cell suspension was diluted with 500 µl 1X Binding Buffer and the suspension was applied to the MS columns. The flow-through was collected. To increase the number of live cells recovered, the column was washed twice with 1 ml of Binding buffer at RT. Flow-through were combined and live cell were pelleted (400 g, 5 min, RT). Cells were then washed twice with PBS containing 0.04% BSA and the final pellet was resuspended in 300 µl of PBS containing 0.04% BSA and passed through a 40 µm cell strainer. Cell were counted and viability was evaluated by trypan blue staining.

The scRNA-Seq was performed using the 10x Genomic platform and the “Single Cell 3' v2” by the Columbia Genome Center. The scRNA-Seq data were aligned using the 10x Genomics package Cell Ranger 2.1.1 on the mm10 mouse genome version. In particular, UMI reads from the different genotypes and biological replicates were counted using “cellranger count” and combined together using “cellranger aggr”. From the downstream analysis, we removed cells in which we detected less than 1000 genes expressed, using “Seurat” R-package (v. 3.0.1). This package was used for all downstream analysis. In details, counts were first normalized (using the normalization method: “LogNormalize”). Linear dimensionality reduction was then performed on the most variable genes after that the data were regress out for cell-cell variability (i.e. batch effect, cell alignment rate, number of detected molecules etc.). All significant principal components were calculated, selected (using 40 PCs) and used to defined cell clusters (with the function “FindClusters”). Clusters were then visualized through UMAP plots using the same number of PCs that were used to define clusters. The different cell types were finally defined by the known expression of genes that have been previously describe to mark specific subpopulation of neuronal and non-neuronal cells (Figures S6B, S6C, and S6D). Genes with average expression < 0.1 normalized UMI counts in the whole dataset were removed from all analyses performed.

Sorting of PFC PNs and parvalbumin expressing INs to identify hypersensitive genomic regions.

—Emx-Cre mice (Jackson lab) and Parv-Cre mice (Jackson lab) were crossed to CRE inducible tdTomato mice (Jackson lab) so that PNs or parvalbumin expression INs expressed the fluorescent protein tdTomato. The PFC of these mice was dissociated as described previously. After dissociation cells were resuspended in sort media supplemented with 100 U/mL DNase I (Worthington Biochemical), 4 mM MgCl₂, and 500 ng/mL DAPI (Invitrogen). These cells were passed through a 40 μm cell strainer, and then FAC sorted. Live cells were selected by gating out DAPI positive cells and positive cells were gated by tdTomato fluorescence. Fifty thousand cell were used to prepare ATAC-seq libraries as described in references (Buenrostro et al., 2015; Monahan et al., 2017). In brief, cells were pelleted (500 g, 5 min, 4°C) and resuspended in ice cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% IGEPAL CA-630). Nuclei were pelleted (1000 g, 10 min, 4°C) and resuspended in transposition reaction mix prepared from Illumina Nextera reagents (for 50 μL: 22.5 μL water, 25 μL 2xTD buffer, 2.5 μL Tn5 Transposase). Transposed DNA was column purified using a Qiagen MinElute PCR cleanup kit (Qiagen) and amplified using barcoded primers and NEBNext High Fidelity 2x PCR Master Mix (NEB). Amplified libraries were purified using Ampure XP beads (Beckman Coulter). ATAC-Seq data were aligned and processed as previously described for ChIP-Seq data.

Luciferase assay.—HEK293T cells were transfected in a 96well plate using Lipofectamine 3000 (ThermoFisher Scientific) following the manufacturer's instruction. We transfected 5ng of pRL Renilla vector, 50 ng of 3X Mef2-Luciferase vector (Addgene #32967), 75 ng of Mef2c CDS (Addgene #32515) and/or Setd1a CDS (Origene #MR215352), Lsd1 CDS (Origene #MR210741), pcDNA3.1 vector. Luciferase assay was performed using the kit Dual-Glo® Luciferase Assay System (Promega, #E2920) following the manufacturer instructions. Measurements were performed using the plate reader SpectraMax M3 (Molecular Devices).

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of axon branching, dendritic spines, dendritic complexity and neuronal numbers.—Quantification, image and statistical analyses cytoarchitecture were performed with LSM Image browser tool, Fiji (Schindelin et al., 2012) and Excel software, blind to the genotype. Quantifications throughout the study are depicted as mean values with error bars indicating the Standard Error of the Mean (s.e.m). The number of independent repeats (N), the statistical test used for comparison and the statistical significance (P values) are specified for each figure panel in the representative figure legend. P values in this manuscript present the following star code: P>0.05 (non-significant), *P < 0.05 , **P < 0.01, ***P < 0.001.

Quantification of axonal morphology in vivo: To analyze the axonal morphology quantitatively, we performed the following procedures for each dataset obtained from the 60-μm coronal sections. For the contralateral axon branch number at P8.5, images were taken with setting as z-series stacks (at 0.5 μm depth intervals) using 40× Zeiss objectives with constant parameters for brightness and contrast. EGFP expressed axons at contralateral layer 1–4 were analyzed by projecting the z series and manually tracing the axon using the

overlay tools of the LSM Image Browser program (Zeiss). Protrusions more than 6 μm long were categorized as axon branches. The axon terminal architecture of each neuron was also characterized by designating the order of a branch (primary, secondary and so on). Once the elaborations of the axon branch tree were outlined, branch numbers were counted manually and analyzed. We compared at least 20 axons from four brains of each genotype in three independent experiments using the Student's *t*-test.

Quantification of cortical axonal morphology in vivo in adult mice.: To analyze and visualize axon branching in adult brain, we performed the following procedures for each dataset obtained from 300- μm coronal sections. Images acquired for the EGFP fluorescence analysis were taken with sequential acquisition setting at $2,048 \times 2,048$ pixel resolution as z-stack (taken at intervals of 1.5 μm) from each sections of frontal cortex using Nikon eclipse Ti2 + W1-Yokogawa spinning disk confocal microscope. Z-stack images for whole contralateral axon column in the somatosensory cortex (S1) were projected by z-project mode with Sum Slices projection type of ImageJ Fiji software. The immunofluorescence in each layer was measured by ImageJ Fiji software and normalized by EGFP expressed cell number in ipsilateral layer 2/3 as described below. All settings of threshold were maintained constant across genotypes. To quantify the number of cells expressing EGFP at the injection site of ipsilateral S1 cortex, images were taken with sequential acquisition setting at $2,048 \times 2,048$ pixel resolution as z-stack (taken at intervals of 2.0 μm) from each sections of frontal cortex using Nikon eclipse Ti2 + W1-Yokogawa spinning disk confocal microscope. The number of cells expressed EGFP were counted using 3D Objects Counter of ImageJ Fiji software.

Quantification of axonal morphology in vitro: For *in vitro* analysis, a plasmid, pCAG-GFP, expressing EGFP driven by the β -actin promoter was transfected into primary neuron cultures to fill the cells and facilitate visualization of all processes. Confocal images of neurons from primary culture were obtained blind to genotype with the LSM510 using a Zeiss 20 \times objective with sequential acquisition setting at 2048×2048 pixels resolution. Images were taken with setting as z-series stacks (at 1.0 μm depth intervals) with constant parameters for brightness, contrast, and pinhole size. Neuronal processes stained for the axonal marker Tau-1 were traced using the overlay tools of the LSM Image browser program (Zeiss). Protrusions more than 8 μm long were categorized as axon branches. The axon architecture of each neuron was also characterized by designating the order of a branch (primary, secondary and so on). Once the elaborations of the axon branch tree were outlined, branch points were counted manually, and the length of primary axon (emanating directly from the soma) and its branches were analyzed. To evaluate the axon length and branching number, all EGFP-expressed cell, which were polarized with single axon (approximately 50 – 80 neurons/cover slip), were analyzed from at least 5 coverslips derived from 3 embryos for each genotype. using the Student's *t*-test.

Quantification of dendritic complexity: To analyze the dendritic complexity of the basal dendrites of L5 neurons in the PrL area of Thy1-GFP/M mice at 6-weeks, images were taken with setting as z-series stacks (at 0.5 μm depth intervals) using 40 \times Zeiss objectives with constant parameters for brightness and contrast. The complexity of the basal dendrites was

analyzed by projecting the z-series to maximum intensity and manually tracing the dendritic tree. Sholl analysis was conducted using concentric radii of 20 μm extending from the center of the soma. Intersections, branches and terminals were counted by distance from the soma and differences evaluated by repeated measures ANOVA.

Quantification of dendritic spines: To analyze the dendritic spine number, images were taken with setting as z-series stacks (at 0.3 μm depth intervals) using 60x Zeiss objectives with constant parameters for brightness and contrast. Z-series stacks were subjected to three-dimensional reconstruction using Zeiss LSM 510 Meta. Spine counts were performed on the second dendritic branch of intact neurons beginning at 25–40 μm away from the soma and extending for 30–60 μm from the origin. Dendritic protrusions were classified as spines (mushroom, long, stubby or thin) or filopodia, based on the parameters previously outlined (Mukai et al., 2008). Differences in the number of spines were evaluated using the Student's *t*-test. Size determination of mushroom spines was carried out using Zeiss LSM image browser overlay tools. The longest straight line in the spine head was counted as the head diameter. The length of the entire spine (including head and stalk) was measured using a bent-line tool and analyzed by Kolmogorov-Smirnov test.

Quantification of neuronal numbers in tissue sections: To quantitate the NeuN and PV positive cell number at 6-weeks, images were taken with setting as a single scan using 20x Zeiss objectives with constant parameters for brightness and contrast. To analyze the VIP-, SST-, and Reelin-positive cell number at 6-weeks, images were taken with setting as z-series (at 3 μm depth intervals) using 20x Zeiss objectives with constant parameters for brightness and contrast.

Electrophysiological analysis: Statistical analyses were performed using SigmaPlot 9.0, Microsoft Excel and GraphPad Prism 4. A two-way repeated-measures ANOVA was used to test differences between genotypes. For simple comparisons of single metrics, unpaired *t*-tests were employed. For *t*-tests, if there was a directional hypothesis, 1-tailed *t*-tests were performed. For AP responses across current steps, 2-way repeated measures ANOVA statistics were employed. In the corresponding figures, data are presented as means \pm s.e.m. *N* indicates number of animals and *n* indicates number of slices or cells from which recordings were made. Asterisks indicate significant differences. All recordings and the majority of data analyses were done blind to the genotype.

Behavioral analysis.—Statistical analyses were performed using SPSS 20.0 and GraphPad Prism 6. Statistical analyses for behavioral tests were performed using the Student's two-tailed *t*-test (T-maze, open field, novel object recognition memory, and fear conditioning) or two-way ANOVA and post hoc pairwise comparisons with Bonferroni correction (social memory). Quantifications throughout the study are represented as the mean value. Error bars indicate the Standard Error of the Mean (s.e.m).

Data processing and analysis of DE genes.—We used RTA (Illumina) for base calling and bcl2fastq (Illumina, version 2.17) for converting BCL to fastq format, coupled with adaptor trimming. We mapped the reads to the reference mouse genome (UCSC/mm10) using STAR ((Dobin et al., 2013) (version 2.5.4b). Read counts and normalization of the

read count data and analysis of differentially expressed genes (DE genes) were performed by using DESeq2 (Love et al., 2014) (version 1.10.2) with default parameters. By performing principle component analysis (PCA) with our data of eight PFC samples (each five *Setd1a*^{+/−} and WT) and read count data from olfactory bulb (OB) and cerebellum, we confirmed that our data of PFC form a single cluster distant from clusters of OB and cerebellum (data not shown).

ChIP-Seq data processing, peak calling and motif discovery.—Reads were aligned to the *Mus musculus* genome version mm10 using bowtie2 v2.2.6 (Langmead and Salzberg, 2012). Poor quality and duplicated reads were removed using respectively samtools v1.3.1 (Li et al., 2009) and Picard (<http://broadinstitute.github.io/picard>). To measure the reproducibility between biological replicates and to avoid arbitrary thresholding biases for both Setd1a and Mef2, we perform IDR following the ENCODE framework (Landt et al., 2012) using MACS v2.1.1 as peak caller (DOI: 10.1186/gb-2008-9-9-r137) and IDR 2.0.2 (Li et al., 2011). We defined good biological replicates when the ratio N1/N2 and Np/Nt was below 2. For all subsequent analysis, biological replicates were merged into single alignment files for each antibody using samtools v1.3.1 (Li et al., 2009) and difference in sequencing depth between the libraries was corrected by normalization for library size using HOMER v4.8.2 (Heinz et al., 2010). Setd1a bound promoter elements were defined as such when they were located within 1 Kb of annotated TSS or they overlapped with H3K4me3; Setd1a bound intergenic peaks were defined as such when they were located in unannotated regions of the genome at least 1 kb away from any annotated TSS; Setd1a bound intronic peaks were defined as such when they overlapped with annotated intron of the *Mus musculus* (version mm10) of the genome. *De novo* motif discovery analysis was performed on feature-separated *Mus musculus*, version mm10, of Setd1a peaks using HOMER v4.8.2 (Heinz et al., 2010) using automatically generated background (sequences randomly selected from the genome, matched for GC% content).

Setd1a binding in promoter, intergenic and intronic regulatory regions.—To identify the properties of Setd1a bound regions in the PFC, we compared the coverage of PFC ChIP-Seq for Mef2 and histone marks that we sequenced (see above). Heatmaps were generated calculating the normalized coverage per each ChIP-Seq using deeptools v2.5.0.1 (Ramirez et al., 2016), 2 kilobases around the center of Setd1a peaks into 50 bp bins. Density plots were generated calculating the coverage of the ChIP-Seq reads 2 kilobases around the center Setd1a peaks into 50 bps bins with deeptools v2.5.0.1 (Ramirez et al., 2016) and plotted using the R-package “ggplot2” and R v3.6 (<http://www.gbif.org/resource/81287>).

Brain region-specific Setd1a binding.—To investigate the Setd1a chromatin binding pattern between different tissues, we compared our PFC epigenetic H3K27ac landscape with different brain tissues using the ENCODE publicly available H3K27ac ChIP-Seq data from the olfactory bulb and the cerebellum (Shen et al., 2012). To make the data analysis comparable between datasets, we reprocessed the ENCODE data as described above. Heatmaps were generated as described above.

Conservation of Setd1a peaks in the Euarchontoglires clade.—Phastcons were obtained using the 59 vertebrate genomes (Phastcons 60ways) (Siepel et al., 2005). We selected data from the Euarchontoglires clade (including 21 vertebrate genomes) and the sources were downloaded from UCSC Genome browser via <http://hgdownload.cse.ucsc.edu/goldenpath/mm10/phastCons60way/>. Setd1a peaks were first normalized by size and heatmaps were generated as described above. Clusters were defined by hierarchical clustering of the data. For cluster 2 and 3 (Figure S3E), we split peaks in half: conserved and non-conserved half. Mef2 motif distribution was calculated using HOMER v4.8.2 (Heinz et al., 2010) with the function “annotatedPeaks” and plotted using the R package “ggplot2”. De novo motif analysis was performed with HOMER v4.8.2 (Heinz et al., 2010) automatic background.

Conservation of murine Setd1a peaks in the human genome.—To evaluate the conservation of Setd1a peaks between mouse and human, we submit the FASTA sequence of 12499 PFC specific Setd1a peaks to the online blast platform <https://blast.ncbi.nlm.nih.gov/Blast.cgi>. We used the megablast algorithm (McGinnis and Madden, 2004) to look for very similar sequences (>95% identity) between the selected species. We found 3813 conserved-Setd1a-bound regions between mouse and human. We performed *de novo* motif analysis on the hg38 version of the human genome using HOMER v4.8.2 (Heinz et al., 2010) and automatic generated background.

scRNA-Seq analysis.—The scRNA-Seq data were aligned using the 10x Genomics package Cell Ranger 2.1.1 on the mm10 mouse genome version. In particular, UMI reads from the different genotypes and biological replicates were counted using “cellranger count” and combined together using “cellranger aggr”. From the downstream analysis, we removed cells in which we detected less than 1000 genes expressed, using “Seurat” R-package (v. 3.0.1). This package was used for all downstream analysis. In details, counts were first normalized (using the normalization method: “LogNormalize”). Linear dimensionality reduction was then performed on the most variable genes after that the data were regress out for cell-cell variability (i.e. batch effect, cell alignment rate, number of detected molecules etc.). All significant principal components were calculated, selected (using 40 PCs) and used to defined cell clusters (with the function “FindClusters”). Clusters were then visualized through UMAP plots using the same number of PCs that were used to define clusters. The different cell types were finally defined by the known expression of genes that have been previously describe to mark specific subpopulation of neuronal and non-neuronal cells (Figures S6B, S6C, and S6D). Genes with average expression < 0.1 normalized UMI counts in the whole dataset were removed from all analyses performed.

ATAC-Seq analysis.—ATAC-Seq data were aligned and processed as previously described for CHIP-Seq data.

Human ATAC-Seq data to evaluate the accessibility of Setd1a conserved regions.—Downloaded ATAC-Seq data from postmortem human brain has been previously described (Fullard et al., 2018). The fastq file were realigned as described previously (see CHIP-Seq analysis section of the methods) on the hg38 genome.

Luciferase assay analysis.—Luciferase values were normalized against Renilla values to account for differences in transfection efficiency. Three biological replicates (each having three technical replicates) on different batch of HEK293 cells were generated. Each condition was normalized on empty vector values. Values are display in Figure 5C and Figure 7D as mean \pm standard error for three biological replicates. Student's unpaired t-test was used to compare groups and test for significant differences.

Interval-based enrichment analysis of Set1a target genes.—To enrich for functional elements in human, we identified conserved PFC-specific Setd1a peaks overlapping with ATAC-Seq from 5 different regions of the human brain (dorsal and ventral PFC, hippocampus, Medial dorsal Thalamus and putamen). By clustering these regions, we identified 1765 regions that are chromatin accessible in the human brain (Figure S4E). Finally, we used a permutation test implemented RegioneR (Gel et al., 2016) package in R to determine if there is any significant enrichment of these elements within genomic signals identified in previous GWAS studies of SCZ and other disorders.

Gene-based enrichment analysis of Set1a target genes among GWAS signals.—For the gene-based enrichment analysis, all conserved, brain-region specific and chromatin accessible ChIP-Seq peaks were first annotated using the PAVIS annotation tool according to their relative positions to known genes in human genome. We then utilized the INRICH analysis tool (Lee et al., 2012), a permutation-based approach that controls for potential biases caused by SNP density, gene density and gene size to analyze the overlap between the human orthologue targets and previously reported GWAS signals to determine if there is any significant overlap. For the INRICH analysis, the reference gene list was downloaded from the INRICH website (<https://atgu.mgh.harvard.edu/inrich/downloads.html>). The reference SNP list of GWAS studies (SNPs on a DNA microarray) was compiled from the table browser (hg19.snpArrayAffy6 and hg19.snpArrayIlluminaHumanOmni1_Quad) of the UCSC genome browser, which covers at least 95% of intervals analyzed. All coordinates were converted to hg38 using the LiftOver tool at the UCSC browser (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>). The default setting of 5000 permutation replicates was used for each trait tested. We prepared the various GWAS signals using the summary association data from the NHGRI (National Human Genome Research Institute) GWAS Catalog (Welter et al., 2014). Specifically, we first downloaded the list of SNPs associated with various human traits from the GWAS Catalog (<http://www.ebi.ac.uk/gwas>, the gwas_catalog_v1.0.1 file, accessed June 2016). For each phenotype, genome-wide significantly associated intervals were defined as follows using the threshold for linkage disequilibrium (LD) used in the SCZ GWAS by PGC ($r^2 > 0.6$) (Schizophrenia Working Group of the Psychiatric Genomics, 2014). We extracted autosomal SNPs with genome-wide significant association ($P < 5 \times 10^{-8}$) and with MAF > 0.05 in the 1000 Genomes Project March 2012 EUR dataset [downloaded as a part of the LocusZoom package (Pruim et al., 2010)] from the full dataset of GWAS Catalog (we denote these as “index SNPs”). An associated genomic locus for each index SNP was defined as the genomic region containing any SNPs that are: 1) in LD with the index SNP at $r^2 > 0.6$ and 2) within ± 10 Mb of the index SNP. SNPs in LD with the index SNP were identified by using PLINK (Purcell et al., 2007) with the 1000 Genomes Project March 2012 EUR dataset. As a

control, we also performed the same analyses on GWAS signals of 15 non-CNS diseases/traits that have similar associated interval size (from 74 intervals to 174 intervals) that are comparable to the size of SCZ GWAS intervals. This procedure of generating independent GWAS signals is a pre-requisite for the INRICH analysis. INRICH analysis was conducted using version 1.1 with its default settings.

Gene-based enrichment analysis of *Setd1a* target genes among LoF DNM-hit genes.—Gene-based enrichment analysis of the human target gene set of *Setd1a* (described above) among LoF DNMs was performed by using DNENRICH (Fromer et al., 2014), a statistical software package for calculating permutation-based significance of gene set enrichment among genes hit by de novo mutations accounting for potential confounding factors such as gene sizes and local trinucleotide contexts DNENRICH analyses were performed with one million permutations using the following input files: 1) gene name alias file: the default dataset included in the DNENRICH package, 2) gene size matrix file: the default dataset included in the DNENRICH package (for RefSeq genes), 3) gene set file: the set of 708 high-confidence *Setd1a* target genes (HGNC symbols available, not including *SETD1A*); all genes were equally weighted, 4) mutation list files: lists of genes hit by LoF DNMs in NDD (SCZ, ID/DD and ASD), congenital heart diseases (CHD) and control subjects in denovo-db (Turner et al., 2017) or Refs. (Xu et al., 2012; Xu et al., 2011; Yuen et al., 2017); all mutations were equally weighted, and 5) background gene file: the default gene list in the gene size matrix. The numbers of LoF DNMs in NDD, SCZ, ID/DD, ASD, CHD and controls were 1,874, 95, 986, 793, 173 and 202, respectively. For analyses restricting the input LoF DNMs to those in genes intolerant to LoF mutations, we used the two metrics, pLI scores (Lek et al., 2016) (threshold = 0.9) and LoF FDR (Petrovski et al., 2013) (threshold = 0.01). Genes for which pLI or LoF FDR was not available were excluded from the background gene list in the corresponding analyses considering pLI or LoF FDR.

GO enrichment and CSEA analysis of *Setd1a* target genes: Analysis of gene ontology (GO) terms enriched among *Setd1a* target genes was performed by using ToppGene with default parameters for four categories (Molecular Function, Biological Process and Cellular Component of GO terms, and Disease terms defined by DisGeNET). The enriched GO terms were visualized with REVIGO Gene Ontology treemap (Supek et al., 2011) (<http://revigo.irb.hr/>). To study the expression pattern of *Setd1a* targets, in the context human brain development, *Setd1a* target genes were imported into the CSEA tool (Dougherty et al., 2010) (<http://genetics.wustl.edu/jdlab/csea-tool-2/>). The output FDR *P* values of CSEA analysis were imported into the MultiExperiment Viewer (ver.4.9.0) as heatmap color code to generate the heatmap view. Genes that overlap with developmental stage and cortical signature genes were summarized from SEA analysis (pSI $p < 0.05$).

DATA AND CODE AVAILABILITY

RNA-Seq (Figure 5 and Figure S6), scRNA-Seq (Figure 5 and S6) and ChIP-Seq (Figure 3, Figure S4 and Figure 7) raw and processed data are available at GEO repository (GSE123652). SCZ GWAS data is publicly available from <http://www.med.unc.edu/pgc/results-and-downloads/>. All other disease GWAS data is public available from <http://>

www.ebi.ac.uk/gwas. LoF DNMs data is publicly available from <http://denovo-db.gs.washington.edu/denovo-db/>.

ADDITIONAL RESOURCES

None

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements.

We thank Dr. Sneha Rao and the Zuckerman Institute imaging facility for help with the in vivo axon imaging and Naoko Haremakei for the maintenance of the mouse colony and technical assistance. This work was partially supported by grants R01MH080234 (to J.A.G), R01DA036894 (to S.L).

References

- Barak O, and Tsodyks M (2014). Working models of working memory. *Curr Opin Neurobiol* 25, 20–24. [PubMed: 24709596]
- Barch DM, and Ceaser A (2012). Cognition in schizophrenia: core psychological and neural mechanisms. *Trends Cogn Sci* 16, 27–34. [PubMed: 22169777]
- Bastos AM, Loonis R, Kornblith S, Lundqvist M, and Miller EK (2018). Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proc Natl Acad Sci U S A* 115, 1117–1122. [PubMed: 29339471]
- Benamer N, Marti F, Lujan R, Hepp R, Aubier TG, Dupin AAM, Frebourg G, Pons S, Maskos U, Faure P, et al. (2018). GluD1, linked to schizophrenia, controls the burst firing of dopamine neurons. *Mol Psychiatry* 23, 691–700. [PubMed: 28696429]
- Bledau AS, Schmidt K, Neumann K, Hill U, Ciotta G, Gupta A, Torres DC, Fu J, Kranz A, Stewart AF, et al. (2014). The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Development* 141, 1022–1035. [PubMed: 24550110]
- Buenrostro JD, Wu B, Chang HY, and Greenleaf WJ (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* 109, 21.29.21–29.
- Chen X, Duan H, Xiao L, and Gan J (2018). Genetic and Epigenetic Alterations Underlie Oligodendroglia Susceptibility and White Matter Etiology in Psychiatric Disorders. *Front Genet* 9, 565. [PubMed: 30524471]
- Choi J, Jang H, Kim H, Lee JH, Kim ST, Cho EJ, and Youn HD (2014). Modulation of lysine methylation in myocyte enhancer factor 2 during skeletal muscle cell differentiation. *Nucleic Acids Res* 42, 224–234. [PubMed: 24078251]
- Crabtree GW, Sun Z, Kvajo M, Broek JA, Fenelon K, McKellar H, Xiao L, Xu B, Bahn S, O'Donnell JM, et al. (2017). Alteration of Neuronal Excitability and Short-Term Synaptic Plasticity in the Prefrontal Cortex of a Mouse Model of Mental Illness. *J Neurosci* 37, 4158–4180. [PubMed: 28283561]
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. [PubMed: 23104886]
- Dougherty JD, Schmidt EF, Nakajima M, and Heintz N (2010). Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. *Nucleic Acids Res* 38, 4218–4230. [PubMed: 20308160]
- Fahrner JA, and Bjornsson HT (2014). Mendelian disorders of the epigenetic machinery: tipping the balance of chromatin states. *Annu Rev Genomics Hum Genet* 15, 269–293. [PubMed: 25184531]

- Fenelon K, Xu B, Lai CS, Mukai J, Markx S, Stark KL, Hsu PK, Gan WB, Fischbach GD, MacDermott AB, et al. (2013). The pattern of cortical dysfunction in a mouse model of a schizophrenia-related microdeletion. *J Neurosci* 33, 14825–14839. [PubMed: 24027283]
- Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, Georgieva L, Rees E, Palta P, Ruderfer DM, et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. *Nature* 506, 179–184. [PubMed: 24463507]
- Fullard JF, Hauberg ME, Bendl J, Egervari G, Cirmaru MD, Reach SM, Motl J, Ehrlich ME, Hurd YL, and Roussos P (2018). An atlas of chromatin accessibility in the adult human brain. *Genome Res* 28, 1243–1252. [PubMed: 29945882]
- Genovese G, Fromer M, Stahl EA, Ruderfer DM, Chambert K, Landen M, Moran JL, Purcell SM, Sklar P, Sullivan PF, et al. (2016). Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. *Nat Neurosci* 19, 1433–1441. [PubMed: 27694994]
- Gogos JA, Crabtree G, and Diamantopoulou A (2019). The abiding relevance of mouse models of rare mutations to psychiatric neuroscience and therapeutics. *Schizophr Res*
- Goldman-Rakic PS (1995). Architecture of the prefrontal cortex and the central executive. *Ann N Y Acad Sci* 769, 71–83. [PubMed: 8595045]
- Gruter T, Wieschollock V, Dubovyk V, Aliane V, and Manahan-Vaughan D (2015). Altered neuronal excitability underlies impaired hippocampal function in an animal model of psychosis. *Front Behav Neurosci* 9, 117. [PubMed: 26042007]
- Gulsuner S, Walsh T, Watts AC, Lee MK, Thornton AM, Casadei S, Rippey C, Shahin H, Consortium on the Genetics of S., Group, P.S., et al. (2013). Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* 154, 518–529. [PubMed: 23911319]
- Hamm JP, Peterka DS, Gogos JA, and Yuste R (2017). Altered Cortical Ensembles in Mouse Models of Schizophrenia. *Neuron* 94, 153–167 e158. [PubMed: 28384469]
- Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38, 576–589. [PubMed: 20513432]
- Hitti FL, and Siegelbaum SA (2014). The hippocampal CA2 region is essential for social memory. *Nature* 508, 88–92. [PubMed: 24572357]
- Hosseini A, and Minucci S (2017). A comprehensive review of lysine-specific demethylase 1 and its roles in cancer. *Epigenomics* 9, 1123–1142. [PubMed: 28699367]
- Kogan JH, Frankland PW, and Silva AJ (2000). Long-term memory underlying hippocampus-dependent social recognition in mice. *Hippocampus* 10, 47–56. [PubMed: 10706216]
- Koike H, Arguello PA, Kvajo M, Karayiorgou M, and Gogos JA (2006). *Disc1* is mutated in the 129S6/SvEv strain and modulates working memory in mice. *Proc Natl Acad Sci U S A* 103, 3693–3697. [PubMed: 16484369]
- Kos MZ, Carless MA, Peralta J, Blackburn A, Almeida M, Roalf D, Pogue-Geile MF, Prasad K, Gur RC, Nimgaonkar V, et al. (2016). Exome Sequence Data From Multigenerational Families Implicate AMPA Receptor Trafficking in Neurocognitive Impairment and Schizophrenia Risk. *Schizophr Bull* 42, 288–300. [PubMed: 26405221]
- Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, et al. (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* 22, 1813–1831. [PubMed: 22955991]
- Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357–359. [PubMed: 22388286]
- Lao Z, Raju GP, Bai CB, and Joyner AL (2012). MASTR: a technique for mosaic mutant analysis with spatial and temporal control of recombination using conditional floxed alleles in mice. *Cell Rep* 2, 386–396. [PubMed: 22884371]
- Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291. [PubMed: 27535533]

- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. [PubMed: 19505943]
- Li Q, Brown JB, Huang H, and Bickel PJ (2011). Measuring reproducibility of high-throughput experiments. 1752–1779.
- Liu S, and Trapnell C (2016). Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Res* 5.
- Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550. [PubMed: 25516281]
- Madison DV, and Nicoll RA (1984). Control of the repetitive discharge of rat CA 1 pyramidal neurones in vitro. *J Physiol* 354, 319–331. [PubMed: 6434729]
- Maes T, Carceller E, Salas J, Ortega A, and Buesa C (2015). Advances in the development of histone lysine demethylase inhibitors. *Curr Opin Pharmacol* 23, 52–60. [PubMed: 26057211]
- Maes T, Mascaro C, Tirapu I, Estiarte A, Ciceri F, Lunardi S, Guibourt N, Perdones A, Lufino MMP, Somerville TCP, et al. (2018). ORY-1001, a Potent and Selective Covalent KDM1A Inhibitor, for the Treatment of Acute Leukemia. *Cancer Cell* 33, 495–511 e412. [PubMed: 29502954]
- Markenscoff-Papadimitriou E, Allen WE, Colquitt BM, Goh T, Murphy KK, Monahan K, Mosley CP, Ahituv N, and Lomvardas S (2014). Enhancer interaction networks as a means for singular olfactory receptor expression. *Cell* 159, 543–557. [PubMed: 25417106]
- McGinnis S, and Madden TL (2004). BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res* 32, W20–25. [PubMed: 15215342]
- Mei Y, Monteiro P, Zhou Y, Kim JA, Gao X, Fu Z, and Feng G (2016). Adult restoration of Shank3 expression rescues selective autistic-like phenotypes. *Nature* 530, 481–484. [PubMed: 26886798]
- Miller EK, Erickson CA, and Desimone R (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16, 5154–5167. [PubMed: 8756444]
- Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, and Shilatifard A (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci U S A* 98, 12902–12907. [PubMed: 11687631]
- Monahan K, Schieren I, Cheung J, Mumbey-Wafula A, Monuki ES, and Lomvardas S (2017). Cooperative interactions enable singular olfactory receptor expression in mouse olfactory neurons. *Elife* 6.
- Mukai J, Dhillia A, Drew LJ, Stark KL, Cao L, MacDermott AB, Karayiorgou M, and Gogos JA (2008). Palmitoylation-dependent neurodevelopmental deficits in a mouse model of 22q11 microdeletion. *Nat Neurosci* 11, 1302–1310. [PubMed: 18836441]
- Mukai J, Tamura M, Fenelon K, Rosen AM, Spellman TJ, Kang R, MacDermott AB, Karayiorgou M, Gordon JA, and Gogos JA (2015). Molecular substrates of altered axonal growth and brain connectivity in a mouse model of schizophrenia. *Neuron* 86, 680–695. [PubMed: 25913858]
- Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V, et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485, 242–245. [PubMed: 22495311]
- O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD, et al. (2012). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485, 246–250. [PubMed: 22495309]
- Petrovski S, Wang Q, Heinzen EL, Allen AS, and Goldstein DB (2013). Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet* 9, e1003709. [PubMed: 23990802]
- Plummer NW, Evsyukova IY, Robertson SD, de Marchena J, Tucker CJ, and Jensen P (2015). Expanding the power of recombinase-based labeling to uncover cellular diversity. *Development* 142, 4385–4393. [PubMed: 26586220]
- Proenca CC, Gao KP, Shmelkov SV, Rafii S, and Lee FS (2011). Slittrks as emerging candidate genes involved in neuropsychiatric disorders. *Trends Neurosci* 34, 143–153. [PubMed: 21315458]
- Quiquempoix M, Fayad SL, Boutourlinsky K, Leresche N, Lambert RC, and Bessaih T (2018). Layer 2/3 Pyramidal Neurons Control the Gain of Cortical Output. *Cell Rep* 24, 2799–2807 e2794. [PubMed: 30208307]

- Rajasekharan S, and Kennedy TE (2009). The netrin protein family. *Genome Biol* 10, 239. [PubMed: 19785719]
- Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, and Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44, W160–165. [PubMed: 27079975]
- Rodriguez-Murillo L, Gogos JA, and Karayiorgou M (2012). The genetic architecture of schizophrenia: new mutations and emerging paradigms. *Annu Rev Med* 63, 63–80. [PubMed: 22034867]
- Saito T (2006). In vivo electroporation in the embryonic mouse central nervous system. *Nat Protoc* 1, 1552–1558. [PubMed: 17406448]
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676–682. [PubMed: 22743772]
- Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, et al. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature* 488, 116–120. [PubMed: 22763441]
- Shilatifard A (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem* 81, 65–95. [PubMed: 22663077]
- Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 15, 1034–1050. [PubMed: 16024819]
- Singh T, Kurki MI, Curtis D, Purcell SM, Crooks L, McRae J, Suvisaari J, Chheda H, Blackwood D, Breen G, et al. (2016). Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nat Neurosci* 19, 571–577. [PubMed: 26974950]
- Stark KL, Xu B, Bagchi A, Lai WS, Liu H, Hsu R, Wan X, Pavlidis P, Mills AA, Karayiorgou M, et al. (2008). Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model. *Nat Genet* 40, 751–760. [PubMed: 18469815]
- Supek F, Bosnjak M, Skunca N, and Smuc T (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6, e21800. [PubMed: 21789182]
- Takata A, Ionita-Laza I, Gogos JA, Xu B, and Karayiorgou M (2016). De Novo Synonymous Mutations in Regulatory Elements Contribute to the Genetic Etiology of Autism and Schizophrenia. *Neuron* 89, 940–947. [PubMed: 26938441]
- Takata A, Xu B, Ionita-Laza I, Roos JL, Gogos JA, and Karayiorgou M (2014). Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. *Neuron* 82, 773–780. [PubMed: 24853937]
- Titley HK, Brunel N, and Hansel C (2017). Toward a Neurocentric View of Learning. *Neuron* 95, 19–32. [PubMed: 28683265]
- Turner TN, Yi Q, Krumm N, Huddleston J, Hoekzema K, HA FS, Doebley AL, Bernier RA, Nickerson DA, and Eichler EE (2017). denovo-db: a compendium of human de novo variants. *Nucleic Acids Res* 45, D804–D811. [PubMed: 27907889]
- Volk DW (2017). Role of microglia disturbances and immune-related marker abnormalities in cortical circuitry dysfunction in schizophrenia. *Neurobiol Dis* 99, 58–65. [PubMed: 28007586]
- Windrem MS, Osipovitch M, Liu Z, Bates J, Chandler-Militello D, Zou L, Munir J, Schanz S, McCoy K, Miller RH, et al. (2017). Human iPSC Glial Mouse Chimeras Reveal Glial Contributions to Schizophrenia. *Cell Stem Cell* 21, 195–208 e196. [PubMed: 28736215]
- Xu B, Ionita-Laza I, Roos JL, Boone B, Woodrick S, Sun Y, Levy S, Gogos JA, and Karayiorgou M (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet* 44, 1365–1369. [PubMed: 23042115]
- Xu B, Roos JL, Dexheimer P, Boone B, Plummer B, Levy S, Gogos JA, and Karayiorgou M (2011). Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat Genet* 43, 864–868. [PubMed: 21822266]
- Yuen, Merico D, Bookman M, J, L.H., Thiruvahindrapuram B, Patel RV, Whitney J, Deflaux N, Bingham J, Wang Z, et al. (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat Neurosci* 20, 602–611. [PubMed: 28263302]

Zucker RS, and Regehr WG (2002). Short-term synaptic plasticity. *Annu Rev Physiol* 64, 355–405.
[PubMed: 11826273]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

Cognitive and circuitry deficits in a mouse model of SETD1A, a schizophrenia risk gene

Reinstating Setd1a or antagonizing LSD1 activity in adulthood rescues deficits

Setd1a binds promoters and enhancers with a striking overlap with Mef2 on enhancers

Evolutionary conserved Setd1a-bound enhancers may regulate psychiatric risk genes

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

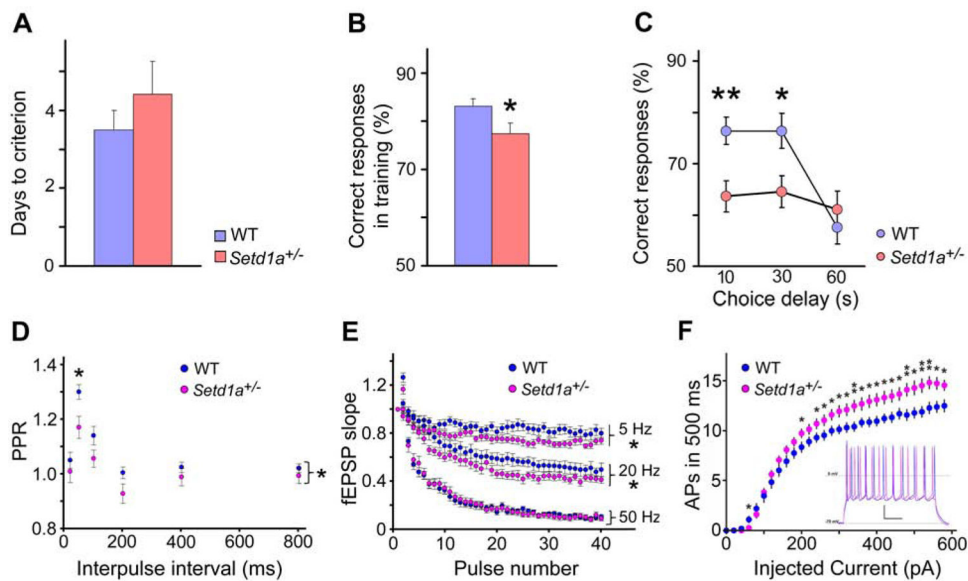


Figure 1. *Setd1a* deficiency impairs WM and alters neuronal short-term plasticity and excitability in PFC.

(A) Days taken to reach criterion on the DNMTTP T-maze task. $P = 0.36$.

(B) Correct responses during training on the DNMTTP T-maze task. $P = 0.041$.

(C) WM performance in a DNMTTP task. Mean percentage of correct responses for WT and *Setd1a*^{+/-} mice. Vertical axis starts at 50%, which represents response accuracy expected by chance. $P = 0.0038$ (10 sec), 0.0172 (30 sec), 0.485 (60 sec). * $P < 0.05$, ** $P < 0.01$.

Student's two-tailed *t*-test, $N = 12$ per genotype (A, B, C).

(D) PPR is significantly decreased in *Setd1a*^{+/-} mice when tested at various ISIs (20, 50, 100, 200, 400 and 800 ms), compared to WT mice (Bonferroni posttests revealed a significant difference at ISI of 50 ms, $P < 0.05$).

(E) Frequency-dependent differences of STD of the fEPSPs in WT and *Setd1a*^{+/-} mice at 5 Hz, 20 Hz and 50 Hz. At 5 and 20 Hz, the STD is significantly greater in *Setd1a*^{+/-} mice compared to WT controls. At 50 Hz, the STD is similar between genotypes. * $P < 0.05$, ** $P < 0.01$.

(F) Summary data of number APs evoked in response to 500 ms currents steps. WT recordings, blue; *Setd1a*^{+/-} recordings, magenta. Starred data represent results of pair-wise *t*-test comparisons at given current steps: * $P < 0.05$, ** $P < 0.01$. *Inset*; *Setd1a*^{+/-} neurons display enhanced excitability (magenta) compared to WT neurons (blue). Representative traces from current-clamp recordings from L2/3 PNs showing AP responses to near-maximal current step (500 pA, 500 ms). Scale bar: 20 mV, 100 ms. Reference levels, $V_{rest} = -70$ and 0 mV.

Histograms show means and error bars s.e.m.

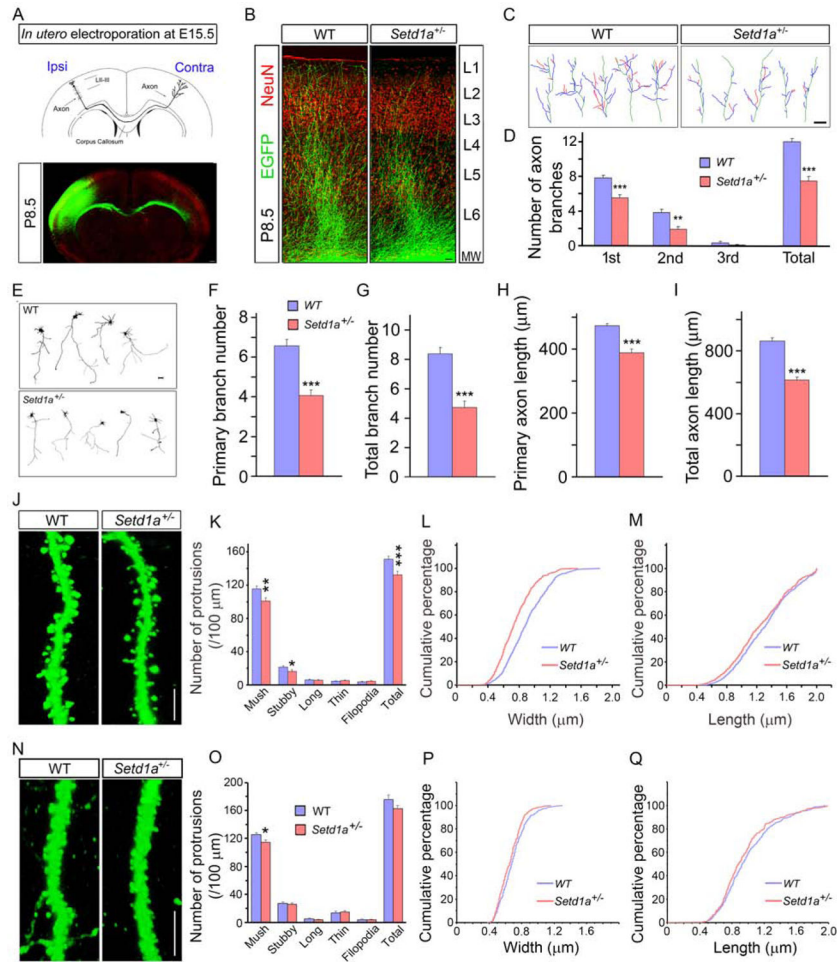


Figure 2. *Setd1a* deficiency disrupts axonal branching and spine density. (A) Illustration of the callosal axon pathway of neurons residing at L2/3, electroperated with EGFP at E15.5 and assayed at P8.5 (upper panel). Lower panel shows EGFP-labeled neurons (green) and NeuN (red). Representative images (lower panel) are composites of more than one image acquired from one brain section under identical scanning parameters. Scale bar, 100 μ m. (B) Representative images of contralateral axon terminal branching of EGFP-labeled neurons from coronal sections of P8.5 brains *in utero* electroperated at E15.5. Images are composites of more than one image acquired from one brain section under identical scanning parameters. Scale bar, 50 μ m. (C) Tracings of representative neurons in *in utero* electroperated brains expressing EGFP at P8.5, from similar sections as in (B). Depicted are 2D projections of the axon terminal in the contralateral cortical L1–4. Primary (blue), secondary (red), and tertiary (pink) branches from axons (green) are highlighted. Scale bar, 20 μ m. (D) Quantitative assessment of contralateral axon branching in L1–4 reveals a reduction of branch number in *Setd1a*^{+/-} mice. (E) Representative images of axonal deficits in *Setd1a*^{+/-} cortical neurons expressing EGFP at DIV5. Scale bar, 50 μ m.

(F-I) Reduction of (F) primary branch points in *Setd1a*^{+/-} neurons (N = 4, n = 315) compared to WT neurons (N = 4, n = 304, *P* < 0.001), (G) total branch points in *Setd1a*^{+/-} (N = 4, n = 315) compared to WT neurons (N = 4, n = 304, *P* < 0.0009), (H) primary axon length at DIV5 in *Setd1a*^{+/-} (N = 4, n = 315) compared to WT neurons (N = 4, n = 304, *P* < 0.0008), (I) total axon length in *Setd1a*^{+/-} (N = 4, n = 315) compared to WT neurons (N = 4, n = 304, *P* < 0.0003).

(J) Alterations in spines of basal dendrites of L5 PNs at the prelimbic area of medial PFC in *Setd1a*^{+/-} mice. Representative, high-magnification images of eGFP-expressing WT (*left panel*) and *Setd1a*^{+/-} neurons (*right panel*). Scale bar, 5 μm.

(K) Distribution of spine morphotypes, as well as total protrusions, in basal dendrites of L5 PNs from *Thy1-GFP*^{+/-}; *Setd1a*^{+/-} mice. Dendritic protrusions were classified as different spine categories (mushroom, long, thin, or stubby) or as filopodia, and counted.

(L and M) Reduction in the width (L), and length (M) of mushroom spines of L5 PNs from *Setd1a*^{+/-}; *Thy1-GFP*^{+/-} PFC relative to WT *Thy1-GFP*^{+/-}.

(N) Alterations in spines of basal dendrites of L2/3 PNs in the prelimbic area of medial PFC in *Setd1a*^{+/-} mice. Representative, high-magnification images of EGFP-expressing neurons from WT (*left panel*) and *Setd1a*^{+/-} (*right panel*) mice. Scale bar, 5 μm.

(O) Reduced density of mushroom spines estimated over 100 μm of dendritic length in basal dendrites of L2/3 PNs in *Setd1a*^{+/-} mice.

(P and Q) Reduction in the width (P) but not length (Q) of mushroom spines of *Setd1a*^{+/-} L2/3 PFC PNs relative to WT neurons.

P* < 0.05; *P* < 0.01; ****P* < 0.001. Student's two-tailed *t*-test (D, F-I, K, O). N: number of independent animals, n: number of neurons. Histograms show means and error bars s.e.m.

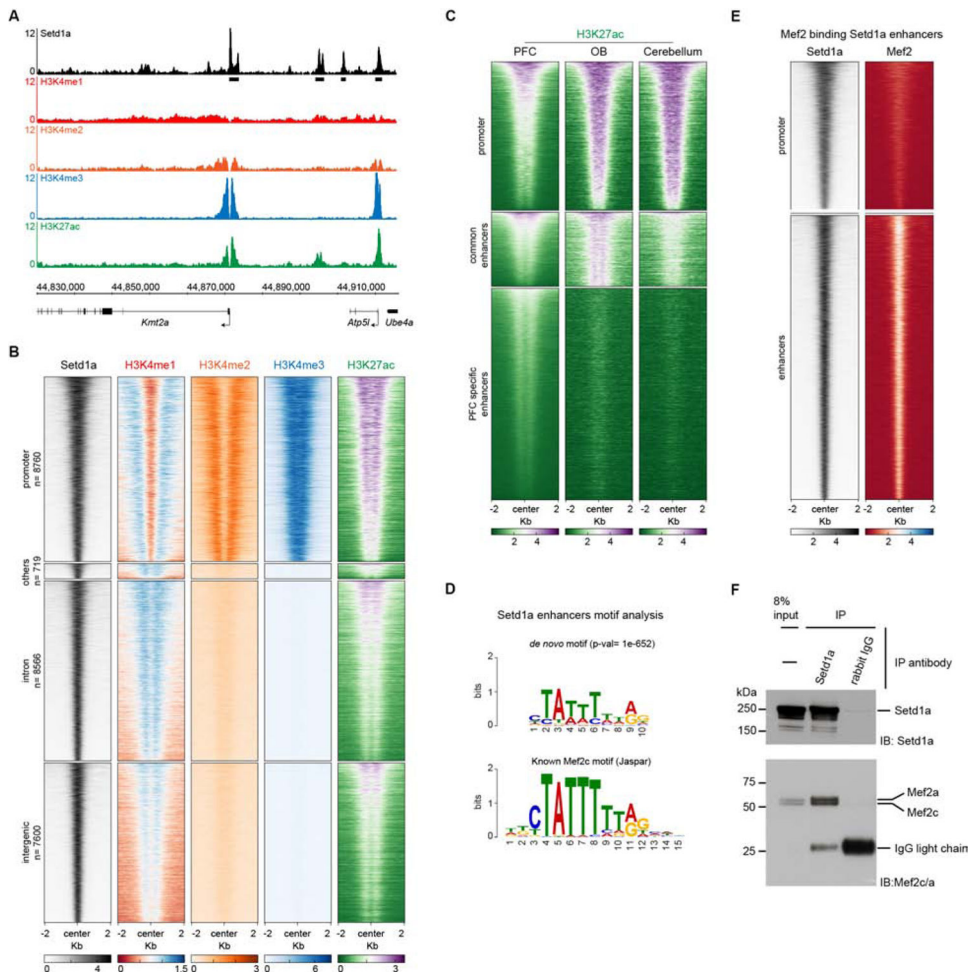


Figure 3. Setd1a binds promoters and enhancers in a complex with Mef2.

(A) Representative ChIP-Seq locus (*Kmt2a* gene). Normalized reads per genomic coverage from PFC ChIP-Seq for Setd1a (black) and chromatin marks (H3K4me1, red; H3K4me2, orange; H3K4me3, blue; H3K27ac, green) are shown. Black boxes below the Setd1a track depict significant peaks that passed all quality checks (see Methods).

(B) Heatmap of listed ChIP-Seq coverage from PFC, 2 kb upstream and downstream of the center of Setd1a peaks ($n = 25,658$). Setd1a peaks were clustered in promoters (*top panel*, defined by the overlap with H3K4me3 peaks) or enhancers (defined by absence of H3K4me3 signal and presence of H3K27ac and H3K4me1 signal) ($n=16,889$). Enhancers were sub-clustered by the annotation of Setd1a peaks. *Top-middle*, peaks in annotated 5'UTRs, coding sequence, transcription end site and 3'UTRs (defined as “others”) ($n=721$); *middle-bottom*, peaks in annotated introns ($n=8567$); *bottom panel*, peaks in intergenic regions ($n=7601$). All heatmaps are sorted in the same order, based on Setd1a and H3K27ac signal.

(C) Heatmap of H3K27ac 2 Kb around Setd1a peaks in PFC, OB and Cerebellum. By clustering, we distinguished 3 different clusters depending on H3K27ac coverage across tissues. While promoter and a subset of enhancers marked as “common enhancers” ($n=4390$) showed H3K27ac in all three analyzed tissues, we identified a third cluster of

enhancers, marked as “PFC specific enhancers” (n=12499), that showed H3K27ac signal only in the PFC.

(D) *De novo* TF motif analysis in Setd1a bound promoters (*top panel*) and enhancers (*bottom panel*). Setd1a bound promoters motif analysis showed enrichment (p -val= 1e-142) for a motif very similar to the known Sp1 motif (*bottom*). Instead, Setd1a bound enhancers showed enrichment (p -val= 1e-852) for the known Mef2c motif.

(E) Heatmap of Setd1a and Mef2 coverage 2 kb around the center of Setd1a peaks in promoters (*top panel*) and enhancers (*bottom panel*).

(F) Co-IP of Setd1a and Mef2a/c from the cortical plate of 6 week old mice. Western blot of 8% input (*left lane*), IP with Setd1a antibody (*middle lane*) or IgG (*right lane*) is shown.

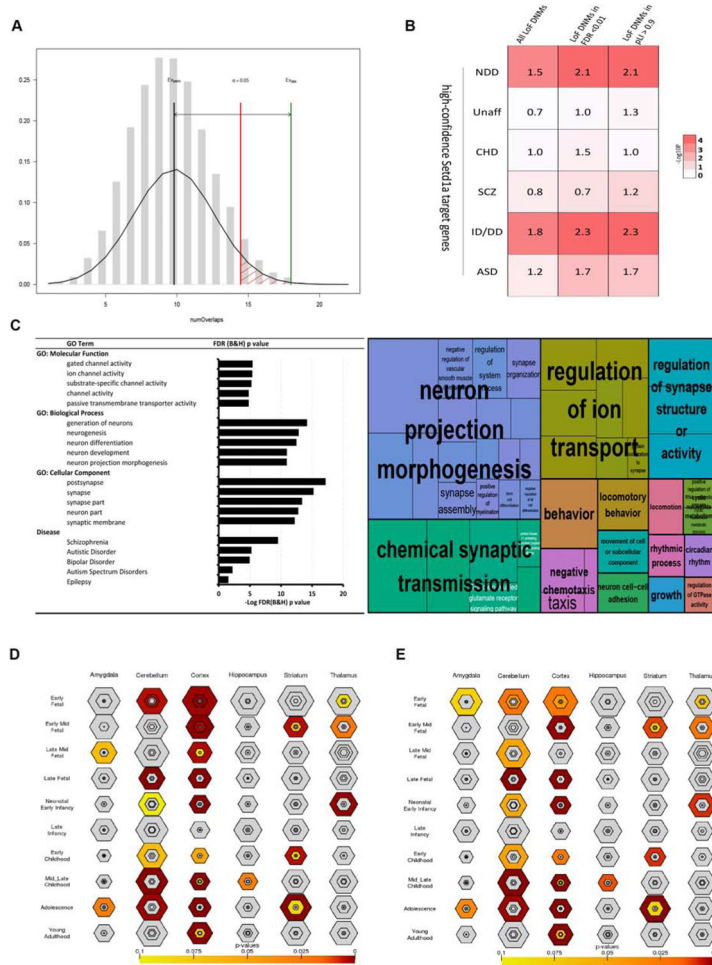


Figure 4. Setd1a targets are enriched for synaptic genes and overlap with established lists of SCZ and NDD risk genes.

(A) Permutation tests ($n=20,000$, randomization method: circularRandomizeRegions) indicate there is significant overlaps between conserved and chromatin accessible Setd1a ChIP-Seq peaks ($P=0.0055$, $Z\text{-score}=2.9$) and established SCZ-GWAS loci identified. X-axis is the number of overlaps, Y-axis is the density of expected number of overlaps determined by permutation. EV_{perm} , Expected average number of overlaps by permutation, EV_{obs} , actual observed number of overlaps. Permutation P value threshold = 0.05.

(B) DNENRICH analysis of the high-confidence Setd1a target genes. Fold changes (observed/expected number of LoF DNMs in the gene-set of interest) are shown in the cells color-coded by p values. Unaff: Unaffected subjects, CHD: congenital heart diseases, SCZ: schizophrenia, ID/DD: intellectual disability/developmental disorders, ASD: autism spectrum disorders, NDD: neurodevelopmental disorders (SCZ+ID/DD+ASD) [FC; P (pLI/LoF-FDR): ID/DD: 2.3/2.3; $1.0 \times 10^{-6}/1.0 \times 10^{-6}$, ASD: 1.7/1.7; 0.002/0.005 and SCZ: 2.3/3.1; 0.065/0.022, total N of LoF DNMs in ID/DD, ASD and SCZ = 986, 793 and 95, respectively].

(C) GO-term analysis of target genes with all conserved accessible Setd1a ChIP-Seq peaks annotated by PAVIS. Enriched GO terms annotated by Toppgene are visualized using REVIGO GO treemap (Supek et al., 2011).

- (D) Spatiotemporal expression pattern of target genes within conserved and accessible Setd1a peaks in human brain. P values are PSI P values as defined in the CSEA tool.
- (E) Spatiotemporal expression pattern of target genes within conserved and accessible Setd1a peaks that do not overlap with the intolerant genes hit by LoF DNMs in NDD defined in (B).

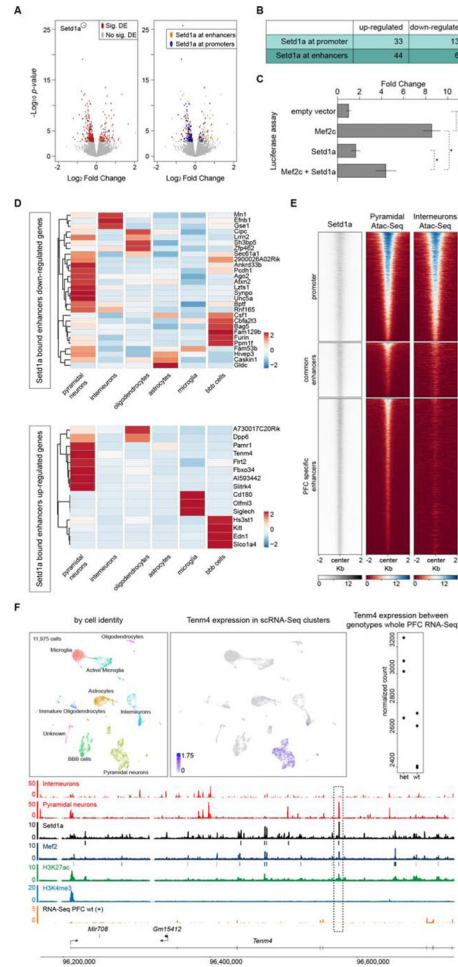


Figure 5. Differentially expressed genes in PFC of *Setd1a*^{+/-} mice and cell-type specific expression of *Setd1a* target genes.

(A-B) Volcano plots of differentially expressed genes in whole-PFC RNA-Seq of *Setd1a*^{+/-} mice are depicted in (A). 342 differentially expressed genes (FDR<5%) are shown in red, while non-significant changes are shown in grey (*left panel*). Of the 342 differentially expressed genes, 271 are *Setd1a* targets (orange and blue, *right panel*). We partitioned these genes in 4 categories based on the binding regions of *Setd1a* in the PFC. We defined ‘*Setd1a* bound promoters up-regulated genes’ and ‘*Setd1a* bound promoters down-regulated genes’ as genes in which promoters are bound by *Setd1a*. We defined ‘*Setd1a* bound enhancers up-regulated genes’ and ‘*Setd1a* bound enhancers down-regulated genes’ as genes in which enhancers, but not promoters, are bound by *Setd1a*. Genes with *Setd1a* binding at both promoters and enhancers, were partitioned into the “*Setd1a*-promoter bound” classes. A summary of the number of genes in each category is shown in (B).

(C) Luciferase reporter assay of *Mef2* responding element when *Mef2c* and/or *Setd1a* are co-transfected transiently in HEK293T cells. While *Mef2c* transfected alone showed 8.5-fold increase activity over the background, when *Mef2c* is co-transfected with *Setd1a*, we observed a significant decreased activity (4.5-fold over background). Values are mean \pm standard error for three biological replicates, normalized to the empty vector. Student’s t-test, * $P < 0.05$, ** $P < 0.01$.

(D) Heatmaps of the expression levels of down-regulated (*top panel*) and up-regulated (*bottom panel*) genes with Setd1a bound enhancers, across different cell types identified by scRNA-Seq (see methods and Figure S6C and S6D). For this analysis we combined the two different identified types of microglia and oligodendrocytes. Heatmaps show the normalized UMI counts per genes of two biological replicates per genotype.

(E) Heatmap of Setd1a ChIP-Seq coverage and ATAC-Seq of sorted PNs and INs, 2 kilobases upstream and downstream of the center of Setd1a bound promoters (*top*) and enhancers (*bottom*). Heatmaps are sorted in the same order, based on the Setd1a signal.

(F) *Tenm4* locus. UMAP representation of 11,975 single cell transcriptomes from the PFC of WT (5,944) and *Setd1a*^{+/-} (6,031) mice, colored by the identified cell-types (see Methods, Figure S5C and S5D), is shown on the *top-left panel*. *Tenm4* is specifically expressed in the scRNA-Seq PN cell cluster (*top-middle panel*) while whole PFC RNA-Seq showed that *Tenm4* is up-regulated in Setd1a mutant mice (*top-right panel*; each dot is the normalized expression of *Tenm4* in each biological replicate per genotype). Representative RNA-Seq data of one biological replicate of WT mice (*orange*), ChIP-Seq of Mef2 (*dark blue*), Setd1a (*black*), H3K27ac (*green*), H3K4me3 (*light blue*) as well as ATAC-Seq from PNs and INs (*light red*) are shown (*bottom panel*). Black and dark blue boxes below Setd1a and Mef2 tracks respectively, represent significant peaks that passed all quality checks (see Methods). Highlighted with a dashed square is the regulatory element that is bound by Setd1a and Mef2, marked by H3K27ac but not by H3K4me3, and accessible in PNs but not in INs.

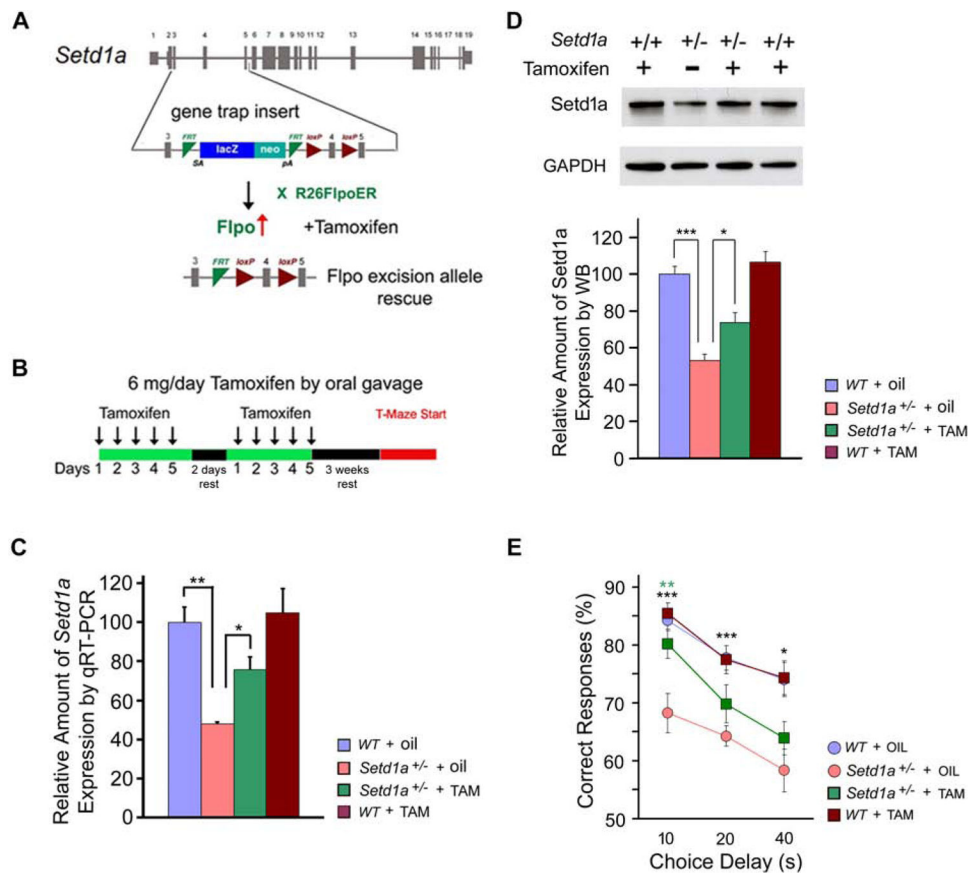


Figure 6. Restoration of *Setd1a* expression in adulthood rescues WM performance.

(A) Restoration of knockout-first allele after TAM treatment. For the rescue experiments, *Setd1a*^{+/-} mice were crossed to *R26FlpoER* mice. FLP (Flpo) induced globally by TAM converts the “knockout-first” allele to a conditional allele, restoring gene activity.

(B) Animals were fed TAM for 5 consecutive days followed by 2 days of rest, then fed for 5 more days followed by 3 weeks of rest before the T-maze assay.

(C) qRT-PCR of *Setd1a* expression in PFC of adult mice, demonstrating the restoration of *Setd1a* mRNA levels following TAM treatment. N = 6 mice each genotype, n = 3 each animal. **P* < 0.05; ***P* < 0.01. Student’s two-tailed *t*-test.

(D) Western blot analysis in lysates from PFC of adult mice, demonstrating the restoration of *Setd1a* protein levels by TAM treatment. N = 4 mice each genotype, n = 3 each animal. **P* < 0.05; ***P* < 0.01, ****P* < 0.001. Student’s two-tailed *t*-test.

(E) WM performance during testing phase. Mean percentage of correct responses for WT and *Setd1a*^{+/-} mice with/without TAM treatment. **P* < 0.05, ***P* < 0.01 (black asterisk: WT + oil versus *Setd1a*^{+/-} + oil, green asterisk: *Setd1a*^{+/-} + oil versus *Setd1a*^{+/-} + TAM).

Histograms show means and error bars s.e.m.

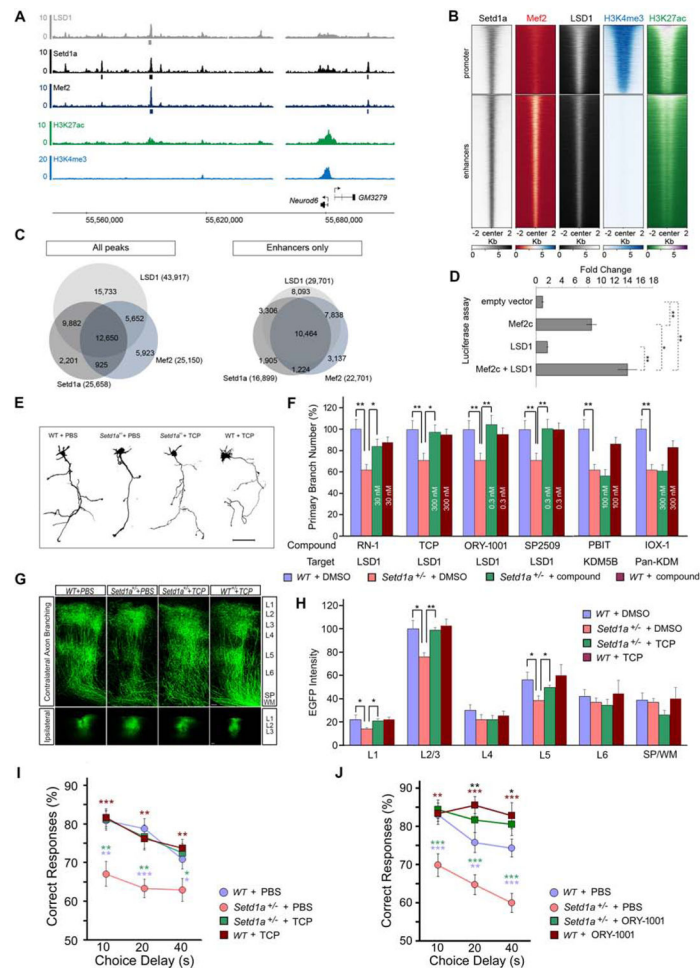


Figure 7. Inhibition of LSD1 activity in adulthood counteracts the effects of *Setd1a* deficiency.

(A) *Neurod6* locus. ChIP-Seq of LSD1 (grey), *Setd1a* (black), *Mef2* (dark blue), H3K27ac (green), H3K4me3 (light blue) is shown. Grey, black and dark blue boxes below *Lsd1*, *Setd1a* and *Mef2* tracks respectively represent significant peaks that passed all quality checks (see methods).

(B) Heatmap of the ChIP-Seq coverage for *Setd1a*, LSD1, *Mef2*, H3K4me3 and H3K27ac 2 kb upstream and downstream of the center of *Setd1a* peaks [promoters (*top panel*) and enhancers (*bottom panel*)]. Heatmaps are sorted in the same order, based on *Setd1a* and H3K27ac signals.

(C) Venn diagram of the overlap between all the peaks (*left panel*) and enhancers only peaks for LSD1, *Setd1a* and *Mef2* ChIP-seq (*right panel*).

(D) Luciferase reporter assay of the *Mef2* responding element when *Mef2c* and/or LSD1 are co-transfected in HEK293T cells. While *Mef2c* alone showed a 8.5-fold increase in activity over the background, *Mef2c* co-transfected with LSD1 results in a significant 14-fold increase in activity. Values are mean \pm standard error for three biological replicates, normalized to the empty vector (see Methods). Student's t-test, * $P < 0.05$, ** $P < 0.01$.

(E and F) *Setd1a*^{+/-} cortical neurons were treated with LSD1 or JmjC domain KDM inhibitors from DIV3 to DIV5. Tracings of representative neurons imaged at DIV5 following TCP or PBS treatment are depicted in (E). The most effective concentrations of each

inhibitor for restoring primary axon branch number in *Setd1a*^{+/-} neurons (0.1 nM – 30 μM) are represented in (F). There are no significant differences between WT + DMSO and WT + compound in any of the comparisons. **P* < 0.05, ***P* < 0.01. Scale bar, 50 μm.

(G) Representative images of contralateral axon-terminal branching of EGFP-labeled neurons from coronal sections of adult brains of mice injected AAV2Synapsine-EGFP virus at 8 weeks (*upper panels*) and viral injection sites in ipsilateral L2/3 (*bottom panels*).

Immuno-histological staining represents EGFP (*green*). Scale bar, 100 μm.

(H) Histograms depicting quantitative assessment of contralateral axon branching in L1 to L6 and white matter in adulthood and indicating that TCP treatment restores contralateral axon branching in L2 to L5 in *Setd1a*^{+/-} mice. WT + PBS: N = 6, *Setd1a*^{+/-} + PBS: N = 7, *Setd1a*^{+/-} + TCP: N = 6, WT + TCP: N = 5. **P* < 0.05, ***P* < 0.01.

(I) Complete rescue of WM performance in a DNMT1 task in *Setd1a*^{+/-} mice treated with TCP. Mean percentage of correct responses for WT and *Setd1a*^{+/-} mice with 3 mg/kg/day TCP or vehicle treatment are depicted. **P* < 0.05, ***P* < 0.01, ****P* < 0.001 (*blue asterisk*: WT + PBS vs *Setd1a*^{+/-} + PBS, *green asterisk*: *Setd1a*^{+/-} + TCP vs *Setd1a*^{+/-} + PBS, *brown asterisk*: WT + TCP vs *Setd1a*^{+/-} + PBS). No significant difference between WT + PBS and *Setd1a*^{+/-} + TCP groups was observed (*Setd1a*^{+/-} + PBS: N = 20, *Setd1a*^{+/-} + TCP: N = 20, WT + PBS: N = 20, WT + TCP: N = 20).

(J) Rescue of WM performance in a DNMT1 task in *Setd1a*^{+/-} mice treated with ORY-1001. Mean percentage of correct responses for WT and *Setd1a*^{+/-} mice with/without 10 μg/kg/day ORY-1001 or vehicle treatment are depicted. **P* < 0.05, ***P* < 0.01, ****P* < 0.001 (*blue asterisk*: WT + PBS vs *Setd1a*^{+/-} + PBS, *green asterisk*: *Setd1a*^{+/-} + ORY-1001 vs *Setd1a*^{+/-} + PBS, *brown asterisk*: WT + ORY-1001 vs *Setd1a*^{+/-} + PBS, *black asterisk*: WT + ORY-1001 vs WT + PBS). No significant difference between WT + PBS and *Setd1a*^{+/-} + ORY-1001 was observed (*Setd1a*^{+/-} + PBS: N = 21, *Setd1a*^{+/-} + ORY-1001: N = 15, WT + PBS: N = 23, WT + ORY-1001: N = 15).

Histograms show means and error bars s.e.m.

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|--|--------------------------|------------------|
| Antibodies | | |
| Rabbit hSETD1A | Bethyl | Cat# A300-289 |
| Rabbit Histone H3 | Abcam | Cat# ab1791 |
| Rabbit Mef2c/a | Abcam | Cat# ab64644 |
| Mouse NeuN | Millipore | Cat# MAB377 |
| Rabbit NeuN | Millipore | Cat# ABN78 |
| Mouse Tau-1 | Millipore | Cat# MAB3420 |
| Rabbit Tbr2 | Abcam | Cat# AB183991 |
| Rabbit GFP | Invitrogen | Cat# A-11122 |
| Mouse Parvalbumin | Swant | Cat# 235 |
| Mouse GFAP | Millipore | Cat# MAB360 |
| Rat PH-3 | Abcam | Cat# AB10543 |
| Rabbit RFP | Rockland | Cat# 600-401-379 |
| Rabbit VIP | Immunostar | Cat# 20077 |
| Mouse Reelin | Millipore | Cat# MAB5364 |
| Mouse SST | Millipore | Cat# MAB354 |
| Rabbit hSET1 | Abcam | Cat# ab70378 |
| Rabbit Mef2 | SantaCruz Biotechnology | Cat# sc-313 |
| Rabbit H3K4me1 | Abcam | Cat# ab8895 |
| Mouse H3K4me2 | Millipore | Cat# 05-1338 |
| Rabbit H3K4me3 | Millipore | Cat# 04-745 |
| Rabbit H3K27ac | Active Motif | Cat# 39133 |
| Rabbit LSD1 | Abcam | Cat# ab17721 |
| Mouse anti-rabbit IgG light chain (HRP) | Abcam | Cat# ab99697 |
| Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor 488 | Thermo Fisher Scientific | Cat# A-11001 |
| Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor 568 | Thermo Fisher Scientific | Cat# A-11004 |
| Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 488 | Thermo Fisher Scientific | Cat# A-11008 |
| Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 568 | Thermo Fisher Scientific | Cat# A-11011 |
| Goat anti-Rat IgG (H+L) Secondary Antibody, Alexa Fluor 488 | Thermo Fisher Scientific | Cat# A-11006 |
| Goat Anti-Rabbit IgG (H+L)-HRP Conjugate | Biorad | Cat# 1706515 |
| Normal Rabbit Serum | Jackson Immuno Research | Cat# 011-000-120 |
| Bacterial and Virus Strains | | |
| rAAV2/hSynapsin/EGFP | Addgene | Cat# 50465-AAV2 |
| Chemicals, Peptides, and Recombinant Proteins | | |
| Lipofectamine 3000 | ThermoFisher Scientific | Cat# L3000008 |

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|--|---|---|
| Lipofectamine 2000 | ThermoFisher Scientific | Cat# 11668027 |
| RN-1 | Tocris | Cat# 10010494 |
| PBIT | Sigma | Cat# SML1058 |
| IOX-1 | Sigma | Cat# SML0067 |
| tranylcypromine | Cayman Chemical | Cat# 10010494 |
| ORY-1001 | Cayman Chemical | Cat# 19136 |
| SP2509 | Cayman Chemical | Cat# 19136 |
| Critical Commercial Assays | | |
| Ovation Ultralow System V2 1-16 | NuGEN technologies | Cat# 0344NB |
| Papain dissociation kit | Worthington | Cat# LK003150 |
| Dual-Glo® Luciferase Assay System | Promega | Cat# E2920 |
| Nuclear Complex Co-IP Kit | Active Motif. Inc. | Cat# 54001 |
| Deposited Data | | |
| ChIP-seq | This paper | GEO: GSE123652 |
| RNA-seq | This paper | GEO: GSE123652 |
| scRNA-Seq | This paper | GEO: GSE123652 |
| ATAC-Seq Emx-CRE and Parv-CRE | This paper | GEO: GSE123652 |
| Deposited Data used in the paper | | |
| ENCODE H3K27ac ChIP | Shen et al., 2012 | GEO: GSE29184 |
| ATAC-Seq human brain | Fullard et al., 2018 | GEO: GSE96949 |
| GWAS | Welter et al., 2014 | http://www.ebi.ac.uk/gwas , the gwas_catalog_v1.0.1 file, accessed June 2016 |
| SCZ GWAS | Schizophrenia Working Group of the Psychiatric Genomics, 2014 | https://www.med.unc.edu/pgc/files/resultfiles/scz2.regions.zip |
| LoF DNMs | Turner et al., 2017 | http://denovo-db.gs.washington.edu/denovo-db/ |
| Experimental Models: Cell Lines | | |
| HEK293T | ATCC | Cat# CRL-3216 |
| Experimental Models: Organisms/Strains | | |
| <i>Setd1a^{tm1a(EUCOMM)Wsi}</i> | EMMA | Cat# EM:06857 |
| Emx1 tm1(cre)Krf/J | Jackson Laboratory | Cat# 005628 |
| Pvalb tm1(cre)Arbr/J | Jackson Laboratory | Cat# 017320 |
| Gt(ROSA)26Sortm14(CAG-tdTomato)Hze/J | Jackson Laboratory | Cat# 007914 |
| C57BL/6J | Jackson Laboratory | Cat# 000664 |
| B6N.129S6(Cg)-Gt(ROSA)26Sortm3(CAG-flpo/ERT2)Alf/J | Jackson Laboratory | Cat# 019016 |
| B6.Cg-Gt(ROSA)26Sortm1.3(CAG-tdTomato,-EGFP)Pjen/J | Jackson Laboratory | Cat# 026932 |
| Oligonucleotides | | |
| Setd1a_F: GGTTATTGATCTGGGCAGGC | Sigma | N/A |
| Setd1a_R: TGACCTGTTTTCAAGCCCTC | Sigma | N/A |

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|--|---|
| CAS_R1_Term: TCGTGGTATCGTTATGCGCC | Sigma | N/A |
| Recombinant DNA | | |
| 3X Mef2-Luciferase vector | Addgene | Cat# 32967 |
| Mef2c CDS vector | Addgene | Cat# 32515 |
| Setd1a CDS vector | Origene | Cat# MR215352 |
| LSD1 CDS vector | Origene | Cat# MR210741 |
| pCAG-GFP vector | Addgene | Cat# 11150 |
| Software and Algorithms | | |
| Megablast | McGinnis and Madden, 2004) | https://blast.ncbi.nlm.nih.gov/Blast.cgi |
| R v3.6 | R Core Team | http://www.gbif.org/resource/81287 |
| deeptools v2.5.0.1 | Ramirez et al., 2016 | https://deeptools.readthedocs.io/en/develop/ |
| <i>Mus musculus</i> genome version mm10 | N/A | N/A |
| <i>Homo sapiens</i> genome version hg38 | N/A | N/A |
| HOMER v4.8.2 | Heinz et al., 2010 | http://homer.salk.edu/homer/ |
| IDR 2.0.2 | Li et al., 2011 | https://github.com/nboley/idr |
| MACS v2.1.1 | Zhang et al. 2008 | https://github.com/taoliu/MACS |
| samtools v1.3.1 | Li et al., 2009 | http://samtools.sourceforge.net/ |
| DESeq2 1.10.2 | Love et al., 2014 | https://bioconductor.org/packages/release/bioc/html/DESeq2.html |
| STAR 2.5.4b | Dobin et al., 2013 | https://github.com/alexdobin/STAR |
| bowtie2 v2.2.6 | Langmead and Salzberg, 2012 | http://bowtie-bio.sourceforge.net/bowtie2/manual.shtml |
| Cell Ranger 2.1.1 | 10X Genomics | https://support.10xgenomics.com/ |
| Seurat v. 3.0.1 | Stuart*, Butler* et al., bioRxiv, 2018 | https://satijalab.org/seurat/faq.html |
| Regioner | Gel et al., 2016 | https://www.bioconductor.org/packages/release/bioc/html/regioner.html |
| Inrich | Lee et al., 2012 | https://atgu.mgh.harvard.edu/inrich/ |
| Revigo | Supek et al., 2011 | http://revigo.irb.hr/ |
| Toppgene | Chen et al 2009 | https://toppgene.cchmc.org/ |
| DNENRICH | Fromer et al., 2014 | https://psychgen.u.hpc.mssm.edu/dnenrich/ |
| CSEA tool | Dougherty et al., 2010 | http://genetics.wustl.edu/jdlab/csea-tool-2/ |
| PAVIS | Huang et al., 2013 | https://manticore.niehs.nih.gov/pavis2/ |
| Fiji | Schindelin et al., 2012 | https://imagej.nih.gov/ij/download.html ; |
| LSM Image Browser | ZEISS | https://www.embl.de/eamnet/html/body_image_browser.html |
| pClamp 10.3 | Molecular Devices | https://www.moleculardevices.com/systems |
| Microsoft Excel 2013 | Microsoft | https://products.office.com/en-us/excel |
| SigmaPlot 9.0 | Systat Software Inc | http://www.sigmaplot.co.uk/products/sigmaplot/sigmaplot-details.php |
| GraphPad Prism | GraphPad software | https://www.graphpad.com/scientific-software/prism/ |

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|----------------------|-----------|---|
| SPSS Statistics 20.0 | IBM Corp. | https://www-01.ibm.com/support/docview.wss?uid=swg24029274 |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript